ABSTRACT

Title of Thesis: FORECASTING ODOR LEVELS FOR BIOSOLIDS PRODUCT BASED ON AMBIENT CONDITIONS

Sirapong Vilalai, Master of Science, 2003

Thesis directed by: Assistant Professor Steven A. Gabriel
Department of Civil and Environmental Engineering

Biosolids are the solid residuals of wastewater treatment that have been treated to reduce pathogens and can be used as fertilizer for agricultural purposes. The major disposal option of biosolids from wastewater treatment plants is land application. Since biosolids odor quality is inconsistent, there is a possibility that odorous biosolids can be distributed to field sites nearby residential areas. The objective of this thesis was to create statistical models using the processing data and ambient conditions at Blue Plains wastewater treatment plant to predict biosolids odor levels. The model shows that $FeCl_3$ addition in the primary process, lime addition, the number of centrifuges out of service and the temperature at the plant are parameters that can be used to predict biosolids odor.

FORECASTING ODOR LEVELS FOR BIOSOLIDS PRODUCT BASED ON

AMBIENT CONDITIONS

by

Sirapong Vilalai

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2003

Advisory Committee:

Assistant Professor Steven A. Gabriel, Chair
Associate Professor Eric A. Seagren
Professor Gregory B. Baecher

DEDICATION

To my parents and family, for their unconditional love, guidance, and support.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Steven A. Gabriel, my advisor, for his kindly guidance thorough the development of this thesis. His technical and editorial advice was essential to the completion of this thesis.

I would like to thank Royal Thai Army for giving me a great opportunity to study master degree in the US and the financial support during my study.

I would like to express my sincere thanks to Chris Peot who gave me an opportunity to work as an intern in DCWASA, permission to use DCWASA data, and useful discussions thorough the development of this thesis. I would like to thank Mark Ramirez and Dorian Tolbert from DCWASA for their advices and explanation about DCWASA process.

I am grateful to Dr. Gregory B. Baecher and Dr. Eric A. Seagren for agreeing to be my committee members, and for reviewing this thesis.

Lastly, thanks Kim Hyunah for encouragement and support.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

This chapter provides an overview of biosolids, a background on the District of Columbia Water and Sewer Authority's Blue Plains wastewater treatment plant, problems under review, and the research objectives.

1.1 Overview of Biosolids in the US

Biosolids are the nutrient-rich production from wastewater treatment plants that can be used as fertilizer for agricultural purposes. Unlike the raw sludge that is the byproduct of the processes used to treat wastewater before discharge to the receiving water bodies, biosolids have been treated to reduce pathogens, and odor. With Regard to the pathogens in biosolids, biosolids can be categorized into two groups: Class A and Class B (Evanylo, 2001). Class A level means that biosolids have passed additional treatment (e.g., high temperature and high pH) and have below-detectable pathogens levels. Class A biosolids can be applied to the land directly without any field restrictions and also can be sold as fertilizer in the market place. Compared to Class A, Class B biosolids have been treated to reduce pathogens so that they will not affect people and the environment but for which field restrictions need to be considered before application.

In the early 1900s, the common way for a wastewater treatment plant to get rid of its sludge was to dump it into oceans, lakes, and rivers (Metcalf & Eddy Inc., 1991). Since the US population in many cities has increased, the environmental impact from wastewater facilities has become much more significant. The growing of algae and the death of animals in the rivers have provoked state and federal standards to be applied to municipal wastewater facilities, such as the limitation of compounds of phosphorous, ammonia and nitrogen, deposited in the rivers. As the result, wastewater treatment plants have to apply more sophisticated treatments to reduce the effluent chemical compounds to meet these regulations. In the US, advanced treatment processes, such as treatment by waste activated sludge, nitrification, and denitrification treatments have been implemented to reduce the chemicals deposited into rivers.

In addition to the purification of effluents from the plants, the disposal options of sludge present another serious problem. Compared to waste liquids that can be discharged directly into rivers after meeting specifications, sludge disposal requires the availability of space and state or federal permission, such as the distance from surface, ground water, or public area. Since the ban of dumping sludge into the ocean (Metcalf & Eddy Inc., 1991), there are three sludge disposal options available: landfills, incineration, and land application (biosolids only). We briefly discuss these options in what follows.

For landfill and surface disposal, sludge is dumped and compressed into the specific area, such as strip mines (Evanylo, 2001). This option is economically attractive since it needs less advanced treatment. One needs to prepare synthetic liners in the landfill space to prevent excessive nutrient leakage into the ground or surface water. For surface disposal, this option needs to leave the land undeveloped at least two years after dumping. Nowadays, it is increasingly difficult to find available space for landfills.

The second option is incineration. Incineration reduces the sludge volume, kills pathogens, and uses less land space than other options (Evanylo, 2001). However, to dispose sludge by incineration, one needs to install special ovens that can produce temperatures higher than 450°F as well as install special equipment to prevent the spreading of ash. Incineration also leaves heavy-metal particles in the oven that may contaminate the air after burning.

Up to now, given the issues of availability and suitability of space for landfills and utilization of nutrients in biosolids, land application seems to be the most appropriate option for biosolids. The benefit of biosolids to agriculture is that farmers can spread the treated biosolids on the farms as fertilizer to help grow the crops as well as reduce fertilizer costs. Biosolids also improve the soil properties in terms of nutrient addition, infiltration, soil compaction, and the ability to retain nutrients. For the benefit of wastewater treatment plants, this approval solves the space-availability problem by delivering biosolids to farms instead of landfill sites.

As this thesis is focused on wastewater treatment and biosolids management by The District of Columbia Water and Sewer Authority (DCWASA), the following sections in this research are related to wastewater treatment and biosolids management for DCWASA.

1.2 Background on the District of Columbia Water and Sewer Authority

The District of Columbia Water and Sewer Authority's Blue Plains treatment plant is the largest advanced wastewater treatment plant in the world. Located in southwest Washington DC, Blue Plains provides wastewater treatment service for more than 2 million users in Washington DC, Montgomery and Prince George's county in Maryland, and Fairfax and Loudoun counties in Virginia. Blue Plains treats more than 370 million gallons of wastewater per day (DCWASA, 2003).

Blue Plains was originally constructed in 1938. From 1938 to 1996, it was known as the District of Columbia Water and Sewer Utility Administration (DCWASA, 2003). The treatment process from 1938 to 1959 composed of screening and primary treatment. The plant applied secondary treatment in 1959 and advanced treatment in 1983. As the population and industry in the DC area increased, the funding from the DC government, which was the income source to operate Blue Plains, was insufficient to maintain the appropriate operating standards, maintenance, and to establish new advanced treatment. On April 18, 1996, the DC government and the US federal government agreed to establish the District of Columbia Water and Sewer Authority as a multi-jurisdictional regional utility. As a result, DCWASA can manage the wastewater fee by itself to improve the infrastructure, the wastewater facilities and wastewater treatment systems at Blue Plains.

The wastewater collecting systems in Washington DC are comprised of the Combined Sewer Systems (CSS) and the Separate Sewer System (SSS). The CSS constitutes about one third of wastewater collecting systems, which was constructed before 1900. This system combines the wastewater from sanitary waste and storm water into the same line. Then, wastewater is sent to DCWASA for treatment during normal weather conditions. When there is a storm in Washington DC and the overflow reaches up to a certain level, the operator will discharge the overflow from the combined system directly into the Potomac River, Anacostia River, and Rock Creek to prevent flooding in city. As a result, the domestic and industrial waste has contaminated the river during storm weather.

The SSS was constructed after 1900 and constitutes about two thirds of the DC wastewater system. The SSS separates wastewater from sanitary sewage from home and industry in one line, and storm water to another line. The sanitary waste is treated in DCWASA and the storm water is discharged directly into rivers. This system reduces the contamination of wastewater into rivers during stormy weather.

1.3 Overview of Wastewater Treatment Process

At the Blue Plains, the wastewater treatment process can be categorized into two parts: the liquid treatment train and the solid treatment train. The liquid treatment train is the set of procedures that purify the wastewater before discharging to the river. The goal of this process is to limit the chemicals loaded into the Chesapeake Bay, such as organic matter, phosphorous, and nitrogen. The other part is the solid treatment train. This process focuses on treating the solids residuals resulting from the liquid treatment train. Biosolids also need to pass regulations concerning pH, pathogens, and odor before delivery to the field sites. The overview of the wastewater treatment process at Blue Plains is shown as follows.

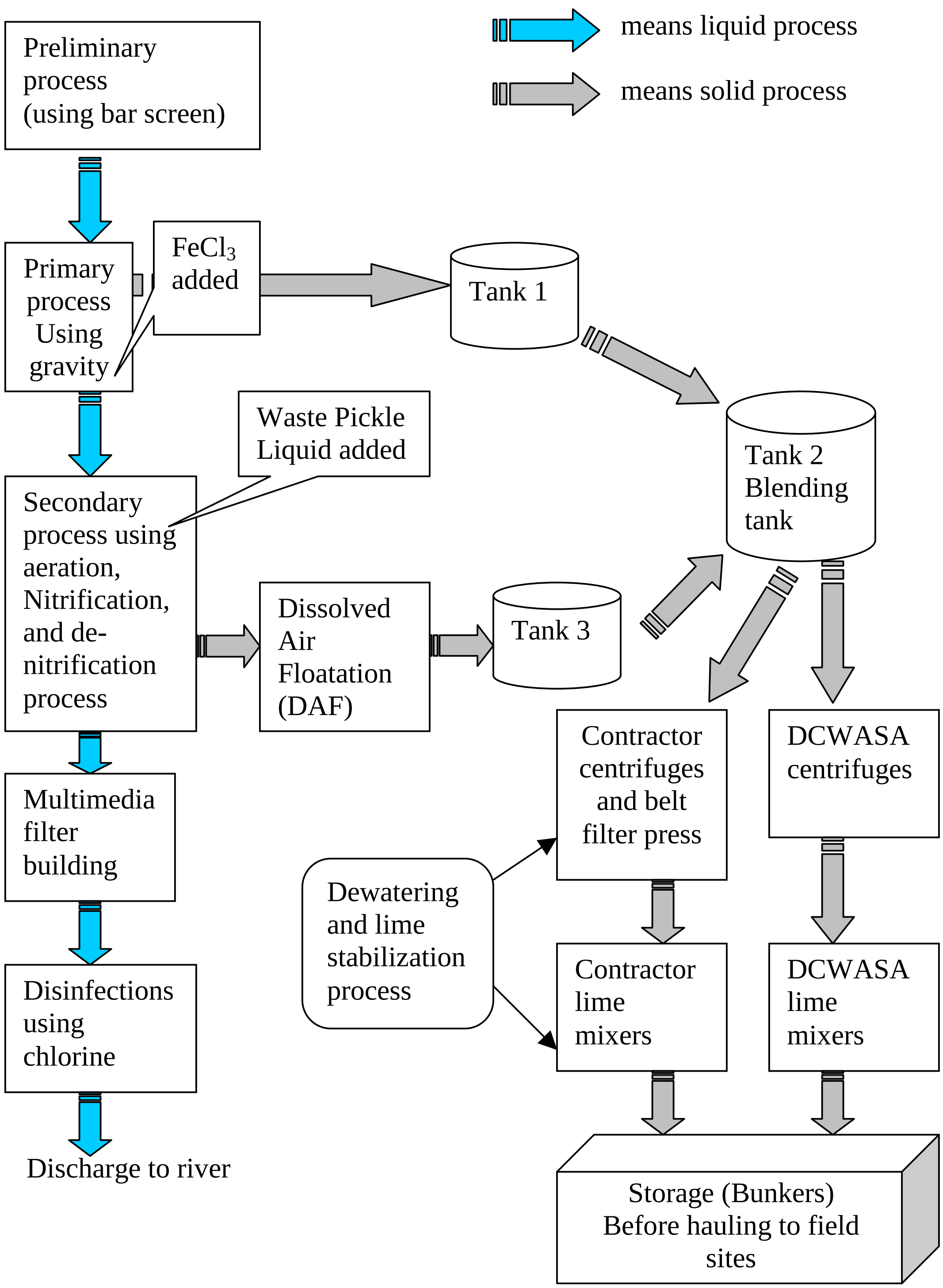


Figure 1.1: Overview of Wastewater Treatment Process (DCWASA, 2003)

1.3.1 Wastewater Treatment Process (DCWASA, 2003)

As this research concentrates on odor released from biosolids, the following paragraphs provide a summary of the wastewater treatment processes at Blue Plains, focused on the solids processes.

1. Preliminary process: The bar screen is the preliminary step to remove large debris, such as boards, branches, trash, or other large particles from the wastewater before passing through the system. The gap between each bar can remove the large particles from the flow. The purpose of this preliminary process is to detect the large particles that can damage the equipment, such as pumps or pipelines. The heavy particles detected by the bar screen will be hauled to the landfill.

After bar screen, the flow is sent to the aerated grit chamber to remove the grit from the wastewater. This process is important because it prevents the abrasion of equipment in the downstream from grit and increases the service life of equipment. The air causes a spiral velocity pattern to settling heavy particles down to the bottom of the tank while the lighter particles eventually are carried out of the tanks to the primary process (Vesilind, 2003). Grit removed from aeration grit chamber will be disposed to nearby landfills by truck.

2. Primary process: The raw wastewater from the preliminary process is pumped to the primary sedimentation tanks. The primary sedimentation tanks are used to slow down the influent and facilitate the settling of suspended solids. Using the gravity in tanks, the suspended solids settle to the bottom of the tanks and scum[1] (e.g. oil and greases.) floats to the surface of the tanks. Both scum and settling sludge are

[1] Scum is defined as the filmy layer that forms on the surface of water.

sent to the degritting and grinding facility to additionally remove grit from the primary sludge. In this process, 45% to 65% of suspended solids are removed from the flow.

3. Secondary process: The secondary aeration tanks form the biological process. Air is blown into in these tanks in order to supply oxygen for microorganisms in the aerobic biological processes. Microorganisms use the oxygen to convert organic wastewater constituents in the flow in the process that produces more microorganisms. The combination of wastewater and microorganisms in aeration tanks is called mixed liquor (Vesilind, 2003). Then, the wastewater from the aeration tanks flows to the secondary sedimentation tanks. In these tanks, a portion of settling sludge at the bottom of the tanks, which we call mixed liquor suspended solids (MLSS), will be pumped back to the aeration tanks to maintain a concentration of microorganisms in the aeration tanks (Vesilind, 2003) while the rest will be wasted and sent to Dissolved air floatation (DAF) thickeners.

After the secondary sedimentation tanks, wastewater flows to the nitrification and denitrification process for ammonia removal. Liquid is sent to multimedia filtration after this process to remove the remaining small particles. Then it passes through the disinfection step using chorine before being discharged into the Potomac river.

For the solids part, the waste activated sludge settles down in the nitrification sedimentation tanks. Some parts are returned to the nitrification reactor (return sludge) while the other part is wasted to the DAF thickeners.

4. DAF thickeners: Sludge from secondary settling tanks, nitrification and de-nitrification settling tanks, which are all activated sludge, is first thickened using dissolved air floatation thickeners. To increase the solids content of sludge, polymer is added to combine all small particles while the air is blown to float these particles to the surface. Then, the chain pad removes all the floating sludge on the surface of tanks and sends to the sludge blending tank system.

5. Blending tanks system: In the blending tank system, sludge from primary and floatation thickeners is mixed together in the blend tank. The primary sludge and waste activated sludge are first stored separately. To discharge sludge from both tanks, the operators calculate the appropriate blend ratio using the influent flow into both tanks.

6. Dewatering and lime stabilization process: After the blending tanks system, the blended sludge from the mixing tank is then dewatered by centrifuges and then lime is added to increase the pH and kill pathogens before sending the sludge to the bunker. Lastly, biosolids will be hauled by truck to the various field sites.

1.4 Problem under Review

Much research in environmental engineering has been conducted in order to study the cause of biosolids odors. At DCWASA, research has been conducted that involve collecting samples from several parts of the plant and then applying laboratory procedures to analyze the characteristics of samples. For example, studies have investigated the amount of odor released if different amounts of lime are incorporated into the samples (Murthy, 2002a). Other studies have investigated the potential of odor released from the sample when the sample container is opened for 30 minutes, 24 hours, and seven days later. Such laboratory experiments allow DCWASA to better understand various wastewater treatment and biosolids production characteristics.

In addition to such laboratory experimentation, the biosolids supervisor can also obtain the actual processing data inside the plant from Processing Control History (PCH), the DCWASA online processing database, as well as the weather data at the plant to see the effect of these data on biosolids odor levels. Up to now, a statistical study of such factors on biosolids odor levels has not been performed.

Combining these aforementioned data sources, we can explore the interaction between the processing data and ambient conditions at the plant with the biosolids odor in the field. For instance, on a day for which there is a small amount of polymer and a small amount of lime added to the sludge, we could analyze the resulting biosolids odor at the field sites. In addition to processing data that operators can control, the ambient data can be taken into account for forecasting purposes. Ambient conditions such as the temperature at the plant could affect microbial activities inside the sludge.

Currently, biosolids management at DCWASA has considered the benefit of odor forecasting models. Even though the process at the plant already meets or exceeds the state and federal specifications, there were cases when the plant had delivered odorous biosolids to fields. In terms of operations, there are more than 10 processes working on biosolids production. All of them have already been evaluated to create products that satisfy environmental regulations. In reality, it is impossible to operate all processes at full capacity due to maintenance problems at Blue Plains. A forecasting model can be used to predict biosolids odors based on the inconsistency of the biosolids production.

1.5. Research Objectives

The objectives of this research are as follows:

1. To understand the interaction of processing variables and ambient conditions at the plant on the resulting biosolids odor levels by season of the year and as well as the entire year by using a correlation analysis.

2. To establish forecasting models to predict the odor levels associated with the processing variables and ambient conditions at the plant by using multiple regression analysis.

We expect that this research will be of benefit in three ways: biosolids management, quality of life to the residents nearby the field sites, and as a part of an optimization model of biosolids distribution for DCWASA currently under development of the University of Maryland (Gabriel and Sahakij, 2003).

For biosolids management, this research enables the biosolids supervisor to better understand the relationship between processing variables or weather variables at the plant and the biosolids odor levels. Using existing processing data, a supervisor could monitor and control the variables that cause odorous biosolids. In some cases, variables not directly controllable are involved, for example, the number of out-of-service mixers or the temperature. In these cases, the supervisor can notify the destination contractors in advance to schedule trucks in anticipation of odorous conditions.

Residents around the field sites will benefit from this biosolids management program in upcoming odor conditions. The biosolids application program is intentionally established to recycle the nutrient-rich products from the wastewater treatment plant for agricultural purposes for free. However, biosolids odors, especially when strong, are inconvenient to farm residents. Additionally, the strong odor on residential areas may be the cause of the opposition or even the ban of biosolids application in some areas. A statistical odor forecasting model will help DCWASA to more efficiently deliver biosolids products with odor taken into account. Thus, the farmers and nearby residents would have a positive attitude to biosolids application and apply biosolids for their agriculture as the purpose of biosolids application.

In future work, DCWASA plans in implementing a multiobjective model to control the biosolids processing and biosolids hauling assignment to reduce unnecessary transportation and chemicals costs while maintaining minimum odor levels. The odor-forecasting model from this research will play a significant role in development of this optimal distribution model (Gabriel and Sahakij, 2003).

The rest of this thesis is organized as follows. In Chapter 2, we discuss the data sets we used for our analysis: the processing data, the field data, and the weather data. In particular we describe variables and their range of values. In addition, we mention why these variables are important in our analysis. Chapter 3 focuses on the descriptive statistics of these variables. It starts by explaining how data were organized for the correlation analysis. Next, analyses using different periods for 2002 are performed. Chapter 4 discusses the multiple regression analyses that we ran. We start with the regression between inspector odor data and significant variables shown from the correlation analysis. Then, we explain several techniques applied to improve adjusted R squared values, such as including interaction variables, dummy variables, etc. Finally, we summarize the best regressions. In Chapter 5, we summarize our finding and describe how DCWASA can use the results for better management of the biosolids odor levels

Chapter 2 Databases

In this chapter we discuss the data set used in the statistical modeling. As shown in Figure 2.1, there are three types of data i.e., DCWASA processing data, field site data, and weather data. All of these data sets are believed to affect biosolids odor levels.

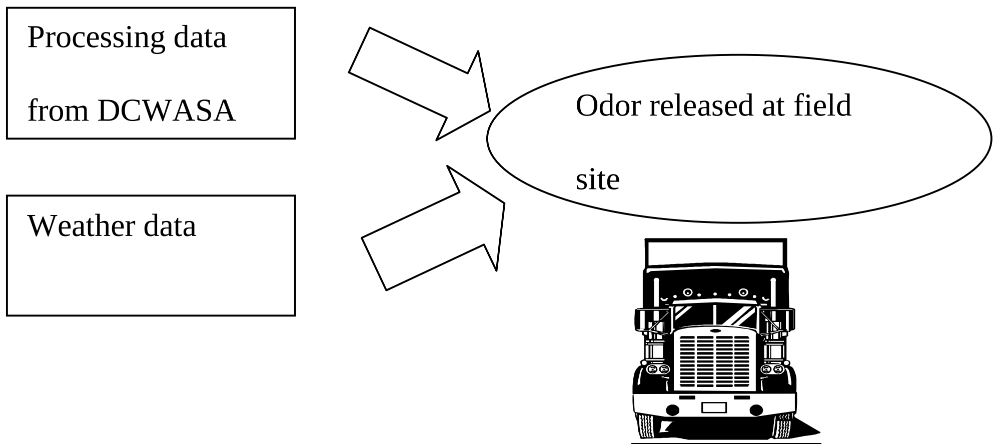


Figure 2.1: Database chart

2.1 Summary of All Data Used in This Research

This section is the summary of all data used in this thesis. It gives an overview of all data used by separating the data into three groups: DCWASA processing data, field data, and weather data. The relevant data within each group are listed as follows.

1. DCWASA processing data.

DCWASA processing data is the data set of all variables from DCWASA operating processes we expected to investigate relative to their reactions with biosolids odor levels. Those data we chose are the following.

1.1 Amount of $FeCl_3$ added (gallons)

- $FeCl_3$ added in primary east tanks
- $FeCl_3$ added in primary west tanks
- Sum of $FeCl_3$ added in primary east and primary west tanks.

1.2 Sludge Blanket depth in secondary settling tank (feet)

- Blanket depth in secondary east tanks
- Blanket depth in secondary west odd tanks
- Blanket depth in secondary west even tanks
- Sum of blanket depth in secondary west odd and west even tanks
- Sum of all blanket depth in secondary tanks

1.3 Waste Pickle liquid added to the secondary process (gallons)

1.4 Blend ratio in blend tank (%)

1.5 Polymer added (lbs of polymer/dry tons of sludge)

- Polymer added to Diffused Air Floatation process (DAF)
- Polymer added to dewatering process
- Sum of DAF polymer and dewatering polymer

1.6 Number of centrifuges in service and out of service

- Number of DCWASA centrifuges in and out of service
- Number of contractor centrifuges in service

- Number of contractor belt filter press in service.

