The National Stormwater Quality Database, Version 1.1

A Compilation and Analysis of NPDES Stormwater Monitoring Information

Alex Maestre and Robert Pitt
Department of Civil and Environmental Engineering
The University of Alabama
Tuscaloosa, AL 35294

Center for Watershed Protection
Ellicott City, Maryland

September 4, 2005

U.S. EPA
Office of Water
Washington, D.C.

Abstract

The National Stormwater Quality Database v. 1.1 (NSQD) contains selected water quality information from the monitoring carried out as part of the U.S. EPA's National Pollutant Discharge Elimination System (NPDES) Phase 1 stormwater permit applications and subsequent permits, during the period of 1992 to 2002. This database contains about 3,765 events from 360 sites in 65 communities from throughout the U.S. For each site, more additional data, including the percentage of each land use in the catchment, the total area, the percentage of impervious cover, the geographical location, and the season, has been included in the database. Information about the characteristics of each event is also included. Total precipitation, precipitation intensity, total runoff and antecedent dry period are also included, if collected. The database only contains information for samples collected at drainage system outfalls; in-stream samples (which were a component of some state programs) were not included in the database, although some outfalls were located in open channel conveyances.

The first phase requirements of the federal stormwater permit program were first published in the Federal Register by the EPA in 1987 and was initially applied to large cities (>100,000 in population), while Phase II of the stormwater permit program was applied to all urban areas as of early 2003. This program requires significant changes in how stormwater is to be managed. Historical approaches only examined drainage issues, while the new regulations also require consideration of water quality issues.

There are a number of commonly accepted notions that are used by stormwater managers and regulators that can have major impacts on local costs and program effectiveness. This research report examines a number of these potential misconceptions to see how well they hold up under a comprehensive set of actual monitoring data collected throughout the U.S. as part of the Phase I stormwater permit program. This research report is mostly comprised of the major sections of the Ph.D. dissertation prepared by Alex Maestre in partial fulfillment of his degree requirements in the Department of Civil and Environmental Engineering at the University of Alabama.

Contents

List of Tables

List of Figures

Acknowledgements

Many people and institutions need to be thanked for their help on this research project. Project support and assistance from Bryan Rittenhouse, the US EPA project officer from the Office of Water, is gratefully acknowledged. The many municipalities who worked with us to submit data and information were obviously crucial and the project could not have been conducted without their help. Finally, the authors would like to thank a number of graduate students at the University of Alabama (especially Veera Rao Karri, Sanju Jacob, Sumandeep Shergill, Yukio Nara, Renee Morquecho, and Soumya Chaturvedula) and employees of the Center for Watershed Protection (Ted Brown, Chris Swann, Karen Cappiella, and Tom Schueler) for their careful work on this project.

Chapter 1: Introduction

The first phase of the federal stormwater permit program was first published in the Federal Register by the EPA in 1987 and was initially applied to large cities (>100,000 in population), while Phase II of the stormwater permit program was applied to all urban areas as of early 2003. This program requires significant changes in how stormwater is to be managed. Historical approaches only examined drainage issues, while the new regulations also required consideration of water quality issues. Unfortunately, some professionals involved with stormwater management may not have an adequate understanding of stormwater characteristics, including its effects, and treatability. As an example, there are a number of commonly accepted notions that are used by stormwater managers and regulators that can have major impacts on local costs and program effectiveness. This research report examines a number of these notions to see how well they hold up under a comprehensive set of actual monitoring data collected throughout the U.S. as part of the Phase I stormwater permit program. This research report also includes a predictive tool that can assist stormwater managers in predicting expected stormwater conditions for local areas.

Researchers from the University of Alabama and the Center for Watershed Protection assembled a large database of stormwater characteristics, the National Stormwater Quality Database (NSQD), as part of an EPA-funded section 104(b)3 project from the Office of Water. This is the largest collection of information on stormwater characteristics ever assembled for US conditions. The research described in this report used this information to test the validity of several commonly accepted notions concerning stormwater, and produced a statistical tool that hopefully can assist stormwater managers and regulators. In addition, many suggestions concerning monitoring strategies for stormwater are summarized, based on the experiences of many of the Phase I permitted communities. The cumulative value of the monitoring data collected over nearly a ten-year period from more than 200 municipalities throughout the country has a great potential in characterizing the quality of stormwater runoff and comparing it against historical benchmarks.

The data set received a comprehensive quality assurance/quality control review, based on reasonableness of data, extreme values, relationships among parameters, sampling methods, and a review of the analytical methods. The statistical analyses were conducted at several levels. Probability plots were used to identify range, randomness and normality. Multivariate analyses were also utilized to characterize significant factors affecting the data patterns. The master data set was also evaluated to develop descriptive statistics, such as measures of central tendency and standard errors. Testing was done for regional and climatic differences, the influences of land use, and the effects of storm size, drainage area and season, among other factors.

This National Stormwater Quality Database (NSQD), in its first version presented here, is not intended for comprehensive characterization purposes for all conceivable situations and to replace the need for all characterization monitoring. Some communities may have obvious unusual conditions, or adequate data may not be available in the database for their region. In these conditions, site specific local outfall monitoring may be needed. In addition, stormwater monitoring will continue to be needed for other purposes in many areas having, or anticipating, active stormwater management programs (especially when supplemented with other biological, physical, and hydrologic monitoring components). These new monitoring programs should be designed specifically for additional objectives, beyond simple characterization. These may include receiving water assessments to understand local problems, source area monitoring to identify critical sources, treatability tests to verify performance of stormwater controls for local conditions, and assessment monitoring to verify the success of local stormwater management approaches (including model calibration and verification). In many cases, however, the resources being spent for conventional outfall monitoring could be more effectively spent to better understand many of these other aspects of an effective stormwater management program.

Report Organization

This report is divided into nine chapters and five appendices. Chapter 2 describes the National Stormwater Quality Database (NSQD). Chapter 3 describes the QA/QC procedures used during the collection of data and creation of the database, including an evaluation of alternative methods to address the presence of non-detected values. Chapter 4 addresses the hypothesis concerning the probability distributions most appropriate for the stormwater constituents. Chapter 5 describes the results of the investigations relating constituent concentrations to main factors and interactions of parameters described in the site description and hydrologic information sections of the database. Chapter 6 presents the results from the "first flush" analysis. Chapter 7 presents detailed results of the statistical tests used to develop predictive models of stormwater characteristics affected by geographical location and land use. Chapter 8 presents an example of how the data in the NSQD can be used to estimate the concentration of stormwater constituents for Maryland and Virginia (the region best represented in the database). Chapter 9 presents the conclusions and recommendations of this research.

Chapter 2: The National Stormwater Quality Database (NSQD) Description

Introduction

The National Stormwater Quality Database (NSQD) was prepared by the University of Alabama and the Center for Watershed Protection under 104(b)3 funding from the U.S. Environmental Protection Agency (EPA). The NSQD is a spreadsheet database and supporting documents describing the monitoring efforts of 65 communities from throughout the U.S. that are larger than 100,000. The monitoring period covered by the NSQD is from 1992 to 2002.

Several efforts have been performed in the past to describe the water quality characteristics of stormwater constituents at different locations. The importance of this EPA-sponsored project is based on the scarcity of nationally summarized and accessible data from the existing U.S. EPA's NPDES (National Pollutant Discharge Elimination System) stormwater permit program. There have been some local and regional data summaries, but little has been done with nationwide data. A notable exception is the Camp, Dresser, and McGee (CDM) national stormwater database (Smullen and Cave 2002) that combined historical Nationwide Urban Runoff Program (NURP) (EPA 1983) data, available urban U.S. Geological survey (USGS), and selected NPDES data. Their main effort had been to describe the probability distributions of these data (and corresponding EMCs, the event mean concentrations). They concluded that concentrations for different land uses were not significantly different, so all their data were pooled into a single urban land use category.

The Clean Water Act (CWA) of 1972 was the first major national regulation in the U.S. requiring control of conventional point source discharges of water pollutants (affecting municipal and industrial discharges). Section 208 also provided the capability to implement stormwater management plans at the regional level. In 1976, the EPA enlarged the planning initiative through the "Section 208: Areawide Assessment Procedures Manual". However, in the late 1970s, some problems arose with the 208 planning projects due to inadequate data and lack of technological development (Whipple, as quoted by Pitt, *et al*. 1999).

Between 1978 and 1983, the EPA conducted the Nationwide Urban Runoff Program (NURP) that examined stormwater quality from separate storm sewers in different land uses (EPA 1983). This program studied 81 outfalls in 28 communities throughout the U.S. and included the monitoring of approximately 2,300 storm events. NURP is still an important reference for water quality characteristics of urban stormwater; however, the collected data poorly represented the southern area of the country and was focused mainly in residential and mixed land use areas. Since NURP, other important studies have been conducted that characterize stormwater. The USGS created a database with more than 1,100 storms from 98 monitoring sites in 20 metropolitan areas. The Federal Highway Administration (FHWA) analyzed stormwater runoff from 31 highways in 11 states during the 1970s and 1980s. Strecker (personal communication) is also collecting information from highway monitoring as part of a current NCHRP (National Cooperative Highway Research Program) funded project. The city of Austin also developed a database having more than 1,200 events.

Other regional databases also exist for U.S. data, mostly using local NPDES data. These include the Los Angeles area database, the Santa Clara and Alameda County (California) databases, the Oregon Association of Clean Water Agencies Database, and the Dallas, Texas, area stormwater database. These regional data are included in the NSQD. However, the USGS and historical NURP data are not included in the NSQD due to lack of consistent descriptive information for the older drainage areas and because of the age of the data from those prior studies. Much of the NURP data is available in electronic form at the University of Alabama's student American Water Resources Association web page at: http://www.eng.ua.edu/~awra/download.htm.

Outside the U.S., there have been important efforts to characterize stormwater. In Toronto, Canada, the Toronto Area Watershed Management Strategy Study (TAWMS) was conducted during 1983 and 1984 and extensively monitored industrial stormwater, along with snowmelt in the Toronto urban area, for example. Numerous other investigations in South Africa, the South Pacific, Europe and Latin America have also been conducted over the past 30 years, but no large-scale summaries of that data have been prepared. About 4,000 international references on stormwater have been reviewed and compiled since 1996 by the Urban Wet Weather Flows literature review team for publication in *Water Environment Research* (most recently by Clark, *et al*. 2001, 2002, 2003, 2004). An overall compilation of these literature reviews is available at: http://www.eng.ua.edu/~rpitt/Publications/Publications.shtml. These reviews include short summaries of the papers and are organized by major topics. Besides journal articles, many published conference proceedings are also represented (including the extensive conference proceedings from the 7th International Conference on Urban Storm Drainage held in Germany in 1996, the 8th International Conference on Urban Storm Drainage held in Sydney, Australia, in 1999, the 9th International Conference on Urban Storm Drainage held in Portland, OR, in 2002, and the Urban Water Systems Modeling conference series for the Toronto meetings organized by Computational Hydraulics, Inc., amongst many other specialty conferences).

In 1987, the amendments to the CWA established a two-phase program to regulate 13 classes of stormwater discharges. Two of these classifications were discharges from large and medium-sized Municipal Separate Storm Sewer Systems. A large MS4 serves an urban population of 250,000 or more, while a medium MS4 serves communities between 100,000 and 250,000. EPA set up a permit strategy for communities complying with NPDES requirements. Monitoring data from this program have been included in some databases. The CDM National Stormwater Runoff Pollution database included 816 NPDES storm events in a database that totals approximately 3,100 events. The Rouge River National Wet Weather Demonstration Program office in Detroit included their NPDES data in their database (Smullen and Cave 2003).

Another important effort has been the development of the National Stormwater Best Management Practices Database (http://www.bmpdatabase.com). This database was created with the purpose to evaluate the performance and effectiveness of stormwater control practices, frequently labeled “best management practices,” or BMP’s. Detention ponds, street cleaning, and hydrodynamic devices are examples of BMPs (ASCE/EPA 2000).

Data Collection

Data from 3,765 storm events at 360 monitoring sites were collected and are stored in version 1.1 of the NSQD. This version contains the results of approximately one fourth of the total number of communities that participated in the Phase I NPDES stormwater permit monitoring activities.

According to the published sampling guidance (40 CFR 122.21) for the permit application, each community was required to sample at least a residential, a commercial and an industrial watershed. At least three samples should be collected every year at each location. Each storm should be at least one month apart and have at least a 3 days antecedent dry period. Only samples from rain events greater than 0.1 inches, and close to the annual mean conditions, were considered valid for the analysis. It was required to collect a composite sample with subsamples collected during the first three hours of the event. An additional grab sample was required during the first 30 minutes of the event to evaluate the “first flush” effect. “First flush” refers to the hypothesis that the concentrations of stormwater constituents are higher at the beginning of the discharge event than during the complete event. Designated states were able to modify some of these sampling requirements to better address local concerns.

Most communities were required to submit annual reports describing the sampling locations and procedures, the equipment, and the quality control and quality assurance (QA/QC) procedures used during the sampling and analysis of the samples, the analytical methods used in the laboratory, and problems encountered during the sample collection. The reports also included the results of the chemical analyses performed by the laboratories.

Figure 1 is a map showing the 65 communities and 17 states included in the first version of the NSQD. The EPA-funded project was intended to focus on the Chesapeake Bay area and parts of the southern U.S. (specifically Birmingham, AL, and Atlanta, GA) as a demonstration of the usefulness of the data. However, it was possible to obtain some data from other parts of the country during the project period and these data were incorporated in the database, allowing some regional analyses. States representing most of the samples included Virginia (24%) and Maryland (13%). The states with low numbers of observations included Pennsylvania, Massachusetts, and Indiana.

Figure 1 also shows the EPA Rain Zones. Each zone corresponds to a geographical region with similar climatic conditions (EPA 1986). There is at least one community per rain zone indicating some geographical representation for the entire country. However, Table 1 indicates that most of the samples were collected west, south and east of the continental part of the country, with few of the large amounts of data from EPA Rain Zone 1 included in the database. EPA Rain Zones 8 and 9 have sparse available data from the Phase I monitoring program, due to few large cities in these areas.

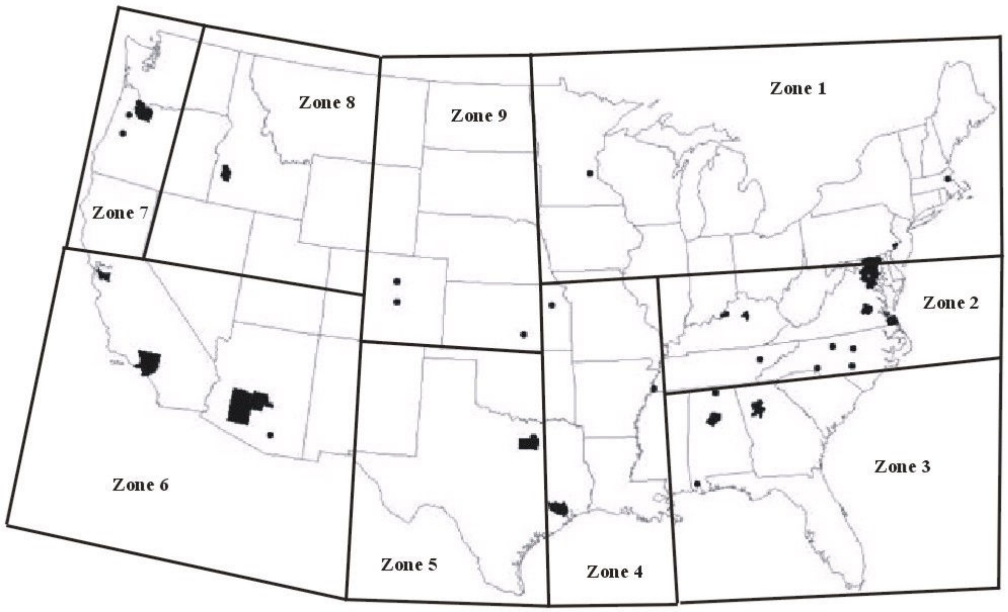


Figure 1. Communities included in the NSQD version 1.1 by rainfall zones

Table 1. Total Samples and Sites by EPA Rain Zone

EPA Rain Zone	Total Samples	Percentage of Samples	Number of Communities	Number of Sites
1	69	1.8	2	12
2	2000	53	28	185
3	266	7.1	8	30
4	212	5.6	4	21
5	485	13	9	33
6	356	9.5	4	30
7	229	6.1	6	28
8	24	0.63	1	4
9	124	3.3	3	17

Each site in the database corresponds to an outfall where the runoff produced in the watershed is discharged. During the monitored events, samples were collected to identify the characteristics of the stormwater being discharged. According to the land use of the watershed, each site was classified as residential, commercial, industrial, open space, freeway, or mixed. When a single land use was not identified for the watershed, then the site was considered mixed, with a predominant land use. Table 2 indicates the total number of sites included in the database, separated by land use.

Table 2. Total Samples and Sites by Land Use

Land use	Number of Sites	Percentage	Number of Events	Percentage
Residential	111	31	1042	28
Mixed Residential	44	12	611	16
Commercial	51	14	526	14
Mixed Commercial	29	8.1	325	8.6
Industrial	54	15	566	15
Mixed Industrial	22	6.1	249	6.6
Institutional	1	0.3	18	0.5
Open Space	10	2.8	49	1.3
Mixed Open Space	13	3.6	168	4.5
Freeways	22	6.1	185	4.9
Mixed Freeways	3	0.8	26	0.7

About one third of the sites included in the database correspond to residential areas, another third is shared by commercial and industrial land uses. The remaining third correspond to freeways, open space, institutional and all the mixed land uses. Several schools were identified in the sites, however only one site was considered 100% institutional.

Summary of U.S. NPDES Phase I Stormwater Data in the NSQD

Table 3 is a summary of selected data collected and entered into the database. The data are separated into 11 land use categories: residential, commercial, industrial, institutional, freeways, and open space, plus mixtures of these land uses. Summaries are shown for the major land use areas and for the total data set combined. The full database includes all of the data. The total number of observations and the percentage of observations above the detection limits are also shown on this summary table. In general, the coefficient of variation (COV) values range from 1.0 to 2.0 for the majority of pollutants across all major land uses.

The following sections describe the structure of the full database and present some findings. The findings presented are focused on specific issues and are illustrated using small portions of the complete database to minimize the effects of other interacting factors (such as using data from a single region and land use to show the effects of sampling methods, for example). Later sections of this report present more comprehensive discussions of the data that do consider interactions of the many factors available in the database.

Database Structure

The database has five major sections: General Information, Items Description, Constituents and Parameters, and the Database itself. In addition, detailed site information along with aerial photographs and topographic maps is provided for each municipality and monitoring location. Each of the sections is a tab in the bottom part of the spreadsheet.

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1

	Area (acres)	% Impervious	Precipitation Depth (in)	Runoff Depth (in)	Conductivity (µS/cm @25°C)	Hardness (mg/L CaCO3)	Oil and Grease (mg/L)	pH	Tempe-rature (C)	TDS (mg/L)	TSS (mg/L)	BOD_5 (mg/L)	COD (mg/L)
Overall Summary (3765)													
Number of observations	3765	2209	3316	1495	685	1082	1834	1665	861	2956	3493	3105	2750
% of samples above detection	100	100	100	100	100	98.7	66.1	100	100	99.0	97.9	96.2	98.4
Median	57.3	50.0	0.48	0.15	121	38.0	4.3	7.5	16.5	80	59	8.6	53
Coefficient of variation	3.7	0.4	1.0	1.9	1.6	1.4	9.7	0.1	0.4	3.4	1.8	7.4	1.1
Residential (1042)													
Number of observations	1042	614	919	372	104	215	483	286	181	814	978	908	748
% of samples above detection	100	100	100	100	100	100	54.9	100	100	99.1	98.3	97.1	98.7
Median	57.3	37.0	0.48	0.10	102	32.0	4.0	7.2	17.0	72.0	49	9.0	54.5
Coefficient of variation	4.8	0.4	1.0	1.5	1.6	1.1	7.8	0.1	0.4	1.1	1.8	1.5	0.93
Mixed Residential (611)													
Number of observations	611	278	491	262	105	168	283	333	137	491	582	549	465
% of samples above detection	100	100	100	100	100	98.2	70.3	100	100	99.2	98.3	94.2	99.6
Median	150.8	44.9	0.53	0.12	112	40.0	4.0	7.50	15.5	86	66	7.8	43
Coefficient of variation	2.1	0.3	0.8	1.3	1.2	1.1	2.6	0.1	0.3	5.2	1.6	1.3	1.2
Commercial (527)													
Number of observations	527	284	462	146	78	156	331	191	98	418	503	452	393
% of samples above detection	100	100	100	100	100	100	71.9	100	100	99.5	95.2	97.6	98.5
Median	38.8	84.5	0.42	0.29	107	36.5	4.6	7.4	16.0	72	43	11.0	58
Coefficient of variation	1.2	0.1	1.0	1.0	1.0	1.1	3.0	0.1	0.4	1.9	2.0	1.1	1.0
Mixed Commercial (324)													
Number of observations	324	237	305	118	59	98	134	156	98	265	297	277	267
% of samples above detection	100	100	100	100	100	99.0	79.9	100	100	99.6	99.7	98.9	99.6
Median	75.0	60.0	0.47	0.28	100	36.0	5.0	7.60	14.5	69.5	54.5	9.0	60
Coefficient of variation	1.4	0.3	1.0	0.9	0.8	1.8	2.9	0.1	0.4	1.9	1.3	1.7	1.0
Industrial (566)													
Number of observations	566	292	482	215	102	132	315	248	140	431	521	455	386
% of samples above detection	100	100	100	100	100	96.2	64.8	100	100	99.5	97.7	95.4	99.0
Median	39.5	75.0	0.50	0.16	139	39.0	4.8	7.50	17.9	86	81	9.0	58.6
Coefficient of variation	1.1	0.3	0.9	1.2	1.3	1.5	11.8	0.1	0.3	3.6	1.6	10.0	1.2

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1 – *Continued*

	Area (acres)	% Impervious	Precipitation Depth (in)	Runoff Depth (in)	Conductivity (µS/cm @25ºC)	Hardness (mg/L CaCO3)	Oil and Grease (mg/L)	pH	Tempe-rature (C)	TDS (mg/L)	TSS (mg/L)	BOD_5 (mg/L)	COD (mg/L)
Mixed Industrial (218)													
Number of observations	218	118	193	117	56	75	72	152	57	186	207	178	175
% of samples above detection	100	100	100	100	100	93.3	80.6	100	100	99.5	100	95.5	98.9
Median	168.0	44.0	0.45	0.29	126	29.3	9.0	7.70	18.0	90	82	7.5	39.9
Coefficient of variation	1.8	0.3	0.9	1.2	0.8	0.6	1.8	0.1	0.3	0.8	1.4	1.8	1.2
Institutional (18)													
Number of observations	18	18	17	14						18	18	18	18
% of samples above detection	100	100	100	100						100	94.4	88.9	88.9
Median	36.0	45.0	0.18	0.00						52.5	17	8.5	50
Coefficient of variation	0	0	0.9	2.1						0.7	0.83	0.7	0.9
Freeways (185)													
Number of observations	185	154	182	144	86	127	60	111	31	97	134	26	67
% of samples above detection	100	100	100	100	100	100	71.7	100	100	99.0	99.3	84.6	98.5
Median	1.6	80.0	0.54	0.41	99	34.0	8.0	7.10	14.0	77.5	99	8	100
Coefficient of variation	1.4	0.13	1.1	1.7	1.0	1.9	0.6	0.1	0.4	0.8	2.6	1.3	1.1
Mixed Freeways (26)													
Number of observations	26		26		21	12	20	17	17	15	23	23	15
% of samples above detection	100		100		100	100	100	100	100	100	100	100.0	100.0
Median	63.1		0.47		353	83	4.5	7.7	16.0	177	88	8.2	47
Coefficient of variation	0.7		0.8		0.6	0.3	1.8	0.1	0.3	0.4	1.1	1.2	0.5
Open Space (49)													
Number of observations	49	37	41	11	2	8	19	19	2	45	44	44	43
% of samples above detection	100	100	100	100	100	100	36.8	100	100	97.8	95.5	86.4	76.74
Median	85	2.0	0.52	0.05	113	150	1.3	7.70	14.6	125	48.5	5.4	42.1
Coefficient of variation	1.5	1.0	1.2	1.4	0.5	0.6	0.7	0.08	0.7	0.7	1.5	0.7	1.5
Mixed Open Space (168)													
Number of observations	168	131	167	93	65	70	90	128	76	148	153	145	145
% of samples above detection	100	100	100	100	100	100	60.0	100	100	99.3	97.4	96.6	96.6
Median	115.4	33.0	0.51	0.10	215	64.2	8.5	7.9	16.0	109	78.0	6.0	34
Coefficient of variation	0.8	0.4	0.8	1.2	1.7	1.3	1.5	0.1	0.3	2.2	1.6	2.7	1.6

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1 – *Continued*

	Fecal Coliform (mpn/100 mL)	Fecal Strepto-coccus (mpn/100 mL)	Total Coliform (mpn/100 mL)	Total E. Coli (mpn/100 mL)	NH3 (mg/L)	N02+NO3 (mg/L)	Nitrogen, Total Kjeldahl (mg/L)	Phospho-rus, filtered (mg/L)	Phospho-rus, total (mg/L)	Sb, total (µg/L)	As, total (µg/L)	As, filtered (µg/L)	Be, total (µg/L)
Overall Summary (3765)													
Number of observations	1704	1141	83	67	1908	3075	3191	2477	3285	874	1507	210	947
% of samples above detection	91.2	94.0	90.4	95.5	71.3	97.3	95.6	85.1	96.5	7.2	49.9	27.1	7.7
Median	5091	17000	12000	1750	0.44	0.60	1.4	0.13	0.27	3.0	3.0	1.5	0.4
Coefficient of variation	4.6	3.8	2.4	2.3	1.4	0.97	1.2	1.6	1.5	1.7	2.6	1.0	2.5
Residential (1042)													
Number of observations	402	257		14	572	889	922	690	926		395		282
% of samples above detection	87.8	87.9		100	82.2	97.6	96.5	83.5	96.8		40.8		7.8
Median	7000	24300		700	0.31	0.60	1.5	0.18	0.31		3.0		0.5
Coefficient of variation	5.2	1.7		1.6	1.1	1.1	1.1	0.9	1.1		2.2		2.5
Mixed Residential (611)													
Number of observations	336	178	26	11	282	531	517	430	552		158		97
% of samples above detection	94.3	97.8	84.6	90.9	58.5	97.9	95.0	83.3	96.2		65.9		11.3
Median	11210	27500	5667	1050	0.39	0.57	1.4	0.13	0.28		3.0		0.3
Coefficient of variation	3.2	2.1	1.3	2.1	1.6	0.78	1.7	1.1	1.7		3.9		2.7
Commercial (527)													
Number of observations	253	201			300	445	469	343	466		235		
% of samples above detection	88.9	92.5			83.3	98.0	97.4	81.0	95.9		33.6		
Median	4600	12000			0.50	0.6	1.5	0.11	0.22		2.3		
Coefficient of variation	3.0	2.7			1.2	1.1	0.9	1.3	1.2		2.9		
Mixed Commercial (324)													
Number of observations	116	95			173	284	276	221	290	89	139		
% of samples above detection	94.8	98.9			67.1	96.8	96.0	93.7	98.6	11.9	45.5		
Median	5400	11900			0.60	0.58	1.4	0.12	0.26	15.0	2.0		
Coefficient of variation	3.0	2.6			1.0	0.7	0.9	2.1	1.5	1.0	1.0		
Industrial (566)													
Number of observations	315	189			272	461	483	344	478	152	255		197
% of samples above detection	87.3	93.7			78.3	96.3	96.3	88.1	96.2	14.5	52.9		10.7
Median	2400	12000			0.42	0.69	1.4	0.10	0.25	3.7	4.0		0.38
Coefficient of variation	5.7	7.0			1.3	0.92	1.1	1.2	1.4	1.4	1.4		2.5

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1 – *Continued*

	Fecal Coliform (mpn/100 mL)	Fecal Strepto-coccus (mpn/100 mL)	Total Coliform (mpn/10 0 mL)	Total E. Coli (mpn/100 mL)	NH3 (mg/L)	N02+NO3 (mg/L)	Nitrogen, Total Kjeldahl (mg/L)	Phospho-rus, filtered (mg/L)	Phospho -rus, total (mg/L)	Sb, total (µg/L)	As, total (µg/L)	As, filtered (µg/L)	Be, total (µg/L)
Mixed Industrial (218)													
Number of observations	79	59	14		99	173	160	179	177		93		
% of samples above detection	98.7	96.9	71.4		30.3	98.8	92.5	84.4	95.5		88.2		
Median	3033	11000	2467		0.58	0.59	1.1	0.08	0.20		3.5		
Coefficient of variation	2.5	2.5	1.5		0.8	0.7	1.5	2.3	1.6		0.9		
Institutional (18)													
Number of observations					18	18	18	17	17				
% of samples above detection					88.9	100	100	82.4	94.1				
Median					0.31	0.6	1.35	0.13	0.18				
Coefficient of variation					0.5	0.6	0.5	0.5	1.0				
Freeways (185)													
Number of observations	49	25	16	13	79	25	125	22	128		61	72	
% of samples above detection	100	100	100	100	87.3	96.0	96.8	95.5	99.2		55.7	50.0	
Median	1700	17000	50000	1900	1.07	0.28	2.0	0.20	0.25		2.4	1.4	
Coefficient of variation	2.0	1.2	1.5	2.2	1.3	1.2	1.4	2.1	1.8		0.7	2.0	
Mixed Freeways (26)													
Number of observations	20	16				22	22	11	22		15		
% of samples above detection	85.0	93.8				100	100	100	100		80		
Median	2600	19000				0.9	2.3	0.03	0.34		3.0		
Coefficient of variation	2.3	1.1				0.7	1.3	0.9	0.7		0.7		
Open Space (68)													
Number of observations	23	22			32	44	45	44	46		19		
% of samples above detection	91.3	90.9			18.8	84.1	71.1	79.6	84.8		31.6		
Median	7200	24900			0.18	0.59	0.74	0.13	0.31		4.0		
Coefficient of variation	1.1	1.0			1.24	0.9	0.9	0.9	3.5		0.4		
Mixed Open Space (168)													
Number of observations	86	75			71	152	123	148	152		88		
% of samples above detection	97.7	100			22.5	97.4	90.2	85.8	96.1		44.3		
Median	3000	21000			0.51	0.7	1.1	0.09	0.25		3.0		
Coefficient of variation	2.3	2.4			1.2	0.8	0.9	1.1	1.1		0.9		

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1 – *Continued*

	Cd, total (µg/L)	Cd, filtered (µg/L)	Cr, total (µg/L)	Cr, filtered (µg/L)	Cu, total (µg/L)	Cu, filtered (µg/L)	Pb, total (µg/L)	Pb, filtered (µg/L)	Hg, total (µg/L)	Ni, total (µg/l)	Ni, filtered (µg/L)	Zn, total (µg/L)	Zn, filtered (µg/L)
Overall Summary (3765)													
Number of observations	2574	389	1598	261	2722	411	2949	446	1014	1430	246	3007	381
% of samples above detection	40.6	30.3	70.2	60.5	87.4	83	77.7	49.8	10.2	59.8	64.2	96.6	96.3
Median	1.0	0.50	7.0	2.1	16	8.0	17.0	3.0	0.20	8.0	4.0	116	52
Coefficient of variation	3.7	1.1	1.5	0.7	2.2	1.6	1.8	2.0	2.5	1.2	1.5	3.3	3.9
Residential (1042)													
Number of observations	695		404		771	90	762	108	275	392	25	784	87
% of samples above detection	31.1		53.2		83.1	63.3	69.4	33.3	6.9	44.1	44.0	96.2	89.7
Median	0.5		4.5		12	7.0	12.0	3.0	0.20	5.6	2.0	73	31.5
Coefficient of variation	3.4		1.2		1.8	2.0	1.9	1.9	0.9	1.2	0.5	1.3	0.8
Mixed Residential (611)													
Number of observations	420	30	193	21	432	29	500	30	115	150	25	515	28
% of samples above detection	34.5	40.0	81.3	52.4	83.8	72.4	78.4	46.7	15.7	60	72.0	92.6	100
Median	0.9	0.30	7.0	2.0	16	5.5	16	3.0	0.20	7.8	5.5	95	48
Coefficient of variation	3.6	0.6	1.5	0.8	1.2	0.9	1.4	0.7	0.8	0.8	0.9	0.9	0.9
Commercial (527)													
Number of observations	379	47	257	27	408	48	399	59	170	242	23	414	49
% of samples above detection	41.7	23.4	60.7	40.7	92.9	79.2	85.5	52.5	6.5	60.3	47.8	99.0	100
Median	0.96	0.30	6.0	2.0	17	7.57	18.0	5.0	0.20	7.0	3.0	150	59
Coefficient of variation	2.7	1.3	1.3	0.6	1.5	0.8	1.6	1.6	0.8	1.2	0.8	1.2	1.4
Mixed Commercial (324)													
Number of observations	188	41	128	27	191	41	244	41		102	26	243	39
% of samples above detection	49.5	34.1	88.3	66.7	93.2	80.5	88.1	63.4		78.4	69.2	98.8	100
Median	0.9	0.35	5.0	2.5	17.5	10	17.0	3.5		5.1	3.5	131.4	73
Coefficient of variation	1.1	0.8	1.1	0.7	3.0	0.6	1.4	0.8		1.3	0.6	1.7	0.8
Industrial (566)													
Number of observations	435	42	250	36	455	42	452	51	199	237	36	473	42
% of samples above detection	49.0	54.8	72.0	55.6	88.6	90.5	75.0	52.9	13.9	61.6	58.3	98.9	95.2
Median	2.0	0.60	12.0	3.0	20.8	8.0	24.9	5.0	0.20	14.0	5.0	199	112
Coefficient of variation	2.2	1.1	1.2	0.7	2.0	0.7	1.9	1.6	2.7	1.0	1.4	1.5	3.6

Table 3. Summary of Available Stormwater Data Included in NSQD, version 1.1 – *Continued*

	Cd, total (µg/L)	Cd, filtered (µg/L)	Cr, total (µg/L)	Cr, filtered (µg/L)	Cu, total (µg/L)	Cu, filtered (µg/L)	Pb, total (µg/L)	Pb, filtered (µg/L)	Hg, total (µg/L)	Ni, total (µg/l)	Ni, filtered (µg/L)	Zn, total (µg/L)	Zn, filtered (µg/L)
Mixed Industrial (218)													
Number of observations	145	25	109	15	150	24	213	25	58	74	15	212	24
% of samples above detection	60.7	92.0	92.7	66.7	90.0	100.0	82.6	92.0	22.4	83.8	100.0	98.6	95.8
Median	1.6	0.60	8.0	2.0	23	6.0	20.0	5.0	0.3	12	5.0	172	2100
Coefficient of variation	1.9	0.6	1.7	0.7	0.8	0.6	1.4	1.0	0.6	0.8	0.6	3.1	1.2
Institutional (18)													
Number of observations							18					18	
% of samples above detection							77.8					100	
Median							5.75					305	
Coefficient of variation							0.8					0.8	
Freeways (185)													
Number of observations	95	114	76	101	97	130	107	126		99	95	93	105
% of samples above detection	71.6	26.3	98.7	78.2	99.0	99.2	100	50.0		89.9	67.4	96.8	99.1
Median	1.0	0.68	8.3	2.3	34.7	10.9	25	1.8		9.0	4.0	200	51
Coefficient of variation	0.9	1.0	0.7	0.7	1.0	1.5	1.5	1.7		0.9	1.4	1.0	1.9
Mixed Freeways (26)													
Number of observations	23		15		23		23					23	
% of samples above detection	56.5		100		100		56.5					100	
Median	0.5		6.0		14		10.0					130	
Coefficient of variation	2.2		1.0		1.0		1.3					0.9	
Open Space (68)													
Number of observations	38		36		39		45					45	
% of samples above detection	55.3		36.1		74.4		42.2					71.1	
Median	0.38		5.4		10		10.0					40	
Coefficient of variation	1.9		1.7		2.0		1.7					1.3	
Mixed Open Space (168)													
Number of observations	107		88		108		155		27	51		156	
% of samples above detection	18.7		81.8		89.8		74.2		14.8	72.5		98.1	
Median	2.0		6.0		9.0		10		0.15	8.0		80	
Coefficient of variation	1.4		1.3		1.0		2.3		0.4	1.1		1.1	

In the General Information tab, the spreadsheet lists the states and municipalities included in the current version of the database. The second tab describes the two main sections of the database: site descriptions and event descriptions. In the items description section, each column in the database is described. The last column in this table shows an example of the value expected in each column. The third tab describes the constituents and parameters included in the database, the number of observations, and the percentage of samples having detected observations. This table is useful to identify those constituents with high percentages of detected values.

The last tab in the database contains the data itself; a matrix of 232 columns by 3,765 rows containing all the data collected and reviewed. Each row represents a storm event for each monitoring location. This part of the table is divided in seven subsections describing the site location, the hydrology of the event and equipment used, and the constituent classifications. Each section of the database is described in the following discussion, with detailed analyses presented in Chapters 4 through 8 of this report.

The following discussion will require a copy of the database for reference. This is available at: http://unix.eng.ua.edu/~rpitt/Research/ms4/mainms4.shtml. Each of the sections and columns included in the spreadsheet will be explained in detail. Summary statistics, probability plots and box and whiskers plots will be used to describe the most important parameters.

Site Description [Columns A through Y]

Column A is an identifier of each storm event stored in the database. It is the table key. Column B describes the site main land use or activity: residential (RE), commercial (CO), industrial (ID), institutional (IS), open space (OS), and freeways (FW). In the case when more than one land use is present, a combination code is used beginning with the land use with the most area in the watershed. For example, if a site was 70% residential and 30% commercial, the site was coded as RE_CO. The percentage of each land use is indicated in the columns J through O.

Column C describes the month of the year when the sample was collected as follows: winter (WI) if the sample was collected in November, December or January; spring (SP) if the sample was collected in February, March, or April; summer (SU) if the sample was collected in May, June or July; and fall (FA) if the sample was collected in August, September, or October. A reasonably uniform number of samples were collected during each of the four periods: about 29% of the samples were collected in the winter, 30% in the spring, 19% in the summer, and 23% in the fall.

Columns D through F indicate the location of the site. LOCATION_ID is the key for sorting the sites, and is a code of eight characters: the first two letters indicate the name of the state, the next four letters is a code for the community, and the last two letters represent the site name. Columns E and F are the name of the community and the name of the site. Column G is the contact information of the person in that community that supplied the database information. Columns H through M are the percentages of the separate land uses in the drainage area, as described in column B.

Column N indicates the total watershed drainage area in acres. Figure 2 shows the distribution of the area by land use. The distribution of the watersheds areas can be considered approximately lognormal. Commercial, industrial, open space, and residential land uses have approximately the same distribution of drainage areas for the monitored outfalls, with a range between ten and one thousand acres. The median monitored watershed area for commercial and industrial sites was about 43 acres, while the median watershed area in residential and open space areas was about 65 acres. Freeways had smaller areas than the other land uses, with median areas being about 2 acres, with a range varying between one and one hundred acres.

Columns O and P list the approximate latitude and longitude of the outfall location in degrees, minutes, and seconds. Most of these coordinates were obtained using the Teraserver website. Column S indicates the EPA Rain Zone location of each site (Figure 1 and Table 1). About 52% of the sites are located in the EPA Rain Zone 2, which contains the Chesapeake Bay region, the main targeted area for this database. Each of the Rain Zones 3 through 7 has about 8% of the total sites. Rain Zones 1 and 9 have each about 3% of the sites. Rain Zone 8 has only one community with four locations, or about 1% of the total number of sites.

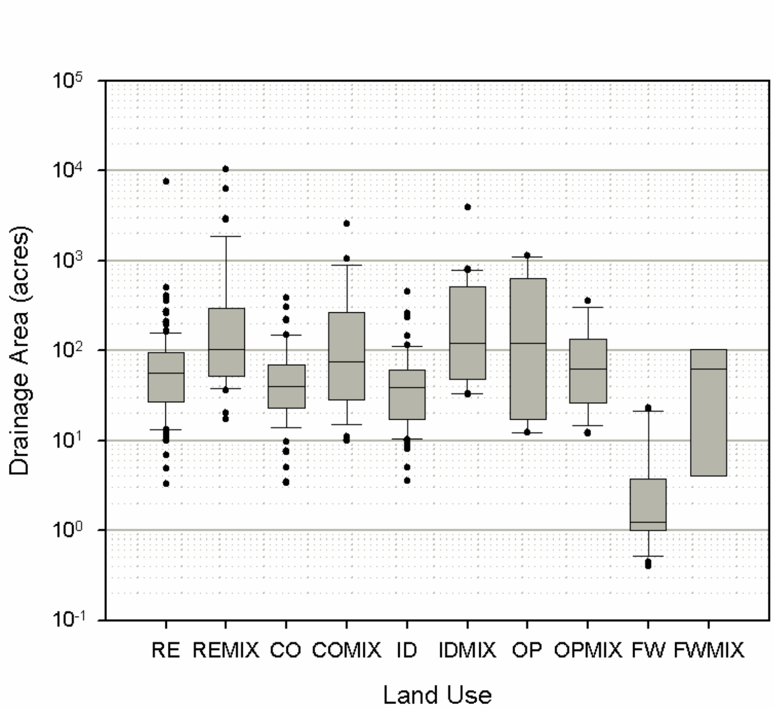


Figure 2. Drainage areas by land use

Column R indicates the total percentage of impervious surfaces reported for each site. Only Newport News, Virginia, contained information describing how the impervious areas were hydraulically connected to the drainage systems. It is expected that a watershed with high levels of impervious (a parking lot for example), is mostly directly connected due to little opportunity for draining to pervious areas. Less water is therefore infiltrated and the stormwater rapidly moves to the connected outfall. About 169 sites (about 47% of the total number of sites) included percentage of impervious surfaces in their annual reports or permit applications. Of this response, about 69 sites were for single or mixed residential areas, 34 sites were single or mixed industrial areas, 34 sites were single or mixed commercial area, 17 sites were single or mixed freeway areas, and 15 sites were single or mixed open space areas.

Figure 3 shows a box and whiskers plot of the reported impervious surface values for the predominant land uses. As expected, the open space sites have the lowest percentage of impervious surfaces (mean about 3.3%), while the mean impervious surface value for the freeway sites is 92%. Industrial and commercial area impervious surface values are higher, with means of 67% and 81% respectively. Residential areas cover almost the complete range, from about 7 to 89%. The impervious surfaces for residential areas are intermediate between the values for open space and the industrial/commercial values, as expected. The mean percentage of impervious areas in residential areas is approximately 41%.

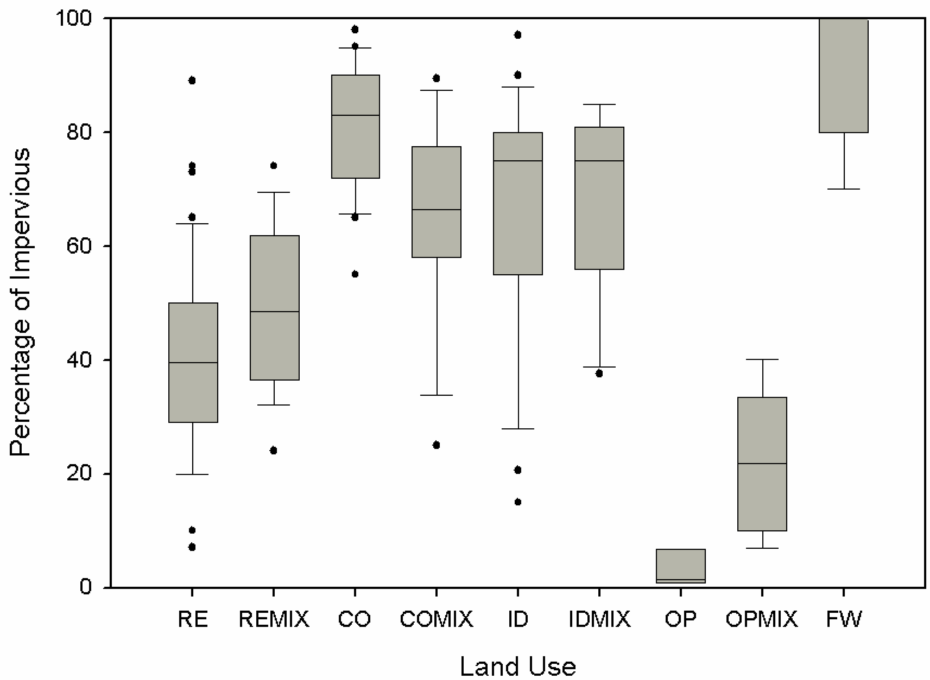


Figure 3. Percentage of impervious surfaces by land use

Column S is a qualifier for the total percentage of impervious surface area in the test watershed, indicating if there was an apparent increase in the percentage of imperviousness during the monitoring period, based on examinations of aerial photographs. Only one site (Pylon Street in Forth Worth, TX) had an apparent increase in the percentage of impervious area during the monitoring period. Column T indicates the volumetric runoff coefficient (Rv), or the ratio between the total runoff depth divided by the precipitation depth for each event. Figure 4 is a scatter plot of the reported percentage of impervious areas and reported Rv. As expected, higher volumetric runoff coefficients are reported for heavily paved areas, such as parking lots or freeways, compared to areas having much more landscaped areas, such as residential areas or parks. However, it is possible that some of the reported Rv values are simply calculated from the percent impervious cover values, and not from monitored rainfall and monitored runoff values.