1.7 Lime addition (lbs of lime per try tons of sludge)

2. Field data

Field data is the data set from farm sites that apply biosolids. Maryland Environmental Service (MES) sends inspectors to the field sites to collect the field data, such as temperature at the field, wind direction, wind velocity, biosolids odor levels, etc. The data we used in this thesis is:

2.1 Inspector's field odor data

3. Weather data

Weather data represent ambient conditions at the plant, such as:

3.1 Minimum temperature, minimum temperature, average temperature (F)

3.2 Snowfall (inches)

3.3 Precipitation (inches)

3.4 Average station pressure (inches of Hg)

3.5 Average daily wind speed (miles per hour)

Table 2.1 is an overview of how we organized all the data described above. All data were taken on a daily basis from 2002 and arranged into four parts: date that the variables were taken from, inspector's field odor data, the set of processing data, and the set of weather data.

Table 2.1: The Organization of All Data for Analysis

		Inspector field odor data	Processing data						Weather data				
Date	day	operator mike	blanket depth secondary east (ft.)	blanket depth secondary west odd (ft.)	blanket depth secondary west even (ft.)	Blanket depth:secondary west odd+secondary west even (ft.)	processing data no.5, 6, 7, ………		Maximum temp. (°F)	Minimum temp (°F)	Average temp. (°F)	weather data no.5, 6, 7, ………	
1/1/02	Tuesday	no data	2.3	2.6	3	5.6	x	x	34	20	27	x	x
1/2/02	Wednesday	5	2.6	2.6	3	5.6	x	x	37	21	29	x	x
1/3/02	Thursday	3.2727273	3.3	3.3	3.4	6.7	x	x	36	30	33	x	x
1/4/02	Friday	6	3.1	3.1	3.5	6.6	x	x	41	27	34	x	x
1/5/02	Saturday	no data	3.1	3.1	2.9	6	x	x	51	26	39	x	x
1/6/02	Sunday	no data	3.3	3.3	3.5	6.8	x	x	39	29	34	x	x
1/7/02	Monday	no data	3.7	3.9	4.7	8.6	x	x	39	30	35	x	x
1/8/02	Tuesday	3	4.1	3.3	3.4	6.7	x	x	40	28	34	x	x
1/9/02	Wednesday	3	3.4	3.2	3.1	6.3	x	xx	49	29	39	x	xx
1/10/02	Thursday	0	3.1	3.1	3.6	6.7	x	x	53	34	44	x	x
1/11/02	Friday	no data	2.8	3.5	3.4	6.9	x	x	51	39	45	x	x
1/12/02	Saturday	no data	2.5	3.1	3.6	6.7	x	x	51	30	41	x	x
1/13/02	Sunday	no data	2.8	2.7	3.7	6.4	x	x	49	36	43	x	x
1/14/02	Monday	5.5263158	2.5	3	3.9	6.9	x	x	51	32	42	x	x
1/15/02	Tuesday	6	2.1	3.1	3.9	7	x	xx	54	36	45	x	xx
x	x	x	x	x	x	x	x	x	x	x	x	x	x
x	x	x	x	x	x	x	x	x	x	x	x	x	x

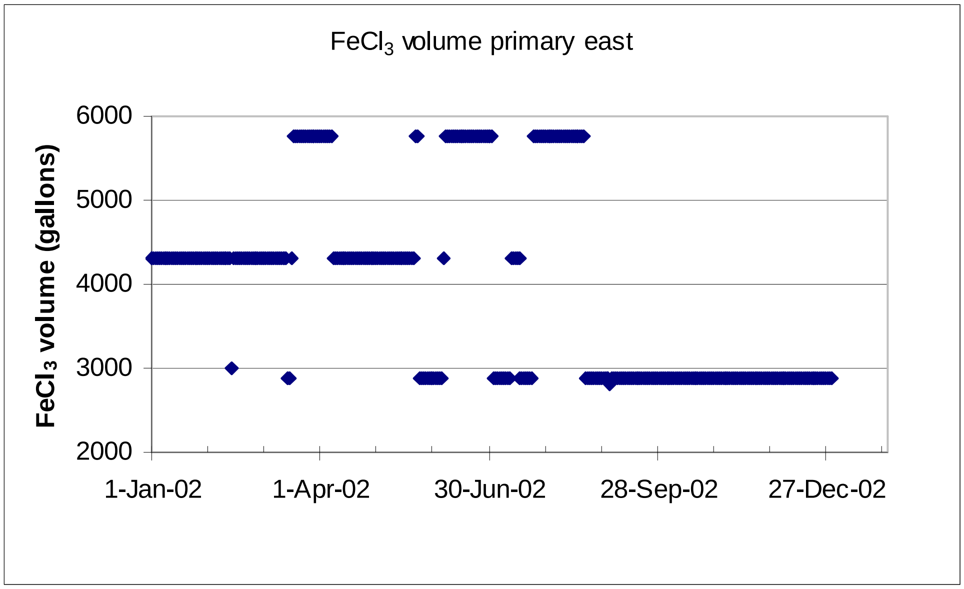
2.2 DCWASA Processing Data

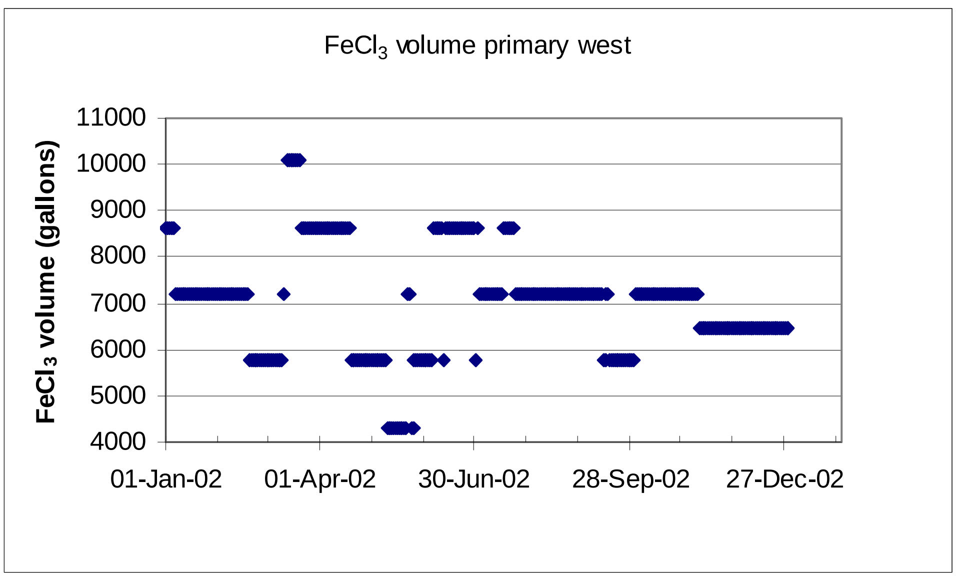
The processing data were expected to be important in affecting the odor released from biosolids. The odor from the sludge is a combination of several factors, such as the constituents of the sludge on that day, the chemical compounds added in the process, the biological activities of microorganisms in the sludge during the liquid and solid phases, the retention time in each phase, and the efficiency of the equipment in the process, such as centrifuges in the dewatering process. For many years, DCWASA has been researching the causes of odor released in wastewater treatment in terms of biological activities and chemicals added. For example, the research has shown that polymer causes the release of trimethylamine (TMA), which causes fishy odors, from biosolids during the lime stabilization process (Murthy, 2001). This research has provided us a reference point in selecting the appropriate parameters for our analysis. In what follows, we describe in more detail each of the processing variables considered.

2.2.1 Amount of $FeCl_3$ Added (gallons)

$FeCl_3$ is a chemical added to the primary settling process to remove phosphorous in wastewater. Phosphorous removal is an important treatment because it can stimulate algae growth in the Chesapeake River. Thus, $FeCl_3$ is added for the precipitation of phosphorous from wastewater (Metcalf & Eddy Inc., 1991). In DCWASA, three $FeCl_3$ variables were available: $FeCl_3$ added to the primary east tanks, $FeCl_3$ added to the primary west tanks and the sum of $FeCl_3$ from the east tanks and the west tanks. Based on decisions with DCWASA staff, these variables were included in the analysis of the correlation between processing data and biosolids odor since it is a major chemical added into primary process (Ramirez, 2003).

Figures 2.2 and 2.3 show the graphs of $FeCl_3$ added to the primary east and west tanks in 2002. When we compared the volume of $FeCl_3$ used from both tanks, we found that the volume of $FeCl_3$ added in the primary west tanks was higher than the east tanks according to the greater number of tanks in the primary west. In 2002, the range of $FeCl_3$ in the primary east tanks was from 2000 to 6000 gallons and the range of $FeCl_3$ in the primary west tanks was from 4000 to 11000 gallons. No study has noted the relationship between $FeCl_3$ added in primary wastewater treatment and biosolids odor levels. Thus, the relationship between $FeCl_3$ addition and odor levels needs to be statistically investigated by using the correlation analysis between the $FeCl_3$ data and the inspector's field odor.


FeCl3 volume primary east
FeCl3 volume (gallons)
6000
5000
4000
3000
2000
1-Jan-02
1-Apr-02
30-Jun-02
28-Sep-02
27-Dec-02


FeCl3 volume primary west
FeCl3 volume (gallons)
11000
10000
9000
8000
7000
6000
5000
4000
01-Jan-02
01-Apr-02
30-Jun-02
28-Sep-02
27-Dec-02

2.2.2 Sludge Blanket Depth in Secondary Settling Tank (feet)

In the secondary treatment, there are three sets of secondary settling tanks: blanket depth secondary west even tanks, blanket depth secondary west odd tanks, and blanket depth secondary east tanks. To use blanket depth in our analysis, we created two more blanket depth variables: first the sum of blanket depth west even and west odd tanks and second the sum of all blanket depth from all tanks. The purpose of the secondary process is to use biological consumption from microorganisms in the aeration tanks to break down the organic matter in the tanks. After secondary aeration tanks, the wastewater is sent to the secondary settling tanks to separate suspended solids from wastewater.

Using gravity, the suspended solids in the settling tanks are settled to the bottom of the tanks. The depth of the suspended solids in this tank is called the blanket depth and can roughly determine the capacity of waste activated sludge in the tanks. At the bottom of the settling tanks, some of sludge is returned to the aeration tanks to balance microorganisms in biological processes (return activated sludge) while the remaining sludge (waste activated sludge) is sent to dissolved air floatation thickeners.

The higher the level of the secondary blanket depth, the greater the biosolids odors. The rationale is that the higher the blanket depth indicates the more retention time for the activated sludge in the settling tanks. The greater retention time of waste activated sludge influences the development of anaerobic conditions for microorganisms in the tank and reduces the oxidation reduction potential (ORP). For example, experiments at DCWASA have shown that the higher blanket depth, the lower the ORP and the greater the production of the sulfur compound (Peot, 2003).

At DCWASA, the level of blanket depth is determined by the balance of the amount of MLSS in aeration tanks (Peot, 2003) and then the excess sludge is wasted to DAF thickeners. However, an insufficient number of DCWASA centrifuges in service, the major problem of DCWASA's dewatering process (Peot, 2003), causes less dewatering capacity. Thus, DCWASA operators have to reduce the wasting of waste activated sludge from sedimentation tanks. The fewer of DCWASA centrifuges operating, the greater the retention time of activated sludge in the sedimentation tanks and the lower ORP in the tanks. As a result, the relationship between the number of centrifuges operating and blanket depth is one of the significant factors to determine biosolids odor.

The following are graphs showing the distribution of blanket depth for all of 2002. Each of the graphs shows the trend of high blanket depth in January and February after that the curve was lower down to May and June, and then climbing up again until it reached the peak point around November and December. According to Figures 2.4, 2.5, and 2.6, the blanket depth ranges vary from 1 foot to 5.5 feet and we consider the values of zero, which are from missing observations, as outliers.

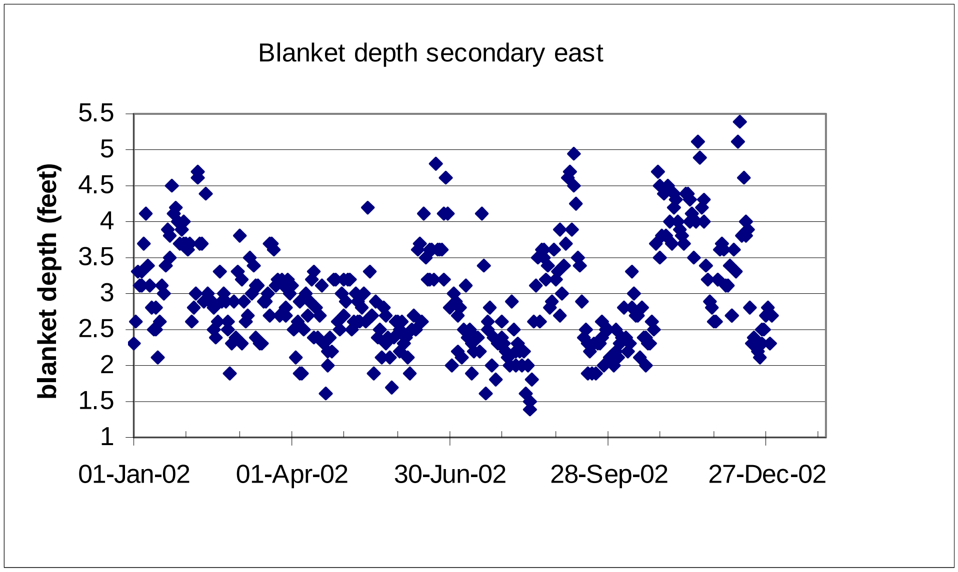


Figure 2.4: Blanket Depth Secondary East Graph

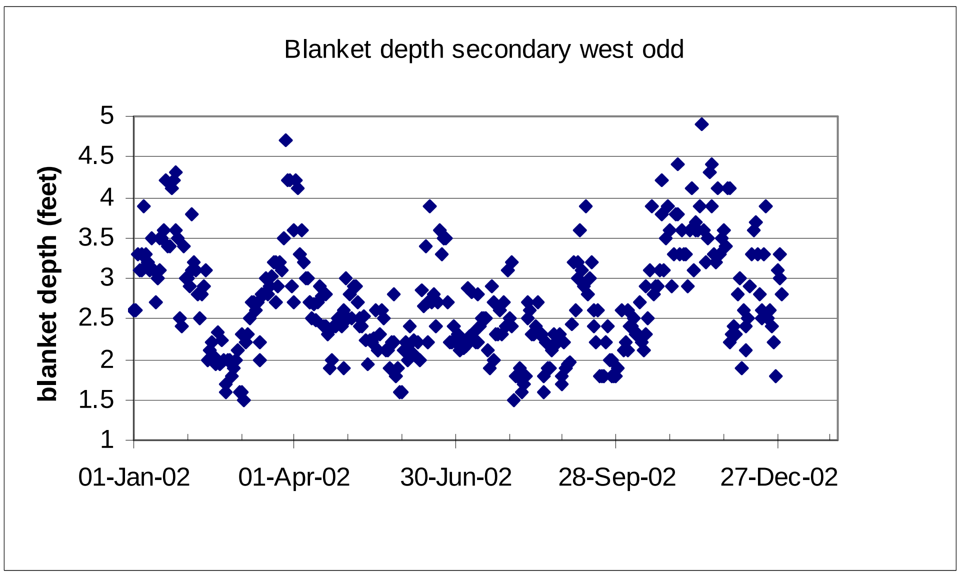


Figure 2.5: Blanket Depth Secondary West Odd Graph

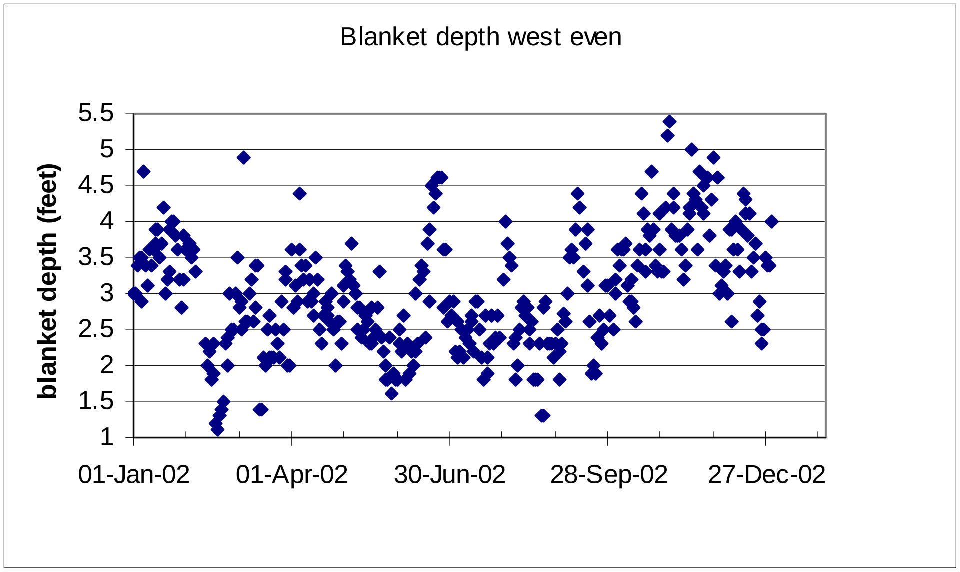


Figure 2.6: Blanket Depth Secondary West Even Graph

2.2.3 Waste Pickle Liquid Added to Secondary Process (gallons)

Waste Pickle Liquid (WPL), a waste product of steel finishing operations, is added to the process for additional removal of phosphorous from wastewater. Even though no research has shown how the WPL correlates with biosolids odor levels, this WPL variable has been added for analysis for the reason that it is a major chemical used in the secondary wastewater treatment process.

From Figure 2.7, the WPL added in the secondary tanks did not present any trends in the graph. The data seem to be randomly added. As a consequence, we added all data in our analysis except those with values of zero, which we treated as outliers.

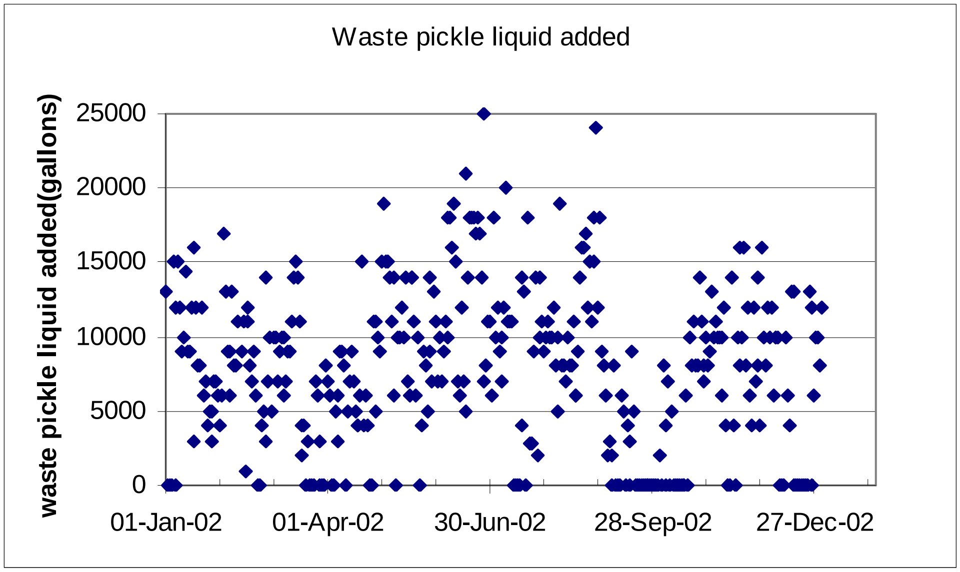


Figure 2.7: Waste Pickle Liquid Added Graph

2.2.4 Blend Ratio in Blend Tank (%)

The blend ratio is the ratio between primary sludge (sludge and scum from the primary process) and waste activated sludge (the combination of sludge from the secondary process, nitrification, and de-nitrification) that were pumped into the blending tank as shown in Figure 2.8.

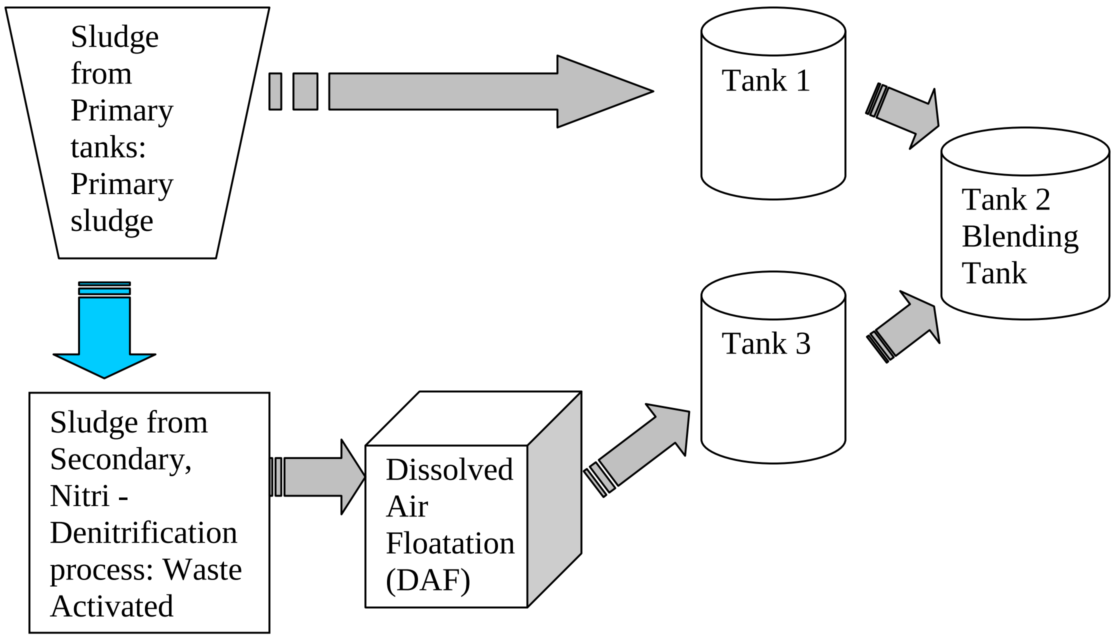


Figure 2.8: Overview of the Sludge Blending System (DCWASA, 2003)

We considered that the sludge and scum from the primary process were the food source for microorganisms in the waste activated sludge. The primary sludge is the highly condensed sludge that has been primarily taken out of the wastewater in primary sedimentation tanks. It contains a high volume of organic matter that can be used by waste activated sludge. In comparison, the waste activated sludge from the secondary, nitrification, and denitrification processes is composed of microorganisms, such as bacteria, and organic matter that can not be removed in the secondary process. Waste activated sludge is pumped from the DAF thickeners to the blend tank number 3. The primary sludge is pumped directly from the primary tanks to blend tank number 1. The blend ratio of blend tank (number 2) is determined by the calculation of influent into tanks number 1 and 3. The blended sludge tank system is the central point to combine all scum and sludge in wastewater system before dewatering and lime stabilization. The formulation of the blend ratio is shown below.

Blend ratio = Sludge pumped from Tank 1/ sludge pumped from Tank 3

The assumption for the blend ratio variable is that the higher the ratio the higher the level of odor. Since the blend tank ratio is the ratio of primary sludge (food) divided by the waste activated sludge (microorganisms), the higher the ratio means that microorganisms have a higher level of food relative to their population. When sludge from the high blend ratio tank is dewatered, it should present more microorganisms than the sludge from the low blend ratio tank. As a result, we assume that the high blend ratio causes more odors.

DCWASA biosolids advisors mentioned that the normal range of this ratio was from 0.3 to 0.70 (Tolbert, 2003). The data out of this range were identified as outliers. As we see from Table 2.9, the blend ratio tended to peak in June, July, August, and September. We can compare the influence of the blend ratio with odor levels in the hot months of the year (June, July, and August) with blend ratio in the cold months (January, February, and March) in Chapter3 as part of the correlation analysis section.

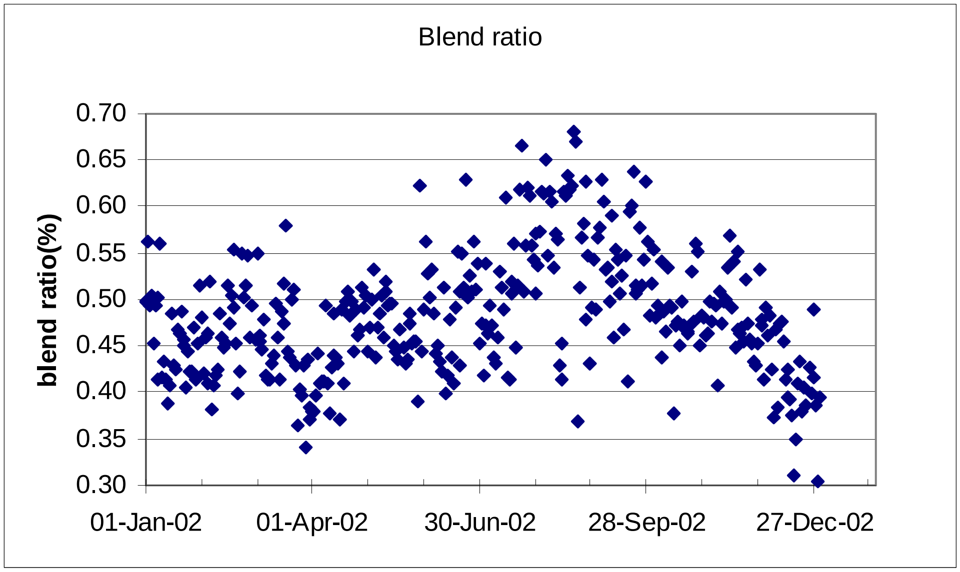


Figure 2.9: Blend Ratio Graph

2.2.5 Polymer Added (lbs/dry tons of sludge)

There are three types of polymer added that we used for our analysis: Dewatering polymer, dissolved air floatation (DAF) polymer and the combination of dewatering and DAF polymer. Polymer is added to both the centrifuge dewatering process and DAF in order to hold the sludge particles together to help the centrifuges to remove water from the sludge. Several studies have shown that polymer added into sludge is a main factor that increases biosolids odor, especially amine production e.g. TMA. For example, the increase of polymer will increase the amount of TMA from biosolids (Murthy, 2001). Also, DAF is expected to be the major source of TMA because of the microbial activity of the waste activated sludge and degradation of the polymer added in DAF resulting in the production of TMA (Kim, 2001a).

Accordingly, we expect that the more polymer added in the DAF and dewatering processes, the greater the odors would be released from biosolids. To use the polymer data in our analysis, biosolids supervisors at DCWASA provided us with the normal data range that they used. The normal range of dewatering polymer is from 10 to 30 lbs per dry ton and DAF polymer is from 80 to 300 lbs per dry ton (Tolbert, 2003). The data outside of this range was considered to be outliers.

From Figures 2.10 and 2.11, dewatering polymer added before the centrifuges had a range between 10-30 lbs per dry ton in 2002 while dewatering polymer added to DAF was between 50-200 lbs per dry ton.

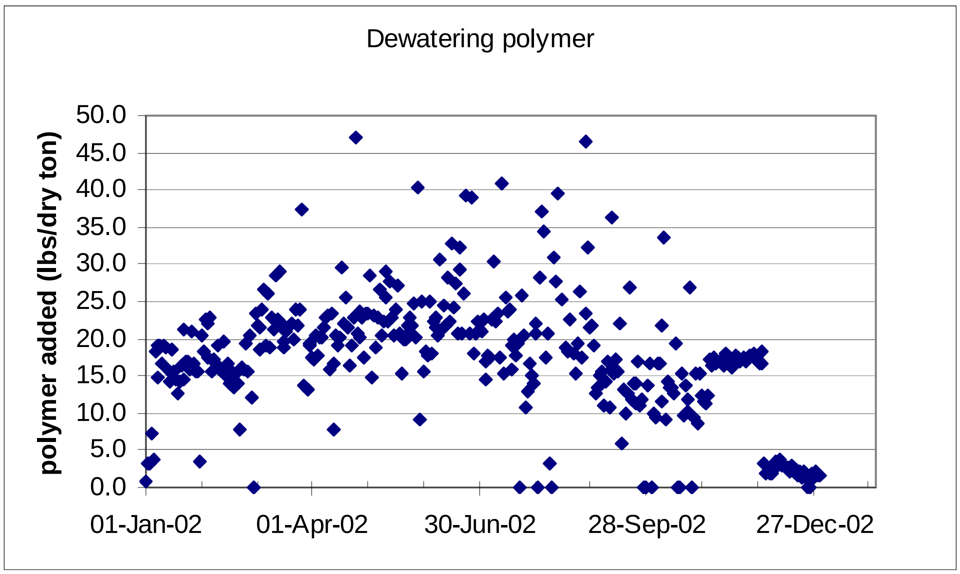


Figure 2.10: Dewatering Polymer for Centrifuges

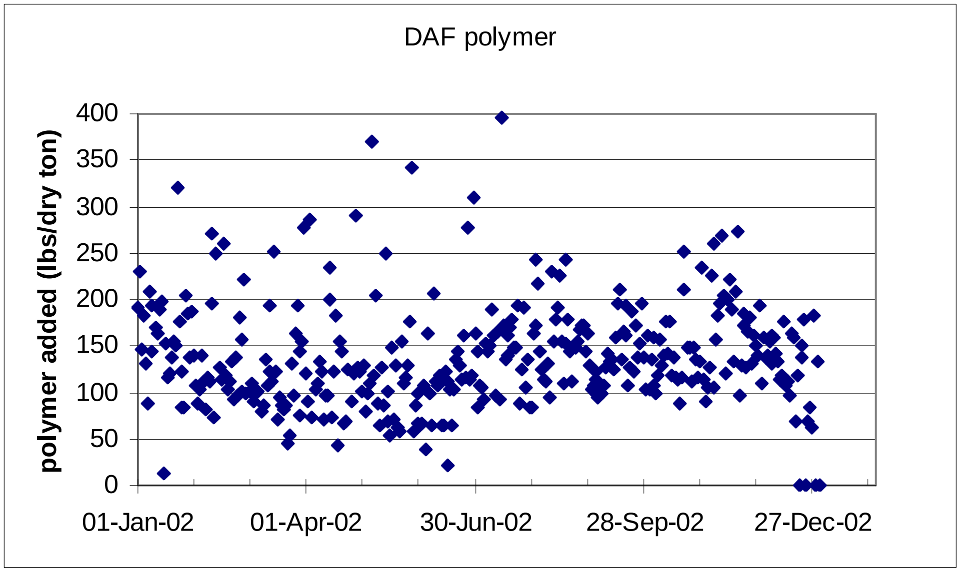


Figure 2.11: DAF Polymer

2.2.6 Number of Centrifuges in Service and Out of Service

The centrifuges in the wastewater treatment processes are used for removing water from the sludge. Sludge from the blend tank needs to have the water part removed to decrease the volume and weight before hauling it to the field in regards to cost and quality considerations. In DCWASA, there are two components of the dewatering process. One is dewatering by DCWASA centrifuges, the other is dewatering by contractor centrifuges and belt filter presses. The contractor assists DCWASA in the dewatering and lime stabilization processes in order to reduce the biosolids loading on DCWASA centrifuges and mixers. The biosolids supervisor decides the working load on the contractor side, at least 150 dry tons or more per day, depending on the working status of DCWASA centrifuges. There are seven centrifuges available in the DCWASA dewatering process and there are two centrifuges and seven belt filter presses available for the contractor.

As discussed in sludge blanket depth section about the relationship between blanket depth and number of centrifuges operating, the greater the number of centrifuges operating the less the retention time of activated sludge in the secondary sedimentation tanks as well as the retention time of blended sludge in the blend tank. Consequently, the greater number of centrifuges and belt filter presses in service the lower the biosolids odor levels.

2.2.7 Lime Addition (lbs per try tons of sludge)

Lime stabilization is the last treatment before biosolids are delivered to the bunker and hauled by truck to the sites. The purpose of lime stabilization is to increase the pH that kills the pathogens for use in agriculture. After lime stabilization, the pH and temperature of the biosolids are increased. Research has shown that decreasing the lime dose increased the biological activities and resulted releasing in reduced sulfur compounds (Murthy, 2001). Also, it has shown that by reducing the polymer addition and incorporating lime dose in the appropriate rate can make a stabilized biosolids with low off-site odor (Murthy, 2001). As a result, we expect the more lime added the less bacteria left in biosolids as long as the lime is well mixed into the sludge. Thus, the odor from biosolids will be decreased.

Table 2.12 shows the distribution of lime addition rates in 2002. The lime addition rate increased from January to the highest point in August. After this point the lime addition rate was lower until November and December. In 2002 the lime addition rates in the cold months (January, February, November, and December) were lower than in the hot months (May, June, July, and August)

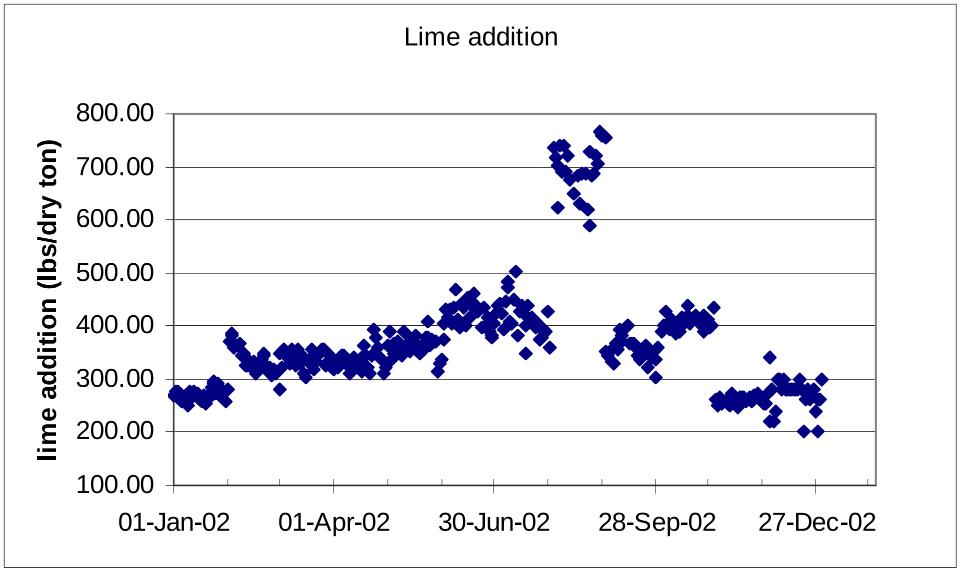


Figure 2.12: Lime Addition Graph

2.3 Field Data

Maryland Environmental Service (MES) provides field data for DCWASA arranged in Geographic Information System (GIS) format. DCWASA uses these data to monitor and update biosolids status for each field site on a monthly basis. To obtain the field data, MES assigns inspectors to field sites after receiving the hauling schedule from the contractors. At the sites, these inspectors: locate where the truck can spread biosolids based on state and federal regulations, and also they collect relevant data, such as: numbers of trucks coming in each day, wind speed, temperature, the odor level of biosolids on truck, etc. Finally, MES summarizes these data in a monthly report to DCWASA.

Table 2.2 is a sample of field data from MES. According to the truck assignment, in one day, there are many inspectors working on several sites in

Maryland and Virginia. Each inspector collected data on that site, such as weather conditions, amount of biosolids hauled biosolids odor, etc.

Table 2.2: The Sample of Field Data.

DateDeliver	**SiteName**	County	Tons of biosolids	Weather condition	Wind Speed (miles/hr)	Wind Direction	Temp Hi/Lo (° F)	Odor	Inspector
1/31/2002	SMITH DAIRY FARM	Caroline, VA	22.5	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.44	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	20.24	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	21.38	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	21.84	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.29	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.32	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.32	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.02	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	22.36	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.68	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.88	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.26	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.29	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.37	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	20.98	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	23.45	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	25.47	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	22.35	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	21.7	S	5	SE	66/55	SL	Carl Burton
1/31/2002	SMITH DAIRY FARM	Caroline, VA	24.85	S	5	SE	66/55	SL	Carl Burton
2/1/2002	SMITH DAIRY FARM	Caroline, VA	23.76	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	24.32	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	24.57	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	24.15	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	23.42	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	22.13	S	18	SW	79/48	N	Bob Heins
2/1/2002	SMITH DAIRY FARM	Caroline, VA	25.4	S	18	SW	79/48	N	Bob Heins

2.3.1 Inspector's field Odor Data

The inspector's field odor data is the odor score corresponding to the biosolids delivered to one particular site on that day. The inspector checks the biosolids odor on the truck and then records it in the field database. In 2002, there were eleven inspectors from MES assigned to field sites. Three or more inspectors were sent to various sites every day depending on the MES assignment.

To obtain the reading of biosolids odor on a particular day, we decided to use an average of all odor scores that one inspector observed on one day, no matter which site it was taken from, to represent the odor score of the biosolids from that inspector on that day. For example, on July 1, 2002, Inspector "Pete" had 20 data points from different field sites. To obtain the odor data of biosolids on July 1, we averaged all 20 data points to represent as Pete's field odor data on July 1. Otherwise, we would have had to track the odor score from each individual site. Since there are approximately 500 sites and given insufficient data for an entire year, this seems unlikely. The inspector odor score might seem to be subjective, however, the inspector's field odor data is the most applicable and capable odor data set possible.

In addition to biosolids odor levels from the inspectors, we tried to search for odor levels using odor-detection equipment, such as Jerome meters, H2S detecting equipment; most of these data were based on laboratory-based experiments. The data were incomplete according to our scope of time for one year. Moreover, odor data from the lab were focused on different types of chemical compounds from the odor. Some concentrated on the odor before lime stabilization, while other concentrated on the odor when they fixed the amount of lime added. Since the lab odor data were not consistent and insufficient, the inspector's field odor data was used in this study.

From the inspector's field, inspectors categorized the associated biosolids smell as "N" for negligible, "SL" for slight, "M" for medium, or "H" for high. To represent these data mathematically, we contacted MES about how to assign score for each odor level. MES used the following numerical values.

N= negligible (0)

SL= slight (3)

M= medium (6)

H= high (9)

Thus we assigned these numbers to be the score of the inspector's field odor data.

All of the inspectors' field odor data were taken by the inspectors who had more than three years of experience. Of the eleven inspectors, we chose six inspectors, Carl, Mike, Wilfred, Cheryl, Pete, and Patrick for our analysis because all of them had at least three of the four categories, such as 0, 3, and 6, corresponding to their field odor data. Only Mike recorded odors levels from all four categories, N, SL, M, and H in his data. We chose the inspectors in this way because we needed a range of values for the later statistical work.

2.4 Weather data

In addition to processing data, we also considered that weather data might have an influence on the biological activity on sludge, especially when extreme temperatures were considered. For instance, it would be interesting to see the reaction between chemicals added in biosolids and odor released when extreme weather conditions, such as snow, or high temperature were presented. Thus, we included weather data in our statistical analysis.

We obtained the weather data from the National Climate Data Center (NCDC) website, the world's largest archive of weather data (http://www.ncdc.noaa.gov/oa/ncdc.html.). We used the weather data from Reagan National Airport Station in DC, the closest station of NCDC to DCWASA, to represent the weather data at the plant. The parameters we used were:

1. Minimum, maximum, and average temperature (F)
2. Snow fall (inches)
3. Precipitation (inches)
4. Average station pressure (inches of Hg[2])
5. Average daily wind speed (miles per hour)

Our assumption about temperature variables (minimum, maximum, and average temperature) was that the higher the temperature, the higher the microbial activity and the higher the biosolids odor. For other variables, we could not make any assumption on their reaction to the biosolids odor in this point but we will use the correlation analysis's result related to these variables to explain their relationship to the biosolids odor.

In Chapter 3, we develop a correlation analysis to understand the role of processing data and weather data on odor levels. In addition, we explain the use of lagged effects for the variables.

---

[2] Hg is the chemical formula for mercury.

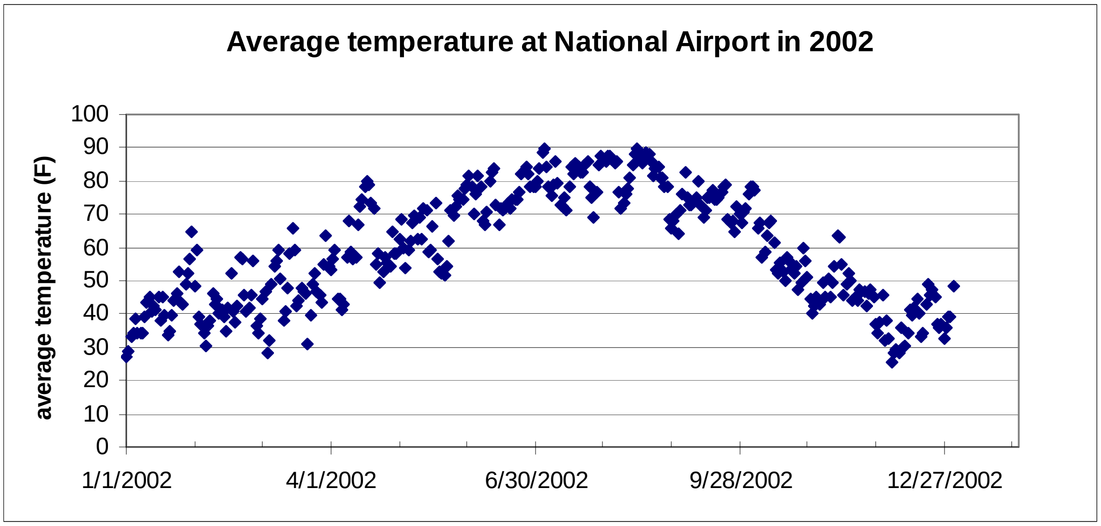
Chapter 3 Descriptive Statistics

In this chapter, we discuss about: 1) the procedures to prepare the data set for correlation analysis: disaggregating the data from 2002 into three periods for analysis, selecting inspectors into each period, setting variables to examine the day lagged effects on odor levels, organizing the selected data into a table for correlation analysis, and 2) the correlation analysis results and discussion.

3.1 Disaggregating Data from 2002 into Three Periods for Analysis

Our assumption on odor levels from biosolids is that the odor levels are the consequence of the constituents in sludge, the treatment processes, and the ambient conditions at the plant. In terms of the temperature aspect, we assume that temperature has significant effects on odor levels, especially when extreme temperature (high or low) is presented. In such cases, the temperature causes some effects on the odor levels by affecting the reactions of the microbial activities. The increasing or decreasing of temperature can change the microbial activity and oxidation reduction potential (ORP) of sludge that can increase or decrease odor as well (Murthy, 2002b). On a high temperature day, the high temperature at the plant obviously adds an additional heat to the process. Research has shown that the changing of oxidation's condition in sludge by the shift of season can affect the generation of reduced sulfur compound (Kim, 2001b). For this reason, temperature is one of the significant factors expected to increase odor levels.

We chose the average temperature for aggregating data from 2002. The reason we chose average temperature rather than maximum or minimum temperature was that the average temperature on any day is the parameter that most likely explains the condition of temperature on that day. Average temperature data from National airport was collected and graphed as shown in Figure 3.1.



In summary, we selected three periods for analysis.

1) The winter period, comprising January, February, and March.

2) The summer period, comprising, June, July, and August.

3) The yearly analysis using all days during 2002.

3.2 Selecting Inspectors into each Period

As discussed in Chapter 2, we chose six inspectors. Each of them had at least three of the four categories of inspectors' field odor. Those inspectors were Carl, Mike, Wilfred, Rob, Pete, and Patrick. In terms of the distribution of inspectors' odor data, many inspectors had scores just only in certain periods in 2002. For example, Wilfred's odor data were presented only from January to May 2002. Therefore, we needed to select appropriate inspectors for each of those 3 periods mentioned in Section 3.1 to produce adequate number of data points. Table 3.1 below shows the number of data points and the descriptions of inspectors' odor data for each period.

Table 3.1: The Description of Inspectors' Odor Data for Each Period

inspectors with 3 categories of odor scores		winter (January to March)			summer season (June-August)			entire year 2002		
inspector	data points in 2002	inspector	data points	description	inspector	data points	description	inspector	data points	description
Carl	150	Carl	45		Carl*	40	all data are "0"	Carl	149	
Mike	159	Mike	46		Mike	54		Mike	158	
Patrick	67	Patrick*	0	no data	Patrick*	0	no data	Patrick	34	
Pete	192	Pete	48		Pete	47		Pete	191	
Rob	156	Rob*	5	all scores are "3"	Rob	56		Rob	156	
Wilfred	41	Wilfred	32		Wilfred*	0	no data	Wilfred	51	

* Not included for analysis in that period.

The selections of inspectors for each period are described as follows:

For the winter season (January to March), Carl, Mike, Pete, and Wilfred's odor data were selected because they had variety of odor scores and also had enough data points (at least more than 30 data points) as shown in Table 3.1. Rob was ignored because of the insufficient data and insufficient variation of his data and Patrick was ignored because missing of data in this period.

For the summer season (June to August), Mike, Pete, and Rob's odor data were selected for the same reason as the winter season. For Carl, even though he had 40 data points, all of his data points had a score "0". Thus, his data were not included for this season.