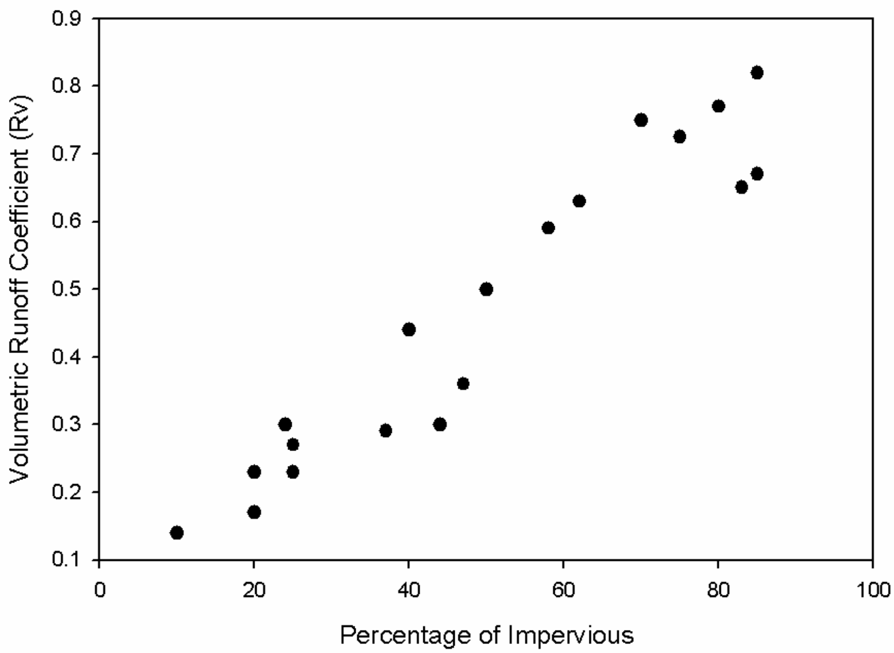


Figure 4. Scatter plot of percentage of impervious and Rv

None of the monitoring agencies reported the TR-55 curve number for the sites. This value is used to estimate the runoff volume using the Soil Conservation Service, SCS (now Natural Resources Conservation Service, NRCS) TR-55 method. Curve numbers (column U) were therefore not examined during this analysis. Only eight sites indicated the year when the land was developed, and these are shown in column V. Because of the low number of observations, this factor also could not be used in the data analyses.

Column W indicates the type of stormwater conveyance reported for the monitored area. This parameter indicates if the site is drained with "curb and gutter" systems typical of areas with high percentages of imperviousness, or if the water is transported beside the road through a grass-lined drainage channel (swales), more common in lower density areas. About 26% of the sites did not report the type of conveyance or it was not possible to identify them using the aerial photographs. Curb and gutter systems were reported for 65% of the sites, while grass swales were reported for 9% of the sites. Grass swales are usually considered a stormwater control, or "BMP," due to their ability to infiltrate large fractions of the runoff before discharge. They may provide some limited concentration reductions of particulate pollutants, but only for the shallowest flows. Detailed analyses are presented in Chapter 5 of this report.

The next column indicates if the site has wet detention ponds. About seven sites (out of the 360 total sites) have a wet pond at the outfall, nine sites have ponds in the watershed, and three sites have ponds in series, all upstream of the monitoring location. Other reported stormwater controls included: dry detention ponds (4 sites), small underground detention storage tanks (2 sites), besides the 32 sites having grass swales as noted above.

The final column in this section (Y) includes important comments that were not assigned to any of the other columns. Typical information in this column is the size of the pipe; if the outlet is a circular (pipe), or a square (box culvert); the number of pipes discharging from the watershed; or if there is a USGS monitoring station at the outfall that reported the data in the NSQD.

Hydrologic Information [Columns Z through AN]

Column Z is the identifier of each storm event stored in the database. It is used as a table sorting "key." Generally, it contains information about the location and the sampling date. Column AA indicates the precipitation depth recorded during the event, in inches. About 3,300 events included this parameter. Precipitation depth, flow volume and similar hydrologic parameters were included in the annual reports or permit applications usually as appendices. During the data collection process, some of these appendices were not copied or located. The highest percentage of events with precipitation by land use was observed in single and mixed freeways (about 99%). The lowest percentage of events with precipitation data was observed in single and mixed residential areas, with 85% of the sites reporting this information. The percentage for the other land uses were: 87% for single and mixed industrial, 90% in single and mixed commercial, 96% in single and mixed open space.

Figure 5 shows the distribution of the available precipitation depth data by land use. The range of precipitation depth varies between 0.01 and 6 inches, indicating that some of the reported events were outside of the range specified by the general monitoring guidance (minimum of 0.1 inches and close to annual average characteristics). The distribution of the rainfall depth data is approximately lognormal, with a median between 0.4 to 0.6 inches. All the land uses have a similar pattern, with approximately the same variance. The mixed freeway category seems to have a narrower range, but they only represent 0.5% of the total events that have precipitation data. Column AB is a qualifier for the precipitation depth data. Some communities collected the data on site, while others used rain gauge data collected from a local airport. Rain gauges located on site are preferred as they are expected to better represent the rainfall conditions that occurred on the monitored site for the monitored event. Twelve percent of the total database events did not include precipitation depth data, 42% of the events were associated with rain data collected on site, 23% of the events did not indicate how the reported rain data was obtained, 7% of the events are associated with rain data from the local airport rain gauge, and the remaining 16% used other methods to determine the event rainfall data, such as regional rain gauges associated with flood monitoring systems.

Columns AC through AF indicate the starting and ending date and time of the event. Column AG indicates the maximum reported 15-minute rain intensity for each event. Events having high rain intensities have high kinetic energies, and it is hypothesized that these events will have increased washoff or erosion of particulate pollutants from watershed surfaces. However, only 1% of the database events reported this parameter. Column AG information was therefore not included in any of the data analyses.

Runoff depth (column AH) is the total volume of stormwater that leaves the monitored watershed during the rain event. For a directly connected paved parking lot, the runoff depth (expressed in inches of runoff for the complete drainage area) is only slightly smaller than the precipitation depth. In contrast, a park having mostly pervious surfaces would record total runoff volumes much smaller than the rain depth because most of the rainwater is infiltrated before it drains from the site. About 36% of the events included runoff data.

Figure 5 also shows the probability plots of runoff depth for each land use. As expected, smaller runoff values were observed in open space and residential areas, while freeways, mixed commercial, and mixed industrial land uses have runoff distributions similar to the rain distributions observed in the precipitation panel. A different pattern was observed for runoff at freeways, which are characterized by their small area and high percentage of impervious cover.

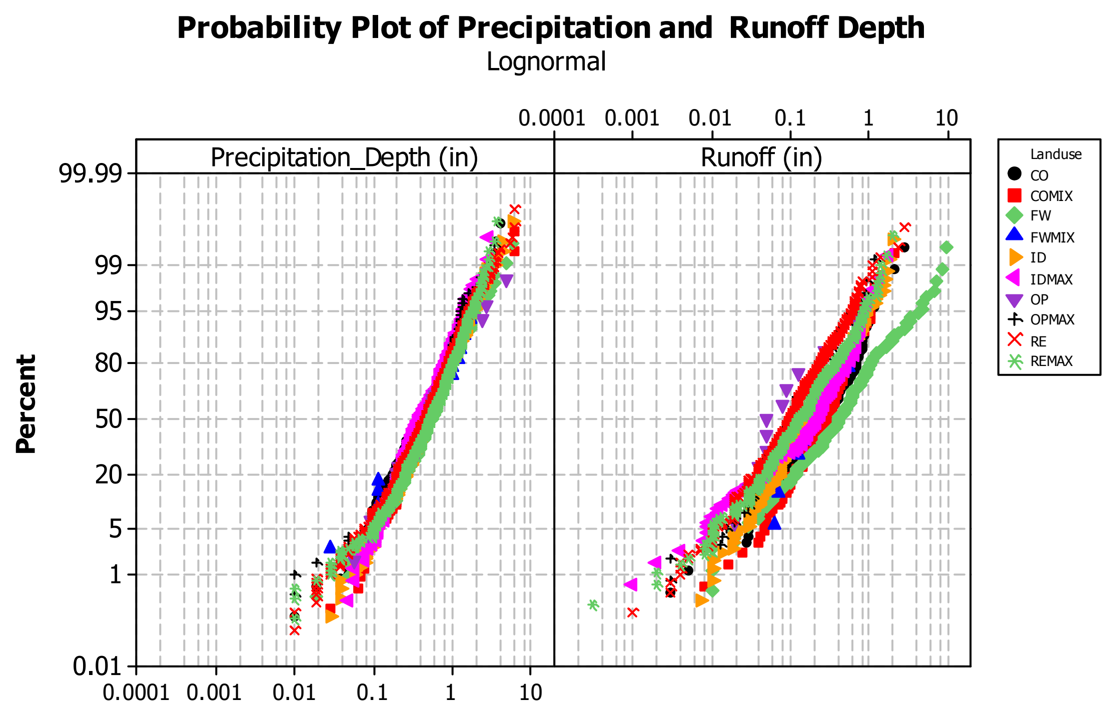


Figure 5. Precipitation and runoff depth by land use

Column AJ indicates if the runoff and precipitation were measured during the complete event or only the first three hours of the storm. The basic NPDES stormwater monitoring guidelines indicates that samples must be collected at least during the first three hours of the event. If the runoff and precipitation were not monitored for the complete event, then site hydrology confusion would occur. Most of the communities recorded the runoff for the complete event, even if monitoring only occurred for three hours. Only Greensboro, Topeka, Chesterfield County, and Fayetteville recorded runoff only for the first three hours of the events.

Column AK indicates if the events were from composite sampling, as required by the Federal Regulations guidance. First flush events were included in the first version of the database, version 1.0. After the paired first flush statistical analyses (see Chapter 6), these first-flush data were removed from the main database to eliminate confusion, leaving only the composite samples in the main database.

Column AL indicates if the composite sample was collected using automatic equipment, or if manual sampling was used. This column can be used to evaluate possible differences in the recorded concentrations due to the sampling method. About 81% of the events were collected using automatic samplers, 10.5% used manual sampling, and about 8.5% of the events did not have any reported sampling method. Detailed analyses concerning the effects of manual versus automatic sampling is discussed in Chapter 5 of this report.

Column AM describes if the collected sample was a flow-weighted or time-weighted composite sample. A flow-weighted composite sample is comprised of several equal volume subsamples that were collected according to the flow rate of the runoff water. The sampler is programmed to collect a subsample for a specified constant flow increment. The total volume in the single composite bottle is therefore proportionate to the total runoff volume associated with the monitored event. A time-weighted composite sample is made up of several equal volume subsamples that were collected at constant periods of time and collected into a single large composite sample bottle. At the end of the event, the total volume of sample in the composite sample bottle is proportionate to the duration of the event. About 73% of the events in the database were collected using flow-weighted composite sampling methods, while only 8% of the events were collected using time-weighted composite sampling methods. No composite sampling method information was available for the remaining 19% of the events.

The last column in this database section describes the number of days without rain prior to the event sampling. It is usually hypothesized that an increase in the number of dry days prior to an event would cause an increase in the constituent concentration. About 38% of the events had this information available. Detailed analyses are presented in Chapter 5 of this report.

Conventional Constituents [Columns AO through BS]

This section of the database contains measurement values for conventional stormwater constituents (conductivity, DO, hardness, oil and grease, pH, temperature, TDS, TSS, BOD_5, COD, fecal coliforms, and fecal streptococcus).

Table 3, presented earlier, contains a summary showing the total number of samples included in the database classified by land use, the percentages of samples detected, the medians, and the coefficients of variation. In general, the lowest concentrations were usually found at open space land uses, followed by residential areas. The highest concentrations were observed at freeway land use sites. Table 4 is a summary contrasting the land uses having the lowest and the highest concentrations of these constituents.

The Mann-Whitney test was used to determine if there is a significant difference between the land uses having the lowest and highest concentrations. As a complement, one-way ANOVA analyses were used to identify if a significant difference existed among any of the land uses. As the number of samples increase, the power of the test also increases. P-values close to zero will indicate that the concentration of at least one land use is statistically different than the other land uses (true for all constituents in Table 4, except for Dissolved Oxygen).

Table 4. Conventional Constituents Summary

Constituent	Land use having the lowest median concentration			Land use having the highest median concentration			Mann-Whitney Test	1-Way ANOVA by Land Use
	n	Land Use	Median	n	Land Use	Median	p-value	p-value
Conductivity (μS/cm)	106	RE	96.5	108	ID	135.5	0	0
Dissolved Oxygen (mg/L)	39	ID	7.3	30	RE	7.8	0.064	0.325
Hardness (mg/L CaCO3)	350	RE	32	139	CO	38.9	0.009	0
Oil and Grease Total (mg/L)	308	RE	3.85	43	FW	8.0	0	0.001
pH (s.u.)	111	FW	7.1	234	ID	7.5	0	0
Temperature (°C)	31	FW	14	140	ID	17.8	0	0
Total Dissolved Solids (mg/L)	854	RE	72	411	ID	92	0	0
Total Suspended Solids (mg/L)	977	RE	49	133	FW	99	0	0
Biological Oxygen Demand (mg/L)	38	OP	5.4	421	CO	11	0	0
Chemical Oxygen Demand (mg/L)	33	OP	42.1	66	FW	100	0	0
Fecal Coliform (colonies/100 mL)	261	ID	2500	21	OP	7200	0.014	0
Fecal Streptococcus (colonies/100 mL)	166	CO	10285	273	RE	24600	0	0.003

Figure 6 contains examples of grouped box and whiskers plots for several constituents for different major land use categories. The freeways sites had the highest reported TSS, COD and oil and grease concentrations. Statistical ANOVA analyses for all land use categories found significant differences for land use categories for all constituents except for dissolved oxygen. Turbidity, total solids, total coliform and total E-coli have not enough samples in each group to evaluate if there is a difference among all land uses. Chapter 5 presents more comprehensive analyses for specific site conditions (considering interactions of land use, geographical location, etc.).

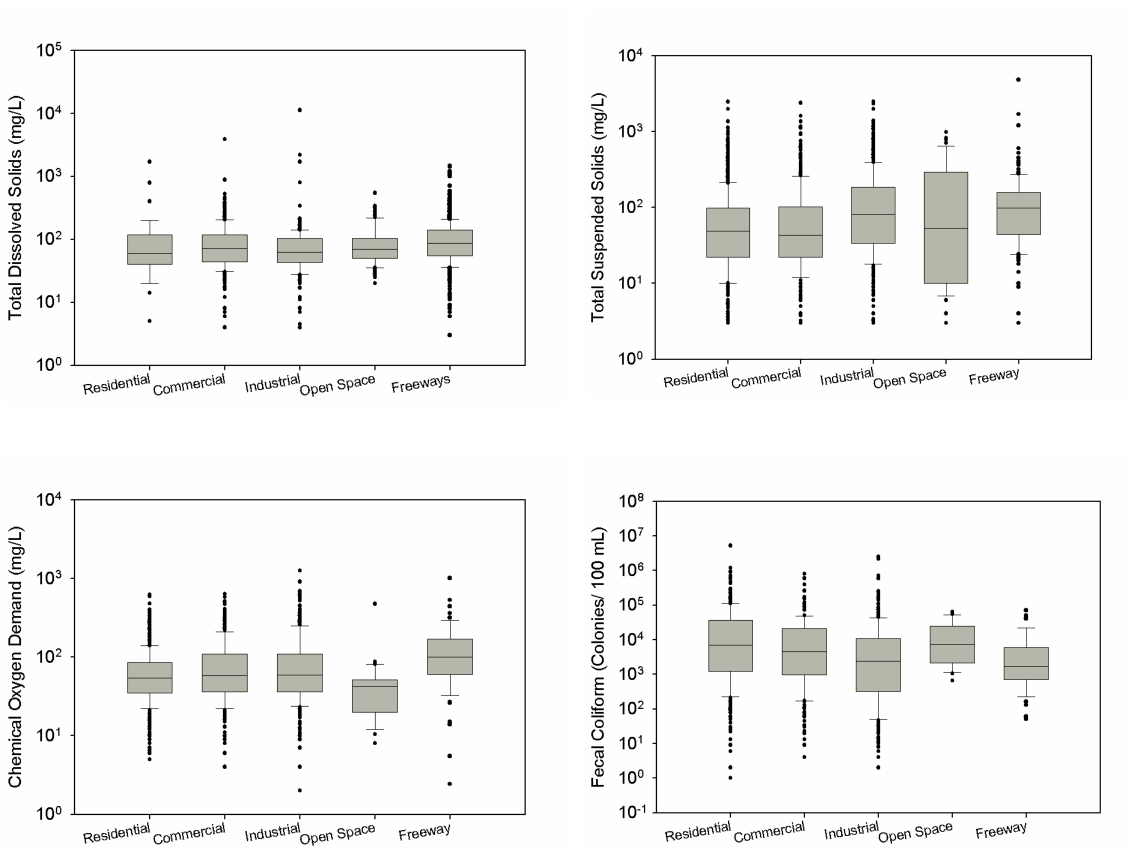


Figure 6. Box and whiskers plots for conventional constituents by single land use

Stormwater temperature depends of many factors, including season, the time of the day, and the types of surfaces in a land use. Column C shows the season of the year when each sample was obtained, the most obvious factor affecting runoff temperature.

Figure 7 shows the water temperatures for each month for the samples collected in the EPA Rain Zones 5 and 6 combined. Similar patterns were observed in the other EPA Rain Zones. Two main periods can be identified in this plot: from February to July the water temperature rises and from August to January the water temperature decreases. Table 4 shows that for almost all conventional constituents, residential and open space land uses have the lowest concentrations, except for pathogen indicators. Industrial and freeway land uses generally have the highest concentrations.

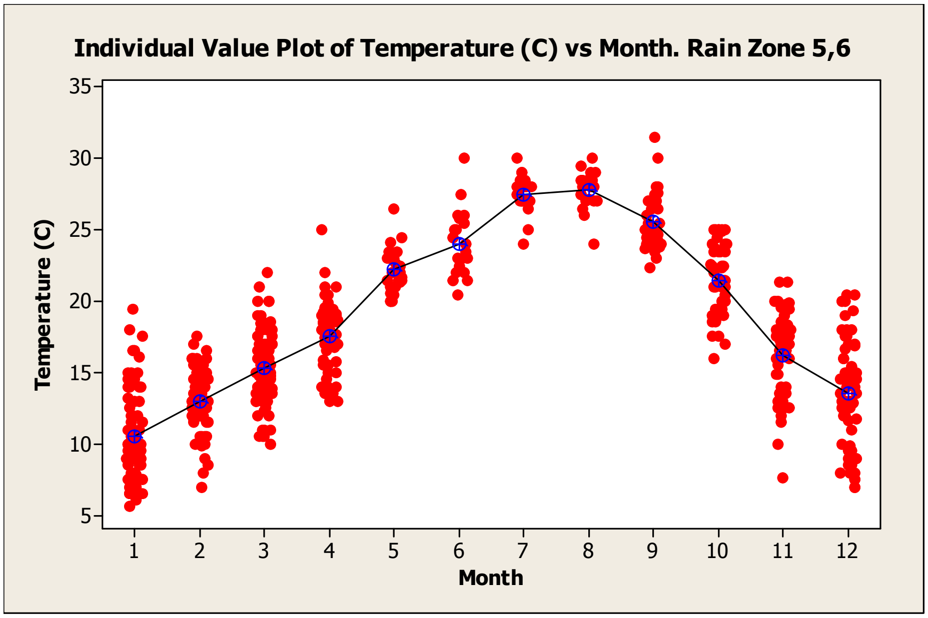


Figure 7. Water temperature in EPA Rain Zones 5 and 6 (line links median values for each month)

Nutrients [Columns BU through CG]

This section in the database contains the compounds associated with nitrogen and phosphorus compounds. Table 5 shows a summary of the land uses having the lowest and highest concentrations for each constituent. Again, the Mann Whitney and ANOVA tests were used to evaluate if there was a significant difference between land uses for these constituents.

In contrast to the conventional constituents, dissolved and total phosphorus have the highest concentrations in residential land uses. There was no significant difference noted for total nitrogen for the different land uses. The median ammonia concentration in freeway stormwater is almost three times the median concentration observed in residential and open space land uses, while freeways have the lowest orthophosphate and nitrite-nitrate concentrations; almost half of the concentration levels that were observed in industrial land uses. Figure 8 shows box plots for TKN, total phosphorus, and nitrite-nitrate for several land uses. It shows that even if there are differences in the median concentrations by a factor of two or three between the land uses, the extreme range of the concentrations within a single land uses can still vary by two or three orders of magnitude. Again, Chapter 5 examines many factors affecting these concentrations, in addition to land use.

Table 5. Nutrients Summary

Constituent	Land use having the smallest median concentration			Land use having the largest median concentration			Mann-Whitney Test	1-Way ANOVA by Land Use
	n	Land Use	Median	n	Land Use	Median	p-value	p-value
Ammonia (mg/L)	485	RE	0.31	69	FW	1.07	0	0
Nitrogen Nitrite-Nitrate (NO_2+NO_3) (mg/L)	24	FW	0.28	429	ID	0.71	0	0.001
Nitrogen Total (mg/L)	63	ID	2.03	81	RE	2.30	0.25	0.698
Nitrogen Kjeldahl Total (TKN) (mg/L)	32	OP	0.74	121	FW	2.00	0	0
Phosphate Ortho (mg/L)	103	FW	0.09	66	ID	0.23	0	0
Phosphorous Dissolved (mg/L)	283	ID	0.11	621	RE	0.17	0	0
Phosphorous Total (mg/L)	427	CO	0.22	933	RE	0.30	0	0

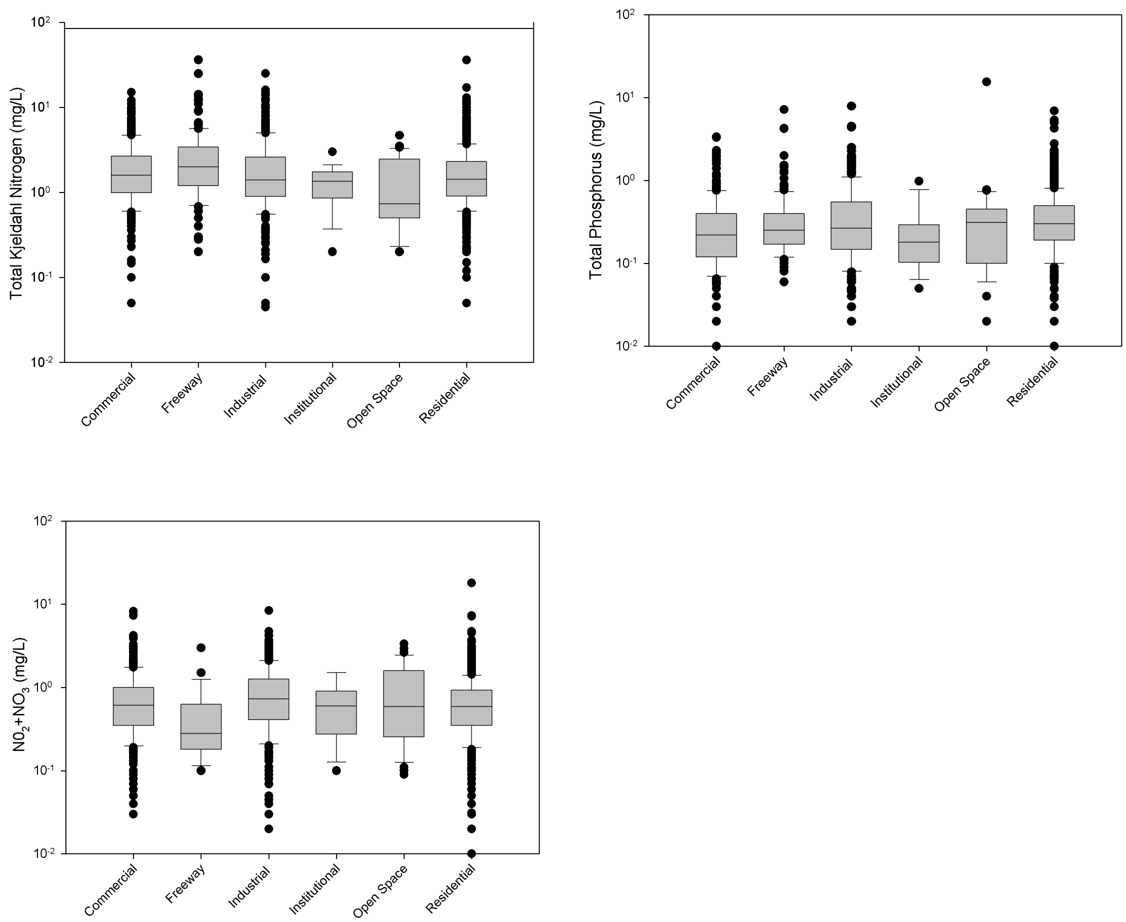


Figure 8. Box and whiskers plots for nutrients by single land use

Metals [Columns CK through EK]

This section in the database contains the metal concentrations. Industrial land uses have higher median concentrations of heavy metals than any of the other land uses, followed by freeways. Table 6 shows the ANOVA results for metals. As expected, open space and residential land uses have the lowest median concentrations. In almost all cases, the median metal concentrations at the industrial areas were about three times the median concentrations observed in open space and residential areas. Arsenic, cadmium, chromium, copper, lead, nickel, and zinc showed significant differences between the extreme land uses at the 1% level of confidence, or less. Other constituents are also included in the database (antimony, beryllium, cyanide, mercury, selenium, silver, and thallium), along with dissolved forms of the metals. Too few observations and large fractions of undetected observations hindered statistical analyses of these other metals.

Table 6. Summary of Metals Concentration

Constituent	Land use having the smallest concentration			Land use having the largest concentration			Mann-Whitney Test	1-Way ANOVA by Land Use
	n	Land Use	Median	p-value	Land Use	Median	p-value	p-value
Arsenic Total (μg/L)	70	CO	2.4	145	ID	4.0	0	0
Cadmium Total (μg/L)	219	RE	0.5	223	ID	1.9	0	0
Chromium Total (μg/L)	241	RE	4.6	186	ID	14.0	0	0
Copper Total (μg/L)	29	OP	10	96	FW	34.7	0	0
Lead Total (μg/L)	19	OP	10	343	ID	26	0	0
Nickel Total (μg/L)	190	RE	5.4	156	ID	16	0	0
Zinc Total (μg/L)	32	OP	40	455	ID	200	0	0

Figure 9 contains examples of grouped box and whiskers plots for lead, copper, and zinc constituents for different major land use categories. The highest lead and zinc concentrations were found in industrial land uses, while the highest copper concentrations were observed at freeways sites.

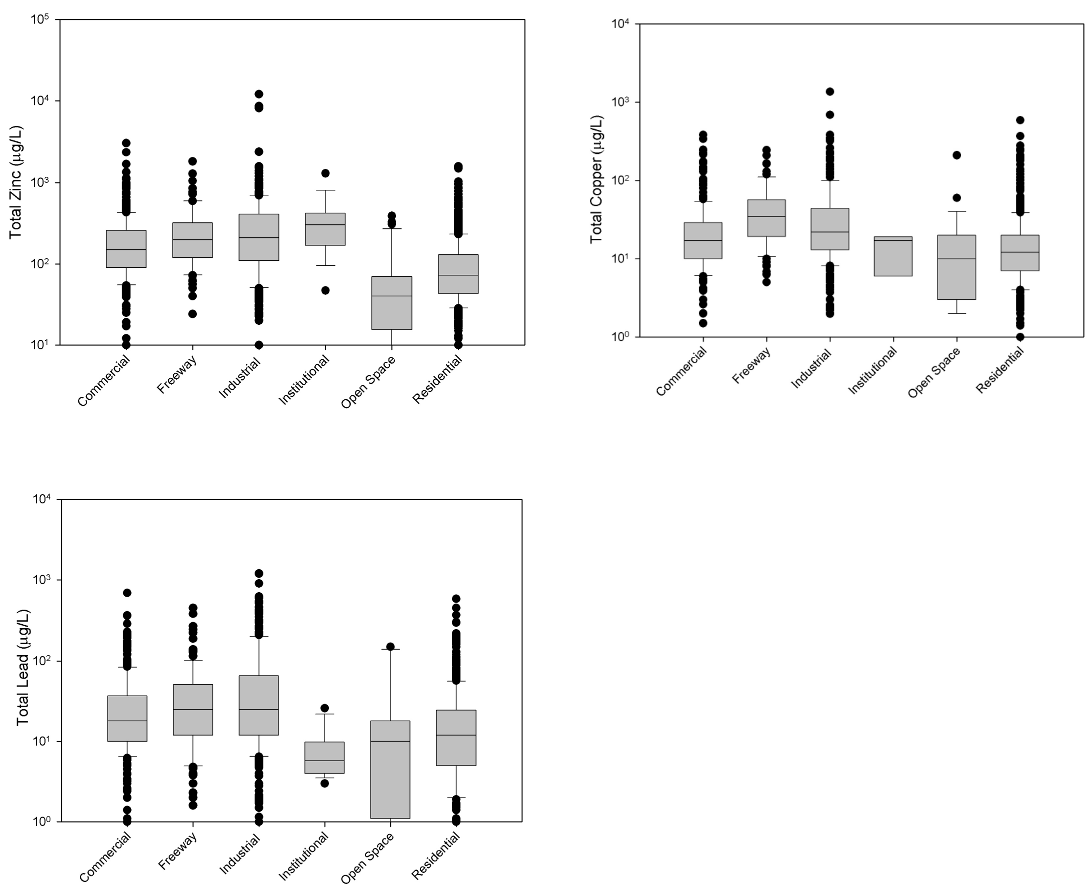


Figure 9. Box and whiskers plots for metals by single land use

Additional Constituents [Columns EM through HW]

These columns contain information for additional constituents that were sampled only during the permit application period (first year of sampling). Some constituents having more than a 30% detection level included: methylenechloride, total petroleum hydrocarbon (TPH), total organic carbon, chloride, nitrate nitrogen, nitrite nitrogen, total organic nitrogen, and iron.

Table 7 shows summaries for these additional constituents that have enough samples to identify significant differences between land uses. Only total petroleum hydrocarbon (TPH) and nitrite nitrogen showed significant differences (at the 5% significance level) between the land uses having lowest and highest median concentrations. The median stormwater TPH concentration in residential areas is almost half the median TPH stormwater concentration at freeway sites.

Table 7. Summary of Additional Constituents

Constituent	Land use having the smallest concentration			Land use having the largest concentration			Mann-Whitney Test	
	n	Land Use	Median	n	Land Use	Median	p-value	Significant α =5%
Total Petroleum Hydrocarbon (mg/L)	36	RE	0.38	20	FW	0.78	0	Yes
Chloride (mg/L) (FW and OP not included)	42	ID	7.1	38	CO	9.5	0.25	No
Nitrogen Nitrate (mg/L) (CO and OP not included)	13	RE	0.69	98	FW	0.84	0.58	No
Nitrogen Nitrite (mg/L) (CO and OP not included)	7	ID	0.07	42	FW	0.17	0.01	Yes
Nitrogen Total Organic (mg/L) (FW not included)	12	RE	0.96	5	CO	1.97	0.19	No
Iron (mg/L)	6	RE	2.99	27	FW	3.60	0.27	No

Site Descriptions and Additional Supporting Information

Supplemental reports were created containing additional information for each community. These site descriptions include (depending on available information) the land use and impervious surfaces for the monitored site, aerial photographs and a topographic map of the area, and descriptions of the sampling procedures and quality control (QA/QC) used during sample collection and analysis. The QA/QC description indicates if blank samples were used during the analysis to check the equipment, the protocols used during the sample collection, and in some cases, the chain of custody of the samples. These supplemental reports also contain descriptions of the sampled parameters, analytical methods, and field instrumentation used by the community.

About 38% of the aerial photographs have better than 1-meter resolution and the remaining photos have 1-meter resolution. The locations of most of the outfalls were included in the database in the Q and R columns (Latitude and Longitude). Table 8 shows the total number of sites with high-resolution aerial photos and with watershed delineations.

Table 8. Additional Site Information

EPA Rain-Zone	Number of Communities	Number of Sites	Sites with high-resolution aerial photos (resolution 0.25 m)	Sites with watershed delineations
1	2	12	0	2
2	28	185	38	18
3	8	30	15	20
4	4	21	15	17
5	9	33	18	0
6	4	30	20	9
7	6	28	19	0
8	1	4	0	0
9	3	17	13	8

Watershed delineations are an important component of the site descriptions by identifying the extent of the contributing area, the different land uses located in the watershed and the sampling location. Only 20% of the sites included their watershed delineations.

Most communities followed the sampling recommendations presented in the Code of Federal Register (40 CFR 122.21), although delegated NPDES state agencies were able to modify the specific requirements to better address local concerns. Almost all communities collected samples at least during the first 3 hours of the event (or the complete event if the duration was shorter). For about 66% of the events, the communities calculated the total runoff for the duration of the total event discharge, but used the concentrations from the shorter monitoring period. Chapter 6 includes a detailed analysis of first-flush concentrations that may indicate the maximum errors that may occur with truncated sampling periods. Seven percent of the events included runoff for only the first three hours of the event. The remaining 25% of the events did not include runoff volume data, or it was not clear if the runoff volume data was obtained during the first three hours, or for the whole event.

Another important monitoring aspect described in the site descriptions is how the composite sample was created. There are two compositing options: flow-weighted and time-weighted. During the time-weighted compositing scheme, subsamples of equal volume were obtained at specific time intervals during the three hour sampling period. All the subsamples were collected in a single bottle, creating the composite sample. In the flow-weighted compositing case, the subsamples were collected for a set flow increment. About 71% of the events were collected using flow-weighted sampling, 5% of the events were collected using time-weighted sampling, and it was not clear how the remaining 24% of the samples were collected. Roa-Espinosa and Bannerman (1995) found that time-weighted composite sampling could be representative of the sampling period, if many subsamples are collected throughout the storm period. Time-weighted compositing is much simpler and less expensive than flow-weighted composite sampling, but may have a slight error in the measured concentrations, compared with the flow-weighted method.

About 62% of the 65 communities represented in the NSQD indicated that they used automatic samplers during their monitoring activities, about 34% did not indicate how they collected their samples, and 4% collected their samples manually. ISCO samplers were the most commonly used automatic sampler, with about 24% of the sites using ISCO 2700, 3700 or 6700 samplers. American Sigma samplers were used at about 12% of the 65 communities. The most common American Sigma sampler models included 800SL, 900AV and 900 MAX. About 69% of the communities did not indicate how, or if, they measured flow, and did not report any flow data. About 20% of the sites used ISCO 3230 or 4230 flow meters. The remaining 11% used other methods to estimate the stormwater discharge volumes.

Problems Encountered during NPDES Stormwater Monitoring

About 58% of the communities also described problems found during the monitoring process and these are summarized in the site summary reports. Some communities reported more than one problem. One of the basic sampling requirements was to collect three samples every year for each of the land use stations. These samples were to be collected at least one month apart during rains having at least 0.1 inch rains, and with at least 72 hours from the previous 0.1-inch storm event. It was also required (when feasible), that the variance in the duration of the event and the total rainfall not exceeded the median rainfall for the area. About 47% of the communities reported problems meeting these requirements. In many areas of the country, it was difficult to have three storm events per year having these characteristics. The second most frequent problem, reported by 26% of the communities, concerned backwater tidal influences during sampling, or the outfall became submerged during the event. In other cases, it was observed that there was flow under the pipe (flowing outside of the pipe, in the backfill material, likely groundwater), or sometimes there was not flow at all. About 12% of the communities described errors related to malfunctions of the sampling equipment. Most of the communities with equipment failures did not report the reasons of the failure. When reported, the equipment failures were due to incompatibility between the software and the equipment, clogging of the rain gauges, and obstruction in the sampling or bubbler lines. Memory losses in the equipment recording data were also periodically reported. Other reported problems were associated with lighting, false starts of the automatic sampler before the runoff started, and operator error due to misinterpretation of the equipment configuration manual.

Sites located on the East coast (Hampton, VA for example) where the hurricane season produces frequent large storms, especially having a high water table, were especially susceptible. Base flows can commonly occur in separate storm drainage systems for a variety of reasons and they may be more important during some seasons than during others. In many cases, they cannot be avoided and should be included in the monitoring program, and their effects need to be recognized as an important flow phase. As an example, Pitt and McLean (1986) found dry weather base flows to be significant sources of many pollutants, even during a comprehensive research project that spent much time surveying the test watersheds to ensure they did not have any inappropriate discharges entering the storm drainage system.

Capturing runoff events within the acceptable range of rain depth was difficult for some monitoring agencies. Rain depth cannot be precisely predicted in many areas of the country. Also, if using rain gauge data from a location distant from the monitoring location, the reported rain depth may not have been representative of the depth that occurred at the site. The rain gauges need to be placed close to the monitored watersheds. This was likely one of the reasons why the runoff depths periodically exceeded the reported rain depths. Rain in urban areas can vary greatly over small distances. The ASCE/EPA (2002) recommended that rainfall gauges be located as close as possible to the monitoring station. In the NSQD, about 7% of the events had site precipitation estimated using rain gauge located at the city airport. About 16% of the events had precipitation depth estimated using their own monitoring network (Hampton Road Sanitation District, for example). Some communities had precipitation networks that were used for flood control purposes for the surrounding area. These networks can be considered better than the single airport rain gauge, but should at least be supplemented with a rain gauge located in the monitored watershed. Another factor that needs to be considered is the size of the watershed. Large watersheds cannot be represented with a single rain gauge at the monitoring station; in those cases the monitoring networks will be a better approach. Large watersheds are more difficult to represent with a single rain depth value.

Many of the monitoring stations lacked flow monitoring instrumentation, or did not properly evaluate the flow data. Accurate flow monitoring can be difficult, but it greatly adds to the value of the expensive water quality data. As noted previously, base flows also need to be properly removed from the event measurements so only direct runoff quantities are reported. It is probably unreasonable to expect to have a permanent flow monitoring station installed at a location where only manual grab samples are being obtained. However, manual flow monitoring can be conducted during manual sampling by carefully noting the flow stage in previously surveyed locations. These observations will need to be obtained during the complete duration of the event.

The three hour monitoring period that most used may have resulted in some bias in the reported water quality data. This limit was likely used to minimize the length of time personnel needed to be at a monitoring location during manual sampling activities. Also, it is unlikely that manual samplers were able to initiate sampling near the beginning of the events, unless they were deployed in anticipation of an event later in the day. A more cost-effective and reliable option would be to have semi-permanent monitoring stations located at the monitoring locations and sampling equipment installed in anticipation of a monitored event. Most monitoring agencies operated three to five land use stations at one time. This number of samplers, and flow equipment, could have been deployed in anticipation of an acceptable event and would not need to be installed in the field continuously.

Some of the site descriptions lacked important information and local personnel sometimes did not have the needed information. This was especially critical for watershed delineations on maps of the area. Also, few of the watershed descriptions adequately described how the impervious areas were connected to the drainage system, one of the most important factors affecting urban hydrologic analyses. In most cases, information concerning local stormwater controls was able to be determined from a variety of sources, but it was not clearly described in the annual reports.

Comparison of NSQD with Existing Stormwater Databases

The NSQD, with 3,765 events (from the 1992-2002 period) represented sites throughout much of the US for most land uses, and for many constituents. It is therefore the most comprehensive stormwater quality database currently available for US stormwater conditions. The historical NURP database (sampling period in the late 1970s and early 1980s) contains the results from 2,300 national stormwater events, while the CDM National Urban Stormwater Quality Database includes the results of approximately 3,100 events (including the NURP data, plus additional data collected by the USGS and about 30 NPDES permits; Smullen and Cave, 2002). Table 9 compares the results of the pooled EMC’s from the NURP (calculated by Smullen and Cave 2002), CDM, and NSQD databases.

The NURP means and medians were computed by Smullen and Cave (2002) using the EPA (1983) data. The CDM and the NSQD results are similar for all constituents, except for lead and zinc. All three databases have similar reported median and mean concentrations for COD and BOD and the nutrients, but are apparently different for TSS and the heavy metals. The pooled mean event mean concentration (EMC) for TSS was 2.3 times larger in the NURP database compared to the NSQD. The largest reduction in mean EMCs was found for lead (7.9 times larger for NURP) followed by copper (7.9 times larger for NURP) and zinc (1.6 times large for NURP).

Table 9. Comparison of Stormwater Databases

Constituent	Units	Source	Event Mean Concentrations		Number of events
			Mean	Median	
Total Suspended Solids	mg/L	NURP	174	113	2000
		CDM	78.4	54.5	3047
		NSQD	79.1	49.8	3404
Biochemical Oxygen Demand	mg/L	NURP	10.4	8.4	474
		CDM[a]	14.1	11.5	1035
		NSQD	10.9	8.6	2973
Chemical Oxygen Demand	mg/L	NURP	66.1	55.0	1538
		CDM	52.8	44.7	2639
		NSQD	71.2	55.6	2699
Total Phosphorus	mg/L	NURP	0.337	0.266	1902
		CDM	0.315	0.259	3094
		NSQD	0.373	0.289	3162
Dissolved Phosphorus	mg/L	NURP	0.100	0.078	767
		CDM[b]	0.129	0.103	1091
		NSQD	0.107	0.078	2093
Total Kjeldahl Nitrogen	mg/L	NURP	1.67	1.41	1601
		CDM	1.73	1.47	2693
		NSQD	1.74	1.37	3034
Nitrite and Nitrate	mg/L	NURP	0.837	0.666	1234
		CDM	0.658	0.533	2016
		NSQD	0.767	0.606	2983
Copper	μg/L	NURP	66.6	54.8	849
		CDM	13.5	11.1	1657
		NSQD	17.8	14.2	2356
Lead	μg/L	NURP	175	131	1579
		CDM	67.5	50.7	2713
		NSQD	24.4	16.5	2250
Zinc	μg/L	NURP	176	140	1281
		CDM	162	129	2234
		NSQD	110	88	2888

Note: a. No BOD_5 for USGS dataset. b. No DP for CDM portion of NPDES dataset

In an effort to recognize why differences were observed between the NURP and NSQD databases, further examinations of two communities that monitored stormwater during both NURP and the Phase I NPDES program were made. As part of their MS4 Phase I application, Denver and Milwaukee both returned to some of their earlier sampled monitoring stations used during the local NURP projects (EPA 1983). In the time between the early 1980s (NURP) and the early 1990s (MS4 permit applications), they did not detect any significant differences, except for large decreases in lead concentrations. Figure 51 compares suspended solids, copper, lead, and zinc concentrations at the Wood Center NURP monitoring site in Milwaukee. The average site concentrations remained the same, except for lead, which decreased from about 450 to about 110μg/L, as expected due to the decrease in leaded gasoline during this period.

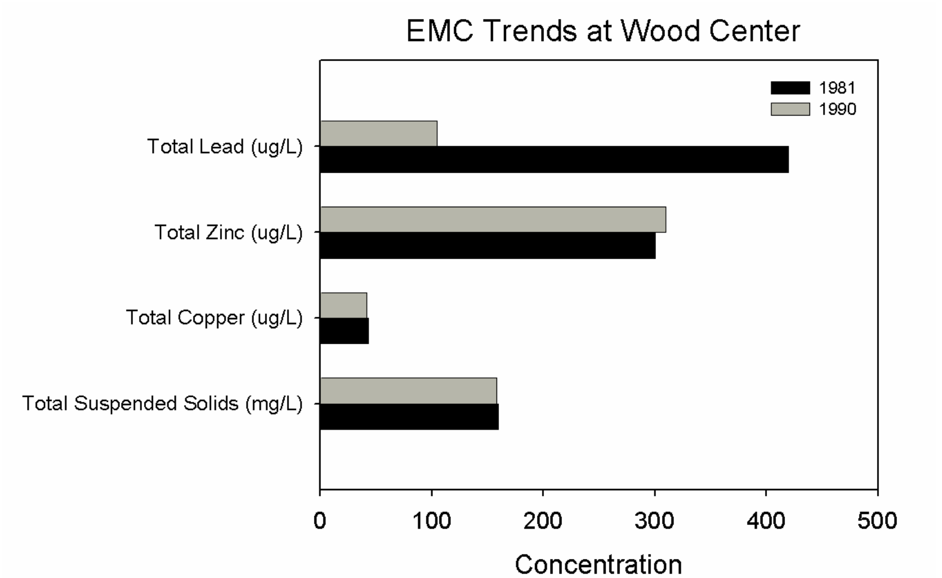


Figure 51. Comparison of pollutant concentrations collected during NURP (1981) to MS4 application data (1990) at the same location (personal communication, Roger Bannerman, WI DNR)

Urban Drainage and Flood Control District performed similar comparisons in the Denver Metropolitan area. Table 43 compares stormwater quality for commercial and residential areas for 1980/81 (NURP) and 1992/93 (MS4 application). Although there was an apparent difference in the averages of the event concentrations between the sampling dates, they concluded that the differences were all within the normal range of stormwater quality variations, except for lead, which decreased by about a factor of four.

Trends of stormwater concentrations with time can also be examined using the NSQD data. A classical example would be for lead, which is expected to decrease over time with the increased use of unleaded gasoline. Older stormwater samples from the 1970s typically have had lead concentrations of about 100 to 500µg/L, or higher (as indicated above for Milwaukee and Denver), while most current data indicate concentrations as low as 1 to 10µg/L.