For the yearly analysis, all inspectors were included. Even though there were a few data points in Patrick and Wilfred's odor data, it was thought that the analysis of their data regardless of season might provide some additional information for our analysis.

3.3 Setting Variables to Observe the Lagged Effects on Odor Levels

It is difficult to identify exactly how long it takes to process biosolids from the beginning state (preliminary process) to the finishing state where biosolids are delivered into the bunkers. In terms of odor levels on biosolids, it's more complicated to identify on which date the processing variables, such as lime addition affect the biosolids odor on the sample. For example, if we have odor data from Pete on February 14, 2002 equal slight, we would want to know when DCWSA added polymer, lime, or $FeCl_3$ into this sample of biosolids and how much DCWASA added. We already have DCWASA processing data in 2002 that provide the status of each processing variable on a daily basis but we can't locate the moment when these variables affects biosolids odor at the plant. In case of weather data, if we believe that the maximum temperature has an effect on biosolids odors, we would want to know how many days back maximum temperature caused biosolids odors on the sample date of February 14.

To prepare our processing and weather data in order to observe lagged effects on odor level, we used the day that biosolids' odor was collected as a reference of processing and weather data. We used the notation "d-i" to denote a variable lagged by i days (relative to odor score). Here i can be {0, 1, 2, 3, 4}.

We chose to analyze the data back from d-0 to d-4 because we believed that this period could capture all the events happened from the first state in the wastewater treatment until the trucks delivered biosolids to the fields. There are several factors that delay the biosolids production. For example, when a number of lime mixers or centrifuges malfunction, DCWASA reduces the biosolids production and assigns the processing work to the contractor who has less capacity. All of these factors contribute to the inconsistency of the process. Thus, using five lagged forms of each variable (d-0 to d-4) we believe that all important events related to biosolids odor can be captured.

Table 3.2 shows how each variable was set in order to capture the lagged effects on odor levels. For example, on January 10, 2002, inspector “Mike” average odor score taken from the field sites on this day was “0” and the maximum temperature (t max) at National airport was 53 ^{o} F. To observe how the odor levels on this day (d-0) were affected by t max on one, two, three, and four days before January 10, 2002, we move the entire d-0 t max column down one, two, three, and four rows according to the number of day lagged, i.e., on d-1 to d-4, respectively. Thus, on January 10, 2002 the t max on d-0 to d-4 relevant to Mike odor level on that day were 53, 49, 40, 39, and 39 respectively.

Table 3.2: How to Set the Lagged

			d-0	d-1	d-2	d-3	d-4
Date	day	mike	t max	t max	t max	t max	t max
1/1/02	Tuesday	no data	34	no data	no data	no data	no data
1/2/02	Wednesday	5	37	34	no data	no data	no data
1/3/02	Thursday	3.2727	36	37	34	no data	no data
1/4/02	Friday	6	41	36	37	34	no data
1/5/02	Saturday	no data	51	41	36	37	34
1/6/02	Sunday	no data	39	51	41	36	37
1/7/02	Monday	no data	39	39	51	41	36
1/8/02	Tuesday	3	40	39	39	51	41
1/9/02	Wednesday	3	49	40	39	39	51
1/10/02	Thursday	0	53	49	40	39	39
1/11/02	Friday	no data	51	53	49	40	39
1/12/02	Saturday	no data	51	51	53	49	40
1/13/02	Sunday	no data	49	51	51	53	49
1/14/02	Monday	5.5263	51	49	51	51	53
1/15/02	Tuesday	6	54	51	49	51	51
x	x	x	x	54	51	49	51
x	x	x	x	x	54	51	49
				x	x	54	51
					x	x	54
						x	x

We applied the same procedures to capture the lagged effects of all processing and weather data sets.

3.4 Organizing The Selected Data into Table for Correlation Analysis

Up to this point, we: 1. selected periods for the correlation analysis: winter, summer, and the entire 2002 data, 2. identified which inspectors were appropriate for each period, and 3. created lagged weather and processing variables. In this section we combine all these procedures into one data table to use in the correlation analysis.

We choose the correlation between Mike field odor data and processing data in the winter period to be the example of how to prepare the data set. First, we pull out all of the processing data and Mike's odor data from the winter period, January to March 2002. Then, create the lagged variables.

Table 3.3 shows how to organize the odor and processing data for the correlation analysis between inspector Mike's odor scores and processing data in the winter period. First we computed all of Mike's odor data in the winter as shown in column three. Then we applied the procedure described in Section 3.3 to create the lagged variables.

Table 3.3: The Sample of Data Set for Field Odor and Processing Data Correlation Analysis

Date	day	mike	all processing data on d-0		all processing data on d-1		all processing data on d-2		all processing data on d-3		all processing data on d-4	
1/1/02	Tuesday	no data	34	20	no data	no data	no data	no data	no data	no data	no data	no data
1/2/02	Wednesday	5	37	21	34	20	no data	no data	no data	no data	no data	no data
1/3/02	Thursday	3.3	36	30	37	21	34	20	no data	no data	no data	no data
1/4/02	Friday	6	41	27	36	30	37	21	34	20	no data	no data
1/5/02	Saturday	0	51	26	41	27	36	30	37	21	34	20
1/6/02	Sunday	3	39	29	51	26	41	27	36	30	37	21
1/7/02	Monday	5.5	39	30	39	29	51	26	41	27	36	30
1/8/02	Tuesday	3	40	28	39	30	39	29	51	26	41	27
1/9/02	Wednesday	3	no data	29	40	28	39	30	39	29	51	26
1/10/02	Thursday	0	53	34	no data	29	40	28	39	30	39	29
1/11/02	Friday	5	51	39	53	34	no data	29	40	28	39	30
1/12/02	Saturday	no data	51	30	51	39	53	34	no data	29	40	28
1/13/02	Sunday	no data	49	36	51	30	51	39	53	34	no data	29
1/14/02	Monday	5.5	51	32	49	36	51	30	51	39	53	34
					51	32	49	36	51	30	51	39
							51	32	49	36	51	30
									51	32	49	36
											51	32

Table 3.4 provides a snapshot of the completed data from 1/1/02 to 1/14/02. The data from 1/1/02 to 1/14/02 from Table 3.3 decreased to five days as shown in Table 3.4 after we deleted our data-missing dates. When we shifted each variable in

creating the lagged variables, we can see a dramatic decrease in the number of data points in the table.

Table 3.4: The Sample of Data Set for Field Odor and Processing Data Correlation Analysis after Deleting Missing Data

Date	day	mike	all processing data on d-0		all processing data on d-1		all processing data on d-2		all processing data on d-3		all processing data on d-4	
1/5/02	Saturday	0	51	26	41	27	36	30	37	21	34	20
1/6/02	Sunday	3	39	29	51	26	41	27	36	30	37	21
1/7/02	Monday	5.5	39	30	39	29	51	26	41	27	36	30
1/8/02	Tuesday	3	40	28	39	30	39	29	51	26	41	27
1/14/02	Monday	5.5	51	32	49	36	51	30	51	39	53	34

3.5 Correlation Approach

The purpose of this analysis was to search for correlations between variables from the inspectors' field odor data and both processing data and weather data. The coefficient's sign of correlation between a pair of variables shows how these variables are correlated to each other. The positive correlation's coefficient means when the value of one variable increases the other variable tends to increase as well. Also, the negative correlation's coefficient means when the value of one variable increases the other variable tends to decrease. The range of correlation coefficient is from -1 to 1. The value closed to -1 or 1 means there is a strong either negative or positive correlation between the pair of variables and the value closed to 0 means there is a weak correlation between the pair of variables (Winston, 1994).

The correlation analysis is useful in regards to the regression work since it is instructive to identify independent variables (e.g., amount of lime used) that correlate strongly with the dependent variable (odor level). When there are a lot of independent variables to choose from (as in the case of the processing data), this is a useful technique for narrowing down the set of candidate variables.

Additionally, such a correlation analysis can identify independent variables which are themselves strongly correlated. Such a situation can indicate the presence of "multicolinearity" (Maddala, 1977) that can lead to incorrectly concluding that certain variables are insignificant (in the statistical sense) when in fact they are actually important. This analysis was performed separately by season for the year 2002 as well as for the entire year.

Several analyses were run as follows:

1. Correlation between inspector's field odor data and processing data from data in 2002, winter, and summer.

2. Correlation between normalized odor data and processing data for data in 2002, summer, and winter.

3. Correlation between inspector's field odor data and weather data from data in 2002, winter, and summer.

In what follows we describe the results of these correlation analyses.

3.6 Correlation between Inspectors' Field Odor Data and Processing Data

The purpose of this analysis was to find processing variables highly correlated with the inspectors' odor for the time frame specified above.

3.6.1 Summary Correlation between Inspectors' Field Odor Data and Processing Data

We summarized the correlation results regardless of periods as shown below

1. Many of the correlation results showed blanket depth parameters had a positive correlation with the inspector's odor scores, i.e., when the blanket depth variables had a correlation coefficient more than 0.3 with the inspector's odor in absolute value. As discussed in Chapter 2, the higher level of the blanket depth in secondary sedimentation tanks, the greater the retention time of the waste activated sludge prior to removal to the DAF thickeners, and the more potential for biosolids odor.

2. Blend ratio variables were negatively correlated with odor. Even though a small number of the correlation results showed strongly negative (coefficient's sign more than absolute 0.3) correlation between blend ratio and inspector's field odor, many correlation results indicated negative correlation between this pair of variables. This is an unexpected sign related to what we expected as discussed in Chapter 2. Namely, the expectation was that the higher the value of the blend ratio the greater the biosolids odor because of more primary sludge (food) relative to waste activated sludge (microorganisms). The causality regarding the relationship between the blend ratio and biosolids odor must be investigated by laboratory experiments to compare the microbial count from the sample of sludge in blend tank with the sample of biosolids on the same day to see the trend of the amount of microorganisms in the blend tank relative to biosolids.

3. Many of the correlation results between lime addition and inspector's odor scores presented highly negative correlation, i.e., when the lime addition variable had a correlation's coefficient value greater than 0.3 with inspector's odor in absolute value. The negative coefficient's sign of lime follows our assumption in Chapter 2 that the more amount of lime added, the lower the biosolids odor levels as well as fewer the microorganisms left after lime mixing.

3.6.2 Correlation between Each Inspector's Field Odor Data and Processing Data for the entire year 2002

The purpose of this section was to observe the relationship between processing data and inspector's odor score regardless of season. All the processing variables in 2002 were arranged as d-0, d-1, d-2, d-3, and d-4 into a table. The inspectors we selected in this period were Pete (191 data points), Carl, Patrick (149 data points), Patrick (34 data points), Wilfred (51 data points), Rob (156 data points), and Mike (158 data points). We delete missing data before running the correlation analysis.

From the Tables 3.5 to 3.10, the correlation analysis results in 2002 are shown. In each table, each row represents the correlation between the inspector's odor score on day d-0 and the relevant processing data from day d-0 to d-4, respectively. We used the abbreviation of processing variables as shown below

B.D. second. E = the blanket depth in secondary east

B.D. second. W. odd = the blanket depth in secondary west odd

B.D. second. W. even = the blanket depth in secondary west even

All B.D.- W. = sum of blanket depth in secondary west

All B.D. = sum of all blanket depth

$FeCl_3$ primary W. = $FeCl_3$ added in primary west

$FeCl_3$ primary E = $FeCl_3$ added in primary east

All $FeCl_3$ = All $FeCl_3$ added

WPL = waste pickle liquid added

Blended ratio = the blended ratio in blended mixing tank

DAF poly. = polymer added in DAF

DEWAT poly. = polymer added in dewatering process

All poly. = all polymer added

WASA # centrif. in service = DCWASA number of centrifuges in service

WASA # centrif. out service = DCWASA number of centrifuges out of service

Contr. # belt. in service = contractor number of centrifuges in of service

Contr. # centrif. in service = contractor number of centrifuges in of service

Lime = lime addition

Table 3.5: The Correlation between Pete's Field Odor and Processing Data in 2002

Pete 2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.- W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primar y E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWA T poly.	All poly.	WASA # centrif. in service	WASA # centrif.o ut service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.25	0.33	0.22	0.30	0.32	0.17	0.08	0.14	-0.20	-0.27	0.12	-0.10	0.13	0.13	0.12	-0.14	-0.51	-0.38
d-1	0.35	0.31	0.17	0.26	0.34	0.13	0.08	0.13	-0.24	-0.20	0.05	-0.08	0.04	0.15	0.10	-0.10	-0.49	-0.33
d-2	0.31	0.34	0.23	0.31	0.35	0.14	0.07	0.12	-0.10	-0.18	0.04	-0.12	0.04	0.13	0.04	-0.04	-0.48	-0.31
d-3	0.31	0.35	0.21	0.30	0.35	0.15	0.07	0.13	-0.19	-0.25	0.10	0.10	0.15	0.16	0.07	-0.13	-0.50	-0.30
d-4	0.35	0.31	0.13	0.23	0.31	0.13	0.08	0.13	-0.09	-0.21	0.08	-0.05	0.13	0.07	0.12	-0.11	-0.49	-0.32

Table 3.6: The Correlation between Carl's Field Odor and Processing Data in 2002

Carl 2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.- W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primar y E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWA T poly.	All poly.	WASA # centrif. in service	WASA # centrif.o ut service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.34	0.36	0.32	0.38	0.43	0.03	-0.02	0.01	-0.11	-0.11	0.19	-0.27	0.11	0.04	0.23	-0.19	-0.35	-0.31
d-1	0.32	0.28	0.23	0.29	0.36	0.04	-0.03	0.01	-0.14	-0.10	0.14	-0.23	0.07	0.12	0.17	-0.18	-0.36	-0.32
d-2	0.30	0.34	0.19	0.29	0.35	0.02	-0.02	0.00	0.06	-0.11	0.06	-0.19	0.06	0.13	0.14	-0.15	-0.33	-0.29
d-3	0.29	0.31	0.24	0.31	0.36	0.03	-0.02	0.01	0.01	-0.12	0.04	-0.17	0.00	0.13	0.16	-0.21	-0.29	-0.29
d-4	0.30	0.39	0.28	0.37	0.40	0.02	-0.01	0.01	-0.10	-0.16	0.00	-0.24	-0.04	0.06	0.23	-0.22	-0.35	-0.29

Table 3.7: The Correlation between Patrick's Field Odor and Processing Data in 2002

Patrick 2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	-0.08	-0.16	-0.13	-0.16	-0.16	-0.51	N/A	-0.51	-0.18	-0.25	0.23	0.90	0.19	0.02	0.02	0.01	-0.21	-0.63
d-1	0.05	-0.41	-0.30	-0.38	-0.24	-0.48	N/A	-0.48	0.33	-0.17	-0.08	-0.39	0.20	0.10	0.03	-0.02	-0.12	-0.56
d-2	-0.01	-0.31	-0.30	-0.33	-0.25	-0.48	0.02	-0.01	0.31	-0.34	-0.13	-0.23	-0.06	0.41	0.28	0.00	-0.12	-0.46
d-3	0.02	-0.22	-0.30	-0.29	-0.20	-0.48	0.02	-0.01	0.29	-0.23	0.16	-0.06	0.17	0.37	0.36	0.07	-0.20	-0.54
d-4	0.11	-0.03	-0.21	-0.13	-0.03	-0.46	N/A	-0.46	0.15	-0.21	-0.10	0.72	-0.11	0.17	0.24	-0.08	-0.11	-0.56

Table 3.8: The Correlation between Wilfred's Field Odor and Processing Data in 2002

Wilfred 2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	-0.17	0.53	0.12	0.41	0.30	0.42	0.47	0.52	0.06	-0.22	0.11	-0.09	0.06	0.27	0.22	0.51	0.00	-0.05
d-1	-0.06	0.46	0.12	0.38	0.31	0.44	0.48	0.51	-0.10	-0.46	0.23	-0.05	0.20	0.34	0.25	0.34	-0.04	0.01
d-2	0.01	0.44	-0.20	0.20	0.18	0.47	0.47	0.51	-0.11	-0.39	0.03	-0.02	0.03	0.34	0.28	0.38	-0.04	0.03
d-3	0.08	0.51	0.00	0.34	0.35	0.49	0.50	0.53	-0.14	-0.39	0.20	-0.11	0.21	0.27	0.06	0.42	-0.04	0.04
d-4	0.11	0.48	0.00	0.29	0.31	0.50	0.48	0.51	-0.15	-0.53	0.13	-0.03	0.14	0.34	0.05	0.43	-0.03	0.10

Table 3.9: The Correlation between Rob's Field Odor and Processing Data in 2002

Rob 2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.06	-0.12	0.01	-0.05	-0.01	-0.08	-0.06	-0.09	0.17	0.14	0.06	0.00	0.04	-0.09	-0.03	0.00	0.18	0.15
d-1	0.08	-0.14	-0.02	-0.08	-0.02	-0.08	-0.06	-0.09	-0.05	0.11	0.07	0.08	0.09	-0.14	0.08	0.00	0.18	0.15
d-2	0.13	-0.09	0.18	0.07	0.12	-0.08	-0.06	-0.09	0.03	0.17	-0.31	-0.07	-0.38	-0.08	0.00	0.00	0.18	0.17
d-3	0.03	-0.06	0.05	-0.01	0.01	-0.08	-0.06	-0.09	0.04	0.26	-0.02	-0.01	-0.03	-0.08	-0.12	-0.10	0.15	0.13
d-4	0.05	-0.15	-0.10	-0.13	-0.07	-0.06	-0.08	-0.08	0.12	0.22	0.01	-0.01	-0.04	-0.09	0.00	-0.13	0.16	0.15

Table 3.10: The Correlation between Mike's Field Odor and Processing Data in 2002

Mike2002

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.02	0.06	0.16	0.13	0.11	0.03	-0.02	0.01	0.17	-0.01	-0.02	0.17	0.01	0.03	0.07	0.23	0.25	-0.04
d-1	0.05	0.10	0.17	0.16	0.16	0.02	-0.03	-0.01	0.09	0.00	-0.01	0.14	0.03	0.12	0.10	0.26	0.22	0.03
d-2	-0.05	0.08	0.03	0.06	0.03	0.03	-0.01	0.01	0.06	-0.03	-0.15	0.12	-0.15	0.06	0.04	0.30	0.20	-0.05
d-3	0.00	0.06	0.02	0.04	0.03	0.05	0.01	0.04	0.16	-0.13	-0.06	0.13	-0.03	0.04	-0.05	0.17	0.28	-0.07
d-4	0.07	0.00	0.03	0.02	0.05	-0.01	-0.06	-0.03	0.11	0.00	0.01	0.18	0.04	0.12	-0.02	0.22	0.28	0.02

The correlation's results between selected inspectors and processing data using data from the entire year 2002 showed some strongly correlated parameters consistent with the assumptions on processing variable to biosolids odor in Chapter 2. Many of blanket depth parameters such as "All B.D." were positively correlated with Pete's, Carl's, and Wilfred's odor scores. This means the higher the level of blanket depth the greater the retention time of microorganisms that will be wasted to the blend tank later and the greater the potential for biosolids odor. In addition, the correlation's results between lime and the inspector's odor scores also showed a negative correlation with Pete's, Carl's, and Patrick's odor scores, which means the more lime added into dewatering sludge the less microorganisms and odor in biosolids. Also, the greater the number of contractor centrifuges in service a indicated negative correlation with Pete's and Carl's odor scores, corresponding to a lower blanket depth in the secondary sedimentation tanks and less biosolids odor.

$FeCl_3$ variables showed the significant correlation with inspector's odor scores. $FeCl_3$ variables were highly correlated to Wilfred's odor scores with a positive sign. This means the more $FeCl_3$ added into primary tanks the greater the biosolids odor scores from inspector Wilfred's field odor.

The blend ratio parameter indicated an unexpected correlation sign. With Wilfred's odor scores, the blend ratio was negatively correlated with biosolids odor. As discussed in Chapter 2, a higher blend ratio was expected to increase the microorganisms left in biosolids as well as odor. This unexpected correlation's result should be investigated as recommended in Section 3.6.1.

3.6.3 Correlation between each Inspector's Field Odor Data and Processing Data Winter 2002

Processing data in January, February, and March have been collected to find the correlation between processing variables and inspector's odor score during the low-temperature period. All processing data were arranged as d-0, d-1, d-2, d-3, and d-4 into a table. The inspectors we selected were Pete (48 data points), Wilfred (32 data points), Carl (45 data points), and Mike (46 data points). Tables 3.11 to 3.14 show the correlation between each inspector's field odor data during winter 2002 with the processing variables from the same period.

Table 3.11: The Correlation between Pete's Field Odor and Processing Data in Winter

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	-0.25	0.06	-0.29	-0.15	-0.21	0.65	0.32	0.62	0.21	0.00	0.11	0.47	0.11	0.10	-0.39	0.12	-0.10	0.34
d-1	-0.11	0.00	-0.26	-0.17	-0.17	0.53	0.36	0.53	0.13	-0.13	-0.19	0.42	-0.16	0.23	-0.41	0.09	-0.10	0.39
d-2	-0.26	-0.08	-0.21	-0.20	-0.24	0.37	0.17	0.36	-0.03	-0.07	0.00	0.34	0.02	-0.03	-0.43	0.10	0.00	0.36
d-3	-0.24	-0.05	-0.32	-0.23	-0.28	0.21	0.04	0.19	-0.11	-0.15	0.08	0.55	0.13	-0.13	-0.45	0.19	0.00	0.35
d-4	-0.16	-0.11	-0.37	-0.29	-0.29	0.13	0.03	0.12	-0.10	-0.20	0.05	0.47	0.09	-0.07	-0.48	0.33	N/A	0.35

Table 3.12: The Correlation between Wilfred's Field Odor and Processing Data in Winter

Wilfred winter

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.11	0.38	0.03	0.27	0.26	0.47	0.07	0.42	0.07	0.13	0.07	-0.08	-0.32	0.07	0.13	0.51	N/A	-0.07
d-1	0.19	0.31	0.16	0.31	0.34	0.48	0.10	0.41	0.14	-0.28	0.00	0.04	-0.17	0.33	0.16	0.26	N/A	0.03
d-2	0.18	0.39	-0.13	0.20	0.25	0.47	0.10	0.38	0.07	-0.10	-0.05	0.08	-0.08	0.30	0.27	0.31	N/A	0.07
d-3	0.37	0.53	0.15	0.43	0.54	0.47	0.14	0.38	0.03	-0.20	0.07	0.10	0.06	0.19	0.24	0.39	N/A	0.05
d-4	0.50	0.45	0.23	0.40	0.56	0.44	0.15	0.36	-0.02	-0.42	-0.12	0.19	-0.11	0.38	0.19	0.42	-0.06	0.12

Table 3.13: The Correlation between Carl's Field Odor and Processing Data in Winter

Carl winter

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.49	0.49	0.61	0.66	0.70	-0.06	-0.23	-0.14	-0.17	-0.08	0.36	-0.38	0.24	-0.10	0.87	-0.05	N/A	-0.91
d-1	0.48	0.46	0.69	0.69	0.71	-0.01	-0.25	-0.10	-0.06	-0.10	0.27	-0.26	0.15	0.09	0.96	-0.04	N/A	-0.91
d-2	0.49	0.59	0.53	0.68	0.72	-0.03	-0.18	-0.09	0.39	-0.14	0.17	-0.25	0.17	0.13	1.00	0.04	0.09	-0.86
d-3	0.47	0.60	0.57	0.73	0.76	-0.02	-0.15	-0.07	0.25	-0.12	0.15	-0.30	0.09	0.14	1.00	-0.07	0.25	-0.86
d-4	0.46	0.68	0.73	0.83	0.82	-0.01	-0.13	-0.06	0.00	-0.20	0.06	-0.40	-0.01	-0.03	0.93	-0.13	N/A	-0.87

Table 3.14: The Correlation between Mike's Field Odor and Processing Data in Winter

Mike winter

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	-0.13	0.01	0.21	0.13	0.04	0.08	0.32	0.20	-0.09	-0.20	0.00	0.09	-0.05	0.16	-0.05	0.19	-0.11	0.23
d-1	-0.09	0.09	0.22	0.18	0.14	-0.04	0.30	0.09	-0.34	-0.16	0.07	0.12	0.10	-0.11	-0.03	0.17	-0.12	0.16
d-2	-0.11	0.13	0.07	0.12	0.05	-0.05	0.21	0.04	-0.18	-0.30	-0.24	0.24	-0.28	0.00	-0.06	0.41	-0.31	0.27
d-3	-0.12	0.01	0.03	0.02	-0.05	0.06	0.28	0.14	0.08	-0.33	-0.05	0.25	-0.08	0.08	-0.15	0.21	0.11	0.26
d-4	0.04	-0.08	0.06	0.00	-0.01	-0.06	0.29	0.05	0.10	-0.23	0.03	0.42	0.12	0.04	-0.09	0.13	N/A	0.30

Using the data in January to March 2002, the correlation's results between inspector's odor scores and processing data showed 3 processing parameters having high correlation with inspector's odor scores with the expected sign as discussed in Chapter 2. First, the blanket depth parameters indicated the positive correlation with Wilfred's and Carl's odor scores. The higher the level of blanket depth the more retention time of waste activated sludge wasted from secondary sedimentation tanks and the more odors on biosolids. Second, dewatering polymer showed highly positive correlation with Pete's odors. Thus, the more polymers added into dewatering process the more sources to stimulate TMA production from microorganisms. Third, lime presented the strongly negative correlation with Carl's odor scores. Using Carl's biosolids odor scores, the more lime added into sludge the fewer microorganisms to produce the biosolids odor.

$FeCl_3$ variables that we could not make an assumption on coefficient's sign showed the positive correlation with Pete's and Wilfred's odor scores. From this result, we can statistically interpret that the more $FeCl_3$ added into primary process the higher Pete's and Wilfred's odor scores in winter period.

3.6.4 Correlation between Each Inspector's Field Odor Data and Processing Data Summer 2002

The purpose of this section was to investigate how much the temperature matters to the relationship between processing data and biosolids odors in the high-temperature period of 2002. Processing variables in June, July, and August were arranged as d-0, d-1, d-2, d-3, and d-4 into a table. The inspectors we selected for the summer were Rob (56 data points), Pete (47 data points), and Mike (54 data points).

From Table 3.15 to 3.17, we see the correlation analysis between each inspector's field odor data during summer 2002 with the processing variables from the same period.

Table 3.15: The Correlation between Mike's Field Odor and Processing Data in the summer

Mike summer

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.30	0.22	0.17	0.20	0.31	0.27	-0.06	0.08	0.23	-0.03	-0.01	0.15	0.05	0.04	-0.06	0.30	0.02	-0.37
d-1	0.41	0.26	0.15	0.22	0.36	0.26	-0.08	0.06	0.40	-0.01	0.03	0.13	0.20	0.30	0.16	0.34	0.01	-0.27
d-2	0.21	0.27	0.02	0.14	0.20	0.25	-0.06	0.07	0.21	-0.09	-0.05	0.00	-0.04	0.17	0.17	0.30	0.03	-0.42
d-3	0.27	0.29	0.03	0.16	0.25	0.25	0.00	0.11	0.27	-0.30	-0.32	-0.03	-0.29	0.29	-0.08	0.10	-0.01	-0.44
d-4	0.27	0.21	-0.03	0.07	0.19	0.14	-0.17	-0.06	0.12	0.01	-0.13	-0.16	-0.17	0.30	-0.06	0.12	0.27	-0.38

Table 3.16: The Correlation between Pete's Field Odor and Processing Data in the summer

Pete summer

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	-0.34	-0.06	-0.04	-0.05	-0.21	-0.16	0.05	-0.04	-0.37	-0.09	0.18	-0.32	0.05	-0.29	0.35	-0.29	0.00	0.17
d-1	-0.15	-0.09	-0.07	-0.10	-0.16	-0.16	0.08	-0.02	-0.17	0.12	-0.03	-0.30	-0.36	-0.29	0.17	-0.04	0.00	0.30
d-2	-0.17	0.01	-0.01	-0.06	-0.12	-0.14	0.06	-0.03	-0.20	0.24	-0.15	-0.23	-0.16	-0.25	0.18	0.15	0.04	0.32
d-3	-0.07	-0.11	0.00	-0.07	-0.09	-0.01	0.08	0.05	-0.17	0.12	-0.08	0.23	0.03	0.09	0.06	-0.07	0.02	0.36
d-4	-0.06	-0.17	-0.25	-0.25	-0.23	-0.25	0.09	-0.06	-0.03	0.20	-0.15	-0.15	-0.08	-0.12	0.22	0.00	0.09	0.34

Table 3.17: The Correlation between Rob's Field Odor and Processing Data in the summer

Rob summer

	B.D. second. E.	B.D. second. W. odd	B.D. second. W. even	All B.D.-W.	All B.D.	$FeCl_3$ primary W.	$FeCl_3$ primary E.	All $Fecl_3$	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.out service	Contr. # belt. in service	Contr. # centrif. in service	Lime
d-0	0.03	0.15	0.16	0.18	0.15	-0.09	-0.04	-0.08	0.14	-0.01	0.04	0.02	-0.03	-0.03	0.07	0.04	-0.02	0.25
d-1	0.01	-0.14	0.04	-0.03	-0.02	-0.06	-0.04	-0.06	-0.11	-0.12	0.16	0.03	0.24	-0.07	0.02	0.02	-0.02	0.10
d-2	0.13	-0.10	-0.11	-0.15	-0.02	-0.07	-0.03	-0.06	-0.17	0.09	-0.18	0.06	-0.26	0.05	-0.15	0.03	0.01	0.16
d-3	-0.02	0.13	-0.08	-0.07	-0.03	-0.09	-0.05	-0.08	0.16	0.14	-0.10	-0.13	-0.11	-0.13	0.02	-0.07	0.03	0.06
d-4	-0.04	-0.04	-0.09	-0.08	-0.08	-0.09	-0.04	-0.08	0.01	-0.02	0.05	-0.19	-0.04	-0.03	-0.07	-0.30	0.03	0.10

The correlation's result between an inspector's odor scores and the processing data during the summer period did not show strong correlation compared with the entire year 2002 and the winter period. The best correlation was from inspector Mike. Some blanket depth parameters indicated a positive correlation with Mike's odor scores ("Blanket depth second. E." and "All blanket depth"), which can be interpreted as the higher the blanket depth the greater retention time for activated sludge in the tanks as well as the more biosolids odor. Also, lime was negatively correlated with Mike's odor scores. The more lime addition to dewatering sludge decreased the biosolids odor by killing the microorganisms left after dewatering.

3.7 Correlation between Normalized Odor Data and Processing Data for Entire Year 2002, Summer Season, and Winter Season.

3.7.1 The Normalized Odor Data

Since there were insufficient odor data for winter and summer, it was necessary to find a way to combine all data from all different inspectors and then to run a correlation analysis. To solve this problem, the data were "normalized" as described below and then the data from different inspectors were combined.

Normalized odor scores were computed as follows using inspector Pete as an example:

Normalized odor data for inspector Pete

= (Pete's odor data from each day – average odor data of Pete from 2002)/ (Standard deviation of Pete odor data from 2002)

Table 3.18 summarizes the average and standard deviation of selected inspectors for normalized correlation analysis using the data from 2002

Table 3.18: Average and Standard Deviation of Inspectors' Field Odor Data

inspector	average odor data in 2002	Standard deviation
Mike	5.156	1.606
Pete	2.774	2.406
Carl Burton	0.237	0.803
Patrick Collins	2.909	1.048
Rob Siers	3.000	0.511
Wilfred Wade	3.078	1.133

For example, if Pete's odor data on January 1, 2002 equaled to "3" signifying slight odor, Pete's normalized odor score would be

$$= (3 - 2.774) / 2.406$$

$$= 0.939$$

Next we averaged the resulting inspectors' normalized odor scores for each day as shown in Table 3.19.

Table 3.19: Normalized Inspectors' Scores

Date:	Norm. Mike	Norm. Pete	Norm. Carl	Norm. Rob.	Norm. Patrick	Norm. Wilfred	Avg. Norm.score
1/1/02	no data	no data	no data	no data	no data	no data	no data
1/2/02	-0.097	no data	3.438	no data	no data	-0.069	1.091
1/3/02	-1.173	no data	no data	no data	no data	no data	-1.173
1/4/02	0.525	no data	no data	no data	no data	-0.069	0.228
1/5/02	no data	no data	no data	no data	no data	no data	no data
1/6/02	no data	no data	no data	no data	no data	no data	no data
1/7/02	no data	0.094	3.438	no data	no data	no data	1.766
1/8/02	-1.343	0.094	no data	no data	no data	no data	-0.624
1/9/02	-1.343	0.094	no data	no data	no data	no data	-0.624
1/10/02	-3.210	0.094	no data	no data	no data	no data	-1.558
1/11/02	no data	0.094	3.438	no data	no data	no data	1.766

By using the normalized odor data, the odor data from all inspectors could then be combined since the scales were similar, thus producing a larger amount of data to use and a greater degree of freedom (Devore, 1995). When this procedure was done, the number of data points from each category was increased as follows.

From Pete's yearly correlation with 192 data points and Carl's yearly correlation with 150 data points, the data points were increased to 2190 data points in the combined inspector's yearly correlation after normalizing.

From Pete's winter correlation with 48 data points, Carl's winter correlation with 45 data points, and Wilfred's winter correlation with 45 data points, the combined inspector's winter correlation data set increased to 540 data points.

From Pete's summer correlation with 47 data points, Mike's summer correlation with 54 data points, the combined inspector's summer correlation data set increased to 542 data points.

3.7.2 The result of Correlation between Normalized Odor Data and Processing Data

We ran correlation analyses between the normalized odor data and the processing data using data for all of 2002, the winter season, and the summer season. The results are shown in Tables 3.20, 3.21, and 3.22 respectively.

All the correlation's results showed no significant variable highly correlated with normalized odor scores. This may be because we mixed different inspectors in the same data set and it reduced the significance of the inspector's odor scores.

Table 3.20: The Correlation between Normalized Field Odor and Processing Data in 2002

Normalized correlation from data 2002(2190 data points)																		
	B.D. second . E.	B.D. second. W. odd	B.D. second . W. even	All B.D.- W.	All B.D.	FeCl3 primary W.	FeCl3 primary E.	All Fecl3	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.o ut service	Contr. # belt. in servic e	Contr. # centrif. in service	Lime
d-0	0.13	0.16	0.15	0.17	0.18	0.06	0.02	0.04	-0.01	-0.08	0.09	-0.02	0.07	0.04	0.09	0.00	-0.13	-0.15
d-1	0.18	0.13	0.11	0.13	0.18	0.05	0.02	0.04	-0.05	-0.07	0.06	-0.03	0.06	0.08	0.10	0.00	-0.13	-0.11
d-2	0.15	0.15	0.11	0.14	0.18	0.05	0.01	0.04	0.02	-0.07	-0.08	-0.06	-0.09	0.09	0.08	0.04	-0.12	-0.11
d-3	0.15	0.16	0.09	0.14	0.17	0.06	0.02	0.05	0.02	-0.07	0.04	0.02	0.05	0.09	0.04	-0.04	-0.11	-0.11
d-4	0.18	0.14	0.05	0.10	0.15	0.05	0.02	0.04	0.03	-0.06	0.03	0.00	0.03	0.05	0.08	-0.05	-0.11	-0.10

Table 3.21: The Correlation between Normalized Field Odor and Processing Data in winter

Normalized correlation from the data on winter (540 data points)																		
	B.D. second . E.	B.D. second. W. odd	B.D. second . W. even	All B.D.- W.	All B.D.	FeCl3 primary W.	FeCl3 primary E.	All Fecl3	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.o ut service	Contr. # belt. in servic e	Contr. # centrif. in service	Lime
d-0	0.13	0.23	0.23	0.28	0.28	0.13	0.07	0.13	0.01	-0.07	0.12	-0.04	0.05	0.03	0.26	0.10	-0.06	-0.20
d-1	0.15	0.21	0.23	0.27	0.29	0.13	0.07	0.12	-0.15	-0.11	0.08	0.00	0.04	0.09	0.21	0.06	-0.07	-0.18
d-2	0.11	0.25	0.11	0.21	0.23	0.07	0.01	0.06	0.06	-0.07	-0.08	0.03	-0.07	0.07	0.17	0.12	-0.08	-0.14
d-3	0.11	0.23	0.14	0.22	0.22	0.09	0.01	0.07	0.06	-0.10	0.07	0.02	0.06	0.04	0.16	0.07	0.13	-0.13
d-4	0.18	0.19	0.14	0.19	0.23	0.07	0.03	0.07	0.02	-0.18	0.08	0.05	0.09	0.01	0.21	0.05	0.01	-0.16

Table 3.22: The Correlation between Normalized Field Odor and Processing Data in summer

Normalized correlation from data on summer (542 data points)																		
	B.D. second . E.	B.D. second. W. odd	B.D. second . W. even	All B.D.- W.	All B.D.	FeCl3 primary W.	FeCl3 primary E.	All Fecl3	WPL	Blended ratio	DAF poly.	DEWAT poly.	All poly.	WASA # centrif. in service	WASA # centrif.o ut service	Contr. # belt. in servic e	Contr. # centrif. in service	Lime
d-0	0.04	0.10	0.10	0.11	0.10	0.02	-0.03	-0.02	0.01	-0.02	0.06	-0.03	0.02	-0.05	0.08	0.01	0.01	0.01
d-1	0.12	0.01	0.05	0.04	0.09	0.02	-0.02	-0.01	0.06	-0.02	0.06	-0.03	0.10	0.03	0.11	0.09	0.01	0.05
d-2	0.08	0.05	0.00	0.02	0.06	0.02	-0.03	-0.01	-0.02	0.07	-0.10	-0.06	-0.14	0.02	0.06	0.10	0.03	0.01
d-3	0.08	0.11	0.01	0.04	0.08	0.04	-0.01	0.01	0.09	0.01	-0.15	0.00	-0.12	0.08	-0.01	-0.03	0.01	0.00
d-4	0.08	0.03	-0.06	-0.02	0.03	-0.05	-0.05	-0.06	0.06	0.05	-0.02	-0.12	-0.05	0.06	0.00	-0.07	0.13	0.02

3.8 Correlation between Each Inspector's Field Odor Data and Weather Data in 2002, Summer Season, and Winter Season.

In addition to the correlation analysis described above between processing data and odor scores, correlation between weather and odor data were also considered. We selected the same inspectors as before except that inspector Rob in 2002 and in the summer because his odor data didn't show much correlation with processing data in these periods. From Tables 3.23 to 3.33, the correlation analysis's results are shown.

The following is a summary of the weather and inspector's odor correlations.

For the entire year, Pete's, Carl's, and Patrick's odor scores were strongly negative correlated with temperature variables. This is different from our assumption that the higher the temperature the higher the biological activity and the more biosolids odor. Another variable showing a significant and negative correlation was the amount of daily snow fall. This coefficient sign was also unexpected because our assumption is the greater the snowfall the less the biological activity and odors. As a result, we did not include these weather variables in the yearly regression analysis.

In the winter period, the temperature variables (maximum, minimum, and average temperature) presented a highly positively correlated coefficient with Pete's odor scores. This followed our assumption that the higher temperature the more biological activity and the greater the biosolids odor.

In the summer period, there was no obvious strongly correlation between weather variables and Pete's or Mike's odor scores.

Table 3.23: Correlation between Pete's Odor Data and Weather Data in 2002

PETE2002

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	-0.52	-0.53	-0.53	0.05	-0.10	0.04	0.02
d-1	-0.53	-0.51	-0.53	0.19	-0.10	0.04	0.11
d-2	-0.53	-0.51	-0.53	0.17	-0.10	0.06	0.11
d-3	-0.47	-0.51	-0.50	0.11	0.10	0.07	0.02
d-4	-0.47	-0.53	-0.51	0.12	-0.10	0.07	0.02

Table 3.24: Correlation between Carl's Odor Data and Weather Data in 2002

CARL 2002

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	-0.28	-0.27	-0.28	0.03	-0.06	-0.18	N/A
d-1	-0.27	-0.29	-0.29	-0.04	-0.05	0.02	N/A
d-2	-0.29	-0.32	-0.31	-0.11	-0.03	0.02	0.29
d-3	-0.32	-0.33	-0.33	-0.02	-0.06	0.02	N/A
d-4	-0.35	-0.33	-0.34	-0.05	-0.06	0.03	0.09

Table 3.25: Correlation between Patrick's Odor Data and Weather Data in 2002

PATRICK2002

	t max	t min	avg. temp.	average daily wind speed	daily precipitati on	average daily station pressure	daily snowfall
d-0	-0.27	-0.45	-0.38	-0.12	-0.07	-0.10	0.02
d-1	-0.37	-0.41	-0.41	0.23	0.03	-0.08	N/A
d-2	-0.48	-0.35	-0.44	0.28	0.13	-0.23	0.38
d-3	-0.34	-0.27	-0.32	0.46	0.14	-0.28	0.47
d-4	-0.24	-0.39	-0.32	0.05	0.10	-0.07	0.07

Table 3.26: Correlation between Mike's Odor Data and Weather Data in 2002

MIKE 2002

	t max	t min	avg. temp.	average daily wind speed	daily precipitati on	average daily station pressure	daily snowfall
d-0	0.19	0.20	0.20	-0.08	0.02	-0.04	0.01
d-1	0.20	0.20	0.20	-0.06	0.10	-0.04	-0.15
d-2	0.20	0.19	0.20	-0.02	0.05	-0.04	0.06
d-3	0.22	0.21	0.22	0.04	-0.07	-0.04	-0.12
d-4	0.25	0.24	0.24	0.03	0.04	-0.05	-0.27

Table 3.27: Correlation between Wilfred's Odor Data and Weather Data in 2002

wilfred 2002

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	0.13	0.17	0.15	0.20	-0.03	0.01	-0.01
d-1	0.24	0.27	0.26	-0.09	0.04	0.01	0.06
d-2	0.23	0.36	0.31	-0.17	0.00	-0.05	-0.01
d-3	0.18	0.33	0.26	-0.04	0.00	-0.08	N/A
d-4	0.18	0.29	0.25	0.05	0.02	0.01	-0.10

Table 3.28: Correlation between Pete's Odor Data and Weather Data in Winter

PETE winter

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	0.11	0.29	0.20	-0.15	0.42	0.10	-0.10
d-1	0.03	0.35	0.18	0.11	0.41	0.10	-0.13
d-2	0.18	0.40	0.31	0.09	0.13	0.24	-0.11
d-3	0.45	0.43	0.47	0.12	-0.01	0.18	-0.10
d-4	0.44	0.23	0.39	0.01	-0.11	0.10	-0.11

Table 3.29: Correlation between Wilfred's Odor Data and Weather Data in Winter

WILFRED winter

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	-0.02	0.10	0.03	0.09	0.11	-0.06	0.06
d-1	0.04	0.12	0.07	-0.29	0.15	-0.06	0.20
d-2	0.00	0.14	0.06	-0.27	0.04	0.19	0.08
d-3	-0.05	0.04	-0.02	-0.11	0.06	0.36	N/A
d-4	0.06	0.12	0.10	0.21	0.02	-0.06	-0.05

Table 3.30: Correlation between Carl's Odor Data and Weather Data in Winter

CARL winter

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	0.00	0.14	0.07	-0.07	-0.12	-0.25	N/A
d-1	0.06	0.07	0.07	-0.23	-0.03	-0.25	N/A
d-2	-0.05	-0.14	-0.10	-0.32	-0.16	-0.11	0.26
d-3	-0.19	-0.14	-0.19	-0.15	-0.17	-0.14	N/A
d-4	-0.28	-0.10	-0.22	-0.19	0.08	-0.31	0.25

Table 3.31: Correlation between Mike's Odor Data and Weather Data in Winter

MIKE winter

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	-0.15	-0.09	-0.13	0.13	0.15	0.22	0.08
d-1	0.04	-0.06	0.00	0.08	0.08	0.23	-0.15
d-2	0.14	-0.01	0.09	0.07	0.04	0.27	0.18
d-3	0.28	0.12	0.24	0.19	-0.14	0.11	-0.13
d-4	0.35	0.27	0.34	0.06	-0.36	-0.08	-0.35

Table 3.32: Correlation between Pete's Odor Data and Weather Data in the summer

PETE summer

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	0.08	0.16	0.11	-0.21	0.00	0.05	N/A
d-1	0.14	0.14	0.14	0.02	-0.12	0.02	N/A
d-2	0.10	0.07	0.09	0.12	-0.11	0.03	N/A
d-3	0.14	0.13	0.14	0.28	-0.13	0.12	N/A
d-4	0.10	0.02	0.07	0.09	-0.12	0.27	N/A

Table 3.33: Correlation between Mike's Odor Data and Weather Data in the summer

MIKE summer

	t max	t min	avg. temp.	average daily wind speed	daily precipitation	average daily station pressure	daily snowfall
d-0	-0.12	-0.06	-0.10	-0.25	0.10	-0.07	N/A
d-1	-0.21	-0.14	-0.19	-0.15	0.14	-0.05	N/A
d-2	-0.21	-0.15	-0.20	-0.10	-0.03	-0.13	N/A
d-3	-0.22	-0.13	-0.20	0.04	0.03	-0.12	N/A
d-4	-0.20	-0.11	-0.17	0.16	0.11	-0.06	N/A

In Chapter 4, we use the correlation analysis described in this chapter to create odor forecasting models by using variables that were strongly correlated with the inspector's odor. In addition, the dummy variables, and the interaction variables will be considered as part of the statistical models.

Chapter 4 Regression Analysis

In this chapter, the results of the correlation analysis from Chapter 3 are used to create the odor forecasting models by applying regression analysis on the variables that showed a high correlation with biosolids odor.