Table 43. Comparison of Commercial and Residential Stormwater Runoff Quality from 1980/81 to 1992/93 (Doerfer, 1993)

Constituent	Commercial		Residential	
	1980 - 1981	1992 - 1993	1980 - 1981	1992 - 1993
Total suspended solids (mg/L)	251	165	226	325
Total nitrogen (mg/L)	3.0	3.9	3.2	4.7
Nitrate plus nitrite (mg/L)	0.80	1.4	0.61	0.92
Total phosphorus (mg/L)	0.46	0.34	0.61	0.87
Dissolved phosphorus (mg/L)	0.15	0.15	0.22	0.24
Copper, total recoverable (µg/L)	27	81	28	31
Lead, total recoverable (µg/L)	200	59	190	53
Zinc, total recoverable (µg/L)	220	290	180	180

The differences found in both the NURP and the NSQD databases are therefore most likely due to differences in geographical areas emphasized by each database. Figure 10 is a national map showing the percentage of events collected in each state as contained in the NSQD database, while Figure 11 shows the percentage of events contained in the NURP database. Half of the events included in the NSQD database were collected in EPA Rain Zone 2 (Maryland, Virginia, North Carolina, Kentucky and Tennessee), while half of the events contained in the

NURP database were collected in EPA Rain Zone 1 (Minnesota, Wisconsin, Michigan, Illinois, New York, Massachusetts and New Hampshire). Only 3% of the events in the NSQD are located in EPA Rain Zone 1, while 50% of the NURP data is from this area. Twenty four percent of the NURP data is located in the Mid-Atlantic and southeast states, while 60% of the NSQD data is from this area (the area that was emphasized for this EPA-funded project). The NSQD is slightly better representative of other parts of the country compared to NURP. As an example, the percentage of the total event data from the west coast is similar for both databases, but the NSQD represents 10 communities with almost 60 different sites, while NURP has only 3 communities and only 7 sites. The total number of sites, communities and events collected in the NURP study are shown in Table 10.

Table 10. Total Events Monitored During NURP by EPA Rain Zones

Rain Zone	Total Events	Percentage of Events	Number of Communities	Number of Sites
1	804	51	12	42
2	324	20	3	10
3	65	4.1	1	5
4	0	0	0	0
5	24	1.5	1	2
6	45	2.8	2	5
7	136	8.6	1	2
8	0	0	0	0
9	188	12	3	12

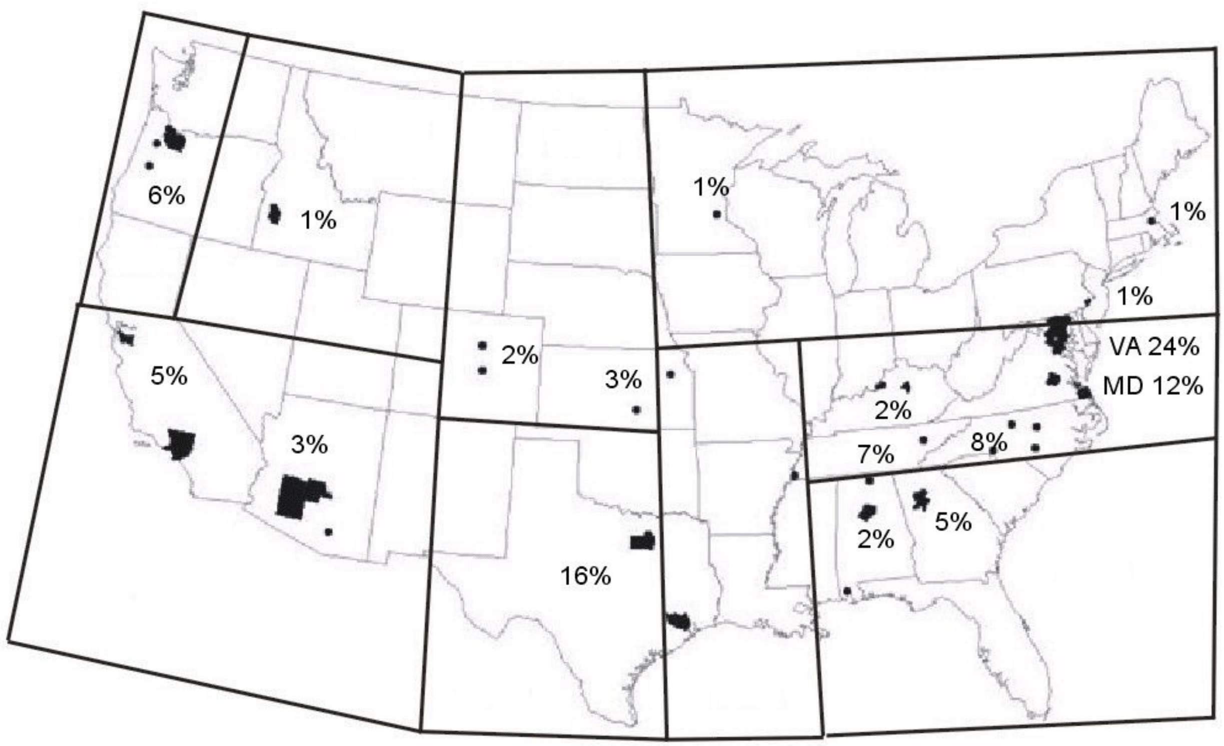


Figure 10. Distribution of collected events using the NSQD database.

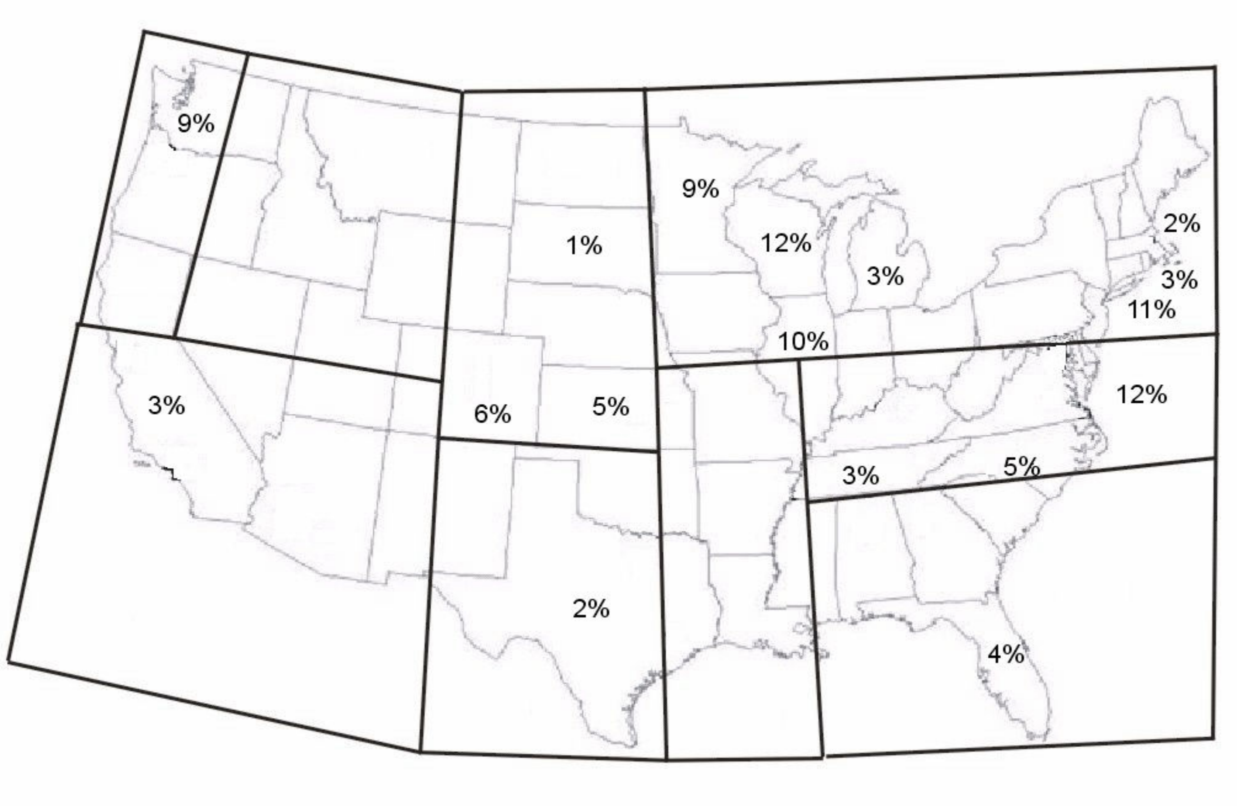


Figure 11. Distribution of collected events using the NURP database.

Figure 12 presents example plots for selected residential area data for different EPA Rain Zones for the country as contained in the NSQD. Rain Zones 3 and 7 (the wettest areas of the country) had the lowest concentrations for most of the constituents, while Rain Zone 1 has some of the highest concentrations.

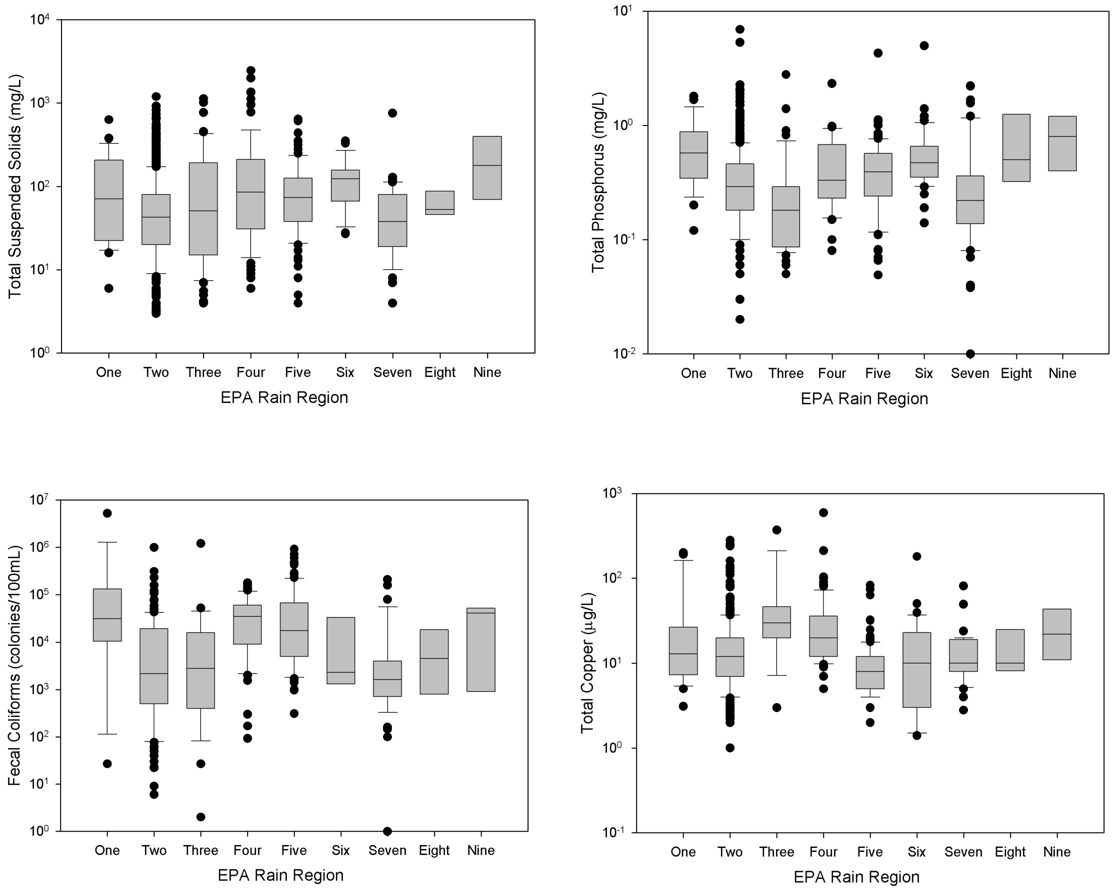


Figure 12. Example of constituents collected in residential land use by EPA Rain Zone

It is likely that the few data from EPA Rain Zone 1 (having relatively high concentrations) in the NSQD and the few data in EPA Rain Zones 2 and 3 (having relatively low concentrations) in NURP are the main reason for the differences in the database summary values.

Land use effects

Another factor that may affect the difference in reported concentrations between the NURP and NSQD databases is the percentage of samples collected for each different land use category. Although each database summarized observed concentrations by land use, having few data from few sites in a land use category reduces the reliability of the estimate. Almost 45% of the NURP database represents residential sites, while residential sites comprise about 30% of the NSQD. The percentage of industrial sites in the NSQD is 15%, while industrial sites in the NURP database represent only 6% of the total. The NSQD contains samples for freeways sites, which are not included in the NURP database. The percentages of mixed land uses and commercial areas are similar for both databases. However, a better representation of open space land uses was observed in the NURP database (10% of the total) compared with the NSQD (3% of the total).

Other Factors

Other factors may influence the differences in reported EMCs in the different databases. Figure 13 shows the probability plot for drainage areas for sites included in the NSQD and NURP databases.

This plot shows that the NURP watersheds are larger than those observed in the NSQD. The median NSQD drainage area was about 50 acres, while it was about twice as large during NURP. The NSQD also has about 10% of the watersheds smaller than 10 acres, representing freeways sites. No literature was found that indicates that there is a relationship between the drainage area and the concentration of stormwater constituents.

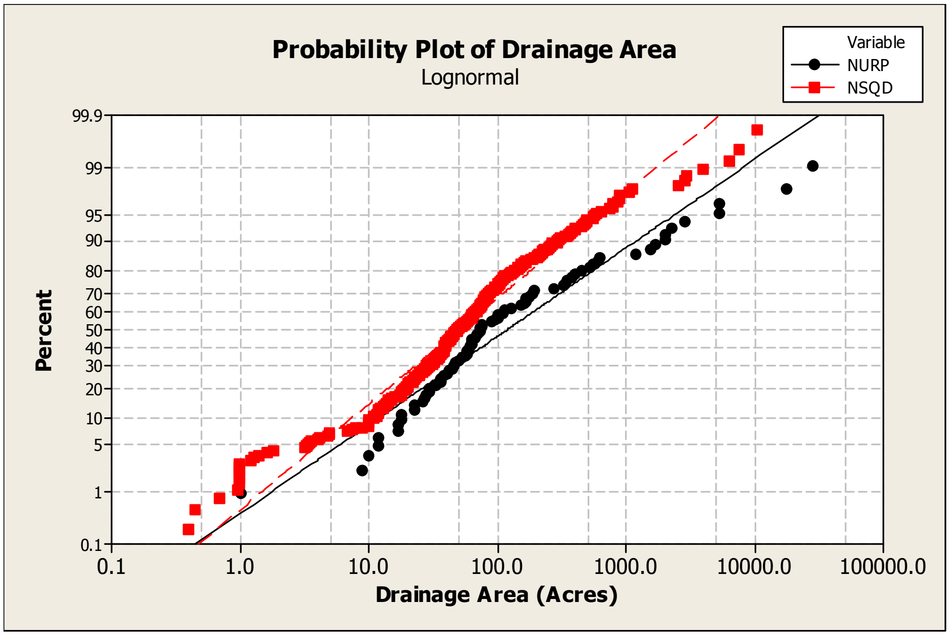


Figure 13. Distribution of collected events using the NURP database

Chapter Summary

This chapter describes the National Stormwater Quality Database. The information collected from the NPDES Phase I stormwater monitoring program was stored in a spreadsheet containing more than 3,700 rows and 250 columns. Each row represents a single monitored event. The main structure of the database is divided into six sections: site descriptions, hydrologic information, conventional constituents, nutrients, metals, and additional constituents. The collected data is grouped into 11 land use categories: residential, commercial, industrial, open space, freeways, mixtures of these land uses, and institutional. Support documents were also created for each community. These documents include aerial photos of the watershed and outfall area (when available), narrative descriptions about the main activities and land uses in the watersheds, sampling and quality control procedures, analytical methods, and equipment used during the collection and analysis of the samples. The last part of the support documents describe the problems that occurred during the collection and analyses of the samples, and meeting discharge permit requirements that specified sampling requirements. This information is useful for interpreting the reported monitoring data and as guidance for future stormwater programs in other communities around the country.

The data from the NSQD was compared with information from the most commonly used stormwater database, the EPA’s Nationwide Urban Runoff Program (NURP) conducted more than 20 years ago. It was observed the concentrations in the NSQD were in general lower than those found during the NURP program. The analysis indicates that the main reason of these differences is the geographical differences represented by the monitoring locations represented in the databases. Most of the samples during the NURP program were collected in the upper Midwest and northeast coast areas of the country, while most of the samples represented in the NSQD were collected in the mid-east coast and southeast areas of the country. The preliminary regional analyses shown in this chapter indicate that southeast areas have lower stormwater concentrations than northeast areas.

Chapter 3: QA/QC Procedures

Introduction

This chapter presents the quality assurance and quality control procedures followed during the creation of the database. These tasks relied on two basic activities: identification of unusual observations and monitoring locations, and the examination of alternative methods to address non-detected pollutant concentration observations (left-censored data).

Quality Control/ Quality Assurance

More than 70 communities were contacted to request information concerning their NPDES Phase I monitoring activities. Communities submitted their reports in either electronic media or on paper. In cases where the data were in electronic form, the data were manipulated with macros and stored in the main Excel spreadsheet. For those communities with data only on paper, the information was typed directly into the spreadsheet.

Once the database was completed, the main table was first reviewed by rows (corresponding to individual runoff events) and then by columns (corresponding to measured constituents). Each row and column in the database was reviewed at least once and compared to information contained in the original reports (when available). For each constituent, probability plots, box and whisker plots, and time series plots were used to identify possible errors (likely associated with the transcription of the information, or as typographical errors in the original reports). Most of the identified errors were associated with the transcription process and, in some cases, errors associated with incorrect units (such as some metal results reported as mg/L when they were really as µg/L).

Additional "logical" plots were used to identify possible errors in the database. A plot of the dissolved (filtered) concentrations against the total concentrations for metals should indicate that the dissolved concentrations are lower than the total forms, for example (Figure 14). Other plots included TKN versus NH_3, COD versus BOD_5, SS versus turbidity and TDS versus conductivity.

In all cases, suspect values were carefully reviewed and many were found to be associated with simple transcription errors, or obviously improper units, which could be corrected. However, about 300 suspect values were removed from the database as they could not be verified. None of the data were deleted without sufficient evidence of a highly probable error. For example, if a set of samples from the same community had extremely high concentrations (in one case, 20 times larger than the typical concentrations reported for other events for the same community) at different sites, but for the same event, this will indicate a very likely error during the collection or analysis of the sample. If just a single site had high concentrations (especially if other related constituents were also high), it would not normally be targeted for deletion, but certainly subject to further scrutiny. If a value was deleted from the database, or otherwise modified, a question mark notation was assigned to the respective constituent in the qualifier column. Appendix B includes all the modifications performed in the database.

In order to calculate the standard deviations for the site quality control tests, each location must have at least two observations. Nine sites were not included in that analysis because they had only one observation. These sites were: ALHUDRAV, KYLXEHL4, KYLXEHL5, KYLXNEL1, MABOA007, ORCCA001, ORODA001, ORODA002, and ORODA004

Many specific statistical methods were used as part of the QA/QC review, in addition to simple data comparisons on multiple generations of data sheets, and logical patterns. The following is one example that was used to identify unusual monitoring locations and to verify the associated data observations with site characteristics.

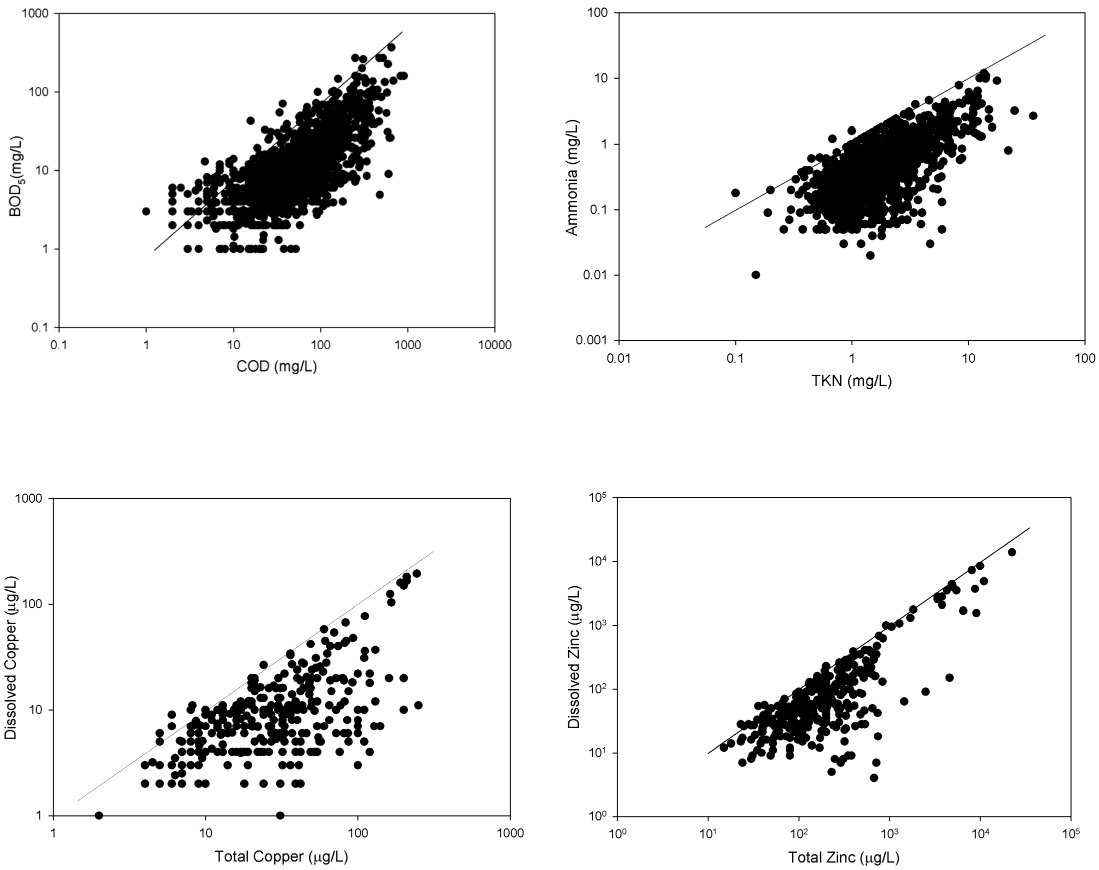


Figure 14. Example scatter plots of stormwater data (line of equivalent concentration shown)

Unusual Monitoring Locations

Box and whisker plots can be used as a preliminary examination of the principal factors and interactions between EPA Rain Zones and land use for any constituent. These plots can also be used to identify sites that do not fit within an established pattern shown by other land use locations from other regions of the country. Figure 23 shows box and whisker plots for residential, commercial, and industrial land uses for EPA Rain Zones 1 through 9. These plots indicate that there are significant differences between EPA Rain Zones and between land uses. Statistical tests also found that the interaction of these two factors was also significant. The median observations by land use have patterns similar to those found during NURP (EPA 1983), and other studies. Residential and open space areas have lower concentrations than commercial and industrial land use areas.

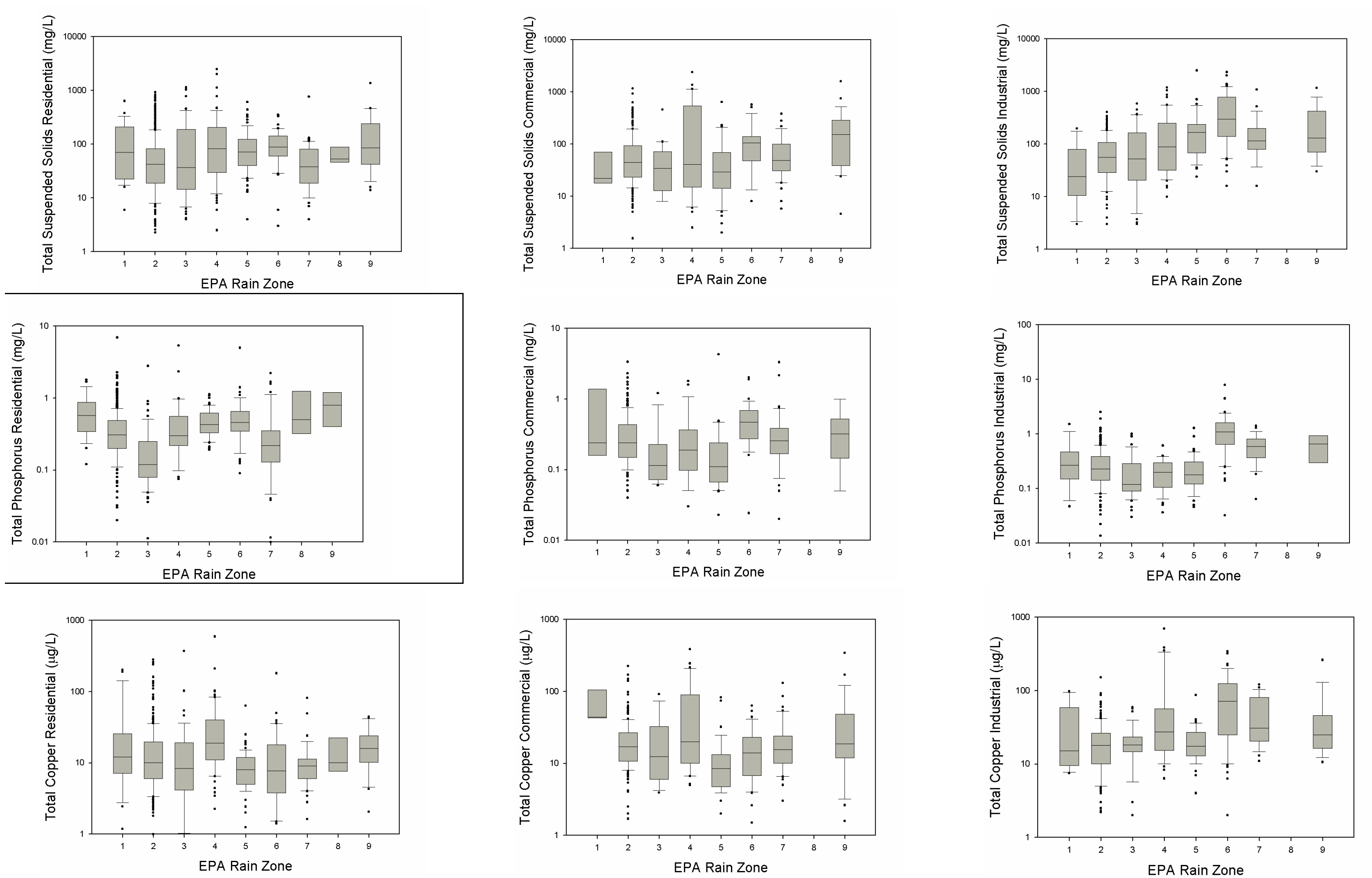


Figure 23. Box and whisker plots for TSS, total phosphorus and total copper by EPA Rain Zone and land use (Two-way ANOVA analyses indicate that both land use and geographical main factors and the interaction are significant at the 0.1% level)

Residential, commercial and industrial areas are the single land uses having the most observations in the database. These three land uses were analyzed separately to identify those sites with different characteristics than the remaining sites in the same land use and EPA Rain Zone. The following is an example using TSS at residential land use sites to demonstrate the method used to detect unusual monitoring sites in the database. Summaries of additional constituents in residential, commercial and industrial land uses are given in Appendix D.

Example Using Single Residential Land Use

The following example explains the steps used to identify unusual locations in the database. This analysis was performed in three steps. First, box and whisker plots we used to identify any site with concentrations unusually high or low compared with the other residential locations. The plot was used to identify preliminary differences between and within EPA Rain Zones. Figure 24 shows that there are some sites in EPA Rain Zone 2 having lower TSS concentrations than the remaining residential sites included in the database. On the other hand, it seems that sites located in EPA Rain Zone 4 have higher concentrations than other groups. The second step was to identify those single residential sites that failed the Xbar and S chart tests for all the observations and by EPA Rain Zone.

A total of 10 Xbar and S charts were created for each EPA Rain Zone and for all the zones combined. An indication of geographical differences is if the Xbar chart using all observations shows clusters close or outside the control limits. The effect will be confirmed if none of the sites failed the Xbar test within EPA Rain Zones. The S chart identifies those sites that have a larger or smaller variation than the overall sites in the set.

Figure 25 shows the Xbar and S chart for the residential land use sites. Six sites have mean TSS values different from the remaining sites in the same group. One important characteristic of this plot is that the control limits change with the number of samples collected at each site. The S chart identifies those sites with standard deviations different than the pooled deviation of the data set. In this case, two sites are outside the control limits. Table 15 shows the sites that failed the Xbar and S chart for all residential sites and for each EPA Rain Zone. Table 15 shows that most of the sites located below the lower control limit were located in North Carolina, Virginia (EPA Rain Zone 2) or Oregon (EPA Rain Zone 7). Sites above the upper control limit were located in Arizona (EPA Rain Zone 6), Kansas (EPA Rain Zone 4), and Colorado (EPA Rain Zone 9).

Xbar plots by EPA Rain Zones also indicate differences within groups. EPA Rain Zones 2, 3, and 4 showed nine sites failing the Xbar test. Six sites out of 54 failed the Xbar chart test in residential land use EPA Rain Zone 2. Each of these sites will be described individually.

The first site was located in Kentland Village (Flagstaff Street), in Prince George County, Maryland (Location_ID = MDPGCOS2, median TSS = 132 mg/L). This site with 63 events has the largest number of observations in the database. An industrial park and a commercial area surrounded this high-density residential site. A special characteristic of this site is the construction of a stadium close of the watershed during the monitoring period.

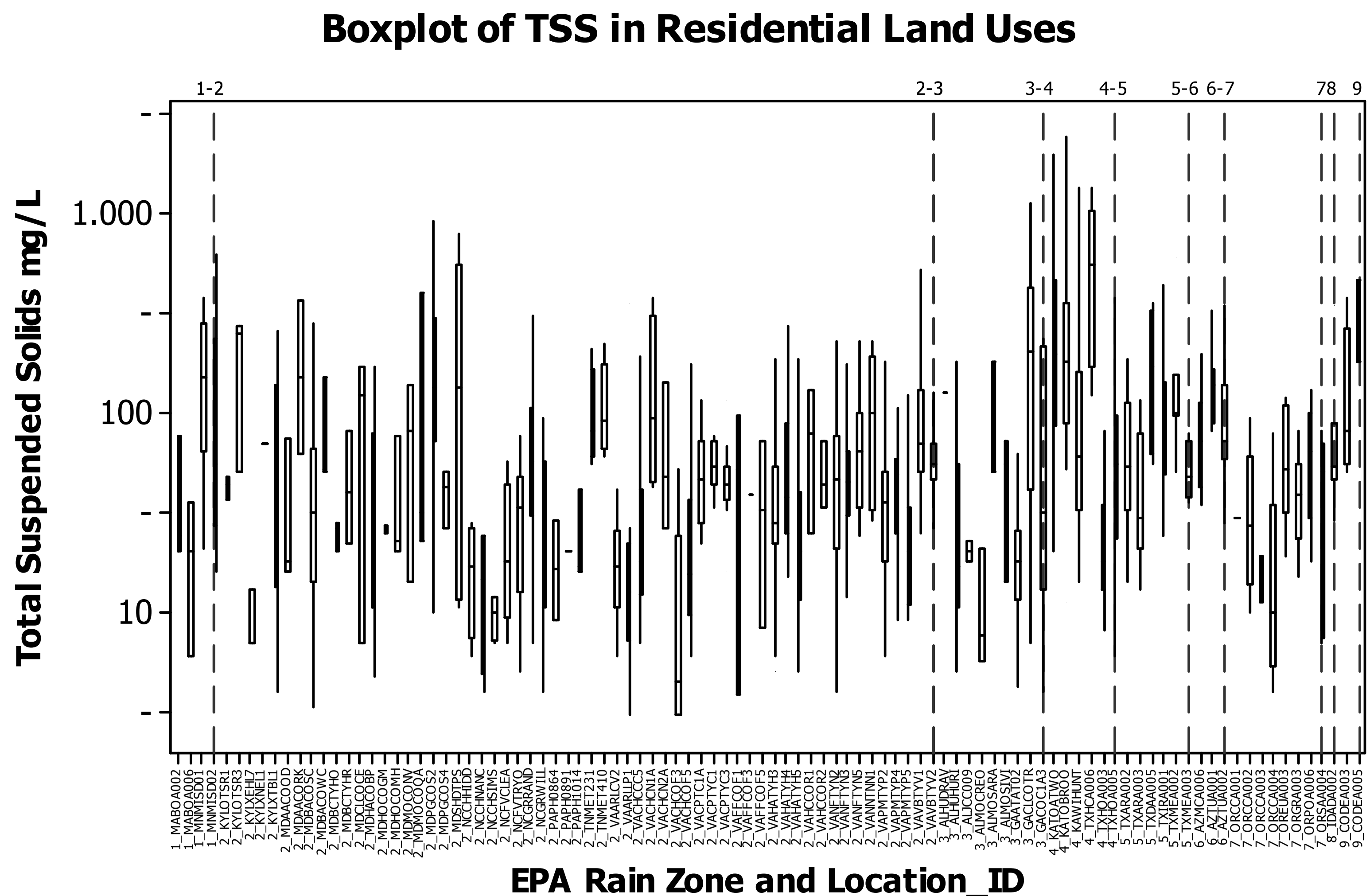


Figure 24. TSS box and whisker plots in residential land use by EPA Rain Zone and location

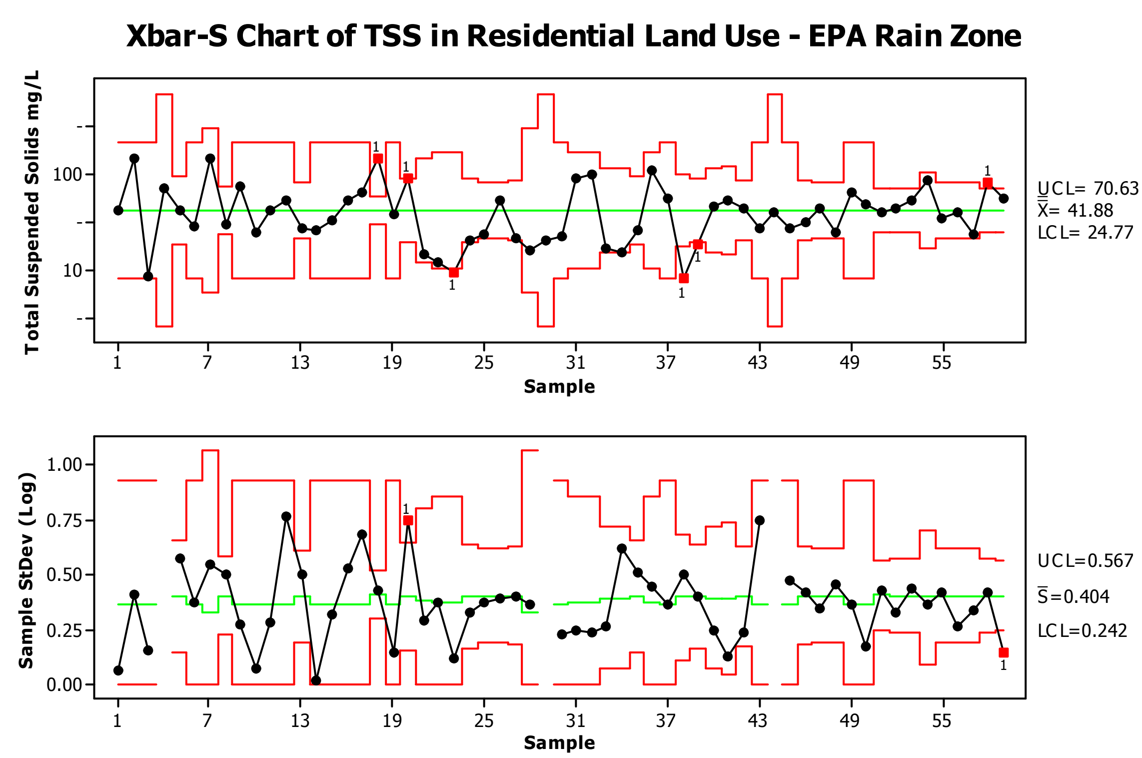


Figure 25. Xbar and S chart for residential land use in EPA Rain Zone 2

Table 15. Sites Failing Xbar and R Chart in Residential Land Uses

Rain Zone	Sites Failing Xbar chart	Sites Failing S Chart
ALL	AZTUA001(H) CODEA005(H) GAATAT02(L) GACLCOTR(H) KATOATWO(H) KATOBROO(H) KYLXEHL7(L) MDPGCOS2(H) MNMISD01(H) NCCHSIMS(L) NCFVCLEA(L) NCFVTRYO(L) NCGRWILL(L) ORCCA004(L) TXHCA006(H) TXHOA003(L) VAARLCV2(L) VAARLLP1(L) VACHCOF3(L) VACHCOF5(L) VAHATYH5(L) VAPMTYP5(L)	VAVBTYV2(L)
1	None	None
2	MDPGCOS2(H) MDSHDTPS(H) NCCHSIMS(L) VACHCOF3(L) VACHCOF5(L) VAVBTYV1(H)	MDSHDTPS(H) VAVBTYV2(L)
3	GACLCOTR(H)	None
4	TXHCA006(H) TXHOA003(L)	None
5	None	None
6	None	None
7	None	None
8	None	None
9	None	None

The second site has 13 observations and was operated by the Maryland State Highway Department (MDSHDTPS, median TSS = 135 mg/L). This 51-acre site is considered 96% single family residential, with 4% agricultural land use. The site is located close to the intersection of two highways. Observed concentrations ranged from 10 mg/L up to 750 mg/L. The highest concentrations were observed in summer and the lowest in spring. Another site in EPA Rain Zone 2 with elevated values has 26 observations and is located close to Bow Creek in Virginia Beach, VA (VAVBTYV1, median TSS = 69). This site is located close to a golf course and is drained by a natural channel.

The site with a standard deviation below the lower control limit (VAVBTYV2) is located next to VAVBTYV1. It has also a high TSS concentration but inside the control limits. A total of 30 samples were collected at VAVBTYV2. The aerial photograph did not indicate any unusual conditions at this site.

In EPA Rain Zone 4, only one site had high concentrations compared with the remaining residential sites. This site (TXHCA006) is located in Harris County, TX. Six samples were collected, having a median TSS of 550 mg/L. This site is also analyzed in Chapter 5 and seems to be affected by flooding or erosion activity. In EPA Rain Zone 3, site GACLCOTR is a new development in Tara Road, Clayton County, and Georgia. Twenty-two samples were collected at this location. The median TSS was 200 mg/L. No unusual conditions were identified when examining the aerial photographs.

Site mean concentrations below the lower control limit in the Xbar chart were located in Virginia, North Carolina and Texas. The two sites located in Virginia are located in Chesterfield County. The first site is located in King Mills Road (VACHCOF3, 10 observations, median TSS = 4 mg/L) and is located in a forested area with less than 20% impervious. The second site (VACHCOF5, 14 observations, median TSS = 15 mg/L) is 50% impervious, but surrounded by a forested area. Only four events were collected at the site between March and August 1993, in Silo Lane, Charlotte, North Carolina (NCCHSIMS, median TSS = 10 mg/L), no unusual characteristics were observed from the aerial photographs. The unusual low concentration site in Houston, Texas is located on Lazybrook Street (TXHOA003, median TSS = 21 mg/L). Freeways (I-610) are located in the north and west part of the watershed. Tall trees surrounding the houses were also observed inside the watershed.

The final step was using ANOVA to evaluate if any EPA Rain Zone was different than the others. The ANOVA table indicated a p-value close to zero, indicating that there are significant differences in the TSS concentration among at least two of the different EPA Rain Zones. The Dunnett's comparison test with a family error of 5% indicate that concentrations in EPA Rain Zones 4 (median TSS= 91 mg/L), 5 (median TSS = 83 mg/L), 6 (median TSS=118 mg/L), 7 (median TSS = 69 mg/L), and 9 (median TSS = 166 mg/L) are significantly higher than the concentrations observed in EPA Rain Zone 2 (median TSS = 49 mg/L).

This same procedure was performed for the following 13 additional constituents in residential, commercial and industrial land use areas: hardness, TSS, TDS, oil and grease, BOD, COD, NO2 + NO3, ammonia, TKN, dissolved phosphorus, total phosphorus, copper, lead and zinc.

Identification of Unusual Sites

The Xbar charts were created for residential, commercial and industrial land uses. In residential areas, 54 sites were identified with at least one constituent out of control. These sites failed when compared with sites in the same EPA Rain Zone. Table 16 shows the sites with more than 4 constituents outside the control limits.

These eight sites were located in EPA Rain Zone 2. Three sites show elevated concentrations, one in all constituents, and another in metals and the third in nutrients. The site located near a golf course in Virginia Beach (VAVBTYV1) shows elevated concentrations in TSS, phosphorus and COD. The site located in Prince George County close to an industrial park (MDPGCOS2), indicated elevated concentrations of total phosphorus, lead and zinc.

The site with the highest number of constituents outside the control limits (10 out of 14 constituents evaluated) was located in Mt. Vernon, Lexington, Kentucky (KYLXTBL1). This site was monitored between 1992 an 1997; it is located close to two high schools and the University of Kentucky. It is interesting that one of the sites having elevated concentrations is located next to one of the sites with a large number of constituents below the lower control limit (VAVBTYV2 is located close to VAVBTYV1). VAVBTYV1 has low concentrations for 6 out of 14 constituents. This indicates that not only can geographical differences be expected; there are also differences between locations in the same EPA Rain Zone. Lead was most frequently found with high concentrations within the same EPA Rain Zone. Eight sites had elevated lead concentrations, while 11 sites had lower concentrations in the same group. The least frequent out-of-bound constituent was oil and grease: none of the sites indicated elevated concentrations of oil and grease when compared with other locations in the same EPA Rain Zone.

Table 16. Sites Failing Xbar Chart in Residential Land Uses

SITE	HA	TSS	TDS	OG	BOD	COD	NO2	NH3	TKN	DP	TP	Cu	Pb	Zn
1MABOA006											L	H		
2KYLOTSR3			H											
2KYLXTBL1	H		H		H	H	H		L	H	H	H	H	H
2MDAACORK													L	
2MDBACOSC							H						L	
2MDBCTYHR									H			H		
2MDCLCOCE														H
2MDHACOBP					L				L		L	L	L	L
2MDHOCOGM					H									
2MDPGCOS2		H									H	L	H	H
2MDSHDTPS		H										L		
2NCCHHIDD							H	H						
2NCCHNANC					L									
2NCCHSIMS		L						H						
2NCFVCLEA			L								L	L		
2NCFVTRYO												L	L	
2NCGRWILL								L		H				
2VAARLCV2					L	L					L		L	
2VAARLLP1			H										L	L
2VACHCN2A												L		L
2VACHCOF3	L	L	L	L		L				L	L	L		
2VACHCOF5		L	L			L					L			
2VACPTSF2												L		L
2VAFFCOF1			L											
2VAHATYH3									L	L				
2VAHATYH5										H				L
2VANFTYN2										H	H			
2VANFTYN3						H							H	
2VANFTYN5													H	
2VAPMTYP2										L			H	
2VAPMTYP4								L		L				
2VAPMTYP5														L
2VAVBTYV1		H				H				H	H			
2VAVBTYV2			H		L		L	L		L		L		L
3GAATAT02												H		
3GACOC1A3					L									
3GACLCOTR		H												
4KATOATWO				L										
4KATOBROO				L										
4TXHOA003		L	H									L		
5TXARA002						L							L	
5TXARA003											H			
5TXDAA005													H	
5TXIRA001													H	
5TXMEA002	H													
5TXMEA003													L	
6AZMCA006												H		
6AZTUA001			H			H								
6AZTUA002													L	
7ORCCA004													L	
7OREUA003	H												H	
7ORGRA003				L										
7ORPOA006	L													
7ORSAA004									L				L	L

Note: H: Site with mean concentrations larger than UCL. L: Site with mean concentrations lower than LCL

In commercial land use areas, six out of 25 locations indicated more than three constituents outside of the control limits (Table 20). Five sites have more than one constituent above the upper detection limit. The site with the largest number was located in Wilhite Drive behind a K-Mart large shopping center in Lexington, Kentucky (KYLXWHL1). This site was monitored between 1992 and 1996. The site indicates elevated nutrients, BOD, hardness and TDS concentrations. The second site was also located in Kentucky. East Land is located in an old commercial area in Lexington (KYLXNEL3). This site has elevated total and dissolved phosphorus concentrations.

Table 17. Sites Failing Xbar Chart in Commercial Land Uses

SITE	HA	TSS	TDS	OG	BOD	COD	NO2	NH3	TKN	DP	TP	Cu	Pb	Zn
2KYLXNEL3								L		H	H			
2KYLXWHL1	H		H		H				H	H	H			
2MDAACOPP							L						L	
2MDHOCODC					L						L			
2MDHOCODC														
2MDPGCOS1		H										H	H	H
2NCGRATHE											H			
2NCGRMERR														L
2VAARLRS3					L	L							L	L
2VACHCCC4	H	L		L	L	L								
2VAHATYH1										L				
2VAHCCOC2														H
2VAPMTYP1										L				
3ALHUMASM													H	
3ALHUWERP													H	
3ALMODAPH					H									
4KATOJACK		H		H								H		H
4TXHOA004				L										
6AZTUA003													H	
7OREUA001	H													
7ORPOA001	L													
9CODEA001												H		H
9CODEA002				H										
9KAWITOWN				L										

Note: H: Site with mean concentrations larger than UCL. L: Site with mean concentrations lower than LCL

A third site having elevated stormwater concentrations was found in Brightseat Road adjacent to Landover Mall in Prince George County, Maryland (MDPGCOS1). This site was monitored between 1992 and 1996. It has elevated TSS, copper, lead and zinc concentrations. A fourth site with elevated stormwater concentrations is located in Topeka, Kansas (KATOJACK). This site is located close to a sand quarry. Median TSS concentrations at this location were close to 600 mg/L. Elevated oil and grease, total lead and total zinc were also found at this location. The last elevated concentration site is located in Denver, Colorado. Cherry Creek at Colfax Avenue (CODEA001) has elevated copper and zinc concentrations. The site is 87% commercial and contains a convention center, hotels and restaurants on 16th Street Mall, the State Capital and other government buildings.

Four out of 25 industrial land use locations indicated more than three constituents with median concentrations outside the upper control limit (Table 18). One site is located in Boston, Massachusetts. The Brighton (MABOA004) watershed drains runoff from warehouses and manufacturing operations associated with mechanical, roofing and electrical activities. According to the site description, there is a large potential for storage of rainfall on rooftops and poorly maintained parking lots and roadways. Extremely high ammonia and TKN concentrations were observed at this location. Another industrial site having high concentrations is located in Greensboro, North Carolina. The site is located at Husband Street (NCGRHUST). Zinc and especially copper concentrations were elevated (median copper = 29 μg/L).

A site located at Santa Fe Shops in Topeka, Kansas (KATOSTFE) had elevated metal concentrations. Railroad activity was present in the watershed. Another industrial site of interest is located on 27th Avenue at the Salt River in Maricopa, Arizona (AZMCA003). It had a median TSS of 668 mg/L. Copper, lead and zinc had extremely high concentrations at this location compared with many other single land uses sites in the database.

Table 18. Sites Failing Xbar Chart in Industrial Land Uses

SITE	HA	TSS	TDS	OG	BOD	COD	NO2	NH3	TKN	DP	TP	Cu	Pb	Zn
1MABOA004								H	H			H		
1MNMISD03												L		
2KYLOTSR2					H									
2KYLXTBL2								L		H	H			L
2MDBACOTC													H	
2MDCHCOIP												L		
2MDPGCOS6		H											H	
2NCCHBREV							H	H						
2NCCHHOSK								H						H
2NCFVWINS														H
2NCGRHUST					H							H		H
2VAARLTC4			H									L	L	L
2VACHCOF1												L		
2VACPTYC5		L	L		L	L	L	L	L	L	L	L		
2VAFFOF10									H				L	
2VAFFOF11					H				H					
2VAHATYH2												H		
2VAVBTYV4		L	L									L		
3ALHUCHIP													H	
3GAATAT01		L										L		
3GACLCOSI										H				
4KATOSTFE												H	H	H
4TNMET211								H						
4TXHCA004								L						
5TXDAA001										H		L		
5TXDAA002		L				L								L
5TXFWA004														H
6AZMCA001							L						L	
6AZMCA003		H					H					H	H	
6AZTUA004												L		
6CAALAL09			L						L					
7ORSAA003					L						L			L
9CODEA007												H		

Note: H: Site with mean concentrations larger than UCL. L: Site with mean concentrations lower than LCL

Non-Detected Analyses

Left-censored data refers to observations that are reported as below the limits of detection, while right-censored data refers to over-range observations. Unfortunately, many important stormwater measurements (such as for filtered heavy metals) have large fractions of undetected values. These missing data greatly hinder many statistical tests. A number of methods have been used over the years to substitute appropriate values for these missing data in order to perform statistical tests:

- ignore the non-detects and report only using the detected values (also report the detection limit and the frequency of missing data). This may be suitable for the most basic summaries of the data.
- replace the non-detects with zero. This is the method suggested by the EPA for reporting discharge quantities associated with discharge permits. This method results in a decreased discharge estimate by assuming that the non-detects are actually associated with no pollutants in the waste stream.
- replace the non-detects with the detection limit. This would result in an increased discharge amount when conducting mass balances.
- replace with half the detection limit. This is usually the most common method used, but still may result in biased results. The biggest problem with any of these set value replacement methods is that a single value is used for each missing data value. This can therefore have dramatic effects on the calculated variance of the data set and makes statistical comparison tests error prone.
- replace with a randomly generated value based on the measured variation of the available observations. This is usually the preferred method as the variation of the data set is preserved, allowing suitable non-paired comparison tests. Paired tests cannot be conducted as there is no knowledge of which values belong with which observation.
- report the actual instrument reading, even if below the "minimum quantification limit" or "method detection limit." This is the best method, from a statistical standpoint, but is rarely available. Most of the detection limits are extremely conservative, especially in comparison with the other errors associated with a monitoring program. The use of "substandard" detection limits enables the use of all statistical tests, however, care must be taken to describe the detection limit methodology and the actual instrumentation errors.

Berthouex and Brown (2002) has an extended discussion of some of these methods applied to environmental analyses. To estimate the problems associated with censored values, it is important to identify the probability distributions of the data in the dataset and the level of censoring. Most of the constituents in the NSQD followed a lognormal distribution (See Chapter 4). Appendix C shows several approaches to analyze censoring observations with single and multiple detection limits. Different comparisons substantiated the conclusion that the non-detected values in the NSQD can be best estimated using the Cohen's maximum likelihood method (a method that randomly generates the missing data based on the known probability distributions of the data), compared to other traditional methods.

The values of the detection limits and their frequencies varied among the different constituents and monitoring locations. This made handling the non-detectable values even more confusing, as each constituent had several detection limits. Therefore, the first step in evaluating the different methods to address censored data was to identify the probability distribution of the dataset. The second step was applying and evaluating the different estimation methods.

Censored Data Distribution

The level of censoring for each constituent was calculated for each land use and site, for 18 selected constituents. These constituents contained low levels of censored values. The National Council of the Paper Industry for Air and Stream Improvement found that for levels of censoring (non-detectable observations) above 60%, the use of any estimation method is not appropriate (NCASI 1995). Table 11 shows the maximum, minimum, and percentage of detected values by constituent for each main land use for the complete dataset. In general, freeway sites have the largest percentage of detected observations, while open space sites have the highest percentage of non-detected observations. This is expected as freeway areas have the highest concentrations and open space areas have the lowest concentrations of most reported constituents.

The constituents having greater than 95% detected observations (of these 18) are conductivity, pH, hardness, TSS, TDS, and COD (except for open space areas). Most of the non-detected observations of these 18 constituents were for oil and grease, dissolved phosphorus, lead, and nickel analyses. The percentage of detected observations for these constituents in open space areas varied between 18% and 75%, while freeways recorded valid values for 89% to 100% of the analyses for the metals.

Residential, commercial and industrial land uses have similar percentages of detected observations for each constituent shown in Table 11. The most frequent detection limit for each constituent was also identified. Because of the duration of the monitoring activities reported in the NSQD, the large number of municipalities involved, and the large number of analytical methods used, each constituent usually had several reported detection limits. The number and percentage of non-detected observations at each detection limit was calculated with respect to the total number of non-detected observations. For example, there are a total of 60 oil and grease observations at freeway sites: 43 detected and 17 non-detected. There were three separate detection limits reported for the non-detected oil and grease observations: < 0.5, < 1 and < 3 mg/L with 1, 2 and 14 observations reported for each, respectively. The frequency distribution of non-detected oil and grease observations at freeways sites was therefore 5.8%, 11.8% and 82.3%, respectively. The results for the remaining land uses and constituents are shown in Table 12. A discussion about the percentage of the detected values and their distributions for each constituent is presented in Appendix D.

Table 11. Percentages of Detected Values by Land Use Category and for the Complete Database

Constituent	Land use*	Total Events	Minimum Detected Concentration	Maximum Detected Concentration	Percentage with detected values
Conductivity (μS/cm)	RE	106	27.3	2020	100
	CO	66	17	894	100
	ID	108	42	1958	100
	OP	2	75	150	100
	FW	86	20	870	100
	TOTAL	685	16.8	5955	100
Hardness (mg/L)	RE	250	3	401	100
	CO	139	1.9	356	100
	ID	138	5.5	888	96.4
	OP	8	11	270	100
	FW	127	5	1000	100
	TOTAL	1082	1.9	1100	98.7
Oil and Grease (mg/L)	RE	533	0.2	2980	57.8
	CO	308	0.8	359	70.8
	ID	327	0.5	11000	65.1
	OP	19	0.5	4	36.8
	FW	60	3	30	71.7
	TOTAL	1834	0.2	11000	66.1
Total Dissolved Solids (mg/L)	RE	861	3	1700	99.2
	CO	399	4	3860	99.5
	ID	412	4.5	11200	99.5
	OP	45	32	542	97.8
	FW	97	12	470	90.0
	TOTAL	2956	3	17900	99.3
Total Suspended Solids (mg/L)	RE	991	3	2426	98.6
	CO	458	3	2385	98.3
	ID	427	3	2490	99.1
	OP	44	3	980	95.5
	FW	134	3	4800	99.3
	TOTAL	3389	3	4800	98.8
BOD (mg/L)	RE	941	1	350	97.6
	CO	432	2	150	97.4
	ID	406	1	6920	95.3
	OP	44	1	20	86.4
	FW	26	2	89	84.6
	TOTAL	3105	1	6920	96.2
COD (mg/L)	RE	796	5	620	98.9
	CO	373	4	635	98.4
	ID	361	2	1,260	98.9
	OP	43	8	476	76.7
	FW	67	2.44	1,013	98.5
	TOTAL	2,750	1	1,260	98.4

RE = residential; CO=commercial; ID=industrial; OP=open space; FW=freeways
Total=total database, all land uses combined, including mixed land uses

Table 11. Percentages of Detected Values by Land Use Category and for the Complete Database - *Continuation*

Constituent	Land use	Total Events	Minimum Detected Concentration	Maximum Detected Concentration	Percentage of detected values
Fecal Coliform (Colonies/100mL)	RE	446	1	5,230,000	88.3
	CO	233	4	610,000	88.0
	ID	297	2	2,500,000	87.9
	OP	23	650	63,000	91.3
	FW	49	50	70,000	100
	TOTAL	1704	1	5,230,000	91.2
Fecal Streptococcus (Colonies/100mL)	RE	305	20	840,000	89.59
	CO	181	20	1,100,000	91.79
	ID	195	22	6,000,000	93.9
	OP	22	160	101,000	90.9
	FW	25	560	130,000	100
	TOTAL	1141	20	6,000,000	94.0
Ammonia (mg/L)	RE	595	0.01	6	81.5
	CO	299	0.02	8	83.3
	ID	253	0.03	10	83.4
	OP	32	0.07	2	18.8
	FW	79	0.08	12	87.3
	TOTAL	1908	0.01	12	71.3
NO2 + NO3 (mg/L)	RE	927	0.01	18	97.4
	CO	425	0.03	8.21	98.1
	ID	417	0.02	8.4	96.2
	OP	44	0.09	3.33	84.1
	FW	25	0.1	3	96.0
	TOTAL	3075	0.01	18	97.3
TKN (mg/L)	RE	957	0.05	36	95.6
	CO	449	0.05	15	96.8
	ID	439	0.05	25	97.3
	OP	45	0.2	5	95.9
	FW	125	0.2	36	71.1
	TOTAL	3191	0.05	66	96.8
Dissolved Phosphorus (mg/L)	RE	738	0.01	2	84.2
	CO	323	0.01	2	81.1
	ID	325	0.02	2	87.4
	OP	44	0.01	1	79.6
	FW	22	0.06	7	95.5
	TOTAL	2477	0.01	7	85.1
Total Phosphorus (mg/L)	RE	963	0.01	7	96.9
	CO	446	0.02	3	95.7
	ID	434	0.02	8	95.9
	OP	46	0.02	15	84.8
	FW	128	0.06	7	99.2
	TOTAL	3285	0.01	15	96.5

Table 11. Percentages of Detected Values by Land Use Category and for the Complete Database - *Continuation*

Constituent	Land use	Total Events	Minimum Detected Concentration	Maximum Detected Concentration	Percentage of detected values
Total Copper (µg/L)	RE	799	1	590	83.6
	CO	387	1.5	384	92.8
	ID	415	1.97	1360	89.6
	OP	39	2	210	74.4
	FW	97	5	244	99.0
	TOTAL	2723	0.6	1360	87.4
Total Lead (µg/L)	RE	788	0.5	585	71.3
	CO	377	1	689	85.4
	ID	411	1	1200	76.4
	OP	45	0.2	150	42.2
	FW	107	1.6	450	100
	TOTAL	2949	0.2	1200	77.7
Total Nickel (µg/L)	RE	419	1	100	45.4
	CO	232	2	110	59.5
	ID	249	1	110	62.7
	OP	38	12	120	18.4
	FW	99	2.8	100	89.9
	TOTAL	1430	1	120	59.8
Total Zinc (µg/L)	RE	810	3	1580	96.4
	CO	392	5	3050	99.0
	ID	432	5.77	8100	98.6
	OP	45	5	390	71.1
	FW	93	6	1829	96.8
	TOTAL	3007	2	22500	96.6

Table 12. Percentages of Non-detected Values for Different Reported Detection Limits by Land Use and for the Total Database

Constituent	Land use*	<**	<0.5	<0.6	<1	<1.2	<1.4	<1.9	<2	<2.24	<2.47	<2.5	<2.9	<3	<3.97	<5	<5.2	<6	<6.5	<7	<8.3	<10	<14	<17
Oil and Grease (mg/L)	RE	0.44	1.78		36.89		2.22		5.78	0.44	0.44	3.56		1.78		39.56		2.67	0.89		0.44	3.11		
	CO		2.22	1.11	34.44		2.22		5.56			3.33	1.11	4.44		40.00		4.44		1.11				
	ID		2.63		51.75		1.75		6.14			3.51				28.95	1.75	2.63				0.88		
	OP				100.00																			
	FW		5.88		11.76									82.35										
	TOTAL	0.16	1.93	0.16	44.69	0.16	2.09	0.16	5.14	0.16	0.16	2.57	0.16	4.02	0.16	32.64	0.32	2.09	0.32	0.16	0.16	2.25	0.16	0.16

*see footnote for Table 11 for definitions of land use categories

** the < sign without a value implies a non-detected value that was not identified

Constituent	Land use	<1	<5	<6	<10
Total Dissolved Solids (mg/L)	RE	14.29	71.43		14.29
	CO				100.00
	ID		50.00	50.00	
	OP		100.00		
	FW	100.00			
	TOTAL	15.00	55.00	5.00	25.00

Constituent	Land use	<0.5	<1	<2	<5	<10
Total Suspended Solids (mg/L)	RE	7.14	7.14		85.71	
	CO	12.50		12.50	62.50	12.50
	ID		25.00		75.00	
	OP		50.00		50.00	
	FW					
	TOTAL	4.76	9.52	4.76	78.57	2.38

Constituent	Land use	<1	<2	<3	<4	<5	<6	<10	<15	<20	<100
BOD (mg/L)	RE	8.70	13.04	8.70	4.35	34.78	8.70	4.35		8.70	
	CO	9.09	18.18	9.09		18.18	9.09	9.09			
	ID	5.26	31.58	5.26		47.37			5.26		
	OP	16.67	50.00	16.67						16.67	
	FW			50.00		50.00					
	TOTAL	33.05	20.34	6.78	3.39	18.64	2.54	1.69	0.85	3.39	1.69

Table 12. Percentages of Non-detected Values for Different Reported Detection Limits by Land Use and for the Total Database - *Continued*

Constituent	Land use	<1	<2	<5	<10	<20	<25
COD (mg/L)	RE	33.33		55.56	11.11		
	CO	50.00		33.33	16.67		
	ID				50.00	25.00	25.00
	OP				20.00	80.00	
	FW				100.00		
	TOTAL	15.91	2.27	15.91	40.91	22.73	2.27

Constituent	Land use	<	<1	<2	<3	<10	<20	<30	<100	<200	<2000	>	>1.6K	>2.4K	>3K	>6K	>12K	>16K	>24K	>30K	>35K	>40K	>60K	>80K	>160K
Fecal Coliform (Colonies/100mL)	RE		7.84	3.92		1.96				3.92			25.49	3.92		1.96		13.73	1.96			1.96	21.57	3.92	7.84
	CO		17.86		3.57	3.57			3.57	3.57			10.71	3.57	3.57			10.71			3.57	3.57	28.57		3.57
	ID		44.44	11.11			2.78	2.78	2.78	5.56			2.78			2.78		8.33					11.11	2.78	2.78
	OP																								100.00
	FW																								
	TOTAL	0.67	22.00	4.00	0.67	2.67	0.67	0.67	1.33	3.33	2.67	0.67	11.33	2.00	2.00	2.67	2.67	13.33	0.67	0.67	0.67	1.33	15.33	2.67	5.33

Constituent	Land use	<1	<2	<4	<200	>16K	>35K	>60K	>80K	>100K	>160K	>240K
Fecal Streptococcus (Colonies/100mL)	RE	21.88	6.25		9.38	15.63	3.13	21.88	6.25	6.25	6.25	3.13
	CO	20.00	20.00	6.67		6.67		46.67				
	ID	50.00	8.33		8.33	25.00		8.33				
	OP					50.00			50.00			
	FW											
	TOTAL	27.94	13.24	1.47	5.88	14.71	1.47	22.06	5.88	2.94	2.94	1.47

Constituent	Land use	**<**	**<0.01**	**<0.02**	**<0.04**	**<0.05**	**<0.1**	**<0.2**	**<0.27**	**<0.3**	**<0.5**	**<1**
Ammonia (mg/L)	RE	30.00		0.91	0.91	13.64	18.18	14.55		1.82	20.00	
	CO	18.00				2.00	22.00	30.00	2.00		26.00	
	ID	23.81				2.38	7.14	28.57			35.71	2.38
	OP	19.23			3.85	7.69					69.23	
	FW							100.00				
	TOTAL	14.63	0.37	0.18	0.55	4.94	9.87	53.75	0.18	0.37	14.81	0.37

Table 12. Percentages of Non-detected Values for Different Reported Detection Limits by Land Use and for the Total Database - ***Continued***

Constituent	**Land use**	<	<0.01	<0.02	<0.03	<0.05	<0.06	<0.1	<0.2	<0.3	<0.5	<1
NO2 + NO3 (mg/L)	RE	16.67		8.33	8.33			37.50	4.17			25.00
	CO	25.00		12.50				50.00		12.50		
	ID	6.25		6.25		6.25		43.75	6.25	6.25		25.00
	OP							100.00				
	FW									100.00		
	TOTAL	15.85	1.22	8.54	4.88	4.88	1.22	35.37	4.88	6.10	3.66	13.41

Constituent	**Land use**	<	<0.01	<0.05	<0.1	<0.2	<0.28	<0.3	<0.5	<1	<1.5
TKN (mg/L)	RE	25.81	3.23		3.23	19.35			25.81	19.35	3.23
	CO	25.00	8.33			16.67			25.00	8.33	16.67
	ID	16.67	5.56			11.11			22.22	38.89	5.56
	OP								100.00		
	FW				50.00					50.00	
	TOTAL	15.71	2.86	0.71	2.86	30.71	0.71	7.86	22.14	13.57	2.86

Constituent	**Land use**	<	<0.001	<0.01	<0.016	<0.02	<0.03	<0.04	<0.05	<0.06	<0.1	<0.12	<0.15	<0.2	<0.5
Dissolved Phosphorus (mg/L)	RE	41.03		0.85	1.71	11.97		1.71	19.66	0.85	12.82	0.85			8.55
	CO	37.70		6.56	1.64	18.03	1.64		11.48	1.64		3.28			18.03
	ID	36.59				4.88			29.27		12.20			2.44	14.63
	OP			11.11					22.22		11.11				55.56
	FW			100.00											
	TOTAL	34.24	0.27	3.80	2.17	13.04	1.90	0.82	14.40	0.82	14.95	0.82	1.09	1.09	10.60

Constituent	**Land use**	<	<0.01	<0.02	<0.03	<0.04	<0.05	<0.06	<0.1	<0.12	<0.15	<0.5
Total Phosphorus (mg/L)	RE	10.00	6.67	3.33		3.33	10.00		30.00	3.33		33.33
	CO	15.79	5.26	10.53						10.53		57.89
	ID	38.89		5.56			11.11	16.67	5.56			22.22
	OP		14.29				28.57		14.29			42.86
	FW				100.00							
	TOTAL	13.91	4.35	9.57	3.48	0.87	7.83	2.61	20.87	2.61	6.09	27.83

Table 12. Percentages of Non-detected Values for Different Reported Detection Limits by Land Use and for the Total Database – *Continued*

Constituent	Land use	<	<0.45	<1	<1.5	<2	<4	<5	<7	<8	<10	<15	<20	<25	<40	<41	<60
Total Copper (µg/L)	RE		0.76	1.53	1.53	3.82	0.76	19.85	0.76	0.76	41.98	2.29	25.95				
	CO	3.57				7.14		14.29			57.14		17.86				
	ID					9.30		9.30		4.65	18.60		48.84	2.33	2.33	2.33	2.33
	OP	20.00									70.00		10.00				
	FW										100.00						
	TOTAL	2.62	0.29	0.87	0.58	4.94	0.87	14.24	0.29	1.16	39.53	0.87	32.56	0.29	0.29	0.29	0.29

Constituent	Land use	<	<0.2	<0.65	<0.7	<1	<1.5	<1.6	<2	<2.5	<3	<4	<5	<6	<7	<9	<10	<15	<20	<25	<30	<40	<42	<50	<53	<55	<60	<100	<200	<250	<500
Total Lead (µg/L)	RE	10.6	0.4	0.4		0.4			2.2	4.0	9.3	1.3	21.7	0.4		7.5	5.8	1.3	4.4	5.8	0.9	3.5	0.4	8.9		0.9	1.3	5.3	3.1		
	CO	38.2			1.8	1.8			3.6	1.8		1.8	9.1			1.8	1.8				3.6	14.6	3.6	1.8				14.6			
	ID	17.5				3.1	1.0	1.0	3.1		3.1		13.4			8.3	2.1	2.1	8.3	5.2		3.1	1.0	9.3	1.0	3.1	2.1	10.3		1.0	1.0
	OP	11.5	3.9						3.9		7.7		11.5									19.2		19.2				23.1			
	FW																														
	TOTAL	14.7	0.3	0.2	0.2	6.2	0.2	0.2	2.6	1.7	4.9	3.3	12.4	0.2	2.0	6.4	14.3	1.1	2.9	4.3	0.9	3.6	0.6	6.1	0.3	1.1	0.9	6.4	2.0	0.2	0.2

Constituent	Land use	<	<0.5	<1	<2	<2.5	<3	<4	<5	<7	<10	<14	<15	<20	<25	<30	<40	<50	<60	<100
Total Nickel (µg/L)	RE	12.66		1.75	7.86	0.87	1.31	1.31	10.92	1.31	6.99		15.72	29.26	2.18	0.44	2.62	3.93	0.87	
	CO	22.34		1.06	6.38		1.06		14.89		6.38		7.45	36.17			2.13	1.06	1.06	
	ID	20.43			3.23		3.23	2.15	9.68	3.23	10.75	1.08	16.13	17.20	4.30	4.30	2.15	1.08		1.08
	OP	16.13			32.26		3.23		6.45					29.03		3.23	3.23	3.23	3.23	
	FW				10.00				60.00					30.00						
	TOTAL	15.30	0.17	1.22	8.17	0.52	1.74	0.87	12.00	1.57	7.65	0.35	14.26	27.30	1.57	1.04	2.26	2.96	0.87	0.17

Constituent	Land use	<	<0.0057	<0.6	<1	<2	<4	<5	<8	<10	<20	<25	<30	<80	<100	<110	<200
Total Zinc (µg/L)	RE	13.79		3.45		10.34		3.45		31.03	13.79			3.45	13.79		6.90
	CO	25.00													50.00		25.00
	ID	16.67					16.67			16.67					16.67	16.67	16.67
	OP	38.46						15.38		15.38	15.38				15.38		
	FW							33.33							66.67		
	TOTAL	15.53	0.97	0.97	5.83	3.88	0.97	3.88	0.97	13.59	17.48	10.68	2.91	0.97	15.53	0.97	4.85

Total lead had the largest number of different detection limits (31 in total) with <10 µg/L as the most frequent censored observation at 14.3%. The constituent with the lowest number of detection limits was TDS, with four levels: < 1, < 5, < 6 and <10 mg/L. Less than 5 mg/L was the most common reported censored TDS observation occurring 55% of the time.

Expected Percentages of Observations at Different Levels of Detection

There are different approved methods to calculate the concentration of a specific constituent in a water sample. *Standard Methods for the Examination of Water and Waste Water* (APHA 1995 and more recent) lists several approved methods for the detection of many of these constituents. The choice of methods presents a problem as these methods have varying features and costs. The objective is usually to select a method with a detection limit that results in useable data for most samples.

The distribution of the data, including the non-detected values, can be used to estimate the percentage of observations that will be detected using different analytical methods. Table 13 shows the expected percentage of observations below a specific detection limit for each of these constituents using the cumulative density function for each constituent and land use. For example, if a stormwater sample is collected at a freeway site and the detection limit of the conductivity method is 100 µS/cm, about 51% of the observations will be not-detects.

Table 13. Percentages of Observations below Specific Concentrations

Constituent	Land use	Percentage of observations smaller than			
		20 µS/cm	100 µS/cm	200 µS/cm	2000 µS/cm
Conductivity (µS/cm)	RE	0	54	84	99
	CO	0	39	82	100
	ID	0	26	72	100
	OP	-	-	-	-
	FW	0	51	85	100
	TOTAL	0	39	73	99

Constituent	Land use	Percentage of observations smaller than				
		1 mg/L	4 mg/L	10 mg/L	160 mg/L	2500 mg/L
Hardness (mg/L)	RE	0	0	5	98	100
	CO	0	4	7	91	100
	ID	0	0	3	95	100
	OP	-	-	-	-	-
	FW	0	0	2	96	100
	TOTAL	0	0.1	3	94	100

Constituent	Land use	Percentage of observations smaller than				
		0.5 mg/L	1 mg/L	2 mg/L	5 mg/L	10 mg/L
Oil and Grease (mg/L)	RE	2	19	31	75	91
	CO	1	11	23	64	87
	ID	1	20	31	66	86
	OP	-	74	-	-	-
	FW	2	5	5	55	75
	TOTAL	0.3	17	29	67	84

Table 13. Percentages of Observations below Specific Concentrations (continued)

Constituent	Land use	Percentage of observations smaller than		
		1 mg/L	5 mg/L	10 mg/L
Total Dissolved Solids (mg/L)	RE	0	0.8	1.5
	CO	0	0	0.5
	ID	0	0	2
	OP	0	0	0
	FW	0	0	0
	TOTAL	0.1	0.7	1.5

Constituent	Land use	Percentage of observations smaller than		
		1 mg/L	5 mg/L	10 mg/L
Total Suspended Solids (mg/L)	RE	0.2	4	11
	CO	0.2	3	9
	ID	0.2	3	5
	OP	0	11	23
	FW	0	2	2
	TOTAL	0.2	3	7

Constituent	Land use	Percentage of observations smaller than		
		1 mg/L	2 mg/L	5 mg/L
BOD_5 (mg/L)	RE	0.2	2	18
	CO	0.2	1	16
	ID	0.2	3	18
	OP	2	11	55
	FW	0	0	31
	TOTAL	1	3	22

Constituent	Land use	Percentage of observations smaller than				
		0.7 mg/L	1 mg/L	5 mg/L	10 mg/L	20 mg/L
COD (mg/L)	RE	0	0.4	1	3	9
	CO	0	1	2	3	7
	ID	0	0	0.5	2	7
	OP	0	0	0	7	37
	FW	0	0	1	4	7
	TOTAL	0	0.2	2	5	13

Constituent	Land use	Percentage of observations smaller than				
		0.01 mg/L	0.05 mg/L	0.1 mg/L	0.2 mg/L	0.5 mg/L
Ammonia (mg/L)	RE	0	3	12	36	71
	CO	0	2	9	28	53
	ID	0	1	7	21	57
	OP	0	11	15	22	93
	FW	0	0	5	20	27
	TOTAL	0.1	2	10	37	65

Table 13. Percentages of Observations below Specific Concentrations (continued)

Constituent	Land use	Percentage of observations smaller than				
		0.01 mg/L	0.05 mg/L	0.1 mg/L	0.2 mg/L	0.5 mg/L
NO2 + NO3 (mg/L)	RE	0	2	5	11	40
	CO	0	1	4	11	40
	ID	0	2	6	11	31
	OP	0	0	18	21	50
	FW	0	0	0	28	72
	TOTAL	0	2	4	10	40

Constituent	Land use	Percentage of observations smaller than				
		0.01 mg/L	0.05 mg/L	0.1 mg/L	0.2 mg/L	0.5 mg/L
TKN (mg/L)	RE	0.1	0.1	0.5	2	6
	CO	0.2	0.2	0.7	2	6
	ID	0.2	0.5	0.7	2	8
	OP	0	0	0	0	44
	FW	0	0	2	2	6
	TOTAL	0.1	0.2	0.6	2	10

Constituent	Land use	Percentage of observations smaller than				
		0.01 mg/L	0.02 mg/L	0.05 mg/L	0.1 mg/L	0.5 mg/L
Dissolved Phosphorus (mg/L)	RE	0.3	3.5	11	32	93
	CO	1	6	21	48	91
	ID	0.3	2.2	16	46	95
	OP	2	7	23	45	93
	FW	5	5	5	14	82
	TOTAL	0.7	4.5	17.5	44.5	94

Constituent	Land use	Percentage of observations smaller than				
		0.01 mg/L	0.02 mg/L	0.05 mg/L	0.1 mg/L	0.5 mg/L
Total Phosphorus (mg/L)	RE	0.2	0.4	1.5	10	28
	CO	0.2	0.6	3	16	82
	ID	0	0.2	3	14	74
	OP	2	2	11	24	80
	FW	0	0	0	3	83
	TOTAL	0.1	0.5	3	12	78

Constituent	Land use	Percentage of observations smaller than				
		2 μg/L	5 μg/L	10 μg/L	20 μg/L	40 μg/L
Total Copper (μg/L)	RE	2.3	14	44	76	92
	CO	0.7	6	26	58	84
	ID	1.2	6	16	46	75
	OP	0	32	54	73	92
	FW	0	0	8	26	58
	TOTAL	1.4	9	31	63	85

Table 13. Percentages of Observations below Specific Concentrations (continued)

Constituent	Land use	Percentage of observations smaller than				
		1 μg/L	3 μg/L	5 μg/L	10 μg/L	50 μg/L
Total Lead (μg/L)	RE	2	14	28	47	88
	CO	0.6	3	8	23	80
	ID	0.7	7	12	24	72
	OP	12	21	33	38	76
	FW	0	3	9	22	72
	TOTAL	2	9	17	36	82

Constituent	Land use	Percentage of observations smaller than				
		1 μg/L	2 μg/L	5 μg/L	10 μg/L	20 μg/L
Total Nickel (μg/L)	RE	1	6	33	55	91
	CO	0.5	3	29	56	92
	ID	0	2	12	33	64
	OP	0	30	39	39	73
	FW	0	1	19	55	84
	TOTAL	0.6	5	26	52	84

Constituent	Land use	Percentage of observations smaller than				
		5 μg/L	10 μg/L	20 μg/L	100 μg/L	200 μg/L
Total Zinc (μg/L)	RE	1	3	7	65	87
	CO	0	0.2	1	28	51
	ID	0.2	0.7	1	24	48
	OP	5	25	35	85	92
	FW	1	2	3	20	51
	TOTAL	0.6	2	4	44	73

Appendix D describes the methods used to analyze censored observations for each constituent. Based on the results presented in Table 13 and these methods, it is possible to estimate the percentage of non-detected observations that can be obtained by constituent and land use. For example, the most frequently reported non-detected ammonia detection limit was 0.2 mg/L. About 37% of the detected and non-detected observations were located below this detection limit. One of the EPA approved methods to measure ammonia has a detection limit close to 0.02 mg/L. If this method was commonly used, the number of non-detected ammonia observations would have been significantly reduced. This is especially evident for metals analyses. Many commercial laboratories use ICP (inductively coupled plasma) procedures for heavy metals, as it is an approved method and generally more efficient than older atomic absorption methods using a graphite furnace. Unfortunately, standard ICP units have greatly reduced sensitivities compared to graphite furnace methods. When filtered heavy metals are to be analyzed, graphite furnace (or ICP-mass spec) methods should be used. It is important that the person conducting a stormwater monitoring program take care in specifying the analytical methods to be used to ensure that most of the data will be usable. Of course, other factors, besides detection limits, must also be considered when selecting analytical methods, including sample preparation, sample storage limits, sample volume needed, safety, cost, disposal problems associated with wastes, interferences, and comparisons with other methods, etc. Burton and Pitt (2002) present a review of many alternative analytical methods that are suitable for stormwater sample analyses.

Effects of Non-detected Observations on Calculating Mean and Standard Deviation Values

The selection of the proper procedure to deal with non-detected values is not an easy task. One option is to ignore the non-detected values and make a statement indicating the percentage of non-detected values found in the dataset. The problem arrives when it is desired to calculate the mean and standard deviation values of a dataset. The presence of non-detected values can strongly bias these parameters, depending on their prevalence. Three methods for dealing with non-detected values were explored during this research: 1) Ignore them; 2) Estimate them with the

Cohen's multi level MLE method for left censored data (NCASI 1995); and 3) replace them with half of the detection limit. In cases were Cohen's method could not be used (i.e. when only two values were detected), half of the detection limit was used as the estimated value to replace the non-detected observations.

Appendix D shows the results for each constituent and land use using the three substitution methods. In general, it was observed that if the censored data were deleted, the mean of the constituent was increased compared to the case where the non-detected values were replaced by half of the detection limit. The same pattern was observed for the standard deviation calculations. The behavior for the coefficient of variation was opposite: the coefficient of variation was reduced when the censored observations were deleted.

When the frequencies of the censored observations were lower than 5%, the means, standard deviations and coefficients of variation were almost identical when the censored observations were replaced by half of the detection limit, or estimated using Cohen's Method. As the percentage of non-detected values increases, replacing the censored observation by half of the detection limit instead of estimating them using the Cohen's maximum likelihood method produces lower means and larger standard deviations.

Effects on Mean, Median and Coefficient of Variation Values at Different Percentages of Censored Observations

As noted above, when the percentage of detected values is high, there are minimal changes in the calculated means, standard deviations, and coefficients of variation for any of the replacement methods. In this discussion, the ratios of the calculated values using the different methods for different frequencies of detection are examined. This analysis identifies the sensitivity of the detection frequencies for each substitution method.

The first task was to evaluate the effect of the substitutions and detection frequencies on the calculated means. When the percentage of detected values is close to 100%, all of the substitution methods produce the same mean, as expected. As the percentage of non-detected values increases, the Cohen's estimated values and half of detection limit methods produces smaller means than if ignored.

Figure 15 is a scatter plot of both ratios (Cohen estimated/ignore and half of the detection limit/ignore) of the calculated mean values. If the scatter plot values formed a line near the 1.0 ratio value, then the "ignore" and the other option would be accurate. If the scatter plot values formed the same line for both of the sets of ratios, then either substitution method would be accurate. The regression equation 3.1 for the Cohen estimated/ignore ratio of calculated mean values has a coefficient of determination of almost 93%. The coefficients in the equation are significant, with a probability that the coefficients are equal to zero smaller than 0.0001.

$$\text{Ratio Mean (Estimated/Ignore)} = 0.316 + 0.0068 * D \qquad (3.1)$$

Where D is the percentage of detected values (0 to 100).

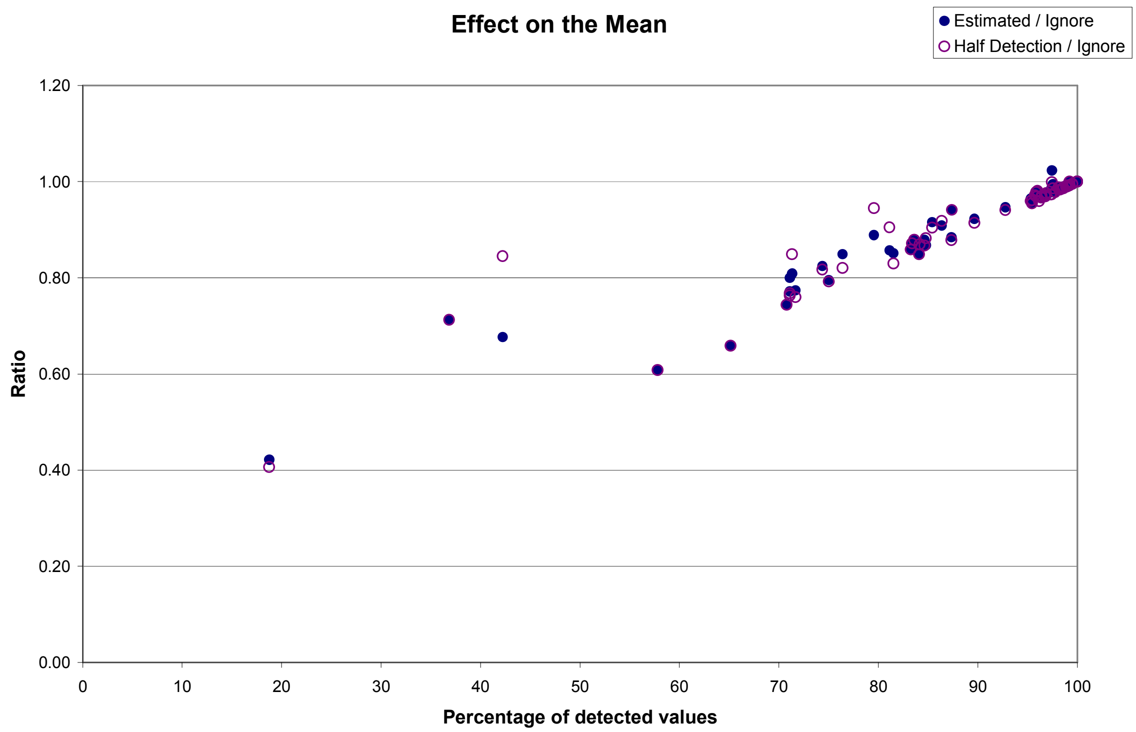


Figure 15. Effects on the mean when using random estimated values versus ignoring the non-detected observations, at different percentage of detected values

For percentages of detected values smaller than 60%, the ratios are located away from the line formed by the other observations. The residual plot of the regression indicates those observations that are most affecting the departure from the regression line. Six observations are considered influential in this plot: oil and grease in open space (most influential), residential and industrial land uses, plus ammonia and lead in open space land uses. The Cook's distance procedure was used to remove the overly influential points in the regression. After removing the influential observations the final regression is therefore:

(3.2) $$\text{Ratio Mean (Estimated/Ignore)} = 0.248 + 0.0075 * D$$

Equation 3.2 indicates that a stormwater dataset having 30% non-detectable observations would have an expected reduction in the calculated mean of 23% when the censored data is appropriately estimated instead of being ignored. The standard deviation of the residuals is 0.014. The coefficient of determination in this case was higher than 96% with no potential or influential points. This equation can be used to estimate the mean of the distribution for data sets with percentages of detected values higher than 60%. When the non-detected observations are replaced by half of the detection limit, the coefficient of determination was reduced to 92% of the actual value. Equation 3.3 describes the relationship between the ratio of the means and the percentage of detected observations.

(3.3) $$\text{Ratio Mean (Half Detection/Ignore)} = 0.250 + 0.0075 * D$$

From the regression of the ratios "estimated/ignore" and "half detection/ignore," replacing by half of the detection limit, or estimating the censored observations using Cohen's method, will produce the same results when the percentage of detected observations is larger than 80%.

The effects on the median are similar to those observed in the mean. When the non-detected values are estimated with Cohen's method instead of ignoring the non-detected values, the regression of the coefficient of determination reduces to 86%.

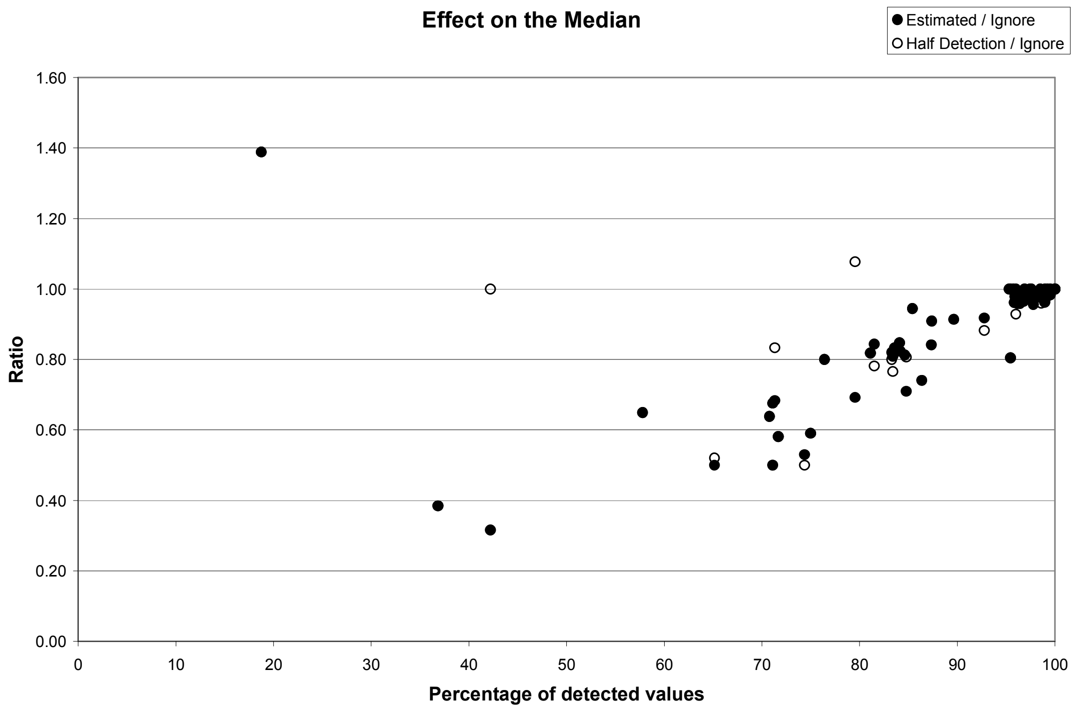


Figure 16. Effect of ignoring the non-detected observations on the median

Equation 3.4 shows the estimated regression line for the median case.

(3.4) $$\text{Ratio Median (Estimate /Ignore)} = -0.326 + 0.0134*D$$

This equation is valid for percentage of detected observations higher than 70%. A reduction of 40% in the median value is expected in a 30% censored dataset when the non-detected observations are estimated using Cohen's method instead of being ignored. The standard deviation of the residuals for this equation is 0.05.

When the censored observations are replaced by half of the detection limit, the coefficient of determination is about 73%. The regression equation for the ratio of the median is therefore not as good in explaining the variability as it was for the mean.

Equation 3.5 shows the calculated regression line for the median when the non-detected values are replace by half of the detection limit.

(3.5) $$\text{Ratio Median (Half Detection/Ignore)} = -0.195 + 0.012*D$$

This equation is valid when the percentage of detected observations is higher than 70%. Replacing the censored observations by half of the detection limit has the same effect on the median as estimating them using Cohen's method, except for dissolved and total phosphorus in open space and lead in residential land uses.

The effects on the calculated standard deviation values also indicate a good correlation between the level of detected observations and the ratio between the "estimate the non-detected or ignore them" values. Figure 17 shows the scatter plot of the median values as a function of the percentage of detected observations. Equation 3.6 presents the estimated regression line of these data.

(3.6) *Ratio Standard Deviation (Estimate/Ignore) = 0.68 + 0.003226*D*

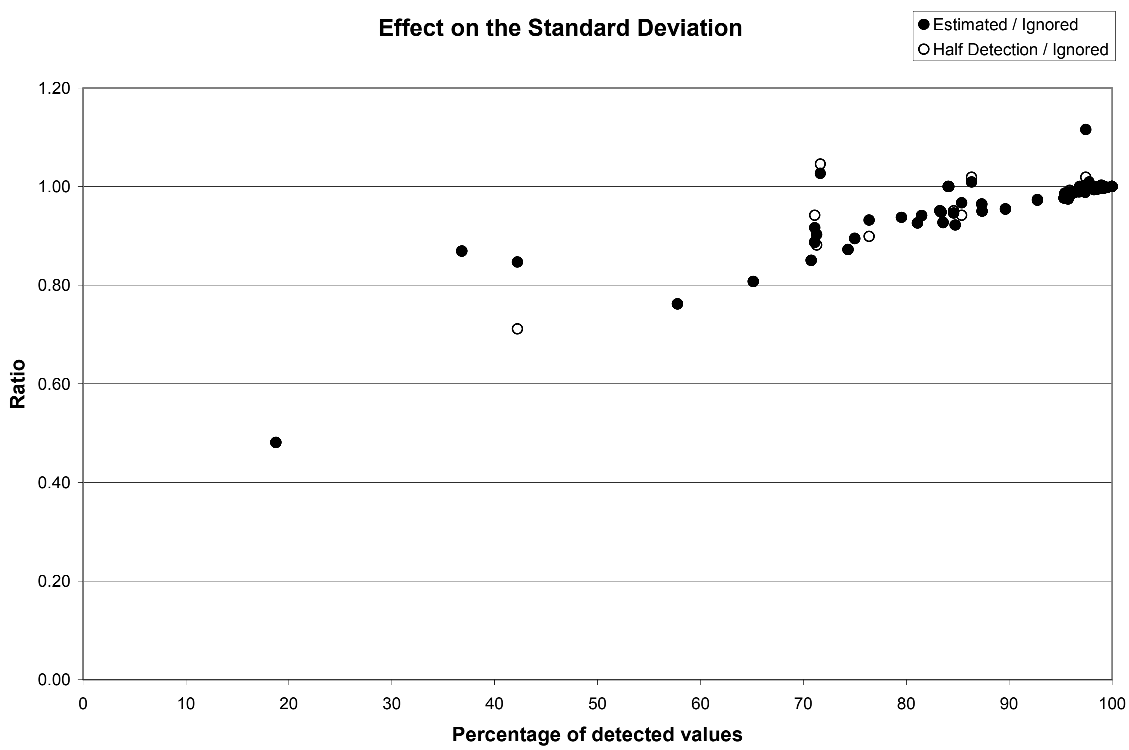


Figure 17. Effect of ignoring the non-detected observations in the median

The regression has a low coefficient of determination (56%) compared to the prior regressions. Oil and grease at freeway sites was considered unusual according to its Cook's distance. The data for this case was examined and no reason was found to eliminate it from the analysis. It was also observed that BOD_5 in commercial land use areas had 3 right-censored observations. Because the Cohen method must be used with left censored observations, these data were eliminated from this analysis. Observations where the percentage of detection was smaller than 70% were not included. Equation 3.7 shows the estimated regression line for those constituents with more than 60% detected observations.