4.1 Objective of Regression Analysis

The odor forecasting models using regression analysis can improve DCWASA management in terms of the prediction of high biosolids' odor as well as provide a more efficient distribution (relative to odor) to the field sites on the day that the model forecasts a high biosolids odor. Currently, the truck assignment of biosolids in DCWASA doesn't consider the biosolids' odor as a factor in the assignment plan. Most of the time, they focus on the availability of the space at each site and the distance from the plant to the site rather than the biosolids' odor. The odor forecasting model can notify the biosolids manager on the day that high biosolids odor is expected to be produced. Consequently, the biosolids manager can adjust a truck assignment's plan in advance to divert high odor biosolids away from sites that have already received this.

4.2 The Requirements on an Appropriate Regression Equation

In this thesis, we used regression analysis approach to study the relationship between the independent variables X (the processing and weather data) and the dependent variable Y (the inspector's biosolids odor scores). Using the concept of least squares, the independent variables were used to form a multiple regression equation to explain the value for the dependent variable. The multiple regression equation predicts the dependent variable Y_i so as to minimize the sum of the squared errors (Maddala, 1977). Initially, we assumed a linear model of the following form:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + ... + \beta_k X_{ik} + u_i$$

where

Y_i = dependent variable (inspector's biosolids odor score) $i = 1, 2, \ldots, n$

β_0 = constant

β_j = coefficient of independent variable j, $j = 1, 2, \ldots, k$

X_{ij} = i^{th} observation for j^{th} independent variable

u_i = error of i^{th} observation

To identify that the regression equation is appropriate in this thesis, we considered three parameters from the regression result: adjusted R square, the correctness of the coefficient's sign for the independent variable, and the P value (related to the t-stats).

1. In each period, we used the adjusted R square of the regression to be one of the criteria to select the best regression in that period. As typically used to explain the fit of regression, R square is the parameter which describes the variation explained by the model. However, the problem with this parameter is that the value of R square will keep increasing when more independent variables are added (Maddala, 1977). Unlike R square, the adjusted R square takes care of this problem. The selected regression equation should present a high adjusted R square value compared to other equations for the same period. We anticipated finding equations with adjusted R square values more than 0.5. However, in some periods where the adjusted R square values from all

regressions were too low we might consider the following parameters in addition to adjusted R square to select the appropriate equation.

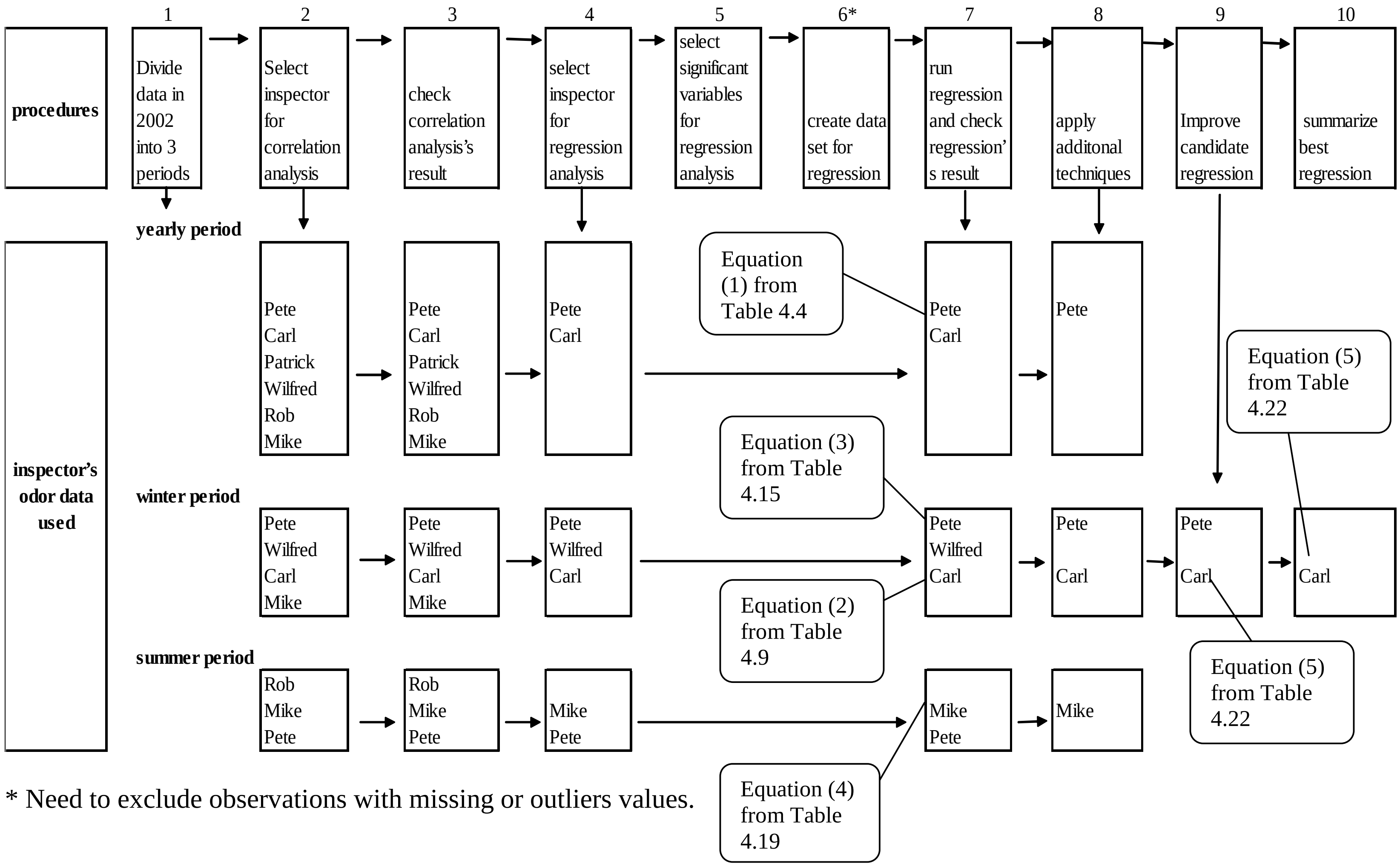
2. Within the regression equation, the coefficient's sign for each independent variable should be explainable by the assumptions stated in Chapter 2. However, for some variables such as $FeCl_3$, for which we couldn't make assumptions in Chapter 2, we can use the correlation analysis results from Chapter 3 to decide the suitable coefficient's sign. Finally, if there are no references for the correct coefficient's sign from both chapters, we choose to select the equation that best meets points 1 and 3.

3. All things being equal, a P value less than 0.05 or at most 0.1 is preferred. For variables showing P values higher than 0.05, one could not reject the null hypothesis that the coefficient of observed variable is equal to zero with 95% confidence. For P values lower than 0.05 we could reject the null hypothesis (Devore, 1995) and only be wrong less than 5% of the time.

As described above, we selected the appropriate regression for each period based on these three parameters: adjusted R square, correct coefficient signs, and statistically significant coefficients at the 5% level or 10% level. (or somewhat above this level when necessary)

4.3 The Procedures to Create Odor Forecasting Models

The procedures to create odor forecasting models were based on the correlation analysis from Chapter 3 and involved the regression analyses in this chapter. We summarize the procedure as shown in Figure 4.1.



* Need to exclude observations with missing or outliers values.

Figure 4.1: The Procedures to Create Odor Forecasting Models

To generate the odor forecasting models, we divided the work into two phases: the correlation analysis phase and the regression analysis phase.

Correlation analysis phase (step 1 to step 3):

1. Disaggregate data in 2002 into three periods: We picked data in 2002 and disaggregated them into three periods: the yearly period, the winter period, and the summer period. As stated in Chapter 3, the yearly period was used in order to observe the relationship between the inspector's odor score and the processing or weather data regardless of the season. Winter and summer periods were used in order to see the relationship between the inspector's odor score and the processing or weather data considering the average temperature when it was too low or too high compared to the average temperature in entire year.

2. Select the inspectors for each period to work on for the correlation analysis: After we disaggregated the data in 2002 into three periods, we selected inspectors for each period considering the number of data points available for that period.

3. Apply correlation analysis to the inspector's odor scores and the processing or weather data in each period: Finally, we ran correlation analyses to see how each inspector's odor scores in that period correlated with the processing and weather data.

Regression analysis phase (step 4 to step 10):

4. Select the inspectors for each period for regression analysis: We selected inspectors whose biosolids' odor scores significantly correlated with either processing or weather data in each period in order to use their odor's scores as the dependent variable in regression. The following are the inspectors we chose in each period.

Yearly period: Pete and Carl.

Winter period: Pete, Carl, and Wilfred.

Summer period: Pete and Mike.

5. Choose significant variables correlated to selected inspector's odor scores in each period: Variables showing significant correlation to the inspector's odor scores were selected to use as independent variables in the regression analysis.

6. Create data set for regression: We needed to create sets of data for each inspector before running regressions by considering the number of data points, the multicollinearity between variables in the same data set, and the chronological consistency.

In terms of the number of data points, the variables that had either "outlier" or missing values can decrease the number of data points for the regression given the need for a complete data set. Thus, we needed to delete the "outlier" and "no data" values from the data set. The more "outlier" and "no data" values in the variables, the greater the number of data points that were excluded from the data set.

In terms of multicollinearity, StatTools, the statistical package we used, cannot run regressions with multicollinearity between independent variables in the same data set. Therefore, we separated variables showing multicollinearity to different data sets before running regression.

In terms of the chronological consistency for the processing variables, to run regression on the selected significant variables, we combined the appropriate variables into a data set taking into account realistic operational considerations. For example, the blanket depth data on day d-0 and lime addition on day d-4 could not be included in the same regression equation because in the real treatment process the blanket depth tanks processing occurs prior to the lime addition process.

7. Run regression: After we obtained all the data sets, we ran regressions on each data set and compared the regression equations based on the key measures discussed above (adjusted R square, coefficient signs, P values).

8. Apply additional techniques: We selected appropriate regressions regarding the discussion on section 4.2 from each period to improve the fitness of the regression equation by using additional techniques such as incorporating dummy variables and interaction variables. The regressions for which we applied these additional techniques are shown below.

Yearly period: Pete's yearly regression.

Winter period: Pete's winter regression, Carl's winter regression.

Summer period: Mike's summer regression.

9. Improve the candidate regression: After applying additional techniques, we selected the significant regressions for final regression's adjustment by adding more processing variables that we thought they might improved the fit of selected regressions. Those regressions were Carl's winter regression and Pete's winter regression.

10. Select the best regression: Finally, the best regression of each period was summarized to use in DCWASA biosolids management.

4.4 Additional Techniques

Two techniques, using dummy variables and allowing for interaction between variables, were applied to explain the part that was unexplained by the existing independent variables in the regressions. Table 4.1 shows the dummy and interaction variables.

Table 4.1: Dummy and Interaction Variables Used in Additional Techniques

dummy variables
When the sum of blanket depths was higher than 10.3 feet
When the sum of Fecl3 added was more than 1290 gallons
When the sum of polymer added was more than 200.05 lbs/dry ton
When lime added was lower than 308 lbs/dry ton
When the average temperature was higher than 77 º F
When the average temperature was lower than 43 º F
When the day that we observed the biosolids odor was Monday
interaction variables
Product of the sum of blanket depth and the sum of Fecl3 added
Product of the sum of polymer added and the blend ratio
Product of the sum of blanket depth and the sum of polymer added

The inspectors associated with good regressions using these techniques were:

Pete (yearly and winter data)

Carl (winter data)

Mike (summer data)

4.4.1 Dummy variables

Dummy variables were created to represent various processing and weather - related conditions. A value of 1 for the dummy variable meant that the condition was true on that particular day whereas a value of 0 meant that it was false. For example, if the condition that “when the maximum temperature was more than 77 º F” was set as a dummy variable, on the day that the maximum temperature was more than 77 º F, the value of the dummy variable was equal to 1, otherwise, it had a value of 0. In our regression analysis, we set six dummy variables for each day:

1. When the sum of blanket depths was higher than 10.3 feet (All B.D.>10.3)

2. When the sum of $FeCl_3$ was more than 1290 gallons (All $FeCl_3$>1290)

3. When the sum of polymer additions was more than 200.05 lbs/dry ton (All poly > 200.05)

4. When lime addition was lower than 308 lbs/dry ton (Lime<308)

5. When the average temperature was higher than 77 º F (avg.temp>77)

6. When the average temperature was lower than 43 º F (avg.temp<43)

7. When the day that we observed the biosolids odor was Monday (when it was Monday)

The following is an explanation for each of these dummy variables and the anticipated coefficient values.

1. “All B.D.>10.3”: This variable was created to capture the days when the sum of blanket depth was higher than usual. The value 10.3 feet came from the 80th percentile of this variable in 2002. We expected that this variable would have a

positive sign related to the increasing of the biosolids odor because the higher the blanket depth, the greater the biosolids odor.

2. “All $FeCl_3$>1290”: This variable was also created to observe the effect of the amount of $FeCl_3$ added when the total amount of $FeCl_3$ added was more than 1290 gallons. The value 1290 came from the 80^{th} of sum of $FeCl_3$ added in 2002. We assumed that this variable should have a positive sign regarding the correlation analysis between inspector’s odor data and processing data in Chapter 3.

3. “All poly > 200.05”: As stated in Chapter 2, polymer causes biosolids odor to increase. Thus, this dummy variable was set to capture the effect of the sum of polymer from the DAF and dewatering processes when more than 200.05 lbs/dry ton were added. This value was the 80^{th} percentile of the data from 2002. This dummy variable was expected to increase biosolids odor and this should have a positive sign.

4. “Lime<308”: As discussed in Chapter 2, the addition of lime can reduce the biosolids odor. This variable was created to observe the effect of lime on biosolids odor when the amount of lime added was lower than usual. The 308 lbs/dry ton came from the 20^{th} percentile of the lime data from 2002. We expected that this variable would have a positive sign and increase the odor levels for biosolids because an insufficient amount of lime was added.

5. “avg.temp>77”: This dummy variable was used in the yearly and the summer regression to analyze the reaction of excessively high-temperature days to DCWASA processes and to biosolids odor levels. The value of 77 º F was the 80^{th} percentile for the average temperature in 2002. We expected that the higher the

temperature, the greater the odor since there would be more favorable conditions for biological activity related to odor.

6. “avg.temp<43”: This dummy variable was used only in the winter regression to investigate the reaction of extremely low-temperature days on DCWASA processes to biosolids odor levels. The value of 43 º F was the 20th percentile for the average temperature in 2002. We expected the extremely low temperature would decrease biological activity and biosolids odor.

7. “when it was Monday”: Normally, contractors stop hauling biosolids during the weekend. Therefore, biosolids odor on Monday should be contrasted to Tuesday through Friday, since biosolids would have been kept in the bunker longer than biosolids on these other days. We expected that the sign of this variable would be positive.

4.4.2 Interaction Variables

Three interaction variables were also created for use in the regression analysis. The interaction variables were:

1. Product of the sum of blanket depth and the sum of $FeCl_3$ added,
2. Product of the sum of polymer added and the blend ratio, and
3. Product of the sum of blanket depth and the amount of polymer added.

These interaction variables were the product of variables that tend to increase odor. Thus, the product of these variables also expected to increase odor. Other interaction variables such as the product of the sum of $FeCl_3$ variable and the WPL added and the product of the polymer and the lime added were also considered but could not be used for regression analysis. In the case of the product of sum of $FeCl_3$

added and the WPL added, there were too many of outliers or missing observation values leading to too small a set of usable observations.

For the product of the polymer and the lime added, it is important to consider the sign of the coefficient for the interaction variable. For example, lime is a chemical that decreases odor and polymer is a chemical that increases biosolids odor. The product of these variables constituting an interaction variable could have a sign that was either positive or negative. An analysis of the correctness of the sign of this interaction variable is not as clear as the interaction variables on number 1, 2, and 3 shown above, which is the product of variables that tend to present the same coefficient sign. For example, in the interaction variable “Product of the sum of blanket depth and the amount of polymer added”, both of blanket depth and polymer variables by themselves tend to increase odor. Therefore, the production of these variables should also increase odor as well.

4.5 Abbreviation of Independent Variables Used in the Regression Analysis

The abbreviations of variables used in the regression analysis are shown below.

B.D. second. E = the blanket depth in secondary east

B.D. second. W. odd = the blanket depth in secondary west odd

B.D. second. W. even = the blanket depth in secondary west even

All B.D.- W. = sum of blanket depth in secondary west

All B.D. = sum of all blanket depth

$FeCl_3$ primary W. = $FeCl_3$ added in primary west

$FeCl_3$ primary E = $FeCl_3$ added in primary east

All $FeCl_3$ = All $FeCl_3$ added

WPL = waste pickle liquid added

Blended ratio = the blended ratio in blended mixing tank

DAF poly. = polymer added in DAF

DEWAT poly. = polymer added in dewatering process

All poly. = all polymer added

WASA # centrif. in service = DCWASA number of centrifuges in service

WASA # centrif. out service = DCWASA number of centrifuges out of service

Contr. # belt. in service = contractor number of centrifuges in of service

Contr. # centrif. in service = contractor number of centrifuges in of service

Lime = lime addition

4.6 The Regression Analysis

In this section, the summary of the best regressions from each period is summarized and discussed. Steps four to eight in Section 4.3, The Procedures to Create Odor Forecasting Models, were applied to the inspector's regression analysis in each period as shown in Figure 4.2.

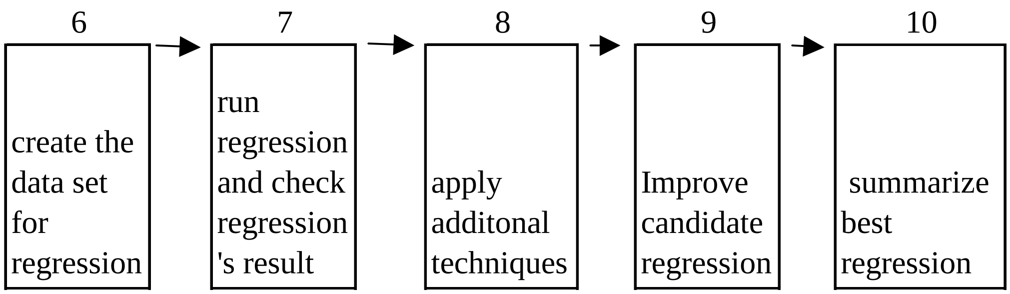


Figure 4.2: The Procedure to Run Regression Analysis

4.6.1 Final Regression Results

Four regressions were selected. All of them showed appropriate P values and coefficient values. The adjusted R square from the best of Carl's winter regression, Pete's winter regression, Mike's summer regression, and Pete's yearly regression were 0.6866, 0.58, 0.3629, and 0.3259, respectively. These best regressions from each period are shown in Table 4.2

Table 4.2: Summary of the Good Regression in each Period

Period	Regression equation	Adjusted R square
Winter regression	Carl's winter regression Y = 5.46243+0.00074X1+O.73544X2-0.02608X3 Y = Carl's odor scores X1 = Fecl3 added to primary east tanks on d-1 (gallons) X2 = number of DCWAS's centrifuges out of service on d-1 X3 = lime addition on d-1 (lbs/dry ton)	0.6866
	Pete's winter regression Y = -42.91+0.1147X1+O.377X2+0.231X3-0.0137X4+0.0014X5 Y = Pete's odor scores X1 = minimum temperature on d-3 X2 = blanket depth in secondary west even tanks on d-3 X4 = polymer addition in dewatering process on d-0 (lbs/dry ton) X4 = lime addition on d-0 (lbs/dry ton) X5 = average station pressure on d-2 (inches of HG)	0.58
Summer regression	Mike's summer regression Y = 19.0589-0.00194X1 Y = Mike's odor scores X1 = Fecl3 added to primary west tanks on d-1 (gallons)	0.3629
Yearly regression	Pete's yearly regression Y = 3.759+0.292X1-2.3513X2 Y = Pete's odor scores X1 = Sum of all blanket depths on d-2 (feet) X2 = number of contractor's centrifuges in service on d-1	0.3259

4.6.2 Yearly Regression Analysis

For the regression for all period of 2002, using appropriate variable identified in the correlation's analysis, we used Microsoft Excel to provide regression results. Two inspectors, Carl and Pete, were selected for the yearly regression based on their high correlated variables presented in Chapter 3 and the number of data points they had, 149 and 191 data points respectively. The best regression in this period was Pete's yearly regression shown as follows:

$$Y = 3.759 + 0.292X_1 - 2.351X_2$$

Where

Y = the Pete's odor score

X_1 = the sum of blanket depth on day d-2 (feet)

X_2 = Contractor number of centrifuges in service on d-1.

Pete's yearly regression was already discussed in section 4.6.1. In the yearly regression analysis, Carl's regression showed only the sum of blanket depth on d-0 as an independent variable with a positive coefficient sign as we expected in Chapter 2. However, this regression gave an adjusted R square value of 0.19 that was too low compared to Pete's yearly regression. Therefore we discussed only Pete's yearly regression.

4.6.2.1 Pete's Yearly Regression

From Pete's correlation analysis in 2002, we picked three processing variables highly correlated with Pete's odor score. These variables were:

1. All blanket depth on d-0, d-2, and d-3.

2. The number of contractor centrifuges in service on d-0 and d-1.

3. Lime addition on d-0 and d-1.

Pete's yearly correlation analysis with weather data in 2002 from Chapter 3 showed only high correlation on temperature variables (maximum, minimum, and average temperature) but with negative signs instead of positive signs. Therefore, we didn't include any weather data for Pete's yearly regression analysis.

Table 4.3 shows six different data sets and six different subsets (e.g. 1.1,…, 1.6) for which we ran regressions. The adjusted R square value of each subset is also shown in this table.

Table 4.3: The Data Set for Pete's Yearly Regression and Adjusted R Squares

	All B.D.	Contr. # centrif. in service	Lime	adj. R^2
set1	d-3	d-1	d-1	
1.1	d-3	d-1		0.3130
1.2		d-1	d-1	0.2980
1.3			d-1	0.1240
1.4		d-1		0.2970
1.5	d-3			0.1290
1.6	d-3	d-1	d-1	0.3110
set2	d-3	d-0	0	
2.1	d-3	d-0		0.3140
2.2		d-0	d-0	0.3050
2.3			d-0	0.1440
2.4		d-0		0.2950
2.5	d-3			0.1290
2.6	d-3	d-0	d-0	0.3180
set 3	d-2	d-1	d-1	
3.1		d-1	d-1	0.2990
3.2	d-2	d-1		0.3259
3.3			d-1	0.1240
3.4		d-1		0.2970
3.5	d-2			0.1360
3.6	d-2	d-1	d-1	0.3230
set 4	d-2	d-0	d-0	
4.1		d-0	d-0	0.3050
4.2	d-2	d-0		0.3220
4.3			d-0	0.1440
4.4		d-0		0.2950
4.5	d-2			0.1360
4.6	d-2	d-0	d-0	0.3250
set 5	d-3	d-1	d-0	
5.1		d-1	d-0	0.3040
5.2	d-3	d-1		0.3130
5.3			d-0	0.1440
5.4		d-1		0.2970
5.5	d-3			0.1300
5.6	d-3	d-1	d-0	0.3140
set 6	d-2	d-1	d-0	
6.1		d-1	d-0	0.3040
6.2	d-2	d-1		0.3260
6.3			d-0	0.1440
6.4		d-1		0.2970
6.5	d-2			0.1360
6.6	d-2	d-1	d-0	0.3260

The appropriate regression with the highest adjusted R square value of 0.3259 came from data set 3.2. The regression equation was:

$$Y = 3.759 + 0.292X_1 - 2.351X_2 \tag{1}$$

Where

Y = the Pete's odor score

X_1 = the sum of blanket depths on day d-2 (feet)

X_2 = number of contractor centrifuges in service on d-1.