(3.7) *Ratio Standard Deviation (Estimate/Ignore) = 0.68 + 0.003226*D*

This equation indicates that for a dataset with 30% censored observations, the standard deviation will be reduced by 9.5% when the non-detected observations are estimated instead of ignored. The standard deviation of the residuals is 0.023. When the censored observations are replaced by half of the detection limits, the coefficient of determination and the equation coefficients were almost the same. Equation 3.8 presents the estimated regression equation for the standard deviation when the censored observations are replaced by half of the detection limits.

(3.8) *Ratio Standard Deviation (Half Detection/Ignore) = 0.6778 + 0.00325*D*

The last parameter examined was the coefficient of variation. The coefficient of determination (69%) for the fitted regression equation was better than for the standard deviation regression, but not as high as for the median and mean regressions. The calculated regression equation is presented as equation 3.9

(3.9) *Ratio Coefficient of Variation (Estimate/Ignore) = 1.53 -0.0053*D*

The standard deviation of the residuals is 0.033. As the number of non-detected observations increases, the coefficient of variation also increases. The regression equation is valid for percentages of detected values higher than 70%. For a data set with 30% censored observations, the expected coefficients of variation using Cohen's method will be 16% higher than if the non-detected values are ignored.

In the case that the censored observations are replaced by half of the detection limits, the coefficient of determination of the resulting equation (equation 3.10) is reduced to 58%. Figure 18 shows the scatter plot for the ratios "estimated/ignore" and "half detection/ignore" for the coefficient of variation.

(3.10) *Ratio Coefficient of Variation (Half Detection/Ignore) = 1.543 -0.0054*D*

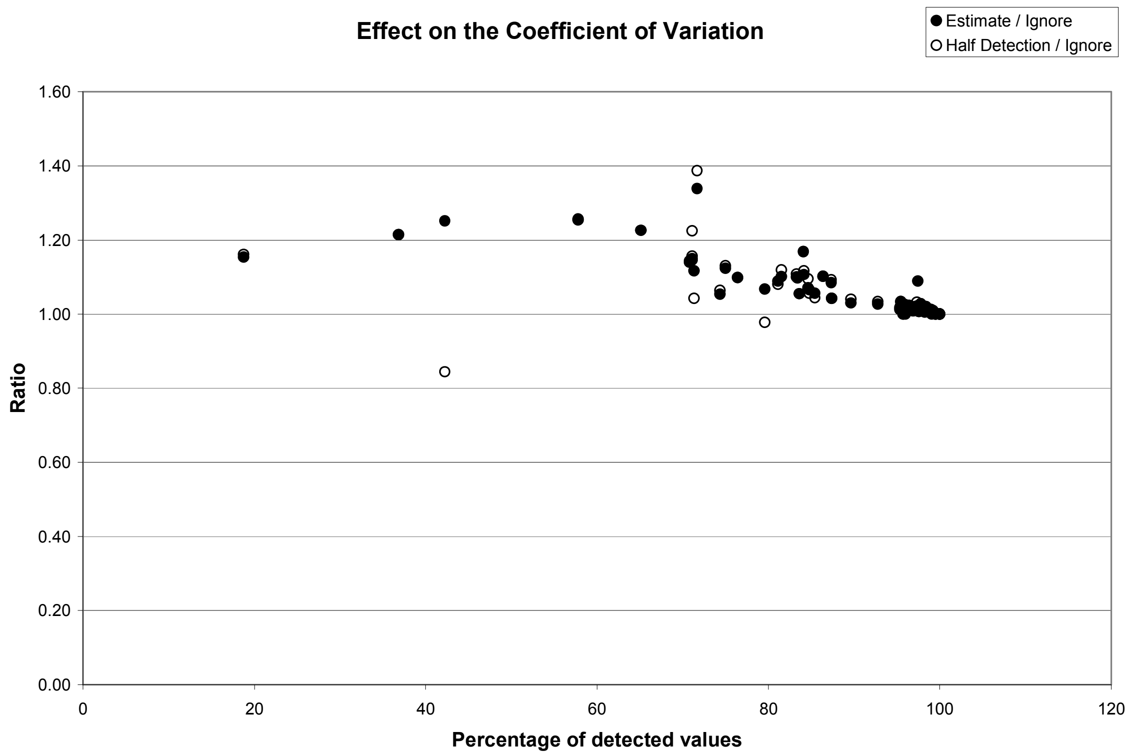


Figure 18. Effect of ignoring the non-detected observations on the coefficient of variation

Total Suspended Solids Analyses at Different Levels of Censoring

To evaluate the effect of the non-detected values in the mean and standard deviation observations at different levels of censoring, one of the constituents with low percentages of non-detected observation (TSS) was trimmed in the lower tail until reduced to 50% of the original distribution. All TSS observations were used during this analysis.

The results are similar to those observed during the analysis of the censoring observations within multiple constituents and land uses. Real mean, median and standard deviation are smaller than the calculated values when censored observations are ignored (Figure 19). The true coefficients of variation are larger than those calculated when the level of trimming is increased.

Table 14. Descriptive Statistics for TSS Truncated at Different Levels

			RATIO			
Total number of samples	% of original samples	Minimum concentration in set (mg/L)	Average	Median	Standard Deviation	Coefficient of Variation
2025	100.00	3	1.00	1.00	1.00	1.00
2015	99.51	3	1.00	1.00	1.00	1.00
1995	98.52	4	0.99	0.98	0.99	1.01
1974	97.48	5	0.98	0.96	0.99	1.02
1954	96.49	6	0.97	0.95	0.99	1.02
1934	95.51	7	0.96	0.95	0.98	1.03
1914	94.52	8	0.95	0.93	0.98	1.03
1873	92.49	10	0.93	0.90	0.97	1.05
1833	90.52	11	0.91	0.87	0.96	1.06
1792	88.49	13	0.89	0.84	0.96	1.07
1752	86.52	15	0.87	0.81	0.95	1.08
1711	84.49	17	0.86	0.78	0.94	1.10
1671	82.52	18	0.84	0.74	0.93	1.11
1630	80.49	20	0.82	0.73	0.92	1.12
1589	78.47	22	0.80	0.71	0.91	1.14
1545	76.30	24	0.78	0.68	0.90	1.15
1496	73.88	26	0.76	0.65	0.89	1.17
1468	72.49	27	0.75	0.64	0.89	1.18
1428	70.52	29	0.73	0.63	0.88	1.20
1387	68.49	31	0.72	0.60	0.87	1.21
1347	66.52	33	0.70	0.58	0.86	1.23
1306	64.49	35	0.68	0.57	0.85	1.24
1266	62.52	37	0.67	0.55	0.84	1.26
1225	60.49	40	0.65	0.53	0.83	1.28
1185	58.52	42	0.63	0.52	0.82	1.29
1144	56.49	44	0.62	0.50	0.81	1.31
1104	54.52	47	0.60	0.47	0.80	1.33
1063	52.49	50	0.58	0.46	0.79	1.35
1023	50.52	52	0.57	0.44	0.78	1.37

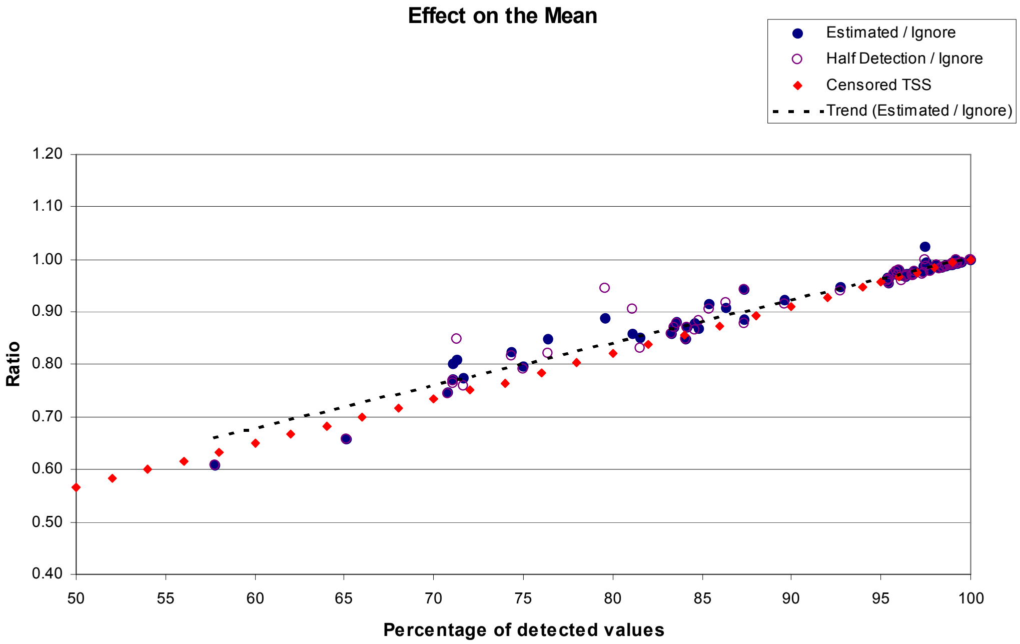


Figure 19. Effect on the mean when TSS observations are truncated

The effect on the mean indicates that when only about 5% of the data is censored or is trimmed, the ratios “replace/ignore,” “estimated/ignore,” or “trimmed/total” observations produced the same results in the mean of the distribution. When the percentage of non-detected observations is increased, the ratios “estimate/ignore” and “half detection/ignore” are higher than the ratio “trimmed/complete” in the TSS distribution. This means that trimming the data set has a larger effect than when the observations are censored. This is explained because for the trimmed/complete ratios, all the censored observations were at one value. In the other case, several detection limits were used during the analysis.

In the previous discussion, it was observed that censored levels less than 30% can be used for predicting simple statistics describing the distribution. The previous figure indicates that levels of censoring close to 45% followed the trend indicated by the ratio “trimmed/complete.” This indicates that even if the regression analysis was recommended for levels of non-detected values smaller than 30%, they can be used for levels of censoring up to 45%.

The effects on the medians are stronger than on the means. When the level of censored observations is close to 30%, the ratio “trimmed/complete” is close to 0.6, compared with 0.75 in the case of the mean (Figure 20). Levels of censoring around 5% do not show the straight-line pattern that was observed with the mean. The trend for censoring levels between 5 and 45% is similar for the “estimated/ignore” ratio; however the dispersion around the trend line is higher.

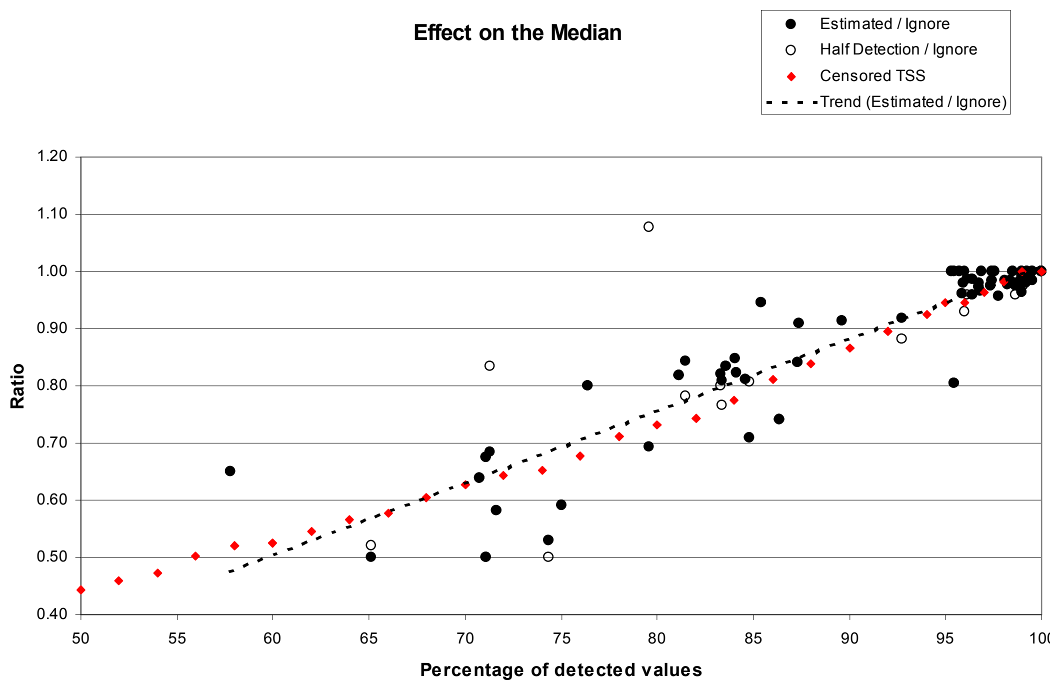


Figure 20. Effect on the median when the TSS dataset is truncated

The effect on the standard deviation of the trimming the TSS is similar to the effect in the mean (Figure 21). When the level of censoring is close to 30%, the ratio "trimmed/complete" is close to 0.85. The dispersion around the trend line is lower than in the median case. When the percentage of non-detected values is lower than 5%, the ratios "estimated/ignore," "half detection/ignore," and "trimmed/complete" are almost the same. For levels of censored observations larger than 15%, the differences among the ratios increase.

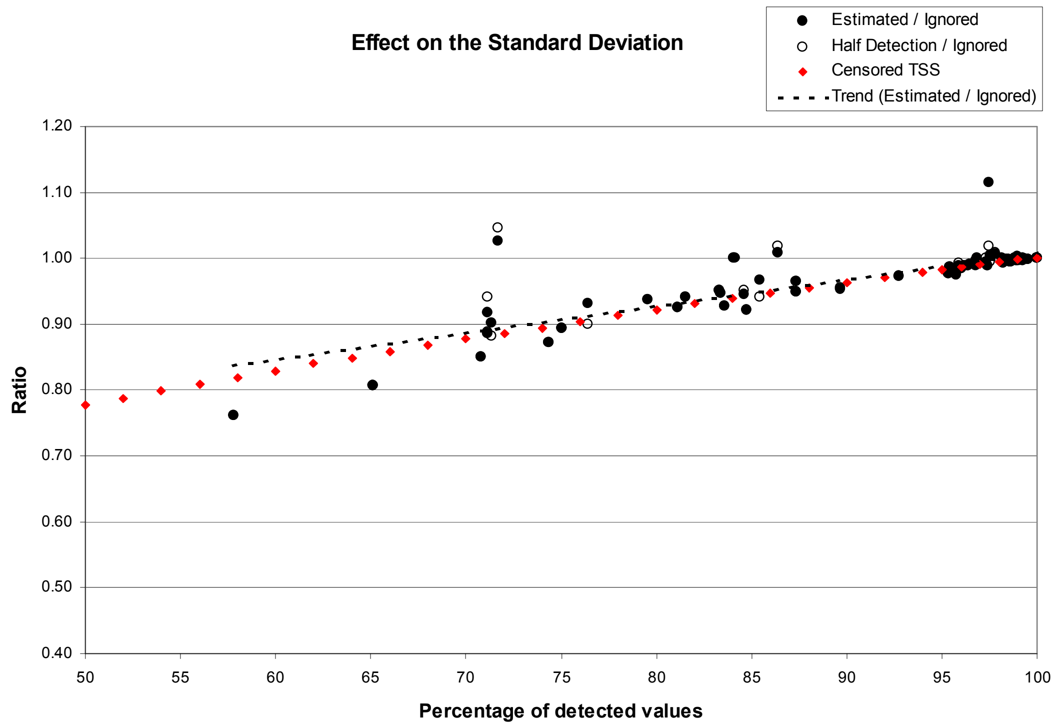


Figure 21. Effect on the standard deviation when the TSS dataset is truncated

The ratio of the effects on the calculated coefficients of variation has a different slope than the previous statistics. As in the mean case when the level of censoring is smaller than 5%, a linear trend between the percentage of detected and the ratio was observed (Figure 22). When the percentage of censored observations is larger than 15%, the differences among the three ratios increase.

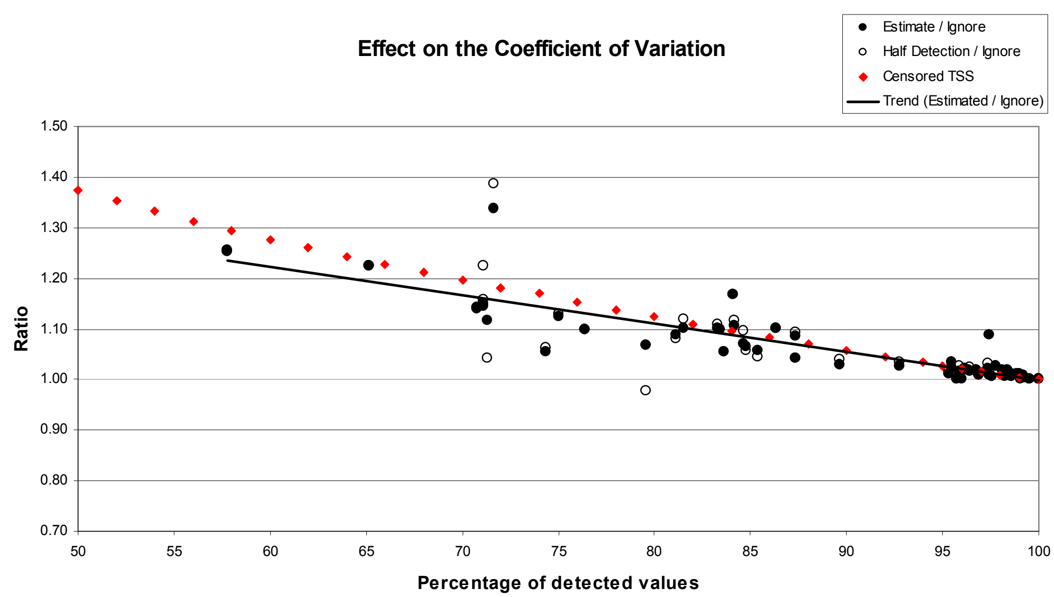


Figure 22. Effect on the coefficient of variation when the TSS dataset is truncated

Summary

The level of censoring observations in a dataset affects the calculated mean, median, standard deviation and coefficient of variation values. As the level of non-detected observations increase, the mean, median and standard deviation are larger than if the censored observations are detected. The opposite behavior is expected for the coefficient of variation. Different laboratories report different detection limits for the same constituents. In many cases, the detection limits are calculated by each laboratory based on their measured repeatability (precision) for a specific laboratory test. Using methods with low precision increases the percentage of non-detected values and the uncertainty of the real mean and standard deviation values.

Open space has the largest number of non-detected observations among land uses. The largest percentages of detected observations were observed in freeways and industrial land uses.

Estimating or replacing by half of the detection limit for levels of censoring smaller than 5% does not have a significant effect on the mean, standard deviation and coefficient of variation values.

Substituting the censored observations by half of the detection limit produces smaller values than when using Cohen's maximum likelihood method. Replacing the censored observations by half of the detection limit is not recommended for levels of censoring larger than 15%.

The censored observations in the database were replaced using estimated values using Cohen's maximum likelihood method for each site before the statistical tests. Because this method uses the detected observations to estimate the non-detected values, it is not very accurate, and therefore not recommended, when the percentage of censored observations is larger than 40%. Table 14 shows those constituents having percentages of non-detected observations smaller than 40% for the three main land uses.

All the methods used in this chapter are approximations to calculate the EMC when censored observations are present. These problems would not exist if appropriate analytical methods were used to analyze the samples. It is very important to select analytical methods capable of detecting the desired range of concentrations in the samples in order to reduce the numbers of censored observations to acceptable levels. Table 3XX summarizes the recommended minimum detection limits for various stormwater constituents to obtain manageable non-detection frequencies (<5%). Some of the open space stormwater measurements (oil and grease and lead, for example), would likely have greater than 5% non-detects, even with the detection limits shown. The detection limits for filtered heavy metals would be substantially less than shown on this table.

Table 3XX. Suggested Analytical Detection Limits for Stormwater Monitoring Programs to Obtain <5% Non-detects

	Residential, commercial, industrial, freeway	Open Space
Conductivity	20 μS/cm	20 μS/cm
Hardness	10 mg/L	10 mg/L
Oil and grease	0.5 mg/L	0.5 mg/L
TDS	10 mg/L	10 mg/L
TSS	5 mg/L	1 mg/L
BOD_5	2 mg/L	1 mg/L
COD	10 mg/L	5 mg/L
Ammonia	0.05 mg/L	0.01 mg/L
NO_2+NO_3	0.1 mg/L	0.05 mg/L
TKN	0.2 mg/L	0.2 mg/L
Dissolved P	0.02 mg/L	0.01 mg/L
Total P	0.05 mg/L	0.02 mg/L
Total Cu	2 μg/L	2 μg/L
Total Pb	3 μg/L (residential 1 μg/L)	1 μg/L
Total Ni	2 μg/L	1 μg/L
Total Zn	20 μg/L (residential 10 μg/L)	5 μg/L

Chapter 4: Stormwater Quality Descriptions Using the Three Parameter Lognormal Distribution

Introduction

Knowing the statistical distribution of observed stormwater data is a critical step in data analysis. The selection of the correct statistical analyses tools is dependent on the data distribution, and many QA/QC operations depend on examining the distribution behavior. However, much data is needed for accurate determinations of the statistical distributions of the data, especially when examining unusual behavior. The comparison of probability distributions between different data subsets is also a fundamental method to identify important factors affecting data observations. Statistical analyses basically are intended to explain data variability by identifying significantly different subsets of the data. The remaining variability that can not be explained must be described. In all cases, accurate descriptions of the data probability distributions are needed. This chapter explores these distributions for the NSQD data.

The Nationwide Urban Runoff Program (NURP) evaluated the characteristics of stormwater discharges at 81 outfalls in 28 communities throughout the U.S. (EPA 1983). One of the conclusions was that most of the stormwater constituent concentration probability plots could be described using lognormal distributions. More recently, Van Buren (1997) also found that stormwater concentrations were best described using a lognormal distribution for almost all constituents, with the exception of some dissolved constituents that were better described with a normal distribution. Beherra (2000) also found that some stormwater constituent concentrations were better described using a lognormal distribution, while others were better described with gamma or exponential distributions. The constituents that were best described with a gamma distribution included total solids, total Kjeldahl nitrogen (TKN), total phosphorous, chemical oxygen demand (COD), barium and copper. The constituents that were best described with an exponential distribution included suspended solids, nitrates and aluminum. In both of these recent studies, fewer than 50 samples (collected at the same site) were available for evaluation.

During the research reported in this chapter, statistical tests were used to evaluate the log-normality of a selection of the constituents in the NSQD database. Statistical descriptions were obtained of each set of data including box and whisker and probability plots for each land use category and for the pooled dataset. It was found in almost all cases that the log-transformed data followed a straight line between the 5th and 95th percentile, as illustrated in Figure 26 for total dissolved solids (TDS) in residential areas.

For many statistical tests focusing on the central tendency (such as for determining the average concentration that is used for mass balance calculations), this may be a suitable fit. As an example, WinSLAMM, the Source Loading and Management Model (Pitt and Voorhees 1995), uses a Monte Carlo component to describe the likely variability of stormwater source flow pollutant concentrations using either lognormal or normal probability distributions for each constituent. However, if the extreme values are of importance (such as when dealing with the influence of many non-detectable values on the predicted concentrations, or determining the frequency of observations exceeding a numerical standard), a better description of the extreme values may be important.

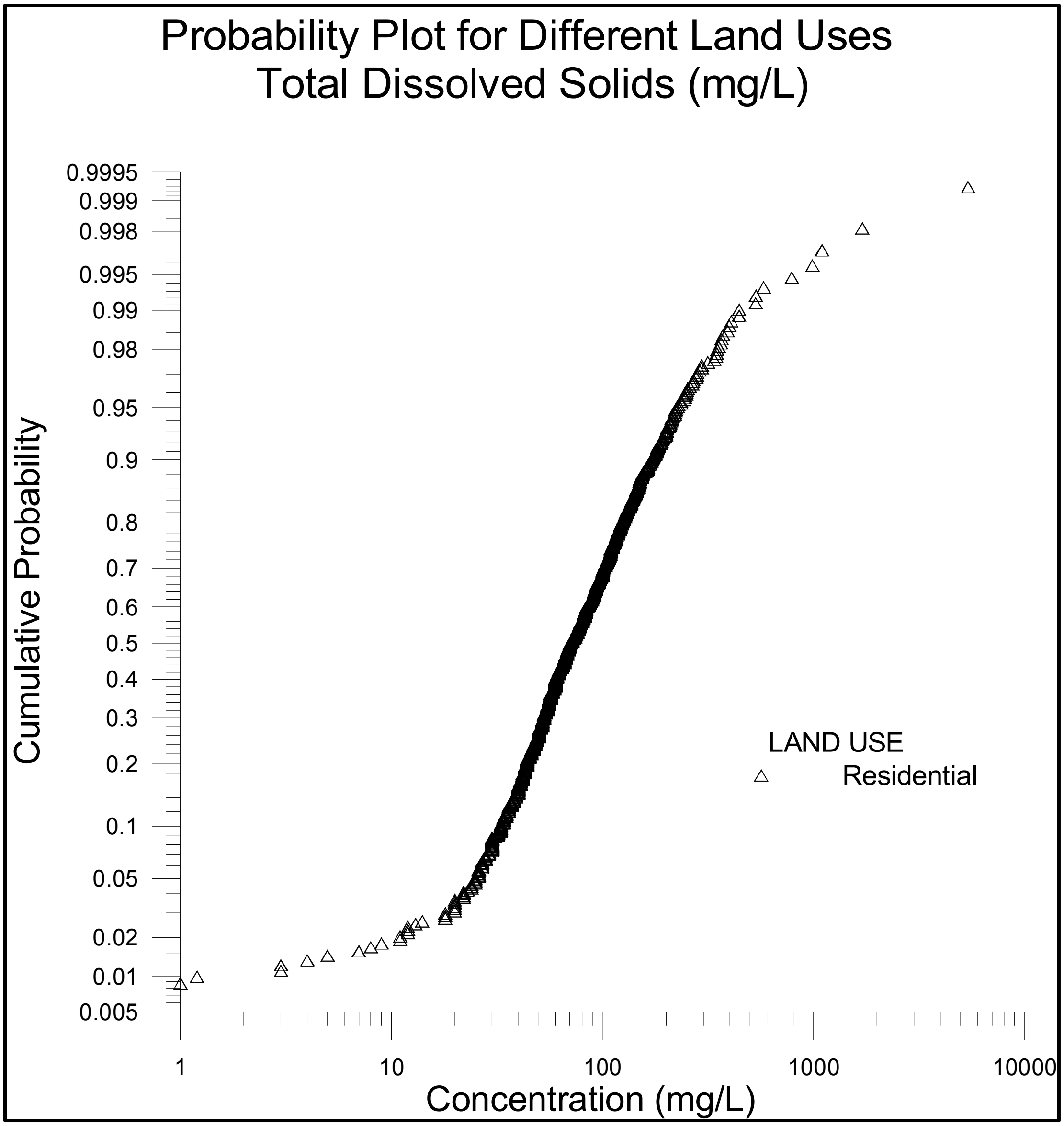


Figure 26. Log-probability plot of total dissolved solids in residential land use

The NSQD underwent an extensive data evaluation process, including multiple comparisons of the all data values in the database to original documents. In some cases, data was available from the local agency in electronic form. These spreadsheets were reformatted to be consistent to the NSQD format. However, it was found that all of the submitted electronic data needed to be verified against original data sheets and reports. When reviewing the NSQD, it was assumed that some of the events in the upper and lower tails of the distributions were caused by errors, most likely due to faulty transcription of the data (such as mislabeling the units for heavy metals or nutrients as mg/L instead of µg/L, for example). Unusual values were verified with the original reports and datasets. While some values (less than 5% of the complete dataset) were found to be in error and were corrected, most of the suspected values were found to be correct stormwater observations. Besides the targeted extreme values, many constituents were also examined in relationship to other related constituents (COD vs. BOD; total metal concentrations vs. dissolved metal concentrations; TKN vs. NH3; TDS vs. specific conductivity; SS vs. turbidity; etc) and unusual behavior was further checked and corrected, as necessary. In some cases, unusual values could not be verified and were therefore eliminated from the dataset, although this was very unusual. After the extensive QA/QC activities and corrections were made to the NSQD, the next step was to conduct a sensitivity analysis to determine the effects of the remaining unusual high and low values on the probability distribution parameters.

The Effects of Unusual High and Low Values on Probability Distribution Parameters

For this evaluation, 10,000 sets of 200 samples each were randomly generated following a lognormal distribution (1, 1), but having differing amounts of extreme values in each data set. For each set, the mean, variance and coefficient of variation were calculated. Two main factors were analyzed using these data: the extreme value factor and percentage of extreme values in each sample. The following percentages of extreme values were selected for evaluation: 0.5, 1, 5, 10, 25 and 50%. For each percentage of extreme values, the following factors were analyzed: 0.001, 0.01, 0.1, 10, 100, 1.000, 10,000, 100,000 and 1,000,000. For example (5%, 100) indicates that in each set, 5% of the data were increased by a factor of 100. The coefficient of variation was then calculated for each set of data. The medians of the coefficients of variation for the 10,000 runs are shown in Figure 27 for each level of extreme values.

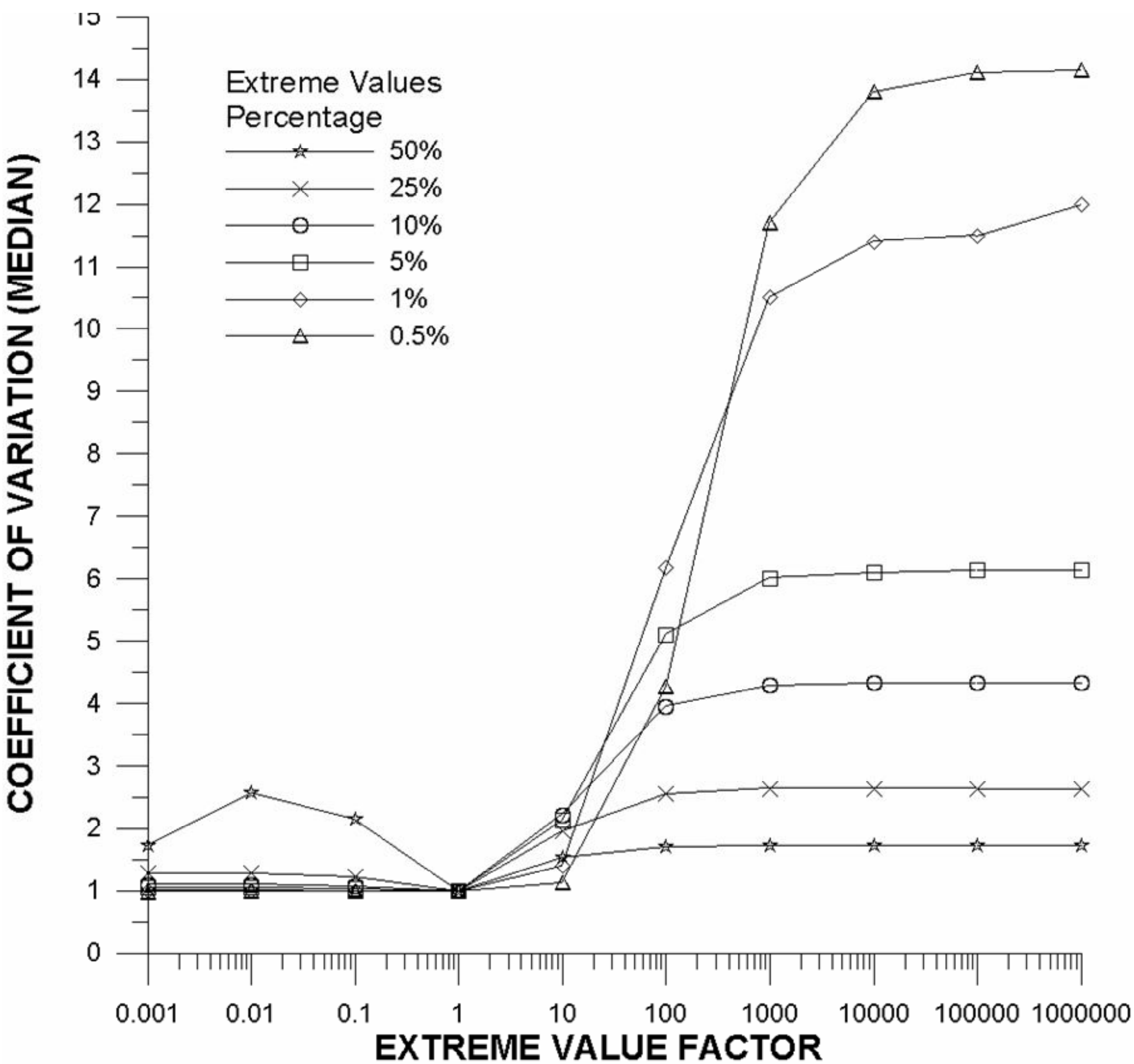


Figure 27. Effect of unusual values on the coefficient of variation (based on LN(1,1))

For a lognormal distribution (1, 1) the coefficient of variation is equal to one. Figure 27 shows how this original value is changed for different amounts of extreme values in the data sets, and for different factors in these extreme values. The horizontal axis represents the factor used in the extreme values. As an example, many of the incorrect extreme values observed in the NSQD for heavy metals were because the units were originally incorrectly reported as mg/L in the submitted information, while the correct units were actually µg/L. This would be an extreme value factor of 1,000. Extreme value factors of 10 were also fairly common and were associated with simple misplacements of decimal points in the data.

Figure 27 also shows that for small error factors (0.1, 0.01 and 0.001) there is not a large effect in the coefficient of variation for percentages smaller than 10%. For larger percentages, the effect in the coefficient of variation is important. When 50% of the data are affected by an error factor of 0.01, the coefficient of variation was increased by almost three times.

High extreme value factors can have a much more important effect on the coefficient of variation. When 10% of the data were increased by a factor of 10, the coefficient of variation was increased almost three times. Notice that affecting 10% of the data by a factor of ten have almost the same effect as affecting 50% of the data by a factor of a hundredth. This effect is reduced when the percentage of elevated values in the dataset is smaller than 10%.

For factors larger than a hundred, the effect on the coefficient of variation is much greater. Very low percentages of elevated values can increase the coefficient of variation by up to 15 times. For example, when only 0.5% of the sample is affected by a factor of a thousand, the coefficient of variation increases almost 12 times more than the correct value. As noted earlier this is important because it is not unusual to find reported values affected by a factor larger than a hundred (See Figure 26). Some of these values can be due to incorrect reporting units, but in many cases they were considered as valid observations because they were supported by similarly high values of other closely related constituents. For factors greater than 10,000 the multiplying value of the coefficient of variation remains stable at the maximum value obtained.

The above analyses indicate that in lognormal distributions, the presence of just a few unusual elevated values is important and can dramatically affect the reported coefficient of variation for the distribution of concentration. This observation is critical in the relatively common case where one or a very few observations are affected by a factor larger than a hundred. In the other extreme, factors smaller than one do not have a large impact on the reported coefficient of variation, except when the percentage of errors is greater than 50%.

The effect of extreme values on the mean and standard deviation was also analyzed. Figure 28 shows the effect of the extreme values on calculated standard deviation. For large extreme value factors (larger than one) the standard deviation increases as the percentage of extreme values increases.

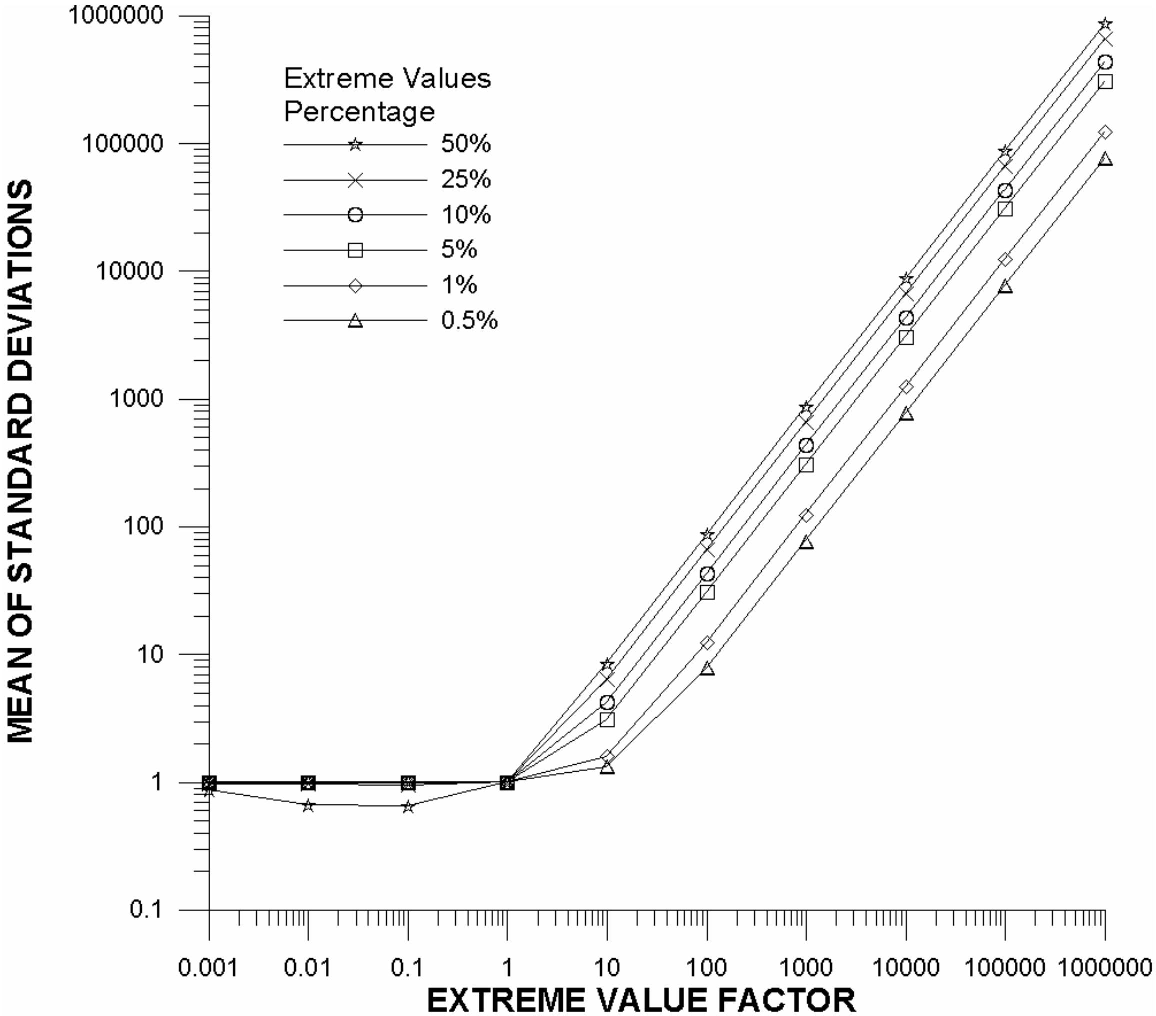


Figure 28. Effect of unusual values on the standard deviation (based on LN(1,1))

Percentages smaller than 25% do not have an important effect on the standard deviation for small extreme value factors. For a specific extreme value factor, changing the extreme value percentages from 0.5% to 50% increases the standard deviation close to 10 times.

The effect of the presence of extreme values on the distribution mean is shown in Figure 29. For small extreme value factors, the mean is reduced almost 80% when the extreme value percentage is close to 50%. This is expected because in a lognormal distribution (1, 1) most of the values are located in the lower tail of the distribution. For extreme value occurrences less than 25%, the mean value is reduced by less than 20%.

Large extreme value factors have much larger effects on the distribution means. As the extreme value percentage increases, the calculated means also increase. If 0.5% of the values are affected by a factor of a hundred, the mean value is doubled. If 50% of the values are affected by the same factor, the mean values are increased by almost 50 times. For factors larger than a thousand, increasing the percentages of extreme values from 0.5% to 50% increases the mean values by up to two orders of magnitude.

These evaluations are important because it points out that for a lognormal distribution, the effects of few elevated values in the upper tail have a much greater effect on common statistics than unusual values in the lower tail. Many stormwater researchers have focused on the lower tail, especially when determining how to handle the detection limits and unreported data. Stormwater constituents usually have unusual values in both tails of the probability distribution. It is common to delete elevated values from the observations assuming they are expendable "outliers". This practice is not recommended unless there is sufficient evidence that the observed values are a mistake. Actual elevated values can have a large effect on the calculated distribution parameters. If these are arbitrarily removed, the data analyses will likely be flawed.

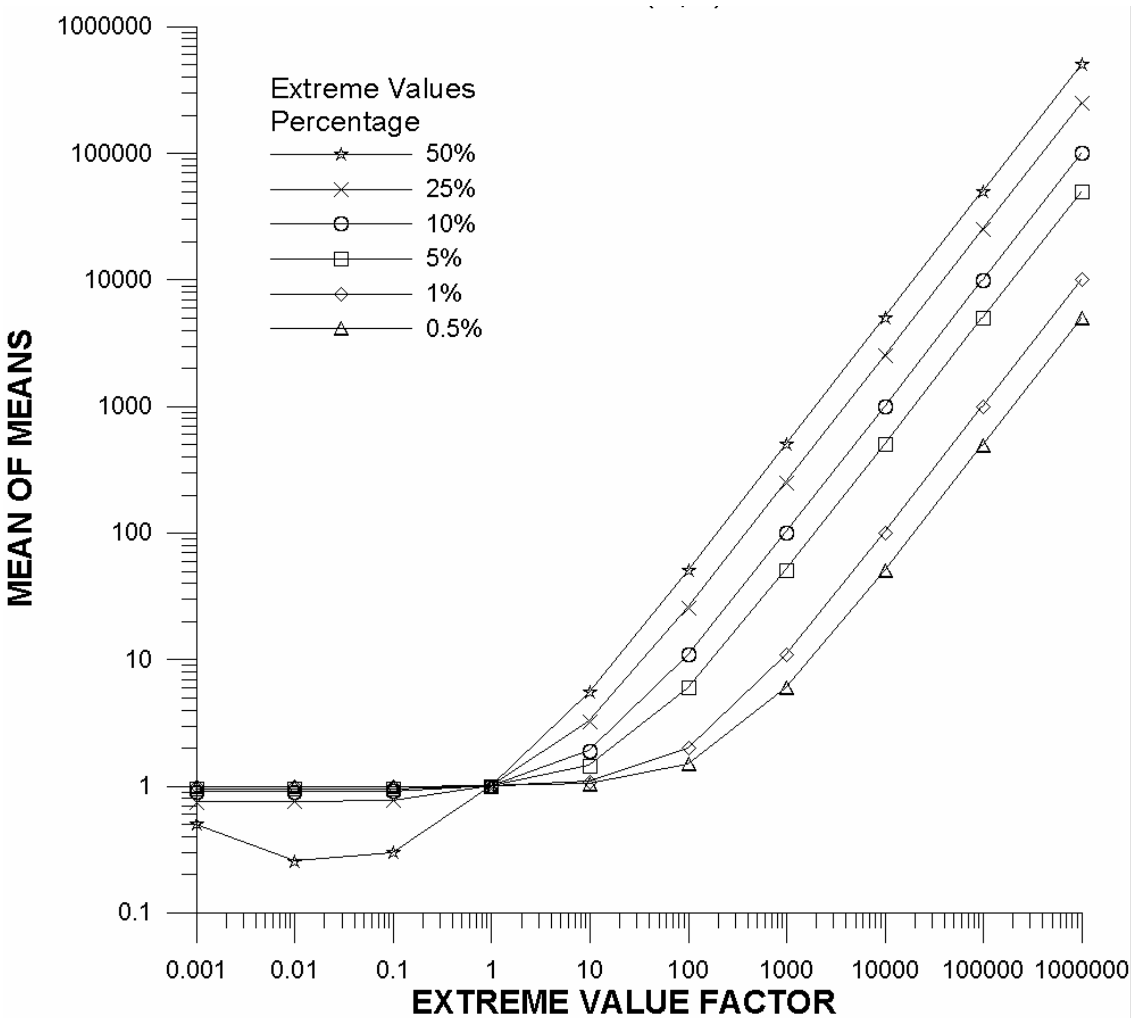


Figure 29. Effect of unusual values on the mean (based on LN(1,1))

Analysis of Lognormality of Stormwater Constituents Parameters

The goodness of fit of twenty nine stormwater constituent probability distributions was evaluated using the Kolmogorov-Smirnov test. Figure 30 shows how the test accepts or rejects the null hypothesis that the empirical and the estimated distributions are the same. If the null hypothesis is valid, then the constituent can be adequately represented by the lognormal distribution. The observations are sorted and a probability is assigned by its rank. The distribution generated by this ranking is known as the empirical distribution. The estimated distribution function is also compared on the same plot. The estimated distribution function is calculated with the mean and standard deviation of the original data. If the distance between the empirical and the estimated distributions is higher than a critical value d_α or D_{max}, the hypothesis of lognormality is rejected. Notice in Figure 30 that the horizontal axis has a logarithmic scale.