Table 4.4 shows the result of regression from data set 3.2.We see that the coefficient of X_1 and X_2 correlate correctly with Pete's odor score as described in Chapter 2. That is the higher the blanket depth the higher the odor level and the greater the number of centrifuges, the lower the odor levels. Also, the P values of the intercept, X_1 and X_2 are lower than 0.05. Thus this equation is the most appropriate one for Pete's yearly regression.

Table 4.4: The Appropriate Pete's Yearly Regression Result from Data Set 3.2

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.578226382
R Square	0.334345749
Adjusted R Square	0.32597274
Standard Error	2.066571395
Observations	162

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	341.0711989	170.5355995	39.93137132	8.87713E-15
Residual	159	679.0440554	4.270717329		
Total	161	1020.115254			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	3.7591	1.2468	3.0150	0.0030	1.2967	6.2215
d-2 All B.D.	0.2919	0.1039	2.8092	0.0056	0.0867	0.4971
d-1 Contr. # centrif. in service	-2.3513	0.3467	-6.7817	0.0000	-3.0360	-1.6665

As discussed in Section 4.4, the additional techniques using the dummy variables and the interaction variables were applied. Table 4.5 shows the dummy variables and interaction variables after we assigned the day lagged of dummy and interaction variables considering the day lagged of independent variables existing on Table 4.4. For example, we used the dummy variables "All B.D.>10.3" and "All $FeCl_3$>1290" for d-2 after considering the existing independent variable "d-2 All B.D." in the regression because we assumed these variables probably occurred on the same day (see Figure 1.1 for details). Similar reassuming applied to other dummy and interaction variables. Note that the dummy variable "AVG. temp>77" was lagged the same number of day as blanket depth since we believed that temperature interacted with microorganisms in the blanket depth to cause odor. Also, the number of days lagged for the “Monday” dummy variable was d-0 since we assumed that the biosolids distributed to the field sites on Monday should have stronger odor than the other days.

Table 4.5: Additional Variables Used on Pete’s Yearly Regression

variables from Pete yearly regression	dummy variables added	interaction variables added
d-2 All B.D.	d-2 All B.D.>10.3	d-2 All B.D. * d-2 All Fecl3
d-1 Contr. # centrif. in service	d-2 All Fecl3>1290	d-1 All poly.* d-1 Blended ratio
d-1 Lime	d-1 All poly.>200.05	d-2 All B.D.*d-1 All poly
	d-1 Lime<308	
	d-2 AVG. temp>77	
	d-0 when it was Monday	

We used two data sets:

Data set 1: Variables from Pete’s yearly regression plus new dummy variables, and

Data set 2: The same data as number 1 as well as interaction variables that were added

Table 4.6 shows the equation after using these additional variables for Pete’s yearly regression. The regression showed an improvement in the adjusted R square value (with acceptable P values). The negative coefficient for the number of contractor centrifuges in service and lime and the positive coefficient for “All B.D.” are reasonable based on the discussion in Chapter 2. The positive coefficient for the interaction variable the “All B.D. and All FeCl3” followed our expectation that these two variables have the potential to increase biosolids odor. However, the sign of the coefficient for the dummy variable “Lime<308” and the dummy variable “when it was Monday” were negative instead of positive, unexpected results.

Table 4.6: The Regression Using the Additional Variables on Pete's Yearly Regression

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.6625	0.4389	0.4062	1.971072

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	6	313.02	52.17	13.4281	< 0.0001
Unexplained	103	400.1678	3.885124		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	8.42127	1.52972	5.50512	0.00000	5.38744	11.45510
d-1 Contr. # centrif. in service	-2.06116	0.47021	-4.38345	0.00003	-2.99372	-1.12860
d-1 Lime	-0.00952	0.00311	-3.06061	0.00282	-0.01569	-0.00335
d-2 All B.D.>10.3	1.43607	0.60377	2.37851	0.01923	0.23864	2.63351
d-1 Lime<308	-1.60839	0.76041	-2.11516	0.03683	-3.11648	-0.10030
d-0 when it was Monday	-1.12810	0.49700	-2.26982	0.02530	-2.11378	-0.14242
d-2 All B.D. * d-2 All Fecl3	0.00002	0.00001	1.70475	0.09126	0.00000	0.00003

In conclusion, the regression equation number one shown in Table 4.4 was taken as the best equation for Pete's yearly regression.

4.6.3 Winter Regression Analysis

As stated in Chapter 3, we used the data from January to March 2002 to represent the winter period. However, when we worked on regression analysis, the number of data points was greatly reduced due to outliers or missing observations. Therefore, we selected two more months: November and December, which also had average temperature lower than 50 ºF (corresponding to our definition of winter) to use in the winter period. We used Palisade's StatTools software[3] to run regressions from all the remaining data sets.

Three inspectors, Carl, Pete, and Wilfred, were selected for regression analysis in winter period. For two inspectors, Carl and Pete, their regression equations showed significant results as will be discussed below. Wilfred's regression showed a lower adjusted R square value and had a fewer number of data points than Carl's and Pete's regressions. Therefore, we didn't include an analysis of his regression in this section.

The inspector showing the best regression in this period was Carl. His regression had an adjusted R square of 0.6866 and was as follows:

$$Y = 5.46243 + 0.00074X_1 + 0.73544X_2 - 0.02608X_3$$

where

Y = Carl's odor data

X_1 = $FeCl_3$ addition in primary east tanks on d-1 (gallons)

X_2 = Number of DCWASA centrifuges out of service on d-1

X_3 = Lime addition on d-1 (lbs/dry ton of biosolids)

4.6.3.1 Carl's Winter Regression

We selected processing and weather variables that were highly correlated with Carl's odor data during the winter period using the results from Chapter 3. Those processing variables were:

[3] Palisade's StatTools software is the statistical software provided by DCWASA late in August 2003 after we finished the yearly regression analysis using Microsoft Excel.

1. All blanket depth on d-3, and d-4

2. $FeCl_3$ addition on primary east tanks on d-1

3. Polymer addition in DAF on d-0 and d-1

4. DCWASA number of centrifuges out of service on d-0 and d-1

5. Lime addition on d-0 and d-1

6. Average daily wind speed on d-1 and d-2

7. Average daily station pressure on d-1 and d-4

8. Daily snowfall on d-2 and d-4

Then we organized these variables into two data sets: data set1 and data set2 as shown in Table 4.7. Data set1 included all variables with an absolute correlation coefficient's value more than 0.5 and data set 2 include all the variables above. Thus, data set1 had more data points (41 data points) than data set2 (33 data points). This was because we selected a fewer number of variables but more highly correlated to avoid deleting too many outlier and missing observation values when we combined all variables into one set. For the WPL variable, there were many outliers in their data so we ignored it to maintain a reasonable number of observations.

Table 4.7: Data Set for Carl's Winter Regression

All variables chosen				data set 1 (41 data points)			data set 2 (33 data points)		
varialble	on day	correl with odor	note.	varialble	on day	correl with odor	varialble	on day	correl with odor
All B.D.	d-3	0.76		All B.D.	d-3	0.76	All B.D.	-	-
	d-4	0.82			d-4	0.82		d-4	0.82
FeCL3 E.	d-1	-0.25		-	-	-	FeCL3 E.	d-1	-0.25
WPL	d-2	0.39	too many outliers in this period	-	-	-	-	-	-
	d-3	0.25		-	-	-	-	-	-
DAF poly.	d-0	0.36		-	-	-	DAF poly.	d-0	0.36
	d-1	0.26		-	-	-		d-1	0.26
WASA # centrif.out service	d-0			WASA # centrif.out service	d-0	0.87	WASA # centrif.out service	-	-
	d-1	0.95			d-1	0.95		d-1	0.95
Lime	d-0			Lime	d-0	-0.905	Lime	-	-
	d-1	-0.907			d-1	-0.907		d-1	-0.907
average daily wind speed	d-1	-0.22		-	-	-	average daily wind speed	d-1	-0.22
	d-2	-0.32		-	-	-		d-2	-0.32
average daily station pressure	d-1	-0.25		-	-	-	average daily station pressure	d-1	-0.25
	d-4	-0.308		-	-	-		d-4	-0.308
daily snowfall	d-2	0.26		-	-	-	daily snowfall	d-2	0.26
	d-4	0.25		-	-	-		d-4	0.25

Tables 4.8 and 4.9 show the appropriate regressions from data set 1 and data set 2. Both of them gave a regression with significant variables (relative to the P value) and the appropriate coefficient signs. Table 4.8 illustrates the regression using only d-1 lime addition as an independent variable with an adjusted R square of 0.43. Table 4.9 shows the regression with the variables $FeCl_3$ added to the primary east tanks, number of DCWASA centrifuges in service, and lime addition as independent variables all lagged one day with an adjusted R square of 0.6866. For data set 2, we used several subsets to run regressions. The reason we ran several subsets was because

of multicollinearity between independent variables in data set 2 and the fact that StatTools could not handle this multicollinearity problem.

Table 4.8: Carl's Winter Regression Using Data Set 1

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.7086776	0.502224	0.4311131	0.983769

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F	p
Explained	1	31.47182	31.47182	33.55667	1.005E-06
Unexplained	39	36.57696	0.9378708		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	9.2073663	1.4709301	6.2595538	0.0000002	6.2321322	12.1826005
d-1 Lime	-0.0246055	0.0042476	-5.7928119	0.0000010	-0.0331971	-0.0160140

Table 4.9: Carl's Winter Good Regression Using Data Set 2

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.8461805	0.7160214	0.6866443	0.7595126

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F	p
Explained	3	42.180168	14.060056	24.373454	4.46356E-08
Unexplained	29	16.728923	0.5768594		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	5.46243	1.84155	2.96621	0.00598	1.69603	9.22883
d-1 Fecl3 E.	0.00074	0.00020	3.73199	0.00082	0.00033	0.00115
d-1WASA # centrif. out service	0.73544	0.26727	2.75173	0.01011	0.18882	1.28206
d-1 Lime	-0.02608	0.00530	-4.91890	0.00003	-0.03693	-0.01524

Data set 2 gives a higher adjusted R square (0.68 vs.0.431). We selected this regression as Carl's winter regression; that is, the equation:

$$Y = 5.46243 + 0.00074X_1 + 0.73544X_2 - 0.02608X_3 \quad (2)$$

where

Y = Carl's odor data

X_1 = $FeCl_3$ addition in primary east tanks on d-1 (gallons)

X_2 = DCWASA number of centrifuges out of service on d-1

X_3 = Lime addition on d-1 (lbs/dry ton of biosolids)

To improve upon this regression equation, additional dummy and interaction variables were added. We selected the dummy and interaction variables as shown in Table 4.10.

Table 4.10: Additional Variables Used to Improve Carl's Winter Regression

variables from Carl winter regression	dummy variables added	interaction variables added
d-1 FeCl3 primary E. d-1 WASA # centrif.out service d-1 Lime	d-1 All B.D.>10.3 d-1 All Fecl3>1290 d-1 All poly.>200.05 d-1 Lime<308 d-1 AVG. temp<43 (winter) d-0 when it was Monday	d-1 All B.D. * d-1 All Fecl3 d-1 All poly.* d-1 Blended ratio d-1 All B.D.*d-1 All poly

We first arranged the data into two sets as follows:

Data set 1: variables from the first and second columns of Table 4.10

Data set 2: Data set 1 plus the interaction variables in column 3 of Table 4.10

From the regression using data set 1, it was observed that the dummy variables did not improve the existing regression equation. For data set 2, the number of data points was reduced from 33 data in the data set 1 to 28 due to outliers and missing data when interaction variables were added. Table 4.11 shows the result of this regression indicating a better adjusted R square. However, the coefficient's sign for the dummy variable "Avg. temp< 43", which we expected to be negative, was positive and the coefficient's sign for the interaction variable "All poly*Blended ratio", which we expected a positive sign, gives a negative sign.

Table 4.11: The Regression from the Improvement on Carl's Winter Regression on Data Set 2

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.9074	0.8234	0.7730	0.657567

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	6	42.34828	7.058047	16.3232	< 0.0001
Unexplained	21	9.080288	0.432395		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	-1.60366	1.386466	-1.1567	0.2604	-4.48697	1.279657
d-1 FeCL3 E	0.000427	0.000213	2.0008	0.0585	-1.7E-05	0.00087
d-1wasa # centrif. out service	0.595885	0.253533	2.3503	0.0286	0.068635	1.123136
d-1 All poly>200.05	1.103172	0.439211	2.5117	0.0203	0.189783	2.016561
d-1 Lime<308	2.163783	0.401481	5.3895	< 0.0001	1.328856	2.998709
d-1 AVG temp<43	0.620727	0.303218	2.0471	0.0534	-0.00985	1.251303
d-1 All poly*d-1Blended ratio	-0.0218	0.010608	-2.0547	0.0526	-0.04386	0.000265

Because the number of data points in data set 2 was changed from the number of data points on Table 4.9 (28 vs. 33), we brought the significant variables from Table 4.9 to run regression on data set 2 in order to prove the fit of these significant variables on a new data set. Table 4.12 shows a higher adjusted R squares than the earlier regression on Table 4.9.

Table 4.12: Using Significant Variables from Carl's Winter Regression Running Multiple Regression Method for Data Set 2

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.8597	0.7391	0.7065	0.747744

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	3	38.00966	12.66989	22.6604	< 0.0001
Unexplained	24	13.41891	0.559121		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	5.83013	1.911358	3.0503	0.0055	1.885281215	9.774979
d-1 FeCL3 E	0.000536	0.00023	2.3306	0.0285	6.13237E-05	0.00101
d-1 wasa # centrif.out service	0.825788	0.266433	3.0994	0.0049	0.275897927	1.375678
d-1LIME	-0.02466	0.005642	-4.3706	0.0002	-0.03630011	-0.01301

In conclusion, the regression equation (2) from Table 4.9 is the best regression of Carl's winter regression. Compared to regression from Table 4.9, the regression shown in Table 4.12 provided the higher adjusted R squares but this regression brought the independent variables from Table 4.9 to run regression on the data set 2 with a smaller number of data points. Therefore, we chose the regression in Table 4.9 because it could explain the Carl's odor levels with the greater number of data points.

4.6.3.2 Pete's Winter Regression

From Pete's correlation analysis in Chapter 3, the following variables were identified:

1. All blanket depth on d-3, and d-4
2. $FeCl_3$ addition on primary east tanks on d-0 and d-1
3. Polymer addition in dewatering process on d-0 and d-1
4. DCWASA number of centrifuges out of service on d-0 and d-1

5. Lime addition on d-0 and d-1

6. Daily precipitation on d-0

7. Average daily station pressure on d-2

8. Maximum temperature on d-3

9. Minimum temperature on d-3

Table 4.13 shows three data sets used in the regression analysis. In the first data set, we took out the dewatering polymer variables on d-0 and d-1 because there were too many outliers in these variables. In data set 2, we used all of the nine variables above and the number of data points dropped from 61 in data set 1 to 46 data points in data set 2.

Table 4.13: Data Set for Pete's Winter Regression Analysis

All variables chosen			data set 1 (61 data points)			data set 2 (46 data points)		
varialble	on day	correl.	varialble	on day	correl.	varialble	on day	correl.
B.D. second. W. even	d-3 d-4	-0.32 -0.37	B.D. second. W. even	d-3 d-4	-0.32 -0.37	B.D. second. W. even	d-3 d-4	-0.32 -0.37
All Fecl3	d-0 d-1	0.65 0.53	All Fecl3	d-0 d-1	0.65 0.53	All Fecl3	d-0 d-1	0.65 0.53
DEWAT poly.	d-0 d-1	0.47 0.42	DEWAT poly.	- -	- -	DEWAT poly.	d-0 d-1	0.47 0.42
WASA # centrif.out	d-0 d-1	-0.39 -0.41	WASA # centrif.out	d-0 d-1	-0.39 -0.41	WASA # centrif.out	d-0 d-1	-0.39 -0.41
Lime	d-0 d-1	0.34 0.39	Lime	d-0 d-1	0.34 0.39	Lime	d-0 d-1	0.34 0.39
daily precipitation	d-0	0.42	daily precipitation	d-0	0.42	daily precipitation	d-0	0.42
average daily station pressure	d-2	0.24	average daily station pressure	d-2	0.24	average daily station pressure	d-2	0.24
t max	d-3	0.45	t max	d-3	0.45	t max	d-3	0.45
t min	d-3	0.43	t min	d-3	0.43	t min	d-3	0.43

Running regressions on data sets 1, the adjusted R square values obtained were 0.14. From data set 2, the adjusted R square value was 0.6356, which is shown in Table 4.14. The P value for each variable in Table 4.14 was lower than 0.1 and the sign of each coefficient was appropriate as stated in Chapter 2. However, this regression included both dewatering polymer variables on d-0 and d-1 in the same regression. Basically, this can be assumed that biosolids odor might be influenced by the amount of polymer added to the dewatering process on day d-1 and d-0. However, we discussed this assumption with biosolids supervisor and he said that this is not true. The dewatering polymer was added into the sludge after blend tank and before centrifuges. In fact, there was no retention point between the blend tank and centrifuges (dewatering process) to cause the multiple-day reaction of dewatering polymer to biosolids odor. Therefore, the effect of polymer to odor is from the amount of polymer addition in one day (Dorian, 2003).To see the effect when only one of these variables, either d-0 or d-1, we ran two additional regressions for data set 2 taking just one of these variables (along with the other ones). The results for the regression with d-1 the one-day lag for polymer showed high P values for the intercept and average daily station pressure (about 0.23) and a adjusted R square of 0.506. The results for the regression with the no lag dewatering polymer showed appropriate P values and the appropriate coefficient sign with an adjusted R square of 0.58. Table 4.15 shows the results of this better regression.

Table 4.14: Pete's Winter Regression on Data Set 2

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.83691477	0.7004263	0.6356537	0.880615

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F	p
Explained	6	64.215773	10.702629	13.2243	4.1303E-08
Unexplained	39	31.563302	0.8093154		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	-45.5047	23.5390	-1.9332	0.0605	-93.1168	2.1074
d-3 t min	0.1146	0.0166	6.8916	0.0000	0.0810	0.1483
d-3 B.D. second. W. even	0.3516	0.2085	1.6864	0.0997	-0.0701	0.7733
d-0 DEWAT poly.	0.2127	0.0592	3.5939	0.0009	0.0930	0.3324
d-1 DEWAT poly.	0.1246	0.0547	2.2753	0.0285	0.0138	0.2353
d-0 Lime	-0.0183	0.0050	-3.6887	0.0007	-0.0284	-0.0083
d-2 average daily station pressure	0.0015	0.0008	1.8444	0.0727	-0.0001	0.0031

Table 4.15: Pete's Winter Regression on Data Set 2 without d-1 DEWAT POLY Variable

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.79165269
R Square	0.62671398
Adjusted R Square	0.58005322
Standard Error	0.94542306
Observations	46

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	60.02608551	12.0052	13.43129	1.0522E-07
Residual	40	35.75299045	0.89382		
Total	45	95.77907595			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-42.91029	24.70844	-1.73667	0.09014	-92.84788	7.02730
d-3t min	0.11474	0.01748	6.56429	0.00000	0.07941	0.15007
d-3 blanket depth secondary west even	0.37761	0.21879	1.72592	0.09208	-0.06458	0.81980
d-0 DEWAT poly	0.23139	0.06159	3.75692	0.00055	0.10691	0.35586
d-0 Lime	-0.01378	0.00478	-2.88234	0.00632	-0.02345	-0.00412
d-2average daily station pressure	0.00140	0.00084	1.66965	0.10280	-0.00029	0.00309

From Table 4.15, we see that the coefficients for each variable as well as the P values are appropriate. Therefore, we selected the regression from Table 4.15 as the best regression for Pete's winter time period. This regression equation is as follows:

$$Y = -42.91 + 0.1147X_1 + 0.377X_2 + 0.231X_3 - 0.0137X_4 - 0.0014X_5 \quad (3)$$

Where

$Y =$ Pete's odor score

$X_1 =$ Minimum temperature on d - 3 (degrees F)

$X_2 =$ Blanket depth in secondary west even tanks on d - 3 (feet)

$X_3 =$ Polymer addition in dewatering process on d - 0 (lbs/dry ton)

$X_4 =$ Lime addition on d - 0 (lbs/dry ton)

$X_5 =$ Average daily station pressure d - 2 (inches Hg.)

Additional dummy and interaction variables were used for data set 2 based on this best regression shown above. The specific variables are indicated in Table 4.16

Table 4.16: Additional Variables Used to Improve Pete's Winter Regression

variables from Pete winter regression	dummy variables added	interaction variables added
d-3 B.D. second. W. even	d-3 All B.D.>10.3	d-3 All B.D. * d-3 All Fecl3
d-0 DEWAT poly.	d-3 All Fecl3>1290	d-0 All poly.* d-0 Blended ratio
d-0 Lime	d-1 All poly.>200.05	d-3 All B.D.*d-0 All poly
d-2 Avg. Daily pressure	d-0 Lime<308	
d-3 Tmin	d-3 AVG. temp<43 (winter)	
	d-0 when it was Monday	

The data was first arranges into two sets as follows:

Data set 1: variables from the first and second columns of Table 4.16 (46 data points), and

Data set 2: Data set 1 plus the interaction variables in column 3 (37 data points).

The regression using data set 1 as shown in Table 4.17, gave a higher adjusted R square value than the regression from Table 4.15 (0.608 vs. 0.58). The independent variable "d-2 Average daily station pressure" was replaced by the dummy variable "d-0 Lime<308" which had a negative coefficient. The negative coefficient of "Lime<308" was not expected because we assumed that the insufficient amount of lime added in dewatering should increase odor levels. Moreover, the P value of constant (intercept) was high.