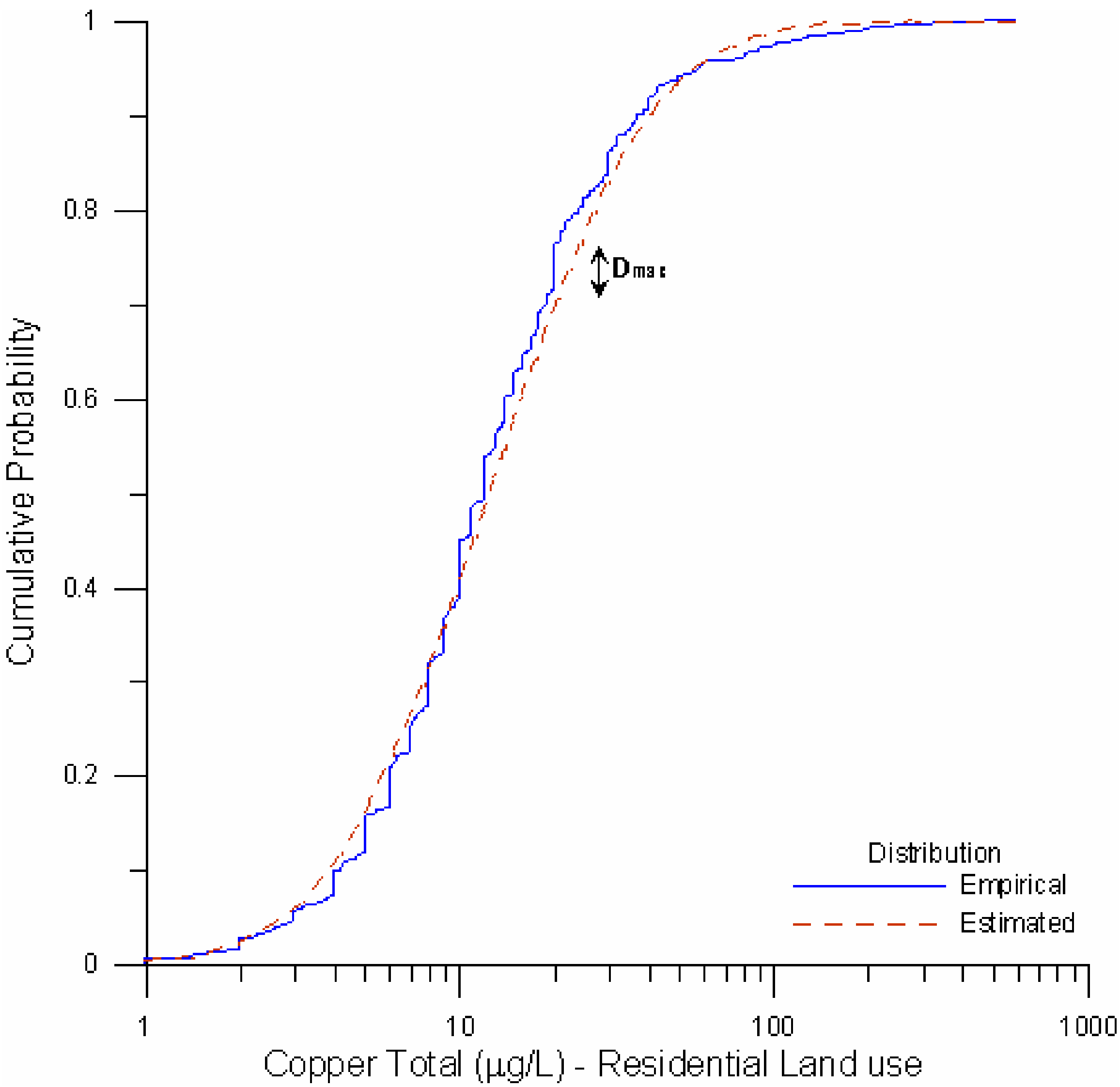


Figure 30. Cumulative and empirical probability distributions of total copper for residential land use data (Goodness of fit test, Kolmogorov-Smirnov)

There are many options to assign probability to a data observation based on ranks. Most methods assign the probability as a percentage of the total range. The probability of the observation is calculated as its rank divided by the number of observations. Kottegoda (1998) suggested that for extreme event analysis, the plotting position can be calculated as:

(4.1)
$$p = \frac{i - 0.5}{n}$$

Where p is the cumulative probability of the observation, i is the rank of the observation and n is the total number of observations. This plotting position was used for the analyses during this research because it does not set the probability of the largest observation as one.

In the Kolmogorov-Smirnov test, the null hypothesis is that the observed data follow a lognormal distribution. If the sample size is small, and the distance between the empirical and the observed distributions is smaller than the critical value D_{max}, the test is interpreted as "there is not enough evidence to reject the hypothesis that the distribution is lognormal." In most cases, the NSQD contains enough samples to be able to accept or reject the null hypothesis with acceptable levels of confidence and power.

The NSQD contains many factors for each sampled event that likely affect the observed concentrations. These include such factors as seasons, geographical zones, rain intensities, etc. These factors may affect the shape of the probability distribution. As more data become available, the critical value D_{max} is reduced in the test. There will always be a specific number of samples that will lead to rejection of the null hypothesis because the maximum distance between the empirical and estimated probability distributions became larger than the critical value D_{max}. The only way to evaluate the required number of samples in each category is using the power of the test. Power is the probability that the test statistic will lead to a rejection of the null hypothesis when it is false (Gibbons and Chakraborti 2003). Masey (1950) states that the power of the Kolmogorov-Smirnov test can be written as:

(4.2)
$$power = 1 - \Pr\left(\frac{-d_\alpha \pm \Delta\sqrt{n}}{\sqrt{F_1(x_0)(1-F_1(x_0))}} < \frac{\{S_n(x_0) - F_1(x_0)\}\sqrt{n}}{\sqrt{F_1(x_0)(1-F_1(x_0))}} < \frac{d_\alpha \pm \Delta\sqrt{n}}{\sqrt{F_1(x_0)(1-F_1(x_0))}} \right)$$

where:

d_α	=	Dmax: critical distance at the level of significance α (confidence of the test),
S_n	=	Cumulative empirical probability distribution,
F_1	=	Cumulative alternative probability distribution,
Δ	=	Maximum absolute difference between the cumulative estimated probability distribution and the alternative cumulative probability distribution.

Massey (1951) also found that for large sample sizes, the power can be never be smaller than

(4.3)
$$power > 1 - \int_{2(\Delta\sqrt{n} - d_\alpha\sqrt{n})}^{2(\Delta\sqrt{n} + d_\alpha\sqrt{n})} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

This reduced expression can be used to calculate the number of samples required to reject the null hypothesis with a desired power. Figure 31 shows the power of the D test for 1%, 5%, and 10% levels of confidence of the test for samples size larger than 35 (Massey 1951). For example, assume that the maximum distance between the alternative cumulative and the estimated cumulative probability distributions is 0.2, and we want an 80% power (0.8) against the alternative at a 5% level of confidence. To calculate the number of required samples, we read that $\Delta(N)^{0.5}$ is 1.8 for a power of 0.8 and 5% level of confidence. Solving for $N = (1.8/0.2)^2 = 81$ samples. If we want to calculate the number of samples when the difference between the alternative cumulative and the estimated cumulative probability function is 0.05, with the same power and level of confidence, then 1,296 samples would be required. When the lines are very close together, it is obviously very difficult to statistically show that they are different, and many samples are needed.

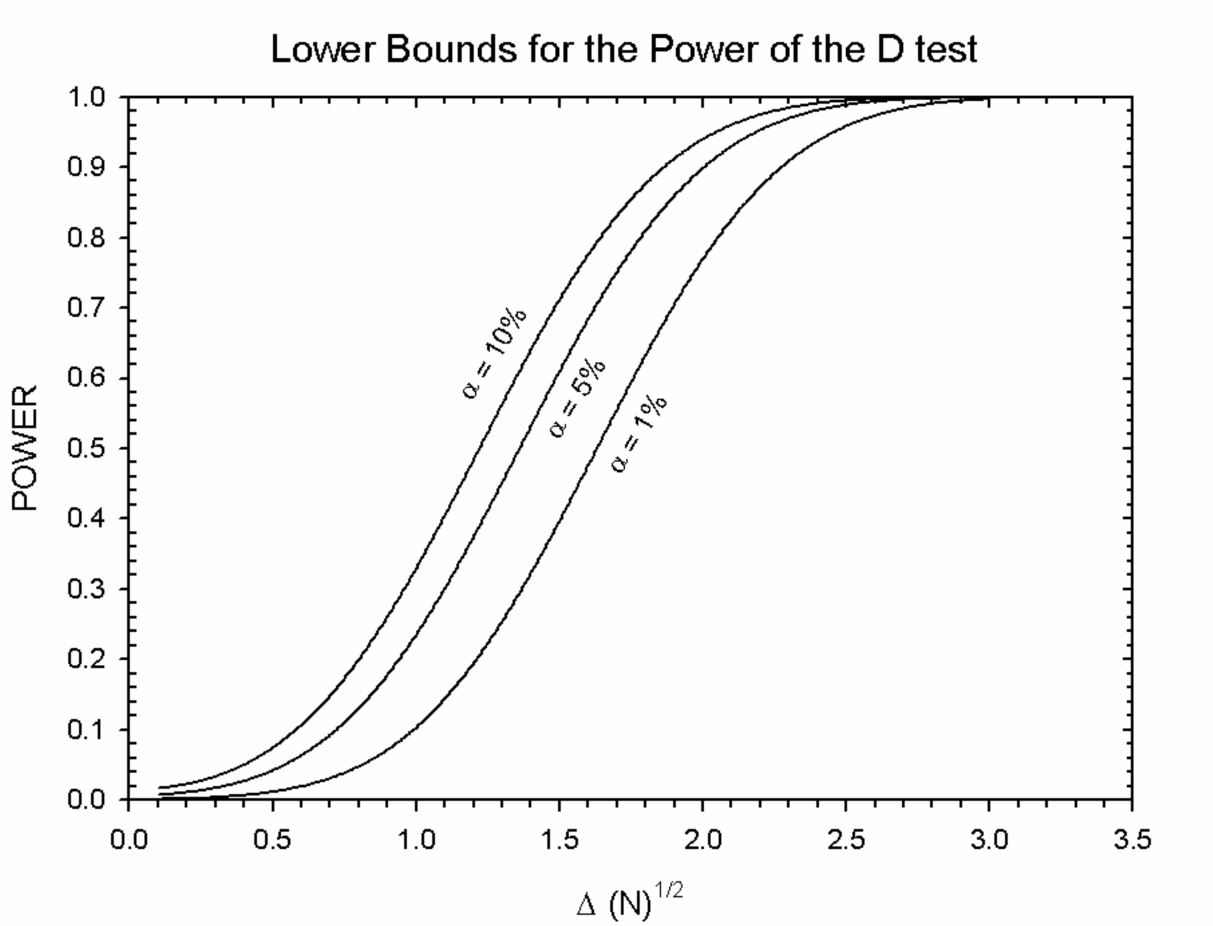


Figure 31. Lower bounds for the power of the D test for α = 1%, 5% and 10% (N>35)

The Kolmogorov-Smirnov test was used to indicate if the cumulative empirical probability distribution of the NSQD residential stormwater constituents can be adequately represented with a lognormal distribution. Table 19 shows the resulting power of the test for D=0.05 and D=0.1, when applied to selected constituents that had high levels of detection in residential land uses.

Table 19. Power of the Test When Applied to Selected Constituents in Residential Land Uses

Constituent	N	Percentage Detected	$\Delta N^{0.5}$ (α=0.05)	Power (D=0.05, β=5%)	$\Delta N^{0.5}$ (α=0.1)	Power (D=0.1, β =10%)
TDS (mg/L)	861	99.2	1.46	0.60	2.92	1
TSS (mg/L)	991	98.6	1.56	0.65	3.12	1
BOD (mg/L)	941	97.6	1.52	0.65	3.04	1
COD (mg/L)	796	98.9	1.40	0.55	2.80	1
NO2+NO3 (mg/L)	927	97.4	1.50	0.60	3.00	1
TKN (mg/L)	957	96.8	1.52	0.65	3.04	1
TP (mg/L)	963	96.9	1.53	0.65	3.06	1
Total Copper (μg/L)	799	83.6	1.29	0.50	2.58	1
Total Lead (μg/L)	788	71.3	1.19	0.40	2.38	1
Total Zinc (μg/L)	810	96.4	1.40	0.55	2.80	1

Table 19 shows that the number of collected samples is sufficient to detect if the empirical distribution is located inside an interval of width 0.1 above and below the estimated cumulative probability distribution. If the interval is reduced to 0.05, the power varies between 40 and 65%. To estimate the interval width, 10 cumulative distributions of 1,000 random data points, having a lognormal (1, 1) distribution, were compared with the estimated cumulative distribution for normal, gamma and exponential distributions. The maximum distance between the cumulative lognormal and the cumulative normal distributions was 0.25. The maximum distance with cumulative gamma (the same for exponential in this case) was 0.28. An interval width of 0.1 was considered appropriate for the analysis.

Another factor that must be considered is the importance of relatively small errors in the selected distribution and the problems of a false negative determination. It may not be practical to collect as many data observations as needed when the distributions are close (such as when the width interval is 0.05). Therefore, it is important to understand what types of further statistical and analysis problems may be caused by having fewer samples than optimal. For example, Figure 32 (total phosphorus in residential area) shows that most of the data fall along the straight line (indicating a lognormal fit), with fewer than 10 observations (out of 933) in the tails being outside of the obvious path of the line.

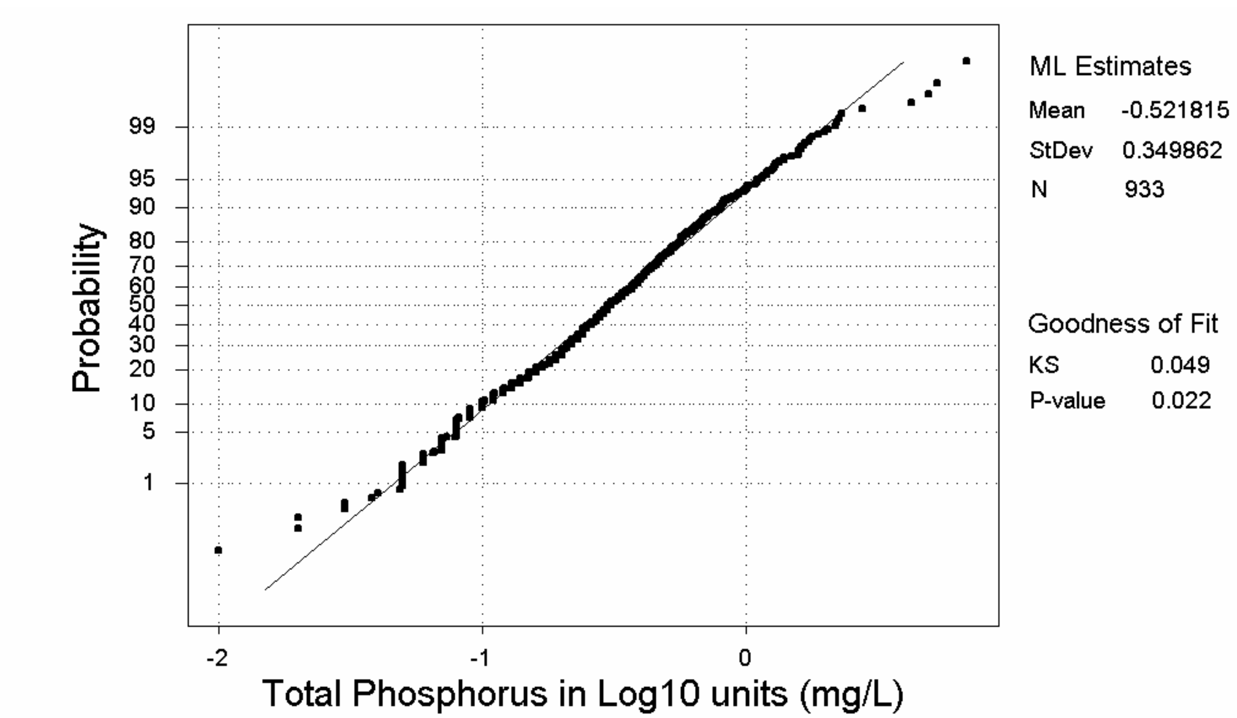


Figure 32. Normality test for total phosphorus in residential land uses using the NSQD

The calculated p-value for the Kolmogorov-Smirnov test is 0.022, indicating that the null hypothesis could be rejected and that there is not enough evidence that the empirical distribution is adequately represented by a lognormal distribution. Notice that errors in the tails are smaller than 0.049. However, the tails are not responsible for the rejection of the null hypothesis (see Figure 33).

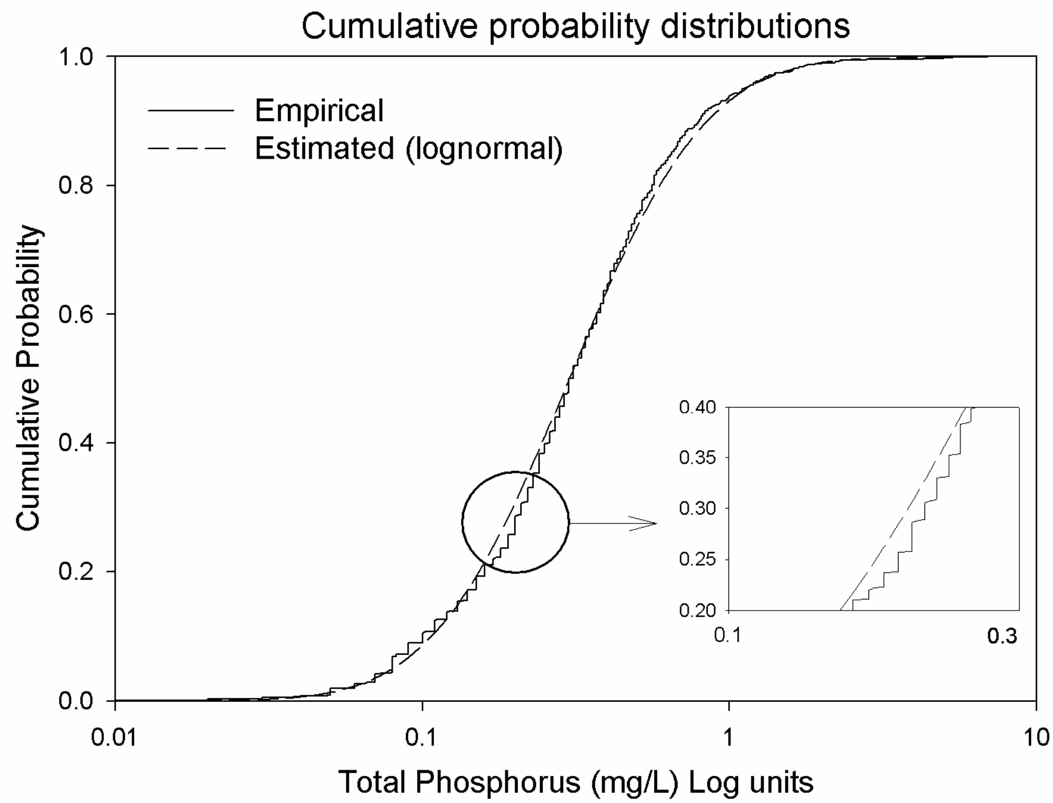


Figure 33. D_{max} was located in the middle of the distribution

In this case, D_{max} is located close to a total phosphorus concentration of 0.2 mg/L (-0.7 in log scale). As in this case, the hypothesized distributions are usually rejected because of the departures in the middle of the distribution, not in the tails. However, as previously pointed out, a small number of observations in the upper tail can change the shape of the estimated cumulative probability distribution by affecting the mean and standard deviation of the data. The methods used previously by Van Buren and Beherra evaluated the probability distributions only using two parameters, the median and the standard deviation. They suggested the gamma and exponential distributions as alternatives to the lognormal for some stormwater constituents. Table 20 shows the comparison for the goodness of fit using the 2-parameter gamma, exponential and lognormal distributions using the method of moments.

Table 20. Comparison of Goodness of Fit for Gamma, Exponential and Lognormal Distributions Using the NSQD v.1.1

CONSTITUENT	PDF	RESIDENTIAL			COMMERCIAL			INDUSTRIAL			OPEN SPACE			FREEWAYS		
		N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value
Conductivity (mS/cm)	Gamma	106 100%	0.381	0	66 100	0.230	0.002	108 100	0.348	0	2 100	-	-	86 100	0.238	0
	Exponential		0.195	0.001		0.237	0.001		0.228	0		-	-		0.232	0
	Lognormal		0.081	0.493		0.100	0.530		0.074	0.619		-	-		0.129	0.113
Hardness (mg/L)	Gamma	250 100%	0.217	0	139 100	0.141	0.008	138 96.4	0.323	0	8 100	0.304	0.458	127 100	0.451	0
	Exponential		0.203	0		0.115	0.067		0.133	0.018		0.369	0.228		0.161	0.003
	Lognormal		0.071	0.166		0.090	0.206		0.080	0.369		0.354	0.268		0.077	0.447
Oil and Grease (mg/L)	Gamma	533 57.8%	0.876	0	308 70.8	0.629	0	327 65.1	0.939	0	19 36.84	0.210	1.080	60 71.7	0.103	0.810
	Exponential		0.514	0		0.304	0		0.697	0		0.265	0.750		0.286	0.002
	Lognormal		0.112	0.001		0.103	0.019		0.098	0.032		0.202	1.127		0.101	0.827
Total Dissolved Solids (mg/L)	Gamma	861 99.3%	0.234	0	399 99.5	0.457	0	413 99.5	0.645	0	45 97.8	0.109	0.698	97 99	0.082	0.553
	Exponential		0.207	0		0.150	0		0.172	0		0.195	0.070		0.171	0.007
	Lognormal		0.050	0.029		0.049	0.303		0.066	0.053		0.120	0.561		0.054	1.136
Total Suspended Solids (mg/L)	Gamma	991 98.6%	0.288	0	458 98.3	0.363	0	428 99.1	0.206	0	44 95.5	0.132	0.464	134 99.3	0.534	0
	Exponential		0.141	0		0.214	0		0.108	0		0.289	0.002		0.168	0.011
	Lognormal		0.032	0.280		0.064	0.053		0.029	0.995		0.113	0.683		0.066	0.627
BOD5 (mg/L)	Gamma	941 97.6%	0.321	0	432 97.5	0.191	0	406 95.3	0.921	0	44 86.4	0.112	0.770	26 84.6	0.272	0.076
	Exponential		0.140	0		0.142	0		0.355	0		0.261	0.011		0.168	0.580
	Lognormal		0.058	0.004		0.054	0.166		0.105	0		0.114	0.746		0.103	1.252
COD (mg/L)	Gamma	796 98.9%	0.129	0	373 98.4	0.137	0	362 98.9	0.216	0	43 76.7	0.373	0	67 98.5	0.163	0.061
	Exponential		0.161	0		0.136	0		0.119	0		0.168	0.312		0.139	0.157
	Lognormal		0.036	0.250		0.038	0.695		0.074	0.040		0.128	0.684		0.107	0.445
Fecal Coliform (Colonies/100 mL)	Gamma	446 88.3%	0.655	0	233 88	0.333	0	297 87.9	-	-	23 91.3	0.179	0.520	49 100	0.239	0.007
	Exponential		0.374	0		0.396	0		0.504	0		0.208	0.324		0.355	0
	Lognormal		0.080	0.013		0.076	0.192		0.051	0.510		0.181	0.503		0.105	0.677
Fecal Streptococcus (Colonies/100 mL)	Gamma	305 89.5%	0.158	0	181 91.7	0.354	0	195 93.8	-	-	22 90.9	0.144	0.869	25 100	0.096	1.262
	Exponential		0.202	0		0.278	0		0.399	0		0.142	0.892		0.164	0.518
	Lognormal		0.077	0.081		0.097	0.091		0.083	0.161		0.181	0.538		0.119	0.990
Ammonia (mg/L)	Gamma	595 81.5%	0.132	0	299 83.3	0.131	0	254 85.8	0.154	0	32 18.7	-	-	79 87.3	0.216	0.003
	Exponential		0.101	0		0.066	0.228		0.071	0.221		-	-		0.105	0.440
	Lognormal		0.044	0.305		0.050	0.589		0.047	0.758		-	-		0.133	0.173
NO2+NO3 (mg/L)	Gamma	927 97.4%	0.197	0	425 98.1	0.147	0	418 96.2	0.080	0.011	44 84.1	0.123	0.654	25 96	0.274	0.055
	Exponential		0.141	0		0.120	0		0.132	0		0.120	0.686		0.177	0.443
	Lognormal		0.070	0		0.040	0.531		0.080	0.011		0.141	0.463		0.139	0.789
TKN (mg/L)	Gamma	957 96.8%	0.203	0	449 97.3	0.127	0	440 95.9	0.195	0	45 71.1	0.169	0.323	125 96.8	0.280	0
	Exponential		0.182	0		0.156	0		0.134	0		0.141	0.556		0.138	0.020
	Lognormal		0.035	0.218		0.042	0.423		0.048	0.292		0.147	0.500		0.074	0.539

* P-values greater than one are used only for comparison. NDet: Number of collected samples and percentage detected

Table 20. Comparison of Goodness of Fit for Gamma, Exponential and Lognormal Distributions Using the NSQD – *Continued*

CONSTITUENT	PDF	RESIDENTIAL			COMMERCIAL			INDUSTRIAL			OPEN SPACE			FREEWAYS		
		N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value	N_{Det}	Dmax	P-value
Dissolved Phosphorus (mg/L)	Gamma	738	0.117	0	323	0.177	0	325	0.200	0	44	0.154	0.127	22	0.449	0
	Exponential	84.1	0.144	0	81.1	0.129	0	87.1	0.135	0	79.5	0.384	0.657	95.5	0.350	0.012
	Lognormal		0.043	0.199		0.075	0.104		0.124	0.682		0.124	0.682		0.170	0.593
Total Phosphorus (mg/L)	Gamma	963	0.184	0	446	0.179	0	434	0.227	0	46	0.666	0	128	0.456	0
	Exponential	96.9	0.129	0	95.7	0.114	0	96.3	0.107	0	84.8	0.320	0.001	99.2	0.187	0
	Lognormal		0.049	0.022		0.038	0.582		0.049	0.273		0.116	0.696		0.085	0.325
Total Antimony (µg/L)	Gamma	288	0.268	0.636	142	-	-	164	0.282	0.045	17	-	-	14	0.423	0.164
	Exponential	2.8	0.417	0.213	2.1	-	-	14.6	0.173	0.473	0	-	-	50	0.465	0.096
	Lognormal		0.233	0.841		-	-		0.096	1.279		-	-		0.419	0.171
Total Arsenic (µg/L)	Gamma	426	0.531	0	213	0.643	0	267	0.291	0	19	0.271	0.828	61	0.125	0.694
	Exponential	42	0.224	0	32.9	0.249	0	54.3	0.141	0.006	31.6	0.462	0.154	55.7	0.266	0.016
	Lognormal		0.154	0		0.164	0.046		0.129	0.016		0.273	0.819		0.149	0.441
Total Beryllium (µg/L)	Gamma	301	0.464	0	163	0.305	0.542	209	0.390	0.002	19	-	-	12	-	-
	Exponential	7.3	0.471	0	4.29	0.530	0.039	10.5	0.539	0	0	-	-	16.7	-	-
	Lognormal		0.200	0.342		0.205	1.108		0.163	0.620		-	-		-	-
Total Cadmium (µg/L)	Gamma	723	0.643	0	358	0.511	0	395	0.445	0	38	0.295	0.051	95	0.110	0.388
	Exponential	30.3	0.358	0	43	0.311	0	49.4	0.237	0	55.3	0.560	0	71.6	0.153	0.083
	Lognormal		0.120	0.004		0.113	0.039		0.083	0.136		0.206	0.338		0.052	1.380
Total Chromium (µg/L)	Gamma	435	0.292	0	235	0.151	0.004	256	0.122	0.008	36	0.252	0.386	76	0.058	1.208
	Exponential	55.4	0.132	0	58.7	0.201	0	72.7	0.067	0.381	36.1	0.272	0.290	98.7	0.176	0.019
	Lognormal		0.069	0.206		0.086	0.262		0.062	0.480		0.180	0.861		0.084	0.685
Total Copper (µg/L)	Gamma	799	0.394	0	387	0.296	0	416	0.408	0	39	0.107	0.226	97	0.451	0
	Exponential	83.6	0.149	0	92.8	0.137	0	89.9	0.177	0	74.4	0.127	0.092	99	0.231	0.090
	Lognormal		0.067	0.005		0.070	0.060		0.080	0.017		0.131	0.742		0.038	1.507
Total Lead (µg/L)	Gamma	788	0.300	0	377	0.297	0	412	0.276	0	45	0.177	0.608	107	0.203	0
	Exponential	71.3	0.173	0	85.4	0.136	0	76.5	0.225	0	42.2	0.389	0.006	100	0.125	0.072
	Lognormal		0.044	0.218		0.057	0.250		0.059	0.223		0.132	1.034		0.039	1.451
Total Nickel (µg/L)	Gamma	419	0.292	0	232	0.260	0	250	0.090	0.159	38	0.164	1.373	99	0.188	0.004
	Exponential	45.3	0.203	0	59.5	0.176	0	62.4	0.111	0.044	18.4	0.261	0.772	89.9	0.227	0
	Lognormal		0.081	0.160		0.056	0.831		0.065	0.525		0.166	1.360		0.091	0.460
Total Selenium (µg/L)	Gamma	318	0.263	0.095	169	0.169	0.952	203	0.434	0.022	19	-	-	16	-	-
	Exponential	6.9	0.254	0.117	7.7	0.174	0.907	5.9	0.256	0.416	21.1	-	-	6.3	-	-
	Lognormal		0.253	0.119		0.196	0.735		0.190	0.841		-	-		-	-
Total Silver (µg/L)	Gamma	406	0.421	0	222	0.143	0.718	287	0.263	0.002	19	-	-	21	-	-
	Exponential	12.6	0.333	0	11.3	0.159	0.563	17.4	0.340	0	5.3	-	-	19	-	-
	Lognormal		0.271	0.001		0.184	0.370		0.146	0.236		-	-		-	-
Total Zinc (µg/L)	Gamma	810	0.244	0	392	0.234	0	433	0.273	0	45	0.180	0.253	93	0.158	0.023
	Exponential	96.4	0.122	0	99	0.141	0	98.6	0.083	0.005	71.1	0.167	0.336	96.8	0.155	0.027
	Lognormal		0.054	0.020		0.040	0.585		0.044	0.389		0.105	0.981		0.063	0.985

* P-values greater than one are used only for comparison. NDet: Number of collected samples and percentage detected

Table 20 shows that for residential, commercial and industrial land uses, the lognormal distribution better fits the empirical data, except for selenium and silver in commercial land uses. In open space land uses, about 50% of the constituents were adequately fitted by the lognormal distribution, 30% by the gamma distribution and the remaining by the exponential distribution. In freeway areas, lognormal distributions better fit most of the constituents, except that fecal streptococcus, total arsenic and total chromium were better fitted by the gamma distribution and ammonia was better fitted by the exponential distribution. Also note in Table 20 that residential, commercial and industrial land uses had larger sample sizes than the other two land uses. It seems that for small sample sizes, gamma and exponential distributions better represent actual stormwater constituent distributions, but once the number of samples increases, the lognormal distribution is best. The few cases were the gamma distribution was a better fit was for NO_2+NO_3 in industrial land uses, and chromium in freeway areas. The exponential distribution better represents total ammonia in freeway areas (with around 70 detected samples) than the other two distribution types.

Other transformations were also tested, such as the square root, and other power functions, but the results were not improved. It was therefore decided to investigate if a three-parameter lognormal distribution function can be used to improve the overall goodness of fit for stormwater constituent probability distributions. As shown in the following section, this third parameter, in some cases, allows a much better fit of the cumulative empirical and estimated probability distributions.

Three Parameter Lognormal Calculations

Goodness of fit was evaluated using 3-parameter lognormal probability distribution. The probability distributions were created for residential, commercial, industrial, open space and freeways land uses. The distribution parameters were calculated using the maximum likelihood and the L-moments methods. The maximum likelihood method requires that it be solved iteratively using three equations (see Appendix C). The results were compared with the 2-parameter standard model and the actual data. The model with the smaller maximum distance between the empirical and the estimated function was selected as the best model. All the calculations were made using only the detected values. In general, the L-moments method provided a better fit for the upper tail of the distribution whereas the maximum likelihood method provided a better fit for the lower tail. Figure 34 shows the three estimated models for TSS in commercial land use areas.

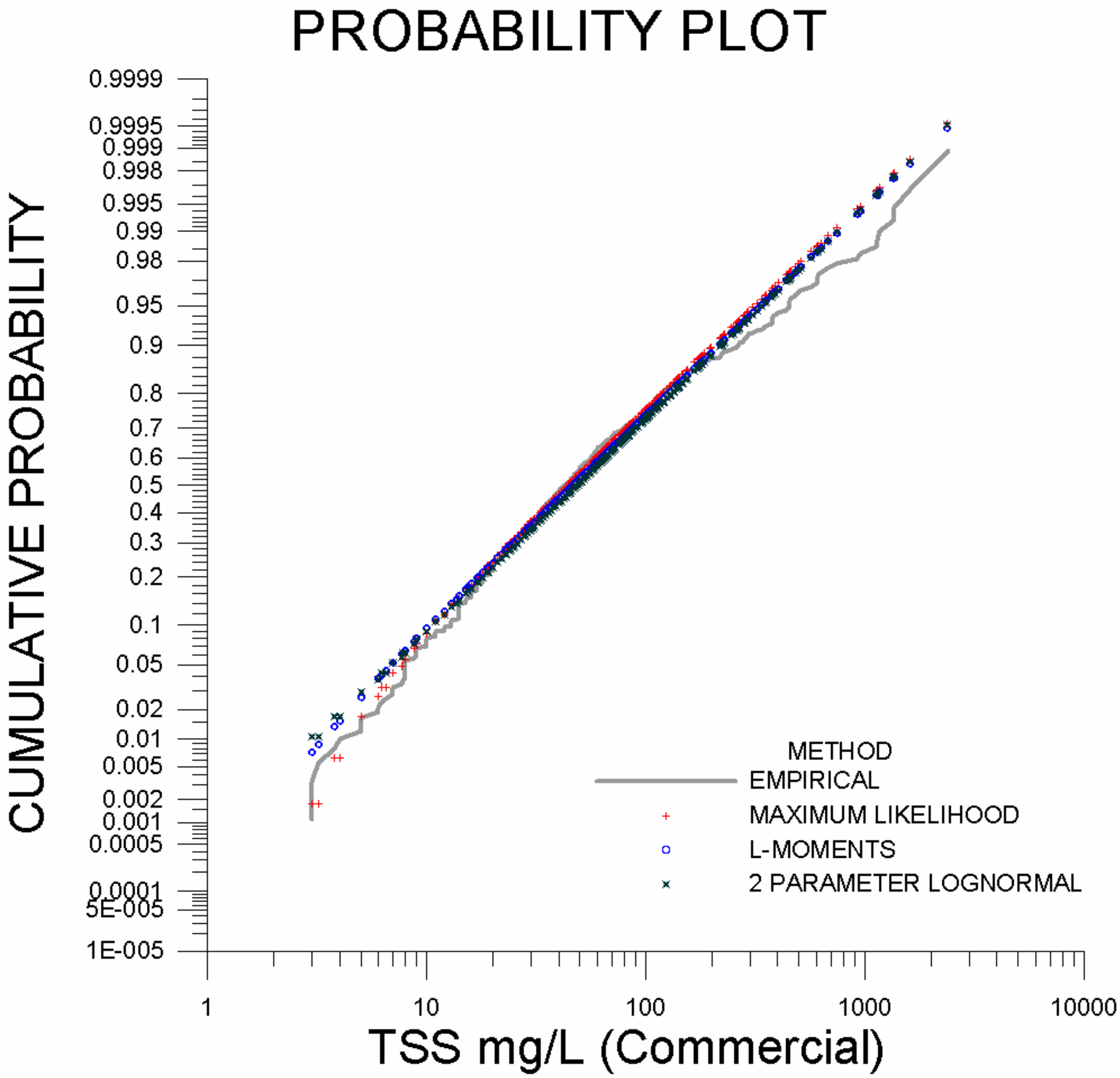


Figure 34. Estimated models for TSS in commercial land uses

In this graph, it is observed that the empirical distribution has higher values in the upper tail compared with any of the three models. In the lower tail, the maximum likelihood method using the 3-parameters better fit the observed values. In this case, the maximum likelihood method was better than the other two models, although none of the methods adequately represented the extreme high values. The L-moments method generally betters fits the upper tail distribution, but typically trims or overestimates the lower tail. Figure 35 shows the results for TDS in industrial land uses. The L-moments better fits the empirical distribution in the upper tail, but it trims any observation smaller than 35 mg/L (almost 20% of the total dataset) in the lower tail. The 2-parameter lognormal and the maximum likelihood method provide better results although both were worse than the L-moments in the upper tail region.

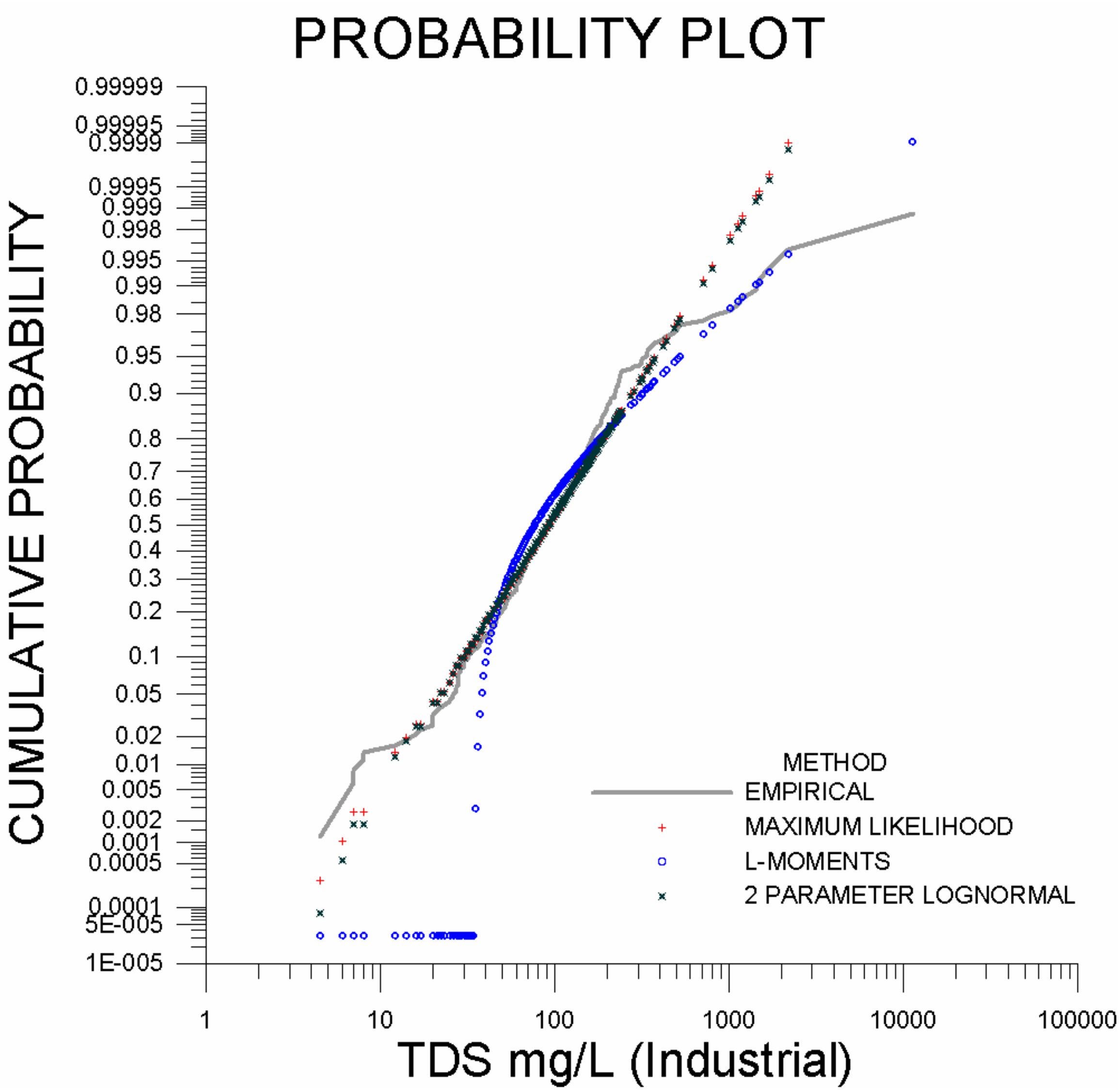


Figure 35. Estimated models for TDS in industrial land use

Table 21 presents the results for 15 constituents in five land uses. For each of the three methods, the p-value was calculated. The higher the p-value, the better is the fit between the empirical and the estimated function. Some of the p-values in the table are larger than one. When the number of samples is large, the p-value is calculated as a chi square distribution with two degrees of freedom. This probability is calculated only with one tail of the chi square distribution. The p-value is two times this probability. The maximum p-value is one, but for effects of comparison this presents two times the probability calculated from a one tail chi square distribution.

The maximum likelihood method with 3-parameters, or the lognormal 2-parameter distribution produced the best descriptions for most of the constituents. For almost all constituents the function estimated by the L-moments method failed the lognormal assumption. Low p-values were obtained because the function was truncated and does not estimate the lower tail of the distribution.

It seems that when the numbers of samples increase, the L-moments method tends to truncate the function. The maximum likelihood method seems to improve the fit of the distribution, but when the number of samples is large, the cumulative estimated probability distribution is far from the cumulative empirical probability distribution, or no convergence is possible during the iteration process.

In commercial, industrial and freeways land uses, the numbers of samples available were between 100 and 500 samples. According to the prior discussion, this number of samples will result in an analysis having a power close or above 0.5. In these cases, most of the better fits were obtained using the L-moments method. In commercial and industrial land uses, more than half of the constituents also had the highest p-values when the L-moments method was used.

In open space areas, there were not many samples available. The small number of samples results in a low power. In this case, the higher p-values results were observed when the 2-parameter lognormal distribution was used. The use of the third parameter in constituents having small numbers of sample observations did not improve the fit of the estimated cumulative probability distribution.

Summary

Most of the stormwater constituents can be assumed to follow a lognormal distribution with little error. The use of the third parameter does not show a significant improvement in estimating the empirical distribution compared with the 2-parameter lognormal distribution. When the number of samples is very large per category (approximately more than 400 samples) the maximum likelihood and the 2-parameter lognormal distribution better fit the empirical distribution. For large sample sizes, the L-moments method usually unacceptably truncates the distribution in the lower tail. When the sample size is small (<100 samples), the use of the third parameter does not improve the fit with the empirical distribution and the 2-parameter lognormal distribution produces a better fit than the other two methods.

The lognormal distribution is a skewed distribution when plotted in real space coordinates. When the sample size is small, the calculated skewness is smaller than the skewness of the real distribution. Insufficient sample sizes are not likely to accurately represent the extreme observations in the actual distribution of the data.