Table 4.17: Pete's Winter Regression with Additional Variables on Data Set 1

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.8126	0.6603	0.6080	0.913419

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	5	61.064415	12.212883	14.0723	< 0.0001
Unexplained	40	34.714661	0.8678665		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	5.36541799	4.1822295	1.2829	0.2069	-3.08718303	13.818019
d-3 t min	0.10816154	0.017308	6.2492	< 0.0001	0.073180865	0.1431422
d-3 B.D. second. W. even	0.43667507	0.2162719	2.0191	0.0502	-0.00042671	0.8737768
d-0 DEWAT poly.	0.16877736	0.064679	2.6095	0.0127	0.038056133	0.2994986
d-0 Lime	-0.0264353	0.0089909	-2.9402	0.0054	-0.04460669	-0.0082639
d-0 Lime<308	-1.7822321	0.8836868	-2.0168	0.0505	-3.56822986	0.0037656

Table 4.18 shows the regression results for data set 2. This regression had a higher adjusted R square than the regression from Table 4.15 (0.6454 vs. 0.580) with the three significant variables: d-3 t min, d-0 DEWAT poly, d-0 lime. However, the P

value for the intercept was fairly high (0.47). Therefore, we preferred the regression shown in Table 4.15.

Table 4.18: Pete's Winter Regression with Additional Variables on Data Set 2

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.8682	0.7538	0.6454	0.873076

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	3	50.644752	16.881584	20.8261	< 0.0001
Unexplained	33	26.749771	0.8105991		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	1.10791589	1.5252388	0.7264	0.4727	-1.99520586	4.2110376
d-3 t min	0.11046495	0.0192501	5.7384	< 0.0001	0.071300288	0.1496296
d-0 DEWAT poly.	0.22636289	0.0582775	3.8842	0.0005	0.107796522	0.3449293
d-0 Lime	-0.0150659	0.0040619	-3.7090	0.0008	-0.02332991	-0.0068018

In conclusion, the regression from Table 4.15 was the best regression for Pete during the winter time period.

4.6.4 Summer Regression Analysis

As stated in Chapter 3, we used the data in June to August 2002 to represent the summer period in the correlation analysis. After we arranged the data set, we found that we didn't have enough observations for some inspectors due to outliers and missing observation values. Therefore, we included the month of September, which also had average temperature higher than 70 ºF (consistent with the other variables chosen). Note that the high temperatures of the summer months might affect biosolids odor levels by increasing the biological activity of the associated microorganisms.

In the summer period, we selected two inspectors (Mike and Pete) for regression analysis from the summer correlation analysis's results. The variables that correlated well with odor were then selected as part of the regression analysis for the summer.

Unfortunately, the weather variables (maximum, minimum, and average temperature) didn't show any significant effect and appropriate correlation with the selected inspectors in the summer period. However, we still believed that temperature variables should indirectly affect the microorganism in secondary tank because of the high correlation on between inspector's odor and blanket depth variables. Eventually, the best result came from Mike's summer regression shown below, with an adjusted R square value of 0.3629.

$$Y = 19.0589 - 0.00194 X_1$$

Where

Y = Mike's odor scores

X_1 = $FeCl_3$ addition in primary west tanks on day d-1 (gallons)

The following section shows the details of the regressions for inspector Mike during this summer period.

4.6.4.1 Mike's Summer Regression

The variables we selected from the correlation analysis for inspector Mike using the summer season were as follows:

1. All blanket depth on d-1 and d-3
2. $FeCl_3$ addition on primary west tanks on d-0 and d-1
3. WPL addition on d-1
4. The ratio in blended tank on d-3
5. Polymer addition in the DAF process on d-3

4. Contractor number of belt filter presses in service on d-0 and d-1

5. Lime addition on d-0 and d-3

6. Average daily wind speed on d-0

With the variables stated above, there was multicollinearity between $FeCl_3$ d-0 and $FeCl_3$ d-1. Thus we created two data sets from the selected variables above and put $FeCl_3$ d-0 into data set1 and $FeCl_3$ d-1 into set 2 with the remain variables included in both data sets; both sets had 31 data points. The best regression was found using data set 1. This regression is shown in Table 4.19.

Table 4.19: Best Summer Regression for Inspector Mike Using Data Set 1

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.619757	0.384099	0.362861	1.602868187

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F	p
Explained	1	46.46505	46.46505	18.08551091	0.000201
Unexplained	29	74.50641	2.569186		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	19.05894	3.471317	5.490407	6.49738E-06	11.9593	26.15859
d-1 FeCl3 primary W.	-0.00194	0.000457	-4.25271	0.000200653	-0.00288	-0.00101

Consequently, we see that the best regression was:

$$Y = 19.0589 - 0.00194 X_1 \tag{4}$$

where

Y = Mike's odor score

X_1 = $FeCl_3$ addition in primary west tanks on day d-1 (gallons)

This regression has an adjusted R square value of 0.3628 with significant P values (less than 0.05). The negative coefficient sign on $FeCl_3$ addition for primary west means that the more $FeCl_3$ added into the primary west tanks during summer, the more the biosolids odor decreased.

To improve upon this regression, we added dummy and interaction variables as shown in Table 4.20.

Table 4.20: Additional Variables Used in Improvement on Mike's summer Regression

variables from Mike summer regression	dummy variables added	interaction variables added
d-1 FeCl3 primary W.	d-1 All B.D.>10.3 d-1 All Fecl3>1290 d-0 All poly.>200.05 d-0 Lime<308 d-1 AVG. temp>77 d-0 when it was Monday	d-1 All B.D. * d-1 All Fecl3 d-0 All poly.* d-0 Blended ratio d-0 All B.D.*d-0 All poly

We created two data sets as follows:

Data set 1: variables from the first and second columns Table 4.20, and

Data set 2: Data set 1 plus the interaction variables from column 3.

Adding the dummy and interaction variables to the equation in Table 4.19 didn't give a better regression. However, the regression from data set 2 gave a higher adjusted R square value 0.6842 (as compared to Table 4.19). Unfortunately, the significant dummy variables added to the regression, "All poly>200.05" and "AVG.temp>77", had negative coefficients instead of positive as expected.

In summary, the best regression equation for the summer season was Mike's summer regression with adjusted R squares 0.3628 shown in Table 4.19.

4.7 Improving Candidate Regressions

Up to this point we used several approaches to search for a good regression (e.g. selected highly correlated independent variables, arranged data sets based on the processing timeframe). The last approach we used was that we selected the processing variables that we expected they could fit the regression and improve adjusted R squares. Figure 4.3 shows the procedures up to this point as well as what was next tried.

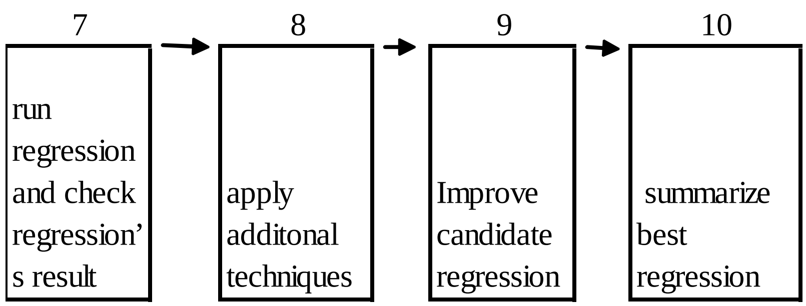


Figure 4.3: Final Step to Find Best Regression

Since Carl's winter regression and Pete's winter regression showed the highest adjusted R square values compared to other inspectors in all periods. These regressions were selected to improve their regression.

4.7.1 Final improvement on Carl's Winter Regression

Table 4.9 shows the best regression for inspector Carl in winter period. It shows the highest adjusted R square compared to other inspectors in all periods. We want to see whether there were any variables we haven't considered but could additionally be used to explain Carl's odor levels. Four data sets were created to this end.

Table 4.21: Data Used to Improvement Carl's Winter Regression

Existing independent variables	Data set added
d-1 FeCL3 primary E. d-1wasa centrif. # out of service d-1LIME	1) d-1 DAF poly 2) d-1 DAF poly, d-1 t max, and d-1 t min 3) d-1 DAF poly, d-2 t max, and d-2 t min 4) d-1 All B.D.

We considered four variables, DAF polymer, the sum of all blanket depth, maximum temperature, and minimum temperature, to additionally explain Carl's odor scores. After tried, the only variable that can improve this regression was maximum temperature on d-2. The result of this regression was shown on Table 4.22. It improved the adjusted R squares value from 0.68 (Table 4.9) to 0.7095 (Table 4.22) with the positive coefficient sign. The detail of the procedures to improve Carl's winter regression was shown in Appendix A.

Table 4.22: The Best Regression from Improvement on Carl's Winter Regression

Summary	Multiple R	R-Square	Adjusted R-Square	StErr of Estimate
	0.8636	0.7458	0.7095	0.731248

ANOVA Table	Degrees of Freedom	Sum of Squares	Mean of Squares	F-Ratio	p-Value
Explained	4	43.93683	10.98421	20.5418	< 0.0001
Unexplained	28	14.97226	0.534723		

Regression Table	Coefficient	Standard Error	t-Value	p-Value	Lower Limit	Upper Limit
Constant	4.315831	1.882495	2.2926	0.0296	0.459714	8.171948
d-1 FeCL3 primary E.	0.000691	0.000193	3.5754	0.0013	0.000295	0.001087
d-1wasa centrif. # out of service	0.728899	0.257344	2.8324	0.0085	0.201753	1.256045
d-1LIME	-0.02644	0.005109	-5.1755	< 0.0001	-0.03691	-0.01598
d-2 t max	0.028029	0.015464	1.8125	0.0806	-0.00365	0.059706

In conclusion, after trying other variables based on the equation two for inspector Carl during the winter period, the improvement of Carl's regression was shown below.

$$Y = 4.3158 + 0.0007X_1 + 0.7289X_2 - 0.0264X_3 + 0.0280X_5 \quad (5)$$

where

Y = Carl's odor data

X_1 = $FeCl_3$ addition in primary east tanks on d-1 (gallons)

X_2 = DCWASA number of centrifuges out of service on d-1

X_3 = Lime addition on d-1 (lbs/dry ton of biosolids)

X_5 = Maximum temperature on d-2 (^{o} F)

4.7.2 Final improvement on Pete's Winter Regression

The data set in Table 4.23 was created to check whether some significant processing variables were missing from the best of Pete's winter regression on Table 4.15. We considered the $FeCl_3$ addition on primary east tanks, number of DCWASA centrifuges out of services, blanket depth in secondary west tanks.

Table 4.23: Data Used for Improvement of Pete's Winter Regression

Existing independent variables	Data set
d-3 t min d-3 B.D.second. W. even d-0 DEWAT polymer d-0 Lime d-2 Average daily station pressure	1) Adding variables Fecl3 primary E. in d-0, d-1, d-2, and d-3 2) Adding WASA # centrif. out sevice on d-0 and d-1 3) Adding B.D. second. W. even on d-0, d-1, and d-2 4) Taking B.D. second. W. even on d-3 and t min on d-3 out

In summary, several tries using the processing variables on Table 4.23 were applied on the best of Pete's winter regression and none of them can improve this regression equation. (see the detail of the procedures to improve Pete's winter regression in Appendix B)

4.8 The Summary of Best Regressions

Several attempts to obtain best regressions have been performed. In each period, we first selected inspectors and variables that were highly correlated with their odor levels. Then the best regression from each period was modified by including the dummy and interaction variables. Up to this step, two best regressions regarding adjusted R squares were Carl's winter regression and Pete's winter regression. As described on step number nine, those two best regressions were taken for the final improvement. The processing variables that we expected to fit the existing independent variables were added into both regression equations in order to check the missing significant variables. Carl's regression were further improved by including one more variable, the maximum temperature on d-2, where increased the adjusted R square to 0.7095 as shown in Table 4.22.

Lastly, we summarized the best regression in each period as shown below.

Yearly period: the best of Pete's yearly regression from Table 4.4.

Winter period: the best of Carl's winter regression from Table 4.22.

Summer period: the best of Mike's summer regression from Table 4.19.

Chapter 5 Conclusions

This chapter concludes the major results from the regression analysis, and describes how the model can work for DCWASA as well as the problems during the processes, and future work.

5.1 The Major Results from the Regression Analysis

In Chapter 4, the best regression equation relative to highest adjusted R squares was obtained from Table 4.22 using the data of $FeCl_3$ addition in primary east tanks on d-1, DCWASA number of centrifuges out of service on d-1, lime addition on d-1, and maximum temperature on d-2 as independent variables to predict the Carl's odor score in winter. The regression equation is shown below.

$$Y = 4.3158 + 0.00069X_1 + 0.7288X_2 - 0.02644X_3 + 0.028X_4$$

where

Y = Carl's odor score

X_1 = $FeCl_3$ addition in primary east tanks on d-1 (gallons)

X_2 = DCWASA number of centrifuges out of service on d-1

X_3 = Lime addition on d-1 (lbs/dry ton of biosolids)

X_4 = Maximum temperature (° F)

The interpretation of the regression coefficients is the following:

The coefficient of "$FeCl_3$ addition in primary east tanks on d-1" can be interpreted as each gallon of $FeCl_3$ that is added into primary east tank one day before the day that Carl records biosolids odors can increase Carl's odor score by 0.00069 units all else being equal.

The coefficient of “DCWASA number of centrifuges out of service on d-1” can be interpreted as each of the DCWASA centrifuge out of service on one day before the day that Carl records biosolids odor can increase Carl’s odor score by 0.7288 units when other independent variables are held constant.

The coefficient of “Lime addition on d-1” can be interpreted as each pound per dry ton of lime added one day before the day that Carl observes biosolids odor can decrease Carl’s odor score by 0.02644 units when other independent variables are held constant.

The coefficient of “Maximum temperature on d-2” can be interpreted as each degree of maximum temperature two days before the day that Carl takes biosolids odor can increase Carl’s odor score by 0.028 units when other independent variables are held constant.

The following are the best regression for summer period and yearly period.

Summer period: the best of Mike’s summer regression with adjusted R squares 0.3628 from Table 4.19.

Yearly period: the best of Pete’s yearly regression with adjusted R squares 0.3259 from Table 4.4.

5.2 How the Models Can Work for DCWASA

The odor forecasting models developed in this thesis have two purposes:

First, they will help DCWASA management to better plan to which sites they want to level biosolids based on expected odor. This will enable them to more equitably allocate the malodorous effects and respond to potential complaints.

Secondly, these models will be used as part of a multi-objective optimization model (Gabriel and Sahakij, 2003) to balance processing and distribution costs with odor impacts. For example, the model will provide a forecast of biosolids odor on a daily basis by using the data of the significant variables in the best regression.

5.3 The Problems during the Statistical Modeling Process

The major problem in this thesis was the inconsistency of the inspectors' odor data. Since we used the odor data from several inspectors, different standards of biosolids odor were created by different inspector's odor perceptions. In the case that there was a discrepancy in the results of the analysis, such as a correlation coefficient's sign among inspectors, it was difficult to select the correct result because there was more than one standard.

The uncertainty in the duration of the process came from the capacity of the equipments. In many occasions there were events that could slow down the biosolids production, such as the broken lime mixer, the broken centrifuges, etc. The broken lime mixer or the broken centrifuges increased the duration of the biosolids production since the contractor, who had less capacity to produce biosolids, had to take responsibility for the biosolids production.

5.4 Future Work

The odor forecasting models developed in this thesis can be further developed by using more accurate measuring methods for biosolids odors. In this thesis, the inspectors' odor data that we used can roughly explain the effect of the processing and weather variables on the biosolids odor according to the several standards from different inspector's odor perceptions, the different of field sites' conditions that different inspectors took samples, etc. Using more accurate equipment to measure biosolids odor at DCWASA in terms of the chemical compounds producing odor can produce more objective biosolids odor data and can measure regularly regardless of inspectors or weather conditions. In addition, the more accurate biosolids odor measurement from equipment can help us to observe more precisely the sensitivity of the biosolids odor when it was affected by various processing and weather data.

Appendix A

The procedures for the final improvement on Carl's winter regression were shown below.

1. Add "d-1 DAF poly": From the best of Carl's winter regression (Table 4.9), we see that the variables "d-1 lime" and "d-1 WASA# centrif out service" were already presented. We added the variable "d-1 DAF poly", which mostly occurred at the same day or at least one day earlier than lime (see details in Chapter 2). We ignored other polymer variables, such as "All poly." and "DEWAT poly." due to outlier and missing data problems.

The result from using this additional polymer-related variable was no improvement as compared to the equation in Table 4.9. In particular, after we added "d-1 DAF poly" variable, it showed that the coefficient was insignificant (P value more than 0.5)

2. Add "d-1 DAF polymer", "d-1 t max", and "d-1t min": For reason similar to part number 1, we wanted to see the effect of temperature's variable (tmax and tmin) and DAF polymer when they were added into the regression to explain Carl's odor scores for the winter time period.

Again no improvement was observed relative to Table 4.9's result. This means that the amount of polymer added in DAF, the maximum temperature, and the minimum temperature on day d-1 cannot explain Carl's odor scores better than the existing independent variables on Table 4.9.

3. Add “d-1 DAF polymer”, “d-2 t max”, and “d-2t min”: We wanted to see the effect of temperature variables (tmax and tmin) one day before the lime was added and DAF polymer’s variable at the same day as the lime addition.

The regression was improved by increasing adjusted R squares from 0.6866 to 0.7095 as shown in Table 4.22.

4. Add “d-1 All B.D”: Considering blanket depth variables which refer to the same day or one day before $FeCl_3$ is added, we chose “d-1 All B.D.”, which had a high correlation (0.71), into the regression. From section 4.6.3.1, we recalled that we had already considered the d-3 and d-4 All B.D. variables but they were not meaningful regression variables.

Even though we added the variable showed significantly correlated with Carl’s odor score (0.71), it didn’t improve the regression on Table 4.9.

Appendix B

The procedures for the final improvement on Pete's winter regression were shown below.

1. Adding variables "$FeCl_3$ primary E." in d-0, d-1, d-2, and d-3: On Table 4.15, blanket depth variable ("d-3 B.D. second. W even") was presented. We assumed that the $FeCl_3$ addition into primary east tanks, which presented as a significant variable in the best of Carl's winter regression on Table 4.9, might explain the Pete's odor during winter as well. Therefore, the "$FeCl_3$ primary E." variables on d-0, d-1, d-2, and d-3 were added into the same data set as the best of Pete's winter regression on Table 4.15 to check their significance.

None of the selected "$FeCl_3$ primary E." variables above were significantly explained the Pete's odor scores in winter. "$FeCl_3$ primary E." variables can significantly explain the Carl's odor scores in winter period but cannot explain Pete's odor scores in winter period.

2. Adding "WASA # centrif. out sevice" on d-0 and d-1: From the best regression of Carl's winter regression, number of DCWASA centrifuges out of service was significantly explained Carl's odor scores. We took this variable to the best of Pete's winter regression to check how this variable explained Pete's odor score.

There was no improvement to add this variable into the best of Pete's winter regression. This significant variable in Carl's winter regression can not explain the Pete's odor scores.

3. Adding “B.D. second. W. even” on d-0, d-1, and d-2: Considering the independent variable “d-3 B.D. second. W. even” in the best of Pete’s regression in winter, this variable was not correlated with the day lagged of “d-0 DEWAT. Poly” and “d-0 Lime” because it was taking such a long period between the blanket depth step on d-3 to the lime and polymer step on d-0 (this means that it took almost 3 days for processing between these steps). Thus we tried adding “B.D. second. W. even”on d-0, d-1, and d-2, which make much more sense than d-3, into regression. However, to add these three variables into data set we needed to delete 2 more data points from the data set of the best of Pete’s winter regression. (from 46 data point to 44 data points).

Unfortunately, none of “B.D.second. W.even” on d-0, d-1, and d-2 can replace the “B.D.second. W. even on d-3”. That means the unreasonable day lagged between “d-3 B.D.second. W. even” and “d-0 Lime” still be existed. This problem made this regression equation lesser appropriate than the good regression on Carl’s winter regression.

4. Taking either “B.D. second. W. even” on d-3 and “d-3 t min” out: To make a good regression of Pete’s winter regression more reasonable, blanket depth variable and minimum temperature on d-3 were taken out to check how change on the regression equation.

The result was the adjusted R squares decreased from 0.58 to 0.16. That means these two variables are significantly explained Pete’s odor score in winter. As shown in number 3, we already tried to keep this variable but using different day lagged concerning the appropriateness of regression. However, none of them can replace the variable “B.D. second. W.even on d-3”.

REFERENCES

DCWASA. 2003. *General Information*. Internet on-line. <http://www.dcwasa.com/about/gen_overview.cmf>

Devore, J.L. 1995. *Probability and Statistics for Engineering and the Sciences*. USA: Duxbury Press.

Evanylo, G.K. 2001. *Land Application of Biosolids for Agricultural Purposes in Virginia.* Virginia Cooperative Extension Publication 452-300.

Gabriel, S.A., Sahakij, P., Peot, C. 2003. *A Multiobjective Optimization Approach for Distribution and Processing of Biosolids Products*. Department of Civil and Environmental Engineering, University of Maryland, College Park.

Kim, H., Murthy, S.N., McConnell, L., Peot, C., Ramirez, M., Strawn, M. 2001a. *Examination of Mechanisms for Odor Compound Generation during Lime Stabilization*, WEFTEC 2001.

Kim, H., Murthy, S.N., McConnell, L., Peot, C., Ramirez, M., Strawn, M. 2001b. *Characterization of Wastewater and Solids Odors Using Solid Phase Microextraction at a Large Wastewater Treatment Plant*. WEF Residuals and Biosolids Management Conference, San Diego, CA

Liu, D. H.F. 2000. *Wastewater Treatment*. USA: Lewis Publishers.

Maddala, G. S. 1977. *Econometrics*. USA: McGraw-Hill, Inc.

Metcalf &Eddy Inc. 1991. *Wastewater Engineering: Treatment Disposal Reuse*. New York: McGraw-Hill.

Murthy, S.N., Sadick, T., Baily, W., Peot, C., Tolbert, D., and Strawn, M. 2001. *Mitigaition of Odors from Lime Stabilized Biosolids.* WEF Residuals and Biosolids Management Conference, San Diego, CA

Murthy, S.N., Kim, H., Glindermann, D., Sadick, T., Peot, C., Bailey, W., McConnell, L., Novak, J. 2002a. *Mechanisms for Odor Generation during Lime Stabilization.* Water Research (submitted).

Murthy, S.N., Peot, C., Ramirez, M., Kim, H., Higgins, M., North, J. 2002b. *Characteristics and Control of Odors from Lime-stabilized and Digested Biosolids.* Annual Meeting for Biosolids Management, Austin, TX

Private communication with Peot, C. in 2003.

Private communication with Ramirez, M. in 2003.

Private communication with Tolbert, D. in 2003.

Vesilind, P.A. 2003. *Wastewater Treatment Plant Design.* USA: Water Environment Federation.

Winston, W.L. 1994. *Operation Research Application and Algorithms*. Belmont, California: Wadsworth Publishing Company.