Table 21. Goodness of Fit for Different Land Uses (p-values Larger than One are Used Only for Comparison)

RESIDENTIAL			Critical	Observed		2 Parameter Log		3 parameter maximum likelihood					3 parameter L-moments				
CONSTITUENT	N	% Detected	$D_{0.05}$	μ	σ	D2	p-value	μ	σ	c	D3m	p-value	κ	α	ξ	D3l	p-value
Conductivity (µS/cm)	106	100	0.132	4.638	0.710	0.081	0.493	4.327	0.919	20.767	0.052	1.133	-1.240	61.461	88.366	0.087	0.398
Hardness (mg/L)	250	100	0.086	3.497	0.706	0.071	0.166	3.539	0.675	-1.114	0.066	0.231	-	-	-	0.635	0
Oil & Grease (mg/L)	533	57.8	0.077	1.428	1.204	0.112	0.001	1.356	1.267	0.164	0.102	0.003	-2.559	2.065	2.482	0.208	0
Fecal Coliform (Col/100 ml)	446	88.3	0.069	8.7205	2.40448	0.080	0.013	8.734	2.367	-2.992	0.078	0.017	-1.929	17283.541	8236.423	0.096	0.001
Fecal Streptococcus (C/100 ml)	305	89.5	0.082	9.8344	1.88029	0.077	0.081	9.907	1.725	-210.924	0.066	0.190	-1.309	38433.290	24834.760	0.077	0.078
Ammonia (mg/L)	595	81.5	0.062	-1.1672	0.9166	0.044	0.305	0.015	0.220	-0.684	0.139	0.000	-	-	-	-	-
DP (mg/L)	738	84.1	0.055	-1.8303	0.85689	0.043	0.199	0.003	0.056	-0.877	0.327	0.000	-	-	-	-	-
Antimony (µg/L)	288	2.8	0.454	1.3554	1.71904	0.233	0.841	6.569	0.023	-700.000	0.368	0.228	-1.117	9.631	5.281	0.270	0.626
Arsenic (µg/L)	426	42	0.102	1.2098	0.85043	0.154	0.000	1.047	0.971	0.356	0.166	0.000	-1.588	1.981	2.632	0.208	0
Beryllium (µg/L)	301	7.3	0.281	0.0283	1.5958	0.200	0.342	-0.423	2.040	0.136	0.186	0.436	-2.063	1.154	0.675	0.227	0.207
Total Cadmium (µg/L)	723	30.3	0.092	-0.3532	1.21891	0.120	0.004	-0.444	1.301	0.033	0.110	0.010	-1.880	0.626	0.511	0.093	0.045
Total Chromium (µg/L)	435	55.4	0.088	1.5794	0.89199	0.069	0.206	1.473	0.980	0.328	0.067	0.236	-1.157	4.003	4.382	0.066	0.242
Total Nickel (µg/L)	419	45.3	0.099	1.7909	0.75669	0.081	0.160	1.601	0.890	0.768	0.083	0.147	-1.078	4.163	5.384	0.084	0.137
Total Selenium (µg/L)	318	6.9	0.281	1.0969	0.83323	0.253	0.119	0.479	1.348	0.876	0.259	0.061	-1.178	2.103	2.577	0.257	0.11
Total Silver (µg/L)	406	12.6	0.19	1.0686	1.3707	0.271	0.001	0.984	1.469	0.089	0.278	0.001	-1.522	3.390	2.767	0.294	0

COMMERCIAL			Observed			2 Parameter Log		3 parameter maximum likelihood					3 parameter L-moments				
CONSTITUENT	N	% Detected	μ	σ	$D_{0.05}$	D2	p-value	μ	σ	c	D3m	p-value	κ	α	ξ	D3l	p-value
Conductivity (µS/cm)	66	100	4.779	0.721	0.167	0.100	0.530	4.736	0.746	3.865	0.097	0.581	-1.011	76.386	108.055	0.093	0.633
Hardness (mg/L)	139	100	3.689	0.988	0.115	0.090	0.206	3.828	0.844	-3.808	0.063	0.653	-0.935	36.911	40.394	0.072	0.474
Oil & Grease (mg/L)	308	70.8	1.609	1.070	0.092	0.103	0.019	1.300	1.358	0.737	0.092	0.062	-1.853	3.648	3.638	0.126	0.002
TDS (mg/L)	399	99.5	4.332	0.791	0.068	0.049	0.303	4.393	0.741	-3.495	0.049	0.289	-1.066	55.106	70.423	0.071	0.035
TSS (mg/L)	458	98.3	3.883	1.180	0.064	0.064	0.053	3.735	1.218	1.988	0.042	0.416	-1.207	55.082	46.245	0.048	0.250
BOD (mg/L)	432	97.5	2.493	0.868	0.066	0.054	0.166	2.302	1.026	1.396	0.040	0.527	-1.002	10.321	11.447	0.044	0.380
COD (mg/L)	373	98.4	4.167	0.865	0.071	0.038	0.695	4.163	0.868	0.194	0.037	0.719	-0.911	54.903	63.127	0.034	0.860
Fecal Coliform (Col/100 ml)	233	88	8.202	2.380	0.095	0.076	0.192	8.191	2.398	1.870	0.077	0.175	-1.768	11370.330	5408.770	0.150	0.000
Fecal Streptococcus (C/100 ml)	181	91.7	8.940	2.061	0.106	0.097	0.091	8.936	2.061	2.494	0.096	0.093	-1.640	16795.600	9532.540	0.056	0.702
Ammonia (mg/L)	299	83.3	-0.706	1.083	0.086	0.050	0.589	-0.697	1.072	-0.002	0.048	0.632	-0.947	0.549	0.522	0.040	0.888
NO2+NO3 (mg/L)	425	98.1	-0.523	0.882	0.067	0.040	0.531	-0.432	0.800	-0.039	0.034	0.837	-0.849	0.510	0.600	0.030	0.954
TKN (mg/L)	449	97.3	0.471	0.828	0.065	0.042	0.423	0.575	0.734	-0.126	0.050	0.228	-0.866	1.246	1.571	0.032	0.816
DP (mg/L)	323	81.1	-2.077	1.016	0.084	0.075	0.104	-2.157	1.092	0.006	0.062	0.273	-1.077	0.127	0.121	0.059	0.315
TP (mg/L)	446	95.7	-1.473	0.881	0.066	0.038	0.582	-1.537	0.935	0.010	0.041	0.466	-0.991	0.196	0.220	0.049	0.264
Arsenic (µg/L)	213	32.9	0.9336	0.92361	0.163	0.164	0.046	0.729	1.098	0.301	0.195	0.010	-1.736	1.565	1.935	0.280	0
Total Cadmium (µg/L)	358	43	0.047	1.309	0.11	0.113	0.039	0.023	1.335	0.011	0.109	0.052	-1.640	1.180	0.879	0.063	0.591
Total Chromium (µg/L)	235	58.7	1.8134	0.71608	0.116	0.086	0.262	1.711	0.787	0.453	0.089	0.225	-0.734	4.462	6.100	0.083	0.295
Total Copper (µg/L)	387	92.8	2.8829	0.90117	0.072	0.070	0.060	2.807	0.965	0.874	0.063	0.117	-1.251	14.103	15.641	0.069	0.067
Total Lead (µg/L)	377	85.4	3.0328	1.03226	0.076	0.057	0.250	3.015	1.049	0.220	0.058	0.225	-1.251	19.633	18.805	0.053	0.329
Total Nickel (µg/L)	232	59.5	1.9782	0.8075	0.116	0.056	0.831	1.668	1.070	1.282	0.089	0.220	-0.969	5.691	6.842	0.076	0.406
Total Selenium (µg/L)	169	7.7	1.366	0.953	0.361	0.196	0.735	0.829	1.492	0.863	0.167	0.670	-0.803	3.940	3.621	0.210	0.638
Total Silver (µg/L)	222	11.3	0.9637	1.35108	0.272	0.184	0.370	1.080	1.174	-0.141	0.182	0.379	-0.911	3.587	3.133	0.165	0.513
Total Zinc (µg/L)	392	99	5.0388	0.84183	0.069	0.040	0.585	5.082	0.803	-4.834	0.039	0.619	-1.021	120.091	144.868	0.052	0.243

Table 21. Goodness of Fit for Different Land Uses (p-values Larger than One are Used Only for Comparison) – *Continued*

INDUSTRIAL			Observed			2 Parameter Log		3 parameter maximum likelihood					3 parameter L-moments				
CONSTITUENT	N	% Detected	μ	σ	$D_{0.05}$	D2	p-value	μ	σ	c	D3m	p-value	κ	α	ξ	D3l	p-value
Conductivity (μS/cm)	108	100	5.011	0.673	0.131	0.074	0.619	4.743	0.848	27.365	0.067	0.767	-1.197	83.673	129.715	0.106	0.174
Hardness (mg/L)	138	96.4	3.794	0.842	0.118	0.080	0.369	3.701	0.914	2.758	0.081	0.346	-1.272	31.221	38.285	0.119	0.047
Oil & Grease (mg/L)	327	65.1	1.623	1.153	0.093	0.098	0.032	1.456	1.298	0.447	0.093	0.048	-3.227	1.068	2.750	0.313	0.000
TDS (mg/L)	413	99.5	4.516	0.870	0.067	0.066	0.053	4.539	0.849	-1.484	0.065	0.064	-1.496	62.528	76.123	0.111	0.000
TSS (mg/L)	428	99.1	4.287	1.200	0.066	0.029	0.995	4.292	1.193	-0.169	0.028	1.023	-1.133	88.174	74.697	0.026	1.119
BOD (mg/L)	406	95.3	2.4121	0.992	0.069	0.105	0.000	2.303	1.085	0.729	0.095	0.002	-2.246	5.718	7.369	0.184	0.000
COD (mg/L)	362	98.9	4.2217	0.91441	0.072	0.074	0.040	4.237	0.899	-0.714	0.076	0.032	-1.096	57.774	63.159	0.046	0.437
Fecal Coliform (Col/100 ml)	297	87.9	7.6064	2.65965	0.084	0.051	0.510	7.574	2.732	1.560	0.055	0.417	-2.356	5638.993	2369.488	0.045	0.688
Fecal Streptococcus (C/100 ml)	195	93.8	9.1491	1.82906	0.101	0.083	0.161	9.190	1.741	-64.211	0.073	0.280	-2.253	11045.202	7378.455	0.199	0.000
Ammonia (mg/L)	254	85.8	-0.7071	1.00903	0.092	0.047	0.758	-0.685	0.985	-0.007	0.049	0.715	-0.864	0.518	0.524	0.046	0.789
NO2+NO3 (mg/L)	418	96.2	-0.3857	0.93116	0.068	0.080	0.011	-0.142	0.703	-0.132	0.043	0.454	-0.689	0.608	0.739	0.045	0.406
TKN (mg/L)	440	95.9	0.4238	0.88424	0.066	0.048	0.292	0.471	0.840	-0.050	0.050	0.239	-1.023	1.256	1.444	0.040	0.502
DP (mg/L)	325	87.1	-2.1766	0.87102	0.081	0.124	0.682	-2.141	0.837	-0.003	0.051	0.450	-1.002	0.093	0.108	0.063	0.211
TP (mg/L)	434	96.3	-1.2683	0.98202	0.067	0.049	0.273	-1.299	1.010	0.005	0.044	0.387	-1.068	0.271	0.271	0.035	0.724
Antimony (μg/L)	164	14.6	1.4793	1.01264	0.269	0.096	1.279	1.275	1.183	0.479	0.113	1.088	-1.334	3.747	3.661	0.150	0.684
Arsenic (μg/L)	267	54.3	1.5218	0.95205	0.113	0.129	0.016	1.19121.2628		0.752	0.128	0.018	-1.069	4.359	4.359	0.116	0.039
Beryllium (μg/L)	209	10.5	-0.3588	1.94765	0.281	0.163	0.620	-0.892	2.658	0.060	0.197	0.362	-2.074	1.346	0.568	0.231	0.191
Total Cadmium (μg/L)	395	49.4	0.7417	1.12552	0.097	0.083	0.136	0.588	1.276	0.161	0.095	0.060	-1.611	1.898	1.686	0.115	0.012
Total Chromium (μg/L)	256	72.7	2.5512	1.06906	0.1	0.062	0.480	2.599	1.015	-0.359	0.059	0.543	-0.911	13.859	13.657	0.050	0.803
Total Copper (μg/L)	416	89.9	3.2275	1.02866	0.07	0.080	0.017	3.179	1.076	0.716	0.073	0.030	-1.343	22.969	22.081	0.057	0.172
Total Lead (μg/L)	412	76.5	3.3651	1.32824	0.077	0.059	0.223	3.333	1.367	0.379	0.057	0.263	-1.374	38.025	27.991	0.054	0.316
Total Nickel (μg/L)	250	62.4	2.8058	0.97034	0.109	0.065	0.525	2.802	0.971	0.042	0.066	0.512	-0.772	17.094	17.834	0.088	0.182
Total Selenium (μg/L)	203	5.9	1.1472	1.27671	0.375	0.190	0.841	6.916	0.014	-1000.000	0.364	0.083	-1.851	2.381	2.025	0.202	0.753
Total Silver (μg/L)	287	17.4	0.113	1.80819	0.192	0.146	0.236	0.158	1.708	-0.010	0.153	0.194	-1.358	2.181	1.394	0.157	0.170
Total Zinc (μg/L)	433	98.6	5.305	0.970	0.066	0.044	0.389	5.359	0.915	-7.026	0.034	0.743	-0.899	198.198	208.408	0.030	0.951

Table 21. Goodness of Fit for Different Land Uses (p-values Larger than One are Used Only for Comparison) – *Continued*

OPEN SPACE			Observed			2 Parameter Log		3 parameter maximum likelihood					3 parameter L-moments				
CONSTITUENT	N	% Detected	μ	σ	$D_{0.05}$	D2	p-value	μ	σ	c	D3m	p-value	κ	α	ξ	D3l	p-value
TDS (mg/L)	45	97.8	4.762	0.744	0.205	0.120	0.561	4.480	0.962	20.400	0.115	0.621	-	-	-	0.759	0
TSS (mg/L)	44	95.5	3.945	1.717	0.21	0.113	0.683	3.672	2.096	2.777	0.095	0.942	-1.184	121.932	72.238	0.173	0.162
BOD (mg/L)	44	86.4	1.6211	0.66954	0.215	0.114	0.746	1.624	0.659	-0.013	0.115	0.733	-0.614	3.421	5.098	0.110	0.801
COD (mg/L)	43	76.7	3.548	0.785	0.231	0.128	0.684	3.323	0.946	5.000	0.151	0.441	-1.221	23.447	-16.159	0.184	0.215
Fecal Coliform (Col/100 ml)	23	91.3	8.9527	1.43209	0.287	0.181	0.503	8.725	1.692	534.506	0.217	0.278	-0.791	13914.660	10684.240	0.187	0.458
Fecal Streptococcus (C/100 ml)	22	90.9	9.6472	1.62819	0.294	0.181	0.538	9.847	1.248	-1070.380	0.165	0.676	-0.599	27514.897	24175.705	0.139	0.921
NO2+NO3 (mg/L)	44	84.1	-0.478	1.07398	0.234	0.141	0.463	-0.378	1.173	0.135	0.141	0.464	-0.579	0.746	0.759	0.122	0.664
TKN (mg/L)	45	71.1	-0.097	0.92495	0.234	0.147	0.500	-0.378	1.173	0.135	0.140	0.569	-0.716	0.918	0.976	0.160	0.389
DP (mg/L)	44	79.5	-2.1844	1.08407	0.224	0.124	0.682	-2.137	1.017	-0.003	0.131	0.606	-0.595	0.135	0.137	0.142	0.484
TP (mg/L)	46	84.8	-1.4264	1.12592	0.215	0.116	0.696	-1.499	1.188	0.009	0.126	0.584	-1.999	0.161	0.171	0.211	0.062
Total Cadmium (μg/L)	38	55.3	0.0291	2.62317	0.287	0.206	0.338	6.921	0.244	-1000.000	0.327	0.022	-1.561	7.207	2.856	0.319	0.028
Total Chromium (μg/L)	36	36.1	-1.79	1.66154	0.325	0.799	0.000	6.926	0.030	-1000.000	0.275	0.279	-1.446	10.225	5.944	0.193	0.756
Total Copper (μg/L)	39	74.4	2.1184	1.20044	0.246	0.131	0.742	6.926	0.035	-1000.000	0.334	0.003	-1.539	8.380	6.792	0.187	0.262
Total Lead (μg/L)	45	42.2	1.8882	1.95431	0.301	0.132	1.034	1.634	2.355	0.188	0.203	0.417	-1.406	16.381	8.735	0.174	0.632
Total Zinc (μg/L)	45	71.1	3.6048	1.19833	0.234	0.105	0.981	3.315	1.527	3.980	0.132	0.658	-1.142	44.515	36.591	0.113	0.885

FREEWAYS			Observed			2 Parameter Log		3 parameter maximum likelihood					3 parameter L-moments				
CONSTITUENT	N	% Detected	μ	σ	$D_{0.05}$	D2	p-value	μ	σ	c	D3m	p-value	κ	α	ξ	D3l	p-value
Conductivity (μS/cm)	86	100	4.586	0.681	0.147	0.129	0.113	4.404	0.795	12.765	0.113	0.226	-1.060	58.592	87.231	0.096	0.411
Hardness (mg/L)	127	100	3.604	0.791	0.121	0.077	0.447	3.485	0.875	3.045	0.077	0.440	-1.395	22.761	30.345	0.114	0.074
Oil & Grease (mg/L)	60	71.7	1.974	0.575	0.207	0.101	0.827	1.395	0.976	2.433	0.161	0.214	-0.543	4.299	7.228	0.098	0.881
TDS (mg/L)	97	99	4.279	0.771	0.139	0.054	1.136	4.342	0.720	-3.500	0.051	1.210	-0.660	56.217	74.574	0.044	1.368
TSS (mg/L)	134	99.3	4.464	1.068	0.118	0.066	0.627	4.513	1.010	-2.484	0.059	0.797	-1.447	76.753	75.337	0.120	0.043
BOD (mg/L)	26	84.6	2.262	0.894	0.281	0.103	1.252	2.013	1.084	1.385	0.099	1.304	-1.212	7.478	8.170	0.117	1.097
COD (mg/L)	67	98.5	4.552	0.980	0.167	0.107	0.445	4.798	0.715	-17.925	0.073	0.999	-0.881	81.577	97.057	0.082	0.817
Fecal Coliform (Col/100 ml)	49	100	7.585	1.716	0.194	0.105	0.677	7.514	1.811	30.900	0.128	0.398	-1.569	3543.87	2053.58	0.108	0.635
Fecal Streptococcus (C/100 ml)	25	100	9.263	1.662	0.264	0.119	0.990	8.854	2.363	540.500	0.211	0.215	-0.896	21116.59	15451.91	0.132	0.833
Ammonia (mg/L)	79	87.3	0.009	1.094	0.164	0.133	0.173	0.141	0.942	-0.083	0.106	0.423	-1.065	1.00	1.01	0.123	0.252
NO2+NO3 (mg/L)	25	96	-1.097	0.868	0.269	0.139	0.789	-1.628	1.301	0.087	0.114	1.073	-1.275	0.241	0.275	0.117	1.038
TKN (mg/L)	125	96.8	0.750	0.887	0.124	0.074	0.539	0.723	0.907	0.038	0.075	0.512	-1.159	1.669	1.907	0.071	0.598
DP (mg/L)	22	95.5	-1.226	1.188	0.287	0.170	0.593	-1.670	1.595	0.053	0.207	0.333	-1.992	0.181	0.190	0.255	0.131
TP (mg/L)	128	99.2	-1.266	0.772	0.121	0.085	0.325	-1.545	0.961	0.051	0.065	0.686	-1.394	0.171	0.231	0.089	0.27
Antimony (μg/L)	14	50	0.967	0.284	0.483	0.419	0.171	4.415	0.008	-80.000	0.397	0.220	0.660	0.563	2.922	0.270	0.719
Arsenic (μg/L)	61	55.7	0.952	0.666	0.227	0.149	0.441	0.290	1.180	0.886	0.106	0.937	-0.679	1.787	2.547	0.096	1.07
Total Cadmium (μg/L)	95	71.6	0.016	0.838	0.165	0.052	1.380	-0.108	0.937	0.081	0.067	1.087	-0.781	0.028	1.574	0.051	1.409
Total Chromium (μg/L)	76	98.7	2.096	0.734	0.157	0.084	0.685	2.165	0.680	-0.450	0.075	0.854	-0.555	6.127	8.575	0.055	1.272
Total Copper (μg/L)	97	99.0	3.525	0.842	0.139	0.038	1.507	3.433	0.915	2.047	0.048	1.295	-0.857	28.576	33.493	0.041	1.443
Total Lead (μg/L)	107	100	3.226	1.164	0.131	0.039	1.451	3.178	1.212	0.587	0.045	1.285	-1.155	28.993	24.965	0.040	1.424
Total Nickel (μg/L)	99	89.9	2.325	0.673	0.144	0.091	0.460	1.989	0.896	2.228	0.062	1.013	-0.960	6.390	9.308	0.073	0.769
Total Zinc (μg/L)	93	96.8	5.273	0.877	0.143	0.063	0.985	5.392	0.757	-17.226	0.054	1.000	-0.920	156.885	189.601	0.064	0.959

Experimental design procedures enable the required sample size to be estimated, according to desired confidence and power of the experimental results. It may be possible, without being able to identify the real skewness, that the best distribution fit could be the gamma or exponential distribution.

The utility of the third parameter has been questioned, especially because one of the objectives in modeling is to be parsimonious. Only in cases where it is important to include the effect of unusual elevated values in the model, is the third parameter recommended. In all the other cases, the use of the 2-parameter distribution is adequate to explain the distribution of most of the contaminants.

When the mean and the standard deviation values are not known, Lilieford's test is recommended to evaluate the goodness of fit to a specific distribution. During this research, the Kolmogorov-Smirnov test was used based on the assumption that the large sample sizes minimized errors associated with small sample sizes and uncertainty in the mean and standard deviation values.

Some constituents (such as TKN, TP, COD and Cu) show an increase in the p-value when the number of samples is acceptable and the 3-parameter lognormal probability distribution is used. The use of the lognormal distribution also has an advantage over the other distributions because it can be easily transformed to a normal distribution.

The few cases where the gamma distribution seems to be a better model was for cases with low counts (constituents in open space or arsenic, chromium and fecal streptococcus in freeways areas; for example). The exponential distribution better fit total ammonia in freeway areas. The remaining constituents were well represented by the lognormal distribution.

The 2-parameter lognormal distribution is considered the most appropriate distribution to represent stormwater constituents. Its use facilitates statistical analyses of the data, because procedures such as ANOVA or regression require the errors to be normally distributed. If the number of observations is small, the use of nonparametric methods will be required, as the distributions cannot be accurately determined. Some nonparametric methods require symmetry in the data distribution. The log transformed constituent concentrations usually satisfy these assumptions.

Chapter 5: Identification of Significant Factors Affecting Stormwater Quality Using the NSQD

Introduction

The normal approach to classify urban sites for estimating stormwater characteristics is based on land use. This approach is generally accepted because it is related to the activity in the watershed, plus many site features are generally consistent within each land use. Two drainage areas with the same size, percentage of imperviousness, ground slope, sampling methods, and stormwater controls will produce different stormwater concentrations if the main activity in one watershed is an automobile manufacturing facility (industrial land use) while the other is a shopping center (commercial land use) for example. There will likely be higher concentrations of metals at the industrial site due to the manufacturing processes, while the commercial site may have higher concentrations of PAHs (polycyclic aromatic hydrocarbons) due to the frequency and numbers of customer automobiles entering and leaving the parking lots.

The results from the previous chapter indicated that there are significant differences in stormwater constituents for different land use categories. This is supported for other databases like NURP (EPA 1983) and USGS (Driver, *et al*. 1985). The main question to be addressed in this chapter is if there is a different classification method that better describes stormwater quality, possibly by also considering such factors as geographical area (EPA Rain Zone), season, percentage of imperviousness cover, type of conveyance, controls in the watershed, sampling method, and type of sample compositing, and possible interactions between these factors.

This chapter presents several approaches to explain the variability of stormwater quality by considering these additional factors. As shown in Chapter 3, ignoring the non-detected observations can adversely affect the mean, median and standard deviations of the dataset, and the resulting statistical test results. Therefore, the calculations presented in this chapter were preceded by substituting the censored observations using the Cohen's maximum likelihood method.

Main Factors Affecting Stormwater Quality

The EPA Rain Zone, percentage of imperviousness, watershed size, land use, type of conveyance, controls in the watershed, sample analysis method, and type of sampling procedures were selected as potential influencing factors affecting stormwater quality for the preliminary analyses in this chapter. Data from sites having a single land use will be used in the basic analyses, while data from the mixed land use sites could be used for verification. The first step was to inventory the total number of events in each of the possible combinations of these factors. The EPA Rain Zone, land use, type of conveyance, type of controls present in the watershed, sampling methods and type of compositing procedures are discrete variables, while percentage of imperviousness and watershed area are continuous variables. The total counts and percentage for each discrete variable option is shown in Table 22.

Table 22. Numbers and percentage of samples by discrete site variable category

Land use	Events	%
Residential	1042	28
Mixed Residential	611	16
Commercial	527	14
Mixed Commercial	324	8.6
Industrial	566	15
Mixed Industrial	249	6.6
Institutional	18	0.48
Open Space	49	1.3
Mixed Open Space	168	4.5
Freeways	185	4.9
Mixed Freeways	26	0.69

EPA Rain Zone	Events	%
1	69	1.8
2	2000	53
3	266	7.1
4	212	5.6
5	485	13
6	356	9.5
7	229	6.1
8	24	0.64
9	124	3.3

Controls	Events	%
Channel Weirs (CW)	30	0.80
Dry Pond (DP)	50	1.3
Detention Storage (enlarged pipe) (DS)	17	0.45
Wet Pond at Outfall (WP)	113	3.0
Wet Pond in Watershed (WP_W)	182	4.8
Wet Pond in Series at Outfall (WP_S)	42	1.1
None	3331	88

Sample Analysis	Events	%
Composite, type not specified	718	19
Flow Composite	2752	73
Time Composite	295	7.8

Type of Conveyance	Events	%
Curb and gutter	2454	65
Grass swale	344	9.1
Not specified	967	26

Sampler	Events	%
Automatic	3055	81
Manual	393	10
Not specified	317	8.4

About 80% of the samples were collected using automatic samplers. It was observed that manual sampling can result in lower TSS concentrations compared to automatic sampling procedures. This may occur, for example, if the manual sampling team arrives after the start of runoff and therefore misses the first flush (if it exists for the site), resulting in reduced event mean concentrations. For those sites using automatic samplers, about 73% of the events were collected using flow-composite samplers, 8% were collected using time-composite samplers, and about 19% did not have any designation available. Flow-composite samples are considered more accurate than time-composite samples when obtaining data for event mean concentrations, unless very large numbers of subsamples are obtained (Roa-Espinosa and Bannerman 1995).

Almost 66% of the events were collected at sites drained with conventional curbs and gutters, 9% were collected at sites having roadside grass swales, and it was not possible to determine the drainage system for about 25% of the samples. Grass swales can reduce the concentrations of suspended solids and metals, especially during low flows. They can also infiltrate large quantities of the stormwater, reducing pollutant mass discharges, runoff volume, and peak flows.

Effects of Stormwater Controls on Stormwater Quality

It is hoped that stormwater controls located in a watershed, or at an outfall, would result in significant reductions in stormwater pollutant concentrations. Figure 36 shows the effects on effluent TSS concentrations when using various controls in residential area watersheds in EPA Rain Zone 2 (Maryland, Virginia, North Carolina, Tennessee and Kentucky), an area having enough samples for an effective statistical analysis. The controls noted for these locations included:

- Channel weir: a flow measurement weir in an open channel that forms a small pool (a very small wet pond).
- Dry pond (DP): a dry detention pond that drains completely between each storm event.
- Wet pond (WP): a wet detention pond that retains water between events, forming a small lake or pond. If the pond is in the watershed but not at the outfall, this will be considered a wet pond inside of the watershed (WPW), which would only treat a fraction of the total stormwater from the site
- Detention storage (DS): Oversize pipes with small outlet orifices, usually under parking lots.

The stormwater monitoring was conducted at the outfalls of the drainage areas, after the stormwater controls. Wet ponds are seen to reduce the TSS concentration in the stormwater more than the other controls (about 78%) compared to the “no control” median value. Detention storage units and dry ponds also reduced the TSS concentrations, but to a smaller extent (about 60% and 37% respectively). Only one site (located in Virginia Beach) had a channel weir control, but that site did not reduce the observed TSS concentrations compared to the “no control” category. The effectiveness of the stormwater controls were evaluated for each constituent separately. The effects of sample analysis method, sampler instrument, and type of conveyance were also examined.

The first step was to identify the suitable subsets that could be examined, based on suitable numbers of samples in each category. The following four land uses and EPA Rain Zones had suitable numbers of sites having controls that could be examined: residential, commercial and industrial areas in EPA Rain Zone 2 and industrial areas in EPA Rain Zone 3. For each group, one-way ANOVA analyses were used to identify if there were any differences in the concentrations of 13 constituents (after log-transformations and substitutions for non-detectable values) for those sites that included different controls. Dunnet’s method was also used to compare sites with each specific stormwater control type with sites without stormwater controls, using a family error rate of 5%. Table 23 shows the results for these analyses for each of these groups.

Tables 23 through 26 show that there are no significant differences between sites with or without wet ponds for all constituents having observations in industrial land uses in EPA Rain Zone 3. Nitrite-nitrate, total phosphorus, total copper and total zinc were significantly lower in concentrations at sites located in EPA Rain Zone 2, having wet ponds before the outfall, compared to sites without stormwater controls. Wet ponds did not reduce the TKN concentrations in any of the four groups. Significant reductions in TSS concentrations were also observed for sites having wet ponds in residential and commercial land uses, but not in industrial land uses.

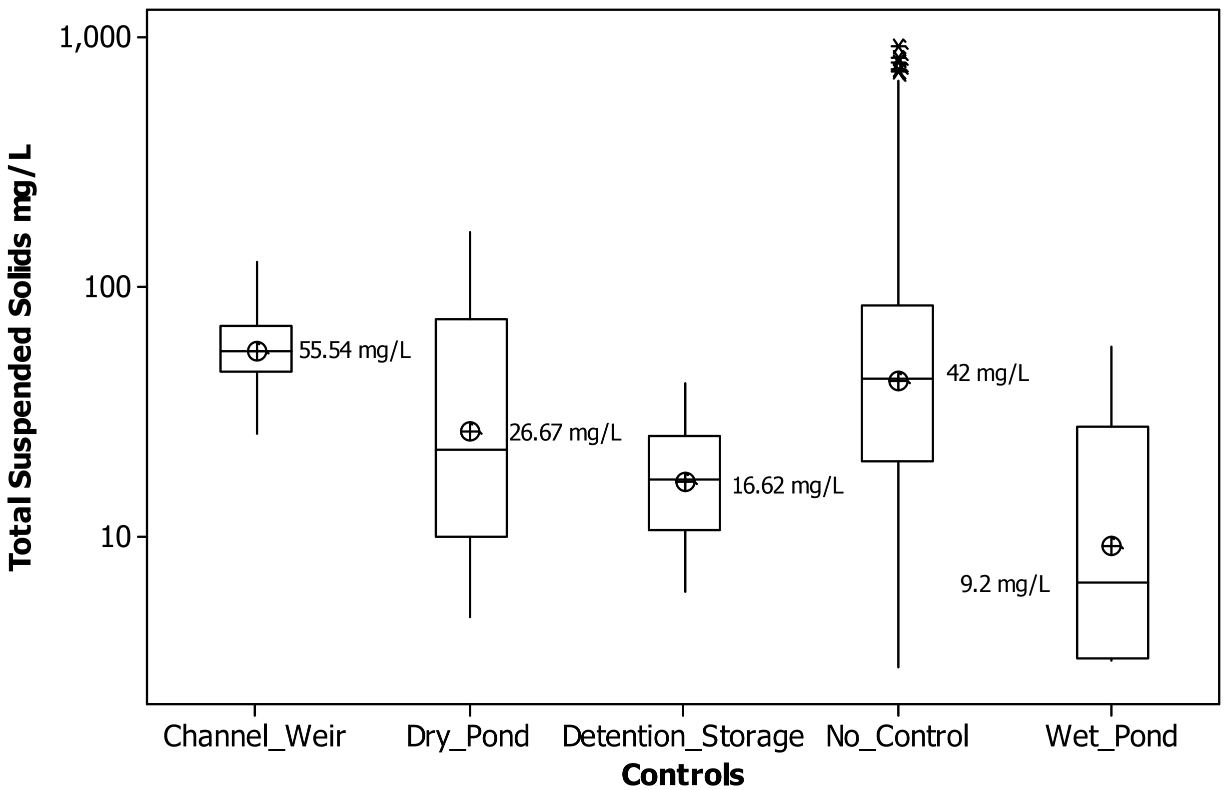


Figure 36. TSS distribution by controls in residential areas and EPA Rain Zone 2 (the cross circles indicate the average concentrations, while the median concentrations are written next to the median bar in the box diagrams)

Table 23. One-Way ANOVA Results by Control Type in Residential Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		***0.024****			0.999			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Weir				3	2.50	=	29	112.80	>	29	55.54	=
DP				3	2.68	=	3	58.88	=	21	26.67	=
DS	7	44.38	=	9	2.19	=	8	98.45	=	9	14.46	<
No Control	61	30.77		202	2.38		424	62.42		559	40.10	
WP	10	66.45	>	13	2.50	=	12	120.39	>	12	9.25	<

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		***0***			***0***			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Weir	29	6.16	<	29	49.02	=	29	0.05	<	29	0.05	<
DP	21	3.44	<	3	33.45	=	3	0.41	=	21	0.59	=
DS	9	3.66	<	9	22.17	<	9	0.40	=	9	0.98	=
No Control	533	11.07		418	56.91		409	0.24		546	0.54	
WP	12	3.10	<	12	24.58	<	12	0.07	<	12	0.28	<

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper μg/L		
p-value		***0.012***			***0***			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Weir	29	1.49	=	29	0.04	<	29	0.23	=	3	2.69	<
DP	21	0.79	<	3	0.15	=	21	0.12	<	21	6.16	<
DS	9	1.38	=	8	0.11	=	9	0.15	<	9	20.75	>
No Control	549	1.34		404	0.14		550	0.30		403	11.01	
WP	12	1.04	=	12	0.03	<	12	0.07	<	4	3.13	<

	Total Lead μg/L			Total Zinc μg/L		
p-value		***0***			***0***	
	n	median	Dunnet	n	median	Dunnet
Weir	3	6.41	<	3	4.11	<
DP	21	1.50	<	21	29.63	<
DS	9	1.16	<	9	103.25	>
No Control	364	7.73		405	67.56	
WP	4	1.00	<	4	10.44	<

Table 24. One-Way ANOVA Results by Control Type in Commercial Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		0.717			0.082			0.477			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
DS	8	58.17	=	8	1.84	=	8	100.69	=	8	19.54	<
No Control	35	58.97		100	4.20		174	74.89		244	48.13	
WP	11	71.80	=	17	2.84	=	26	89.99	=	26	19.47	<
WPW	9	47.11	=	13	3.36	=	13	71.12	=	13	16.85	<

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		***0***			***0***			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
DS	8	4.44	<	8	27.18	<	8	0.30	=	8	1.18	=
No Control	241	14.66		174	73.62		174	0.39		242	0.60	
WP	26	7.06	<	26	35.99	<	26	0.13	<	26	0.48	=
WPW	12	5.41	<	13	23.88	<	13	0.16	<	13	0.22	<

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper μg/L		
p-value		0.057			***0***			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
DS	8	1.04	=	7	0.09	=	8	0.16	=	8	14.14	=
No Control	241	1.59		161	0.11		238	0.25		194	17.53	
WP	26	1.19	=	25	0.05	=	26	0.13	<	6	5.57	<
WPW	13	1.03	=	13	0.03	=	13	0.08	<	4	6.00	<

	Total Lead μg/L			Total Zinc μg/L		
p-value		***0***			***0***	
	n	median	Dunnet	n	median	Dunnet
DS	8	1.61	<	8	82.57	<
No Control	194	16.41		197	188.02	
WP	7	4.90	<	7	44.26	<
WPW	4	2.49	<	4	39.68	<

Table 25. One-Way ANOVA Results by Control Type in Industrial Land Use, Rain Zone 2

	Hardness mg/L			**Oil and Grease mg/L**			**TDS mg/L**			**TSS mg/L**		
p-value		none			***0***			none			0.693	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
No Control				81	3.85					205	51.96	
WP				37	1.43	<				29	48.05	=

	BOD_5 mg/L			**COD mg/L**			**Ammonia mg/L**			**$NO_2 + NO_3$ mg/L**		
p-value		0.466			none			none			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
No Control	200	10.63								197	0.61	
WP	29	9.30	=							29	0.30	<

	TKN mg/L			**Dissolved Phosphorus mg/L**			**Total Phosphorus mg/L**			**Total Copper μg/L**		
p-value		0.166			none			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
No Control	198	1.22					200	0.23		150	16.00	
WP	29	0.98	=				29	0.09	<	29	7.38	<

	Total Lead μg/L			**Total Zinc μg/L**		
p-value		0.353			***0***	
	n	median	Dunnet	n	median	Dunnet
No Control	142	11.16		157	180.01	
WP	29	8.66	=	29	60.44	<

Table 26. One-Way ANOVA Results by Control Type in Industrial Land Use, Rain Zone 3

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		None			None			0.112			0.281	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	n	median	Dunnet
No Control							44	69.53		44	48.35	
WP							25	49.84	=	25	70.40	=

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.221			0.395			0.165			0.193	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	n	median	Dunnet
No Control	44	6.41		44	37.00		3	0.12		30	0.57	
WP	23	5.14	=	25	43.06	=	25	0.03	=	25	0.40	=

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper μg/L		
p-value		0.807			0.191			0.438			0.106	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	n	median	Dunnet
No Control	43	1.18		39	0.07		43	0.16		38	16.66	
WP	25	1.12	=	25	0.06	=	25	0.19	=	25	12.58	=

	Total Lead μg/L			Total Zinc μg/L		
p-value		0.454			0.608	
	n	median	Dunnet	n	median	Dunnet
No Control	31	8.49		38	143.28	
WP	25	6.73	=	25	156.93	=

Note. The bold, italicized probability values indicate “statistically significant” findings at the 0.05 level, or better.
Dunnet test compared if sites with control produces larger concentrations “>”, smaller concentrations “<” or not statistically difference “=” than sites without control at a family error of 5%.
“None” indicates no samples were collected for this constituent in the group.

Dry ponds were only available for evaluation in the residential land use category in EPA Rain Zone 2. No significant differences were found for TSS or nitrite-nitrate for sites having dry ponds. However, significant reductions of BOD_5, TKN, total phosphorus, total copper, total lead and total zinc were noted.

Some communities have installed detention-storage facilities (enlarged pipes) under parking lots to reduce runoff flow rates. More than 400 of these underground pipes are located in Arlington, Virginia, for example. A significant reduction in the TSS, BOD_5, COD, total lead, and total zinc concentrations were observed at sites with these underground devices. On the other hand, these controls did not indicate a significant difference in the concentrations of nutrients (ammonia, nitrite-nitrate, TKN, dissolved phosphorus and total phosphorus), compared to comparable sites not having stormwater controls. A conflicting situation was observed in EPA Rain Zone 2 for total zinc for sites having underground enlarged pipes. Zinc concentrations at residential land uses were significantly higher, while zinc concentrations at commercial areas were significantly lower, compared to sites with no stormwater controls. It is possible that the sites having elevated zinc concentrations used galvanized metal enlarged pipe systems.

Sampling Method Effects on Stormwater Concentrations

The use of manual or automatic sampling is a factor that is sometimes mentioned as having a possible effect on the quality of the collected samples. Manual sampling is usually preferred when the number of samples is small and when there are not available resources for the purchase, installation, operation, and maintenance of automatic samplers. Manual sampling may also be required when the constituents being sampled require specific handling (such as for bacteria, oil and grease, and volatile organic compounds) (ASCE/EPA 2002). Automatic samplers are recommended for larger sampling programs, when better representations of the flows are needed, and especially when site access is difficult or unsafe. In most cases, where a substantial number of samples are to be collected and when composite sampling is desired, automatic sampling can be much less expensive. Automatic samples also improve repeatability by reducing additional variability induced by the personnel from sample to sample (Bailey 1993). Most importantly, automatic samplers can be much more reliable compared to manual sampling, especially when the goal of a monitoring project is to obtain data for as many of the events that occur as possible, and sampling must start near the beginning of the rainfall (Burton and Pitt 2002).

Residential, commercial and industrial sites located in EPA Rain Zone 2 were used to evaluate any significant differences between the two sampling methods. One-way ANOVA analyses were used to identify any statistical differences between the two groups. Dunnet’s test was used to compare manual sampling against automatic sampling. Tables 27 through 29 show the results from these ANOVA analyses.

Table 27. One-Way ANOVA Results by Type of Sampler in Residential Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		***0***						***0.004***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	23	51.9		All manual			318	65.4		420	45.5	
Manual	28	22.4	<				66	50.0	<	78	19.2	<

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.162			***0***			0.909			***0.005***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	396	11.3		312	62.2		310	0.229		410	0.51	
Manual	78	9,8	=	66	36.4	<	66	0.233	=	75	0.66	>

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		***0.048***			0.308			***0***			***0.025***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	410	1.40		302	0.136		416	0.325		256	11.57	
Manual	78	1.16	<	63	0.120	=	73	0.230	<	77	8.80	<

	Total Lead µg/L			Total Zinc µg/L		
p-value		***0***			***0.02***	
	n	median	Dunnet	n	median	Dunnet
Automatic	247	9.74		256	73.71	
Manual	71	4.14	<	76	53.22	<

Table 28. One-Way ANOVA Results by Type of Sampler in Commercial Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		***0***			***0.009***			0.25			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	23	97.86		70	4.75		123	76.36		179	52.29	
Manual	12	22.34	<	19	2.30	<	18	60.80	=	24	20.55	<

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.189			***0.003***			0.569			0.137	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	178	14.86		123	79.74		123	0.359		178	0.55	
Manual	23	11.70	=	18	44.02	<	18	0.433	=	23	0.75	=

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		0.117			0.554			***0.003***			***0.001***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	177	1.63		113	0.097		176	0.261		127	20.27	
Manual	24	1.21	=	17	0.115	=	23	0.157	<	23	11.80	<

	Total Lead µg/L			Total Zinc µg/L		
p-value		0.422			0.404	
	n	median	Dunnet	n	median	Dunnet
Automatic	130	17.62		130	208	
Manual	20	13.66	=	23	168	=

Table 29. One-Way ANOVA Results by Type of Sampler in Industrial Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			0.723			0.362			0.402	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic				62	3.68		128	73.2		171	51.45	
Manual				14	4.10	=	10	100.0	=	19	62.82	=

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.112			0.371			***0***			***0.021***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Automatic	166	9.65		127	55.02		122	0.243		163	0.558	
Manual	19	13.47	=	10	67.68	=	10	1.54	>	19	0.904	>

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper μg/L		
p-value		***0.008***			0.870			***0.056***			0.797	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	N	median	Dunnet
Automatic	164	1.135		109	0.091		166	0.214		108	15.66	
Manual	19	1.944	>	10	0.086	=	19	0.315	=	22	14.97	=

	Total Lead μg/L			Total Zinc μg/L		
p-value		0.908			***0.028***	
	n	median	Dunnet	n	median	Dunnet
Automatic	109	11.27		115	156	
Manual	16	10.83	=	22	233	>

Note. The bold, italicized probability values indicate "statistically significant" findings at about the 0.05 level, or better.
Dunnet's test compared if sites with controls produce larger concentrations ">", smaller concentrations "<" or not statistically difference "=" than sites without controls at a family error of 5%.
"None" indicates no samples were collected for this constituent in the group.

Tables 27 through 29 indicated that BOD_5 and dissolved phosphorus measurements are not affected by differences in sampling methods used in residential, commercial or industrial areas in EPA Rain Zone 2. In residential and commercial land uses, TSS and COD concentrations obtained using automatic samplers were almost twice the concentrations obtained when using manual sampling methods. Median total phosphorus concentrations were about 50% higher using automatic samplers, while no effects were noted for other nutrients. Figure 37 contains box and whisker plots comparing automatic versus manual sampling methods in residential land uses in EPA Rain Zone 2. TSS, total copper and total zinc have lower concentrations using manual sampling compared with automatic sampling (p-values of 0, 0.025 and 0.02 respectively). The opposite pattern was observed for nitrate-nitrate; manual sampling shows higher concentrations than samples collected with automatic samples (p-value of 0.005).

In industrial land uses, the pattern was found to be opposite. Ammonia, nitrate-nitrite, TKN and total zinc indicated higher concentrations when using manual sampling methods compared to using automatic samplers. Concentrations for these constituents were almost twice as high when using manual sampling, except for ammonia that was almost six times higher when manual sampling was used compared to automatic sampling methods. These elevated concentrations were observed in industrial sites located in Fairfax County Virginia, Howard County Maryland and the city of Charlotte in North Carolina. Sites with controls were not included in this analysis of the effects of sampling method.

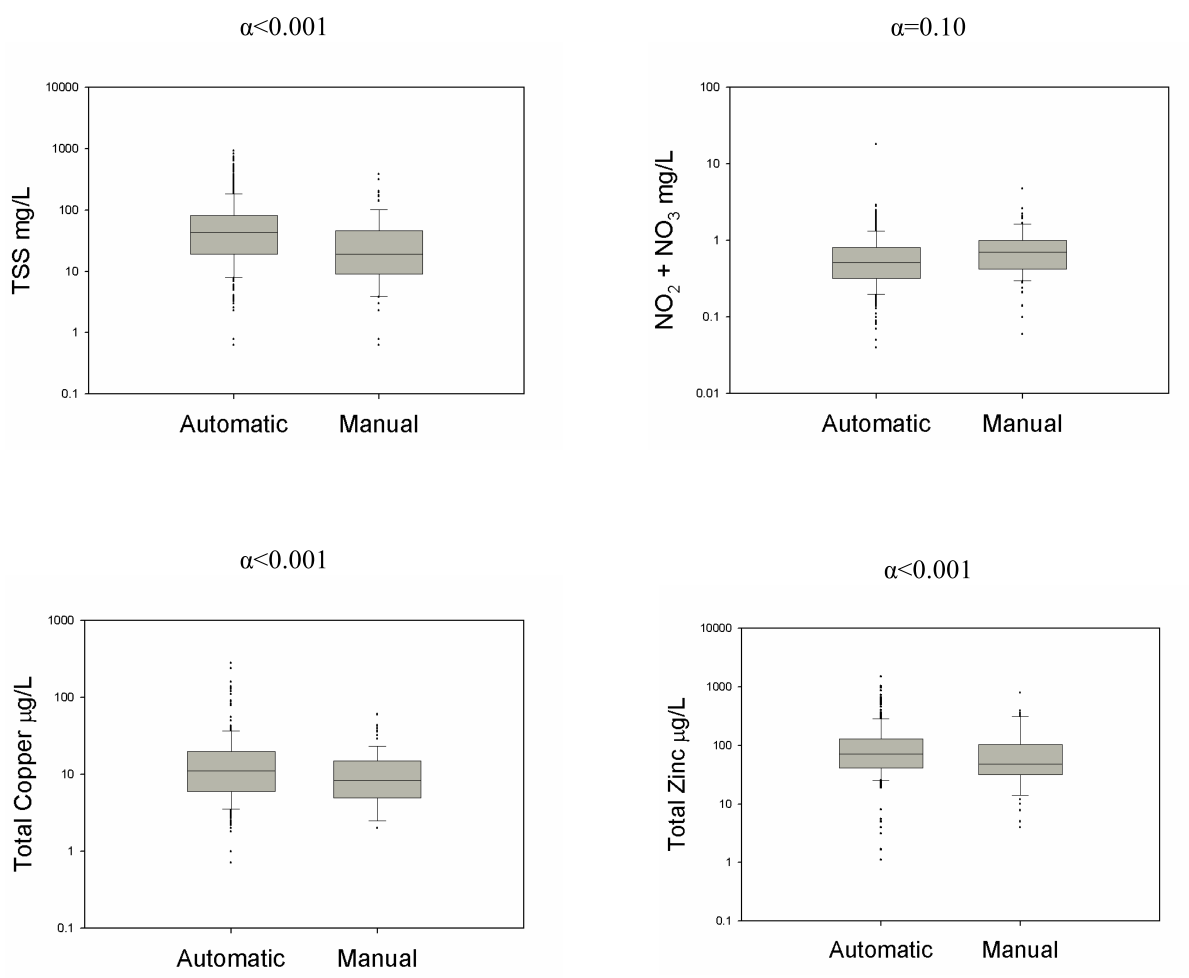


Figure 37. Comparison of reported concentrations in residential land use and EPA Rain Zone 2 for automatic vs. manual sampling methods

Sample Compositing Procedures

Time and flow-weighted composite options were also evaluated in residential, commercial, and industrial land uses in EPA Rain Zone 2 and in industrial land uses in EPA Rain Zone 3. With time-compositing, individual subsamples are combined for even time increments. As an example, automatic samplers can be programmed to collect a subsample every 15 minutes for deposit into a large composite bottle. An automatic sampler can also collect discrete subsamples at even time increments, keeping each sample in a separate smaller sample bottle. After the sampled event, these samples can be manually combined as a composite. With flow-weighted sampling, an automatic sampler can be programmed to deposit a subsample into a large composite bottle for each set increment of flow.

The Wisconsin Department of Natural Resources conducted a through evaluation of alternative sampling modes for stormwater sampling to determine the average pollutant concentrations for individual events (Roa-Espinosa and Bannerman 1995). Four sampling modes were compared at outfalls at five industrial sites, including: flow-weighted composite sampling, time-discrete sampling, time-composite sampling, and first flush sampling during the first 30 minutes of runoff. Based on many attributes, they concluded that time-composite sampling at outfalls is the best method due to simplicity, low cost, and good comparisons to flow-weighted composite sampling (assumed to be the most accurate). The time-composite sampling cost was about 25% of the cost of the time- discrete and flow-weighted sampling schemes, but was about three times the cost of the first flush sampling only. The accuracy and reproducibility of the composite samples were all good, while these attributes for the first flush samples were poor. Burton and Pitt (2001) stress that it is important to ensure that acceptable time-weighted composite sampling include many subsamples. Any sampling scheme is very inaccurate if too few samples are collected. Samples need to be collected to represent the extreme conditions during the event, and the total storm duration. Experimental design methods can be used to determine the minimum number of subsamples needed considering likely variations. It is more common to now include the use of "continuous" water quality probes at sampling locations, with in-situ observations obtained every few minutes. Unfortunately, these details were not available for the NSQD sampling sites; some sites may have had too few subsamples to represent the storm conditions, while others may have had sufficient numbers of subsamples. Also, most of the NSQD samples only represented the first 3 hours of runoff events. If events were longer, the later storm periods were likely not represented. These issues are discussed more in the next subsection.

One-way ANOVA tests were used to evaluate the presence of significant differences between these two composite sampling schemes. Dunnet's comparison test was used to evaluate if concentrations associated with time-compositing were larger or lower than concentrations associated with flow- compositing. Tables 30 through 33 show the results of these tests.

Table 30. One-Way ANOVA Results by Sample Compositing Scheme in Residential Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none						0.229			***0***	
	n	median	Dunnet	n	median	Dunnet	n	Median	Dunnet	n	median	Dunnet
Flow composite				No composite			351	64.02		398	36.08	
Time composite							14	76.90	=	80	90.30	>

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.785			0.416			***0***			0.097	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	379	11.04		348	56.28		345	0.24		388	0.52	
Time composite	78	10.75	=	14	47.93	=	14	0.62	>	80	0.60	=

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		0.215			0.832			***0***			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	391	1.30		334	0.139		392	0.292		228	9.99	
Time composite	80	1.46	=	14	0.132	=	80	0.426	>	85	16.89	>

	Total Lead µg/L			Total Zinc µg/L		
p-value		***0***			***0***	
	n	median	Dunnet	n	median	Dunnet
Flow composite	222	5.94		227	50.77	
Time composite	85	19.62	>	85	142	>

Table 31. One-Way ANOVA Results by Sample Compositing Scheme in Commercial Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			Few samples			Few samples			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite										163	38.18	
Time composite										30	135.6	>

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.563			Few samples			Few samples			0.875	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	162	13.43								163	0.583	
Time composite	30	14.56	=							30	0.567	=

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		0.637			Few samples			0.118			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	163	1.47					161	0.242		115	14.91	
Time composite	30	1.36	=				30	0.194	=	30	36.34	>

	Total Lead µg/L			Total Zinc µg/L		
p-value		***0***			***0***	
	n	median	Dunnet	n	median	Dunnet
Flow composite	115	11.96		115	156	
Time composite	30	52.23	>	30	408	>

Table 32. One-Way ANOVA Results by Sample Compositing Scheme in Industrial Land Use, Rain Zone 2

	Hardness mg/L			**Oil and Grease mg/L**			**TDS mg/L**			**TSS mg/L**		
p-value		none			Few samples			0.076			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite							101	68.5		116	44.2	
Time composite							9	132.9	=	40	84.6	>

	BOD_5 mg/L			**COD mg/L**			**Ammonia mg/L**			**$NO_2 + NO_3$ mg/L**		
p-value		0.861			0.519			***0***			0.488	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	112	9.67		100	53.93		96	0.25		109	0.547	
Time composite	39	9.94	=	9	63.04	=	9	1.11	>	39	0.614	=

	TKN mg/L			**Dissolved Phosphorus mg/L**			**Total Phosphorus mg/L**			**Total Copper μg/L**		
p-value		0.672			0.601			0.338			0.070	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	109	1.06		82	0.087		111	0.208		72	15.75	
Time composite	40	1.13	=	9	0.074	=	40	0.242	=	40	21.27	=

	Total Lead μg/L			**Total Zinc μg/L**		
p-value		***0.001***			0.930	
	n	median	Dunnet	n	median	Dunnet
Flow composite	66	9.34		72	189.7	
Time composite	40	22.23	>	40	186.8	=

Table 33. One-Way ANOVA Results by Sample Compositing Scheme in Industrial Land Use, Rain Zone 3

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			none			***0.012***			0.103	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite							16	93.93		16	22.44	
Time composite							20	43.11	<	20	52.38	=

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.265			0.088			none			***0.006***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Flow composite	16	5.32		16	33.38					9	0.248	
Time composite	20	7.25	=	20	50.36	=				19	0.844	>

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		***0.002***			***0***			0.556			0.143	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	N	median	Dunnet
Flow composite	15	0.64		11	0.035		15	0.161		15	12.56	
Time composite	20	1.63	>	20	0.116	>	20	0.189	=	20	18.19	=

	Total Lead µg/L			Total Zinc µg/L		
p-value		0.279			0.163	
	n	median	Dunnet	n	median	Dunnet
Flow composite	10	9.75		15	108.9	
Time composite	19	6.08	=	20	161.4	=

Note. Dunnet's test compared if sites with control produces larger concentrations ">", smaller concentrations "<" or not statistically difference "=" than sites without control at a family error of 5%.
"None" indicates no samples were collected for this constituent in the group.

Tables 30 through 33 show that no significant differences were observed for BOD_5 concentrations using either of the compositing schemes for any of the four categories. A similar result was observed for COD except for commercial land uses in EPA Rain Zone 2, where not enough samples were collected to detect a significant difference. TSS and total lead median concentrations in EPA Rain Zone 2 were two to five times higher in concentration when time-compositing was used instead of flow-compositing.

Nutrients in EPA Rain Zone 2 collected in residential, commercial and industrial areas showed no significant differences using either compositing method. The only exceptions were for ammonia in residential and commercial land use areas and total phosphorus in residential areas where time-composite samples had higher concentrations. Metals were higher when time-compositing was used in residential and commercial land use areas. No differences were observed in industrial land use areas, except for lead. Figure 38 shows box and whiskers plots for TSS using both methods.

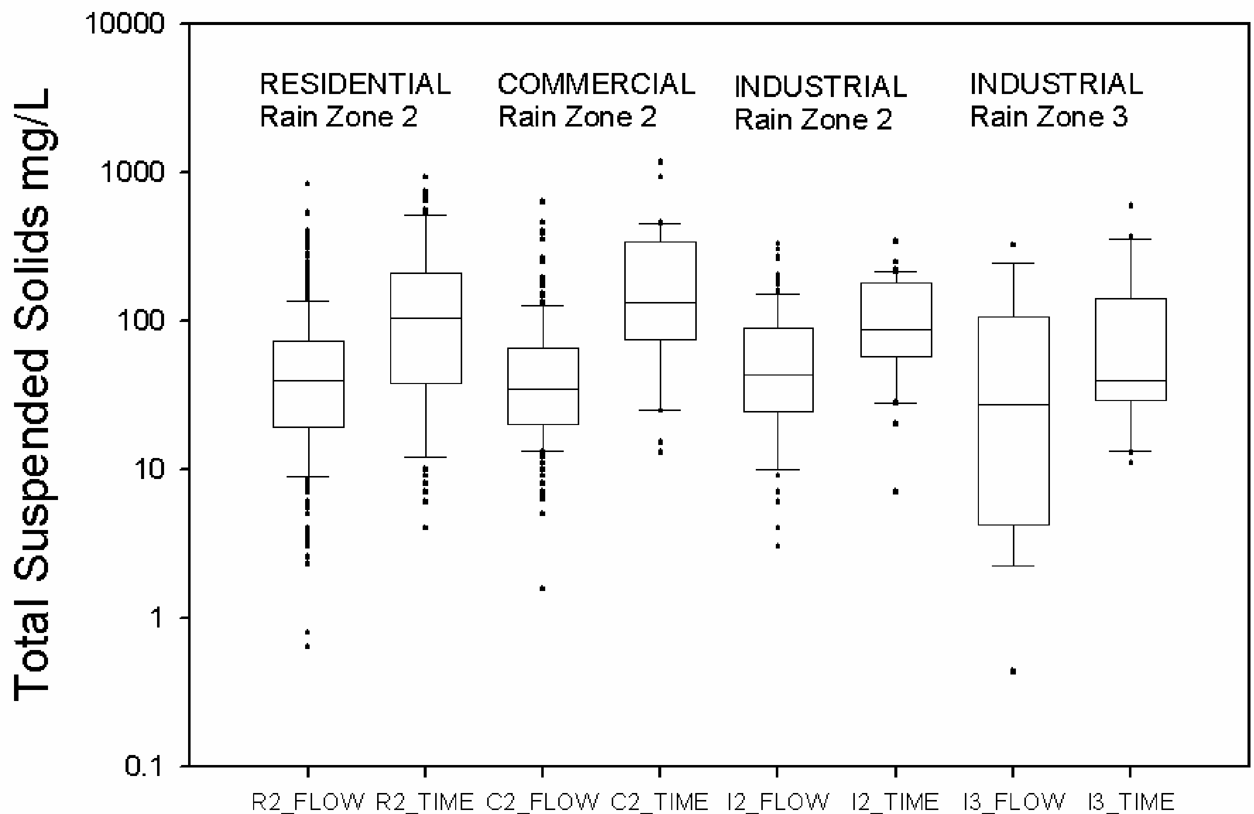


Figure 38. Comparisons between time- and flow-composite options for TSS

Sampling Period during Runoff Event and Selection of Events to Sample

Another potential factor that may affect stormwater quality is the sampling period during the runoff event. Automatic samplers can initiate sampling very close to the beginning of flow, while manual sampling usually requires travel time and other delays before sampling can be started. It is also possible for automatic samplers to represent the complete storm, especially if the storm is of long duration, as long as proper sampler setup programming is performed (Burton and Pitt 2001). However, automatic samplers are not capable of sampling bed load material, and are less effective in sampling larger particles (>500 µm). Manual sampling, if able to collect a sample from a cascading flow, can collect from the complete particle size distribution. Bed load samples and special floatable capture nets may be needed to supplement automatic samplers to obtain information for the complete range of solids.

The NPDES stormwater sampling protocols only required collecting composite samples over the first three hours of the event instead of during the whole event. Truncating the sampling before the runoff event ended may have adversely affected the measured stormwater quality.

Selecting a small subset of the annual events can also bias the monitoring results. In most stormwater research projects, the goal is to sample and analyze as many events as possible during the monitoring period. As a minimum, about 30 samples are usually desired in order to adequately determine the stormwater characteristics with an error level of about 25% (assuming 95% confidence and 80% power) (Burton and Pitt 2001). With only three events per year required per land use for the NPDES stormwater permits, the accuracy of the calculated EMC is questionable until many years have passed. Also, the three storms need to be randomly selected from the complete set of rains in order to be most statistically representative, not just for a narrow range of rain depths as specified in the NPDES sampling protocol.

Flagstaff Street, in Prince George MD, had the most events collected for any site in the NSQD. They collected 28 events during two years of sampling (1998 and 1999). A statistical test was made choosing 6 events (three for each year) from this set, creating 5,600 different possibilities. Figure 39 shows the histogram of these possibilities. The median TSS of the 28 events was 170 mg/L, with a 95% confidence interval between 119 and 232 mg/L. Only 60% of the 5,600 possibilities were inside this confidence interval. Almost half (40%) of the possibilities for the observed EMC would therefore be outside the 95% confidence interval for the true median concentration if only three events were available for two years. As the number of samples increase, there will be a reduction in the bias of the EMC estimates. In Southern California, Leecaster (2002) determined that ten years of collecting three samples per year was required in order to reduce the error to 10%.

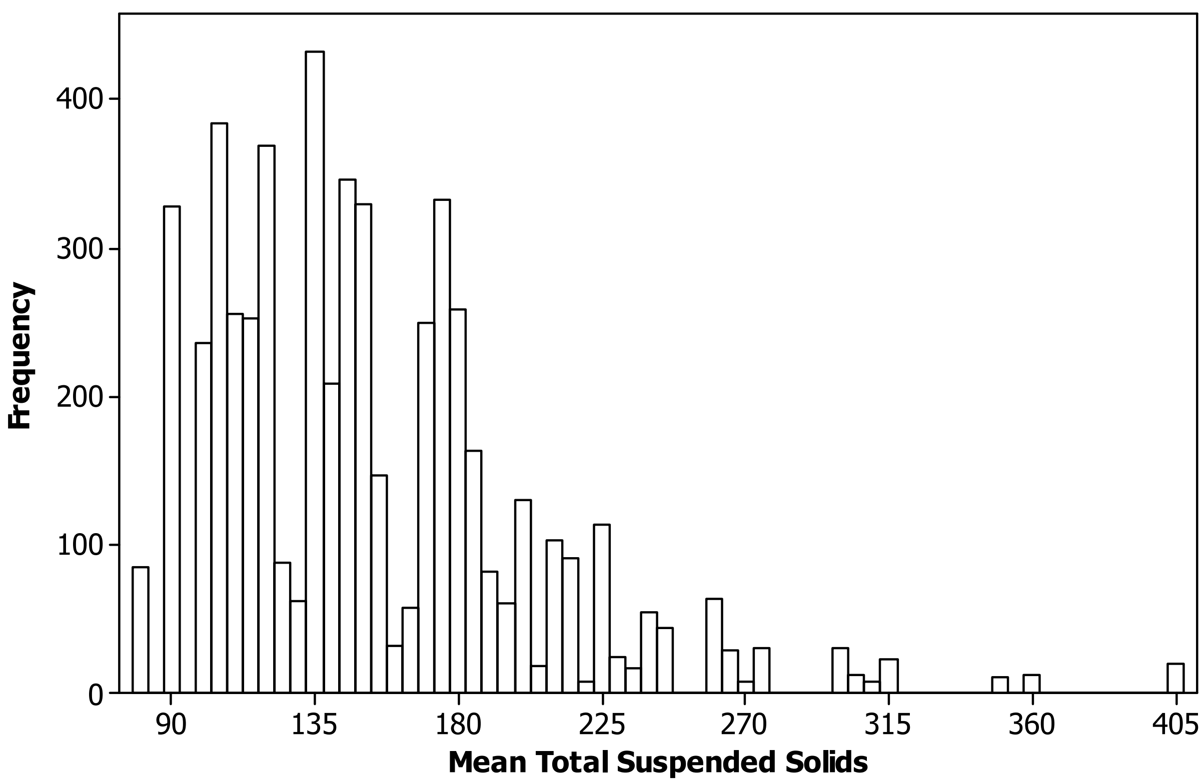


Figure 39. Histogram of possible TSS concentrations in Flagstaff Street based on collecting three samples per year for two years (the measured median TSS concentration was 170 mg/L)

Type of Conveyance

Almost all of the samples in the NSQD were collected using automatic samplers and flow compositing. Statistical tests investigating the effects of the type of conveyance only used information from flow-weighted composite samples to reduce potential errors associated with other sampling schemes, as discussed above. Grass swales are considered to be effective stormwater controls compared to conventional curb and gutter stormwater collection systems. Grass swales are commonly found in residential areas with low levels of imperviousness, especially in low density residential areas. NSQD data from residential and mixed residential sites in Virginia, Georgia, and Texas were used to compare stormwater concentrations in areas drained by grass swales and by concrete curbs and gutters.

Historical swale performance tests usually focused on pollutant mass discharges and not concentrations. Swales normally infiltrate significant amounts of the flowing water, resulting in large mass discharge decreases. Most swales operate with relatively deep water, and any “filtering” benefits of the grass (and hence concentration reductions) are usually minimal. Very shallow flows in swales do have particulate pollutant concentration reductions, but these are rarely observed during moderate to large flows (Nara and Pitt 2005).

One-way ANOVA analyses were used to identify any significant differences in stormwater pollutant concentrations between watersheds drained with grass swales or with curbs and gutters. Dunnett’s test was used to determine if grass swales produced different concentrations than curbs and gutters. The results are shown in Tables 34 through 37.

Table 34. One-Way ANOVA Results by Type of Conveyance in Residential Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			0.824			none			none	
	n	median	Dunnet	n	median	Dunnet	n	Median	Dunnet	n	median	Dunnet
Curb and Gutter				59	3.11							
Grass Swale				7	2.95	=						

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		none			none			none			none	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter												
Grass Swale												

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		none			none			none			***0***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter										82	10.67	
Grass Swale										7	3.11	<

	Total Lead µg/L			Total Zinc µg/L		
p-value		0.112			***0.002***	
	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	77	11.7		82	59.46	
Grass Swale	7	5.67	=	7	17.85	<

Table 35. One-Way ANOVA Results by Type of Conveyance in Industrial Land Use, Rain Zone 2

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			none			***0***			***0.023***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter							67	45.5		69	37.52	
Grass Swale							77	184	>	7	97.70	>

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		***0***			***0.035***			0.492			none	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	67	6.84		66	50.16		61	0.223				
Grass Swale	5	39.98	>	7	85.64	>	7	0.285	=			

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		none			***0.012***			0.468			0.905	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter				50	0.07		64	0.174		20	13.0	
Grass Swale				4	0.23	>	7	0.232	=	7	12.36	=

	Total Lead µg/L			Total Zinc µg/L		
p-value		none			0.447	
	n	median	Dunnet	n	median	Dunnet
Curb and Gutter				20	225.7	
Grass Swale				7	188.4	=

Table 36. One-Way ANOVA Results by Type of Conveyance in Residential Land Uses, Rain Zone 3

	Hardness mg/L			Oil and Grease mg/L			TDS mg/L			TSS mg/L		
p-value		none			none			***0.049***			0.425	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter							11	94.06		12	19.2	
Grass Swale							6	47.84	<	6	29.6	=

	BOD_5 mg/L			COD mg/L			Ammonia mg/L			$NO_2 + NO_3$ mg/L		
p-value		0.749			***0.027***			none			none	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	11	7.56		11	29.36							
Grass Swale	5	6.63	=	5	67.27	>						

	TKN mg/L			Dissolved Phosphorus mg/L			Total Phosphorus mg/L			Total Copper µg/L		
p-value		0.17			0.324			0.319			0.007	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	11	1.22		8	0.07		12	0.22		11	19	
Grass Swale	6	0.94	<	6	0.04	=	6	0.14	=	6	5	<

	Total Lead µg/L			Total Zinc µg/L		
p-value		0.154			0.781	
	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	9	12.9		11	49.5	
Grass Swale	6	4.20	=	6	43.0	=

Table 37. One-Way ANOVA Results by Type of Conveyance in Industrial Land Use, Rain Zone 3

	Hardness mg/L			**Oil and Grease mg/L**			**TDS mg/L**			**TSS mg/L**		
p-value		none			none			0.134			***0.014***	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter							10	76.74		10	9.68	
Grass Swale							6	131.6	=	6	91.2	>

	BOD_5 mg/L			**COD mg/L**			**Ammonia mg/L**			**$NO_2 + NO_3$ mg/L**		
p-value		0.461			0.446			none			none	
	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	10	4.68		10	29.40							
Grass Swale	6	6.61	=	6	41.26	=						

	TKN mg/L			**Dissolved Phosphorus mg/L**			**Total Phosphorus mg/L**			**Total Copper μg/L**		
p-value		0.299			0.077			0.460			0.098	
	n	median	Dunnet	n	median	Dunnet	N	median	Dunnet	N	median	Dunnet
Curb and Gutter	9	0.515		5	0.046		9	0.138		9	8.57	
Grass Swale	6	0.885	=	6	0.027	=	6	0.202	=	6	22.32	=

	Total Lead μg/L			**Total Zinc μg/L**		
p-value		0.157			***0.007***	
	n	median	Dunnet	n	median	Dunnet
Curb and Gutter	4	4.86		9	72.86	
Grass Swale	6	15.5	=	6	198.9	>

Note. Dunnet's test determined if sites with controls produce larger concentrations ">", smaller concentrations "<" or not statistically difference "=" than sites without control at a family error of 5%.
"None" indicates no samples were collected for this constituent in the group.

Total lead and total phosphorus did not have any significant differences in concentrations when comparing the two conveyance systems in both land use areas. Total copper concentrations from residential land uses in EPA Rain Zones 2 and 3 were lower when grass swales were used instead of curbs and gutters. No copper concentration differences were observed at industrial land uses having different conveyance systems.

Figure 40 shows box and whiskers plots for TSS in industrial land uses, EPA Rain Zones 2 and 3 and residential areas in EPA Rain Zone 2. The median concentrations in industrial land uses were smaller in locations where curbs and gutters were used compared to sites having grass swales. The statistical tests did not identify a significant difference between the median concentrations in residential areas in EPA Rain Zone 3 (the residential boxes have much more overlap than for the industrial sites).

Concentration Effects Associated with Varying Amounts of Impervious Cover

The reported values for imperviousness do not reflect the amount of pavement and roofs that are not directly connected to the drainage system. Directly connected impervious areas (DCIA) are also referred to as effective impervious areas (EIA). For example, imagine a park with a single paved basketball court surrounded by turf; the area of the court will be counted as part of the total impervious area, but would not be considered as part of the effective impervious area. The runoff from the paved court would likely be totally infiltrated by the grass and will not be discharged to the drainage system. In this case, even if we have a value for "total imperviousness," the "effective percentage of imperviousness" is zero.

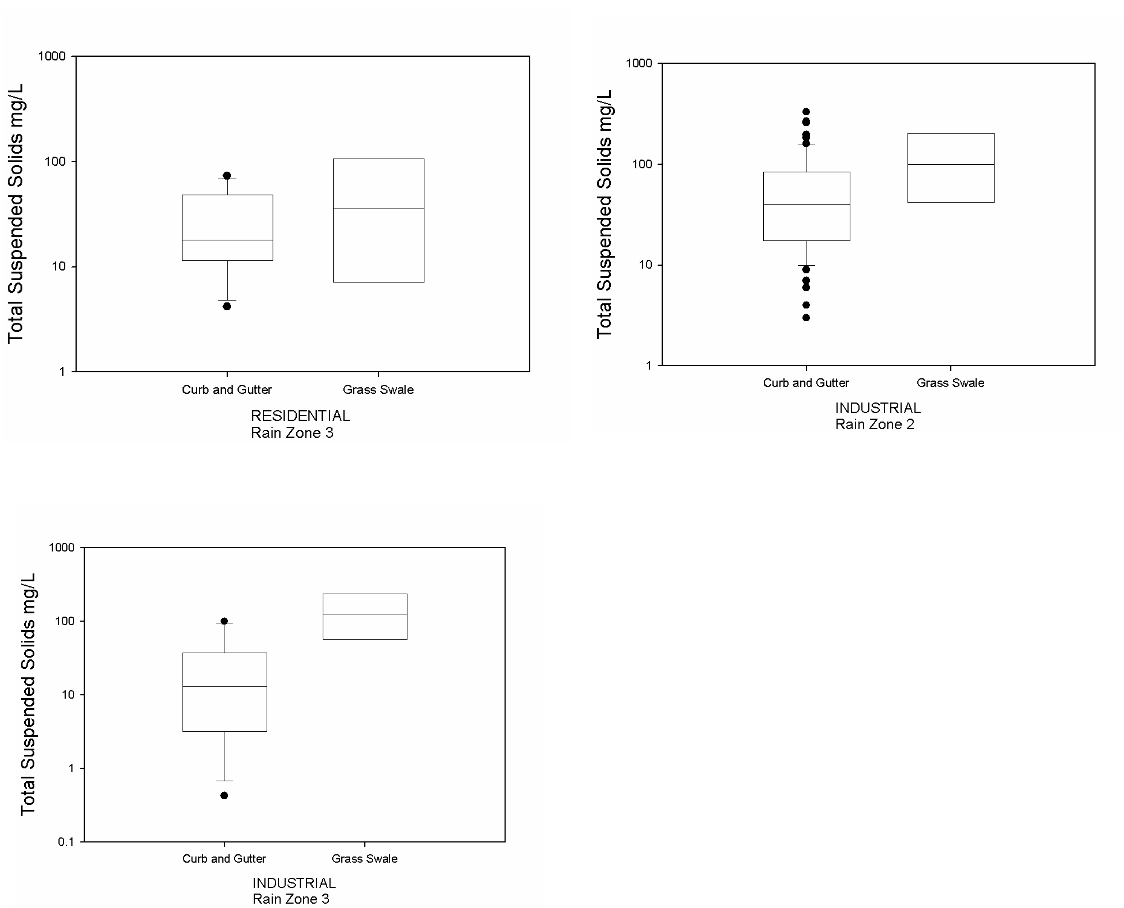


Figure 40. TSS concentration by type of conveyance (Significant differences were observed in industrial land uses)

It is therefore difficult to compare database concentrations with the imperviousness values due to these potential uncertainties in the actual effective imperviousness. Figure 41 is an example plot of the percent imperviousness values of different land uses for COD. Each vertical set of observations represent a single monitoring location (all of the events at a single location have the same percent imperviousness). The variation of COD at any one monitoring location is seen to vary greatly, typically by about an order of magnitude. These large variations will make trends difficult to identify. All of the lowest percentage imperviousness sites are open space land uses, while all of the highest percentage imperiousness sites are freeway and commercial land uses. This plot shows no apparent trend in concentration that can be explained by imperviousness. However, it is very likely that a significant and important trend does exist between percent effective imperviousness and pollutant mass that is discharged. While the relationship between imperviousness and concentration is not clear, the relationship between effective imperviousness and total runoff volume is much stronger and more obvious as the non-paved areas can infiltrate much water.

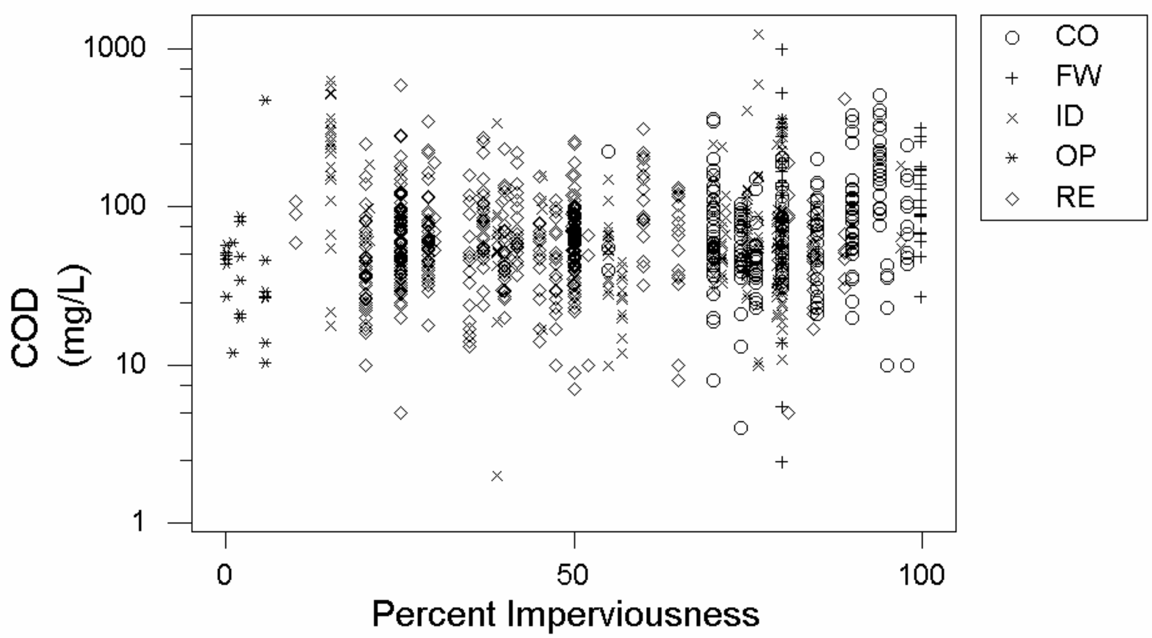


Figure 41. Plot of COD concentrations against watershed area percent imperviousness values for different land uses (CO: commercial; FW: freeway; ID: industrial; OP: open space; and RE: residential)

One important feature in the percentage of imperviousness is that most of the residential sites have low levels of imperviousness, while commercial and industrial sites usually have high percentages of imperviousness. Figure 42 shows the mean TSS concentration for residential, commercial and industrial land uses in the database. Only four of the monitored residential watersheds have percentage imperviousness values larger than 60%. Two commercial sites have less than 60% imperviousness, with the remaining commercial sites above this value. Analyses concerning the effects of impervious cover on stormwater concentrations for each land use separately are difficult as there are limited ranges of impervious cover within each land use category.

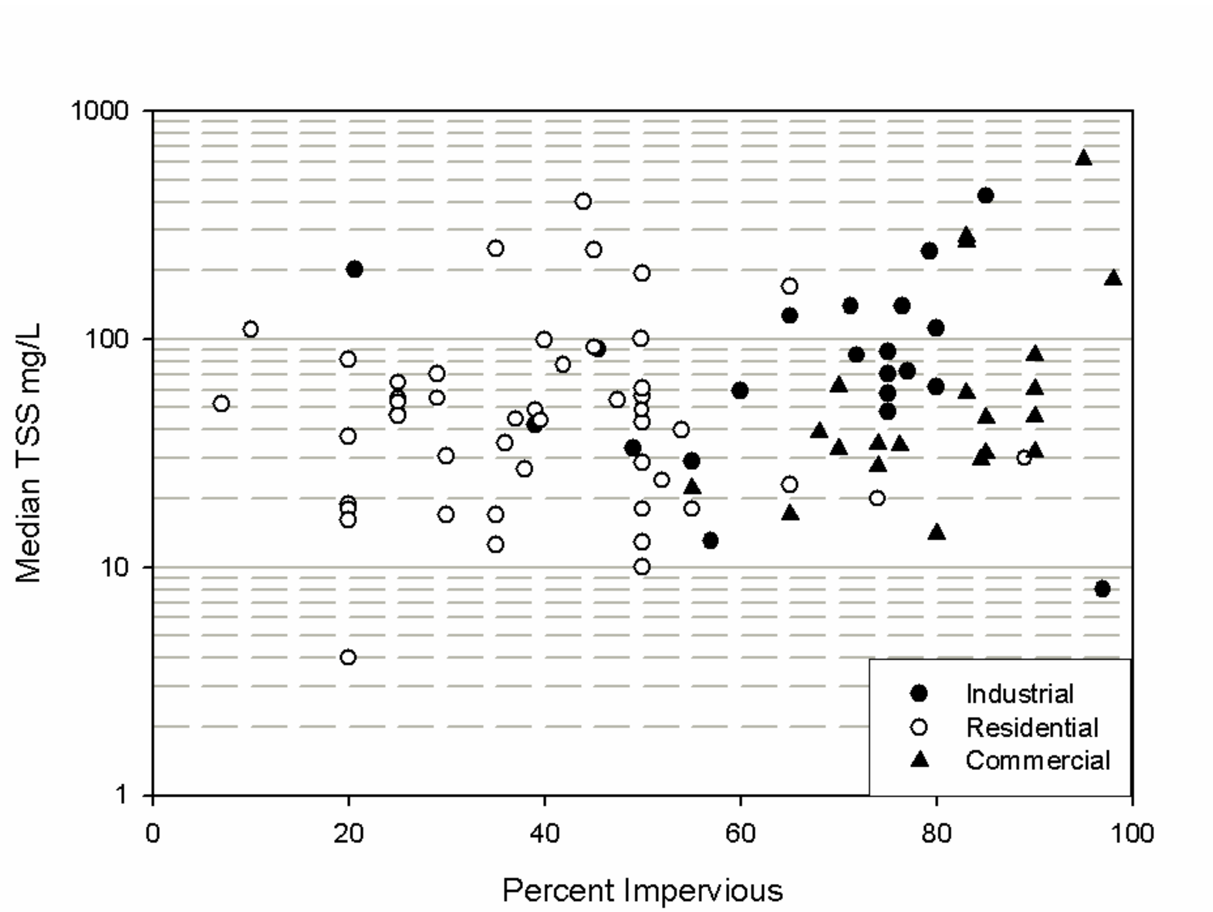


Figure 42. TSS concentrations by impervious cover and single land use

Regression analyses were used to identify possible relationships between constituent concentrations and the percentage of imperviousness for residential land use data. Table 38 shows the results from these regression analyses. Residential land uses in EPA Rain Zone 2 were examined during these analyses. Median concentrations from sites using automatic, flow-weighted samplers, and not having any controls and with curb and gutter conveyance systems were selected for analyses. Data from the site KYLOTSR3 were not used during these analyses because sewage disposal facilities were located in the test watershed. Solids and heavy metal median concentrations were higher at this location than for the remaining residential sites in the same Rain Zone.

Only nitrate-nitrite indicated a significant regression relationship between percentage of imperviousness and constituent concentration for these sites, as shown in Figure 43. In this case, the slope was negative, indicating a reduction in the concentration as the level of imperviousness increased. One possible explanation is that the nutrients are associated with landscaped areas and the use of fertilizers which all decrease with increasing impervious areas. This does not indicate that the total mass of nitrate-nitrite will be reduced. The load of this constituent depends on the total runoff volume that is discharged during the event. As the percentage of imperviousness increases, the runoff volume also increases due to lack of infiltration. Even if the concentration is shown to decrease, the total mass discharged may still increase with increasing amounts of pavement or roofs. There was not enough evidence to indicate a relationship between concentration and percentage of imperviousness for the other 11 constituents examined.

Table 38. Regression of Median Concentrations by Percentage of Impervious in Residential land Use, EPA Rain Zone 2

		Constant		Impervious			
Constituent	**n**	**Coefficient**	**p-value**	**Coefficient**	**p-value**	**R^2 adjusted**	**Significant at 0.05 level?**
TDS mg/L	10	71.94	0.002	-0.386	0.446	0	Not significant
TSS mg/L	10	74.44	0.002	-0.715	0.172	0.121	Not significant
BOD_5 mg/L	10	8.74	0.117	0.076	0.619	0	Not significant
COD mg/L	10	53.94	0.027	0.332	0.578	0	Not significant
Ammonia mg/L	10	0.319	0.052	-0.002	0.639	0	Not significant
NO_3-NO_2 mg/L	9	0.756	0	-0.009	***0.013***	0.556	Significant
TKN mg/L	9	1.817	0.003	-0.016	0.247	0.069	Not significant
DP mg/L	10	0.237	0.033	-0.003	0.349	0	Not significant
TP mg/L	10	0.561	0.002	-0.006	0.13	0.171	Not significant
Cu □g/L	11	16.51	0.005	-0.140	0.225	0.065	Not significant
Pb □g/L	11	46.64	0.336	-0.337	0.767	0	Not significant
Zn µg/L	11	98.13	0.027	-0.572	0.542	0	Not significant

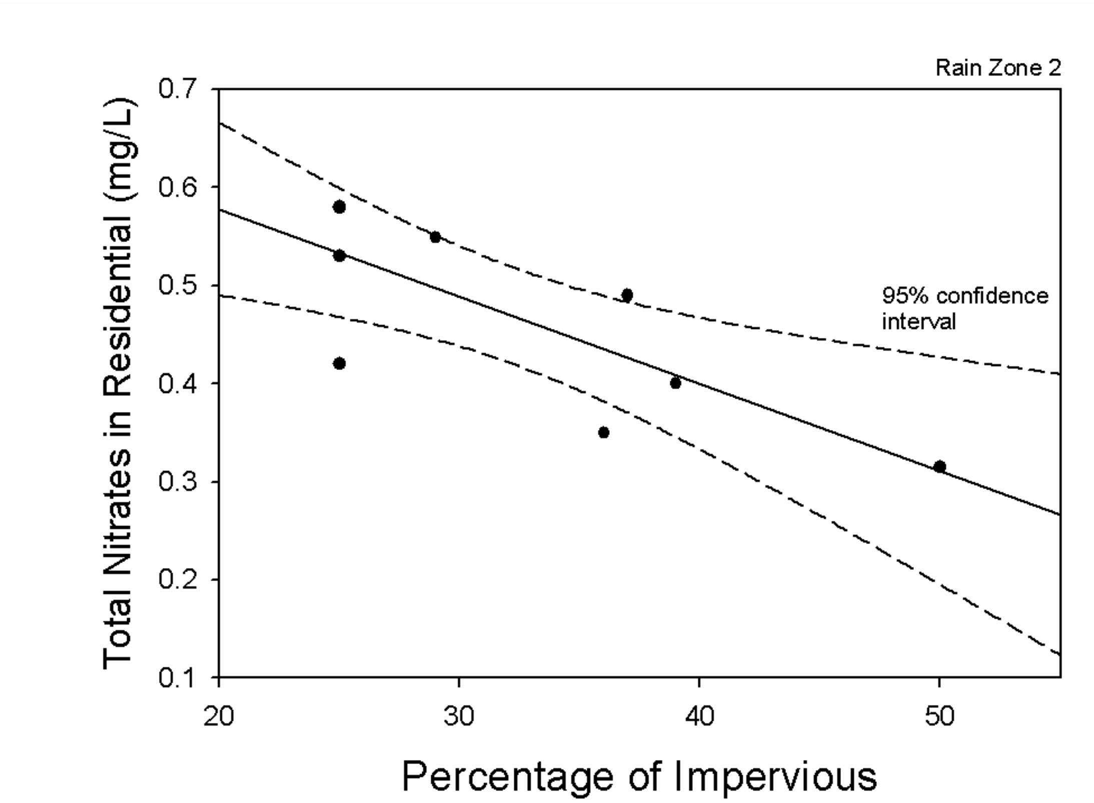


Figure 43. Total nitrates regression at different percentages of impervious

The same regression analysis was performed for commercial and industrial land uses in EPA Rain Zone 2. The results of the regression analyses are shown in Table 39.

Table 39. Regression of Median Concentrations by Percentage of Impervious in Commercial and Industrial land use, EPA Rain Zone 2

		Constant		Impervious			
Constituent	n	Coefficient	p-value	Coefficient	p-value	R^2 adjusted	Significant at 0.05 level?
TDS mg/L	5	-4.80	0.854	0.821	0.103	0.523	Not significant
TSS mg/L	5	-22.01	0.406	0.805	0.097	0.541	Not significant
BOD_5 mg/L	5	-1.80	0.879	0.153	0.410	0	Not significant
COD mg/L	5	1.41	0.968	0.748	0.215	0.268	Not significant
Ammonia mg/L	5	-0.05	0.906	0.005	0.439	0	Not significant
NO_3-NO_2 mg/L	5	0.01	0.985	0.007	0.438	0	Not significant
TKN mg/L	5	-0.84	0.467	0.030	0.140	0.426	Not significant
DP mg/L	5	-0.02	0.858	0.001	0.516	0	Not significant
TP mg/L	5	-0.10	0.649	0.004	0.271	0.168	Not significant
Cu μg/L	5	4.26	0.759	0.089	0.679	0	Not significant
Pb μg/L	6	15.69	0585	-0.021	0.961	0	Not significant
Zn μg/L	6	247.9	0.269	-0.949	0.765	0	Not significant

None of the median stormwater constituents in commercial and industrial areas seem to be affected by changes in impervious cover. There is not enough evidence to indicate a significant relationship between constituent concentration and percentage of imperviousness. More samples will be required to identify those regression relationships.

Seasonal Effects on Stormwater Quality

Another factor that may affect stormwater quality is the season when the sample was obtained. If the few samples collected for a single site were all collected in the same season, the results may not be representative of the whole year. The NPDES sampling protocols were designed to minimize this effect by requiring the three samples per year to be separated by at least 1 month. The few samples still could be collected within a single season, but at least not within the same week. Seasonal variations for residential stormwater data are shown in Figure 44. These variations are not as obvious as the land use or geographical variations, except for bacteria which appear to be lowest during the winter season and highest during the summer and fall (a similar conclusion was obtained during the NURP, EPA 1983, data evaluations). The database does not contain any snowmelt data, so all of the data corresponds to rain-related runoff only.

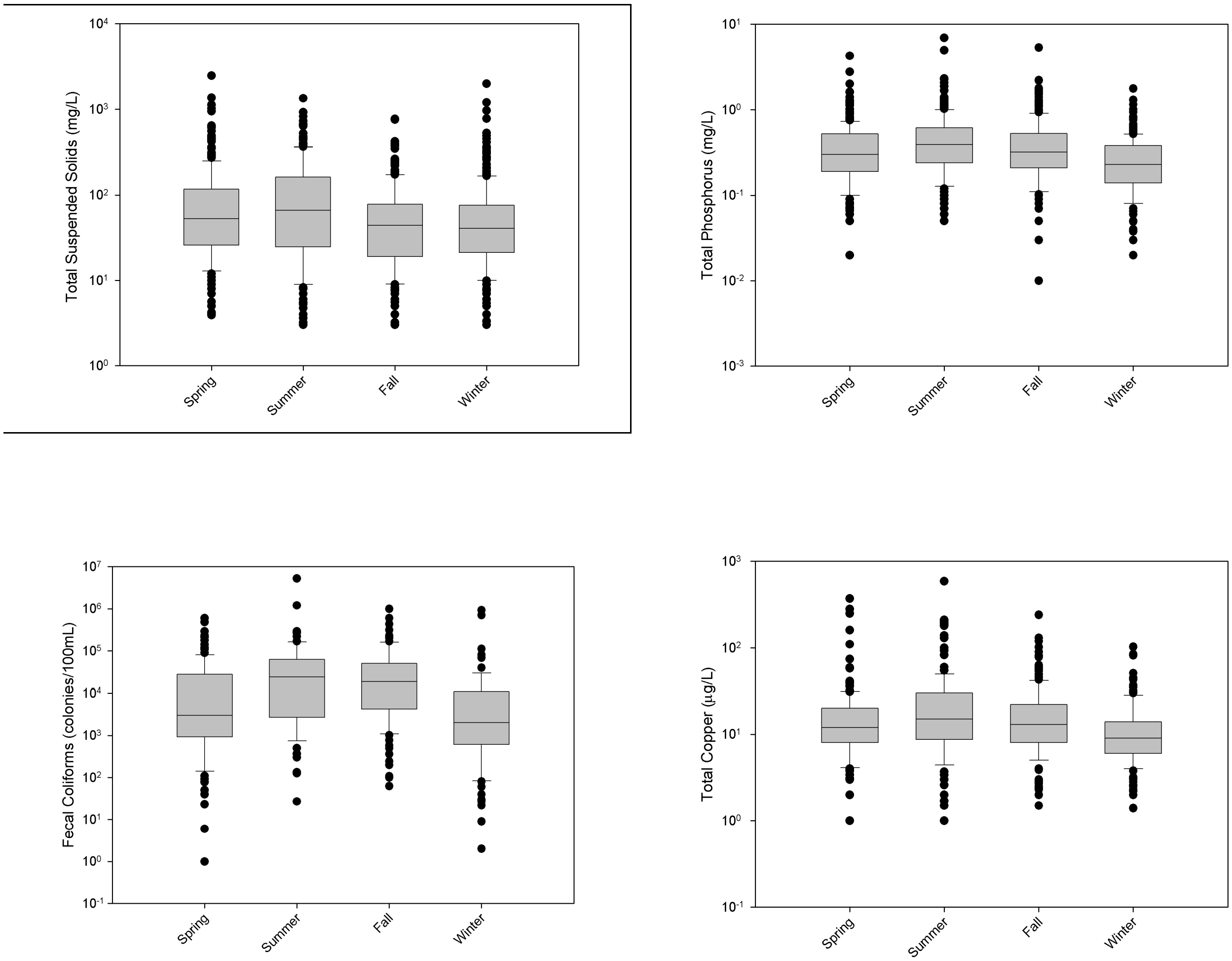


Figure 44. Example residential area stormwater pollutant concentrations sorted by season

Precipitation Depth Effects on Stormwater Quality

A common assumption is that higher runoff concentrations are associated with smaller rain events. While this has been shown to be true during controlled washoff studies (Pitt 1987), or for sheetflows taken from relatively small paved areas during rains (see Chapter 6 discussion about first flush observations), this has not been frequently detected for samples collected at outfalls for areas having a mixture of surfaces and for typical random periods of high rain intensities. Figure 45 contains several scatter plots showing concentrations plotted against rain depth. There are no obvious trends of concentration associated with rain depth for the NSQD data.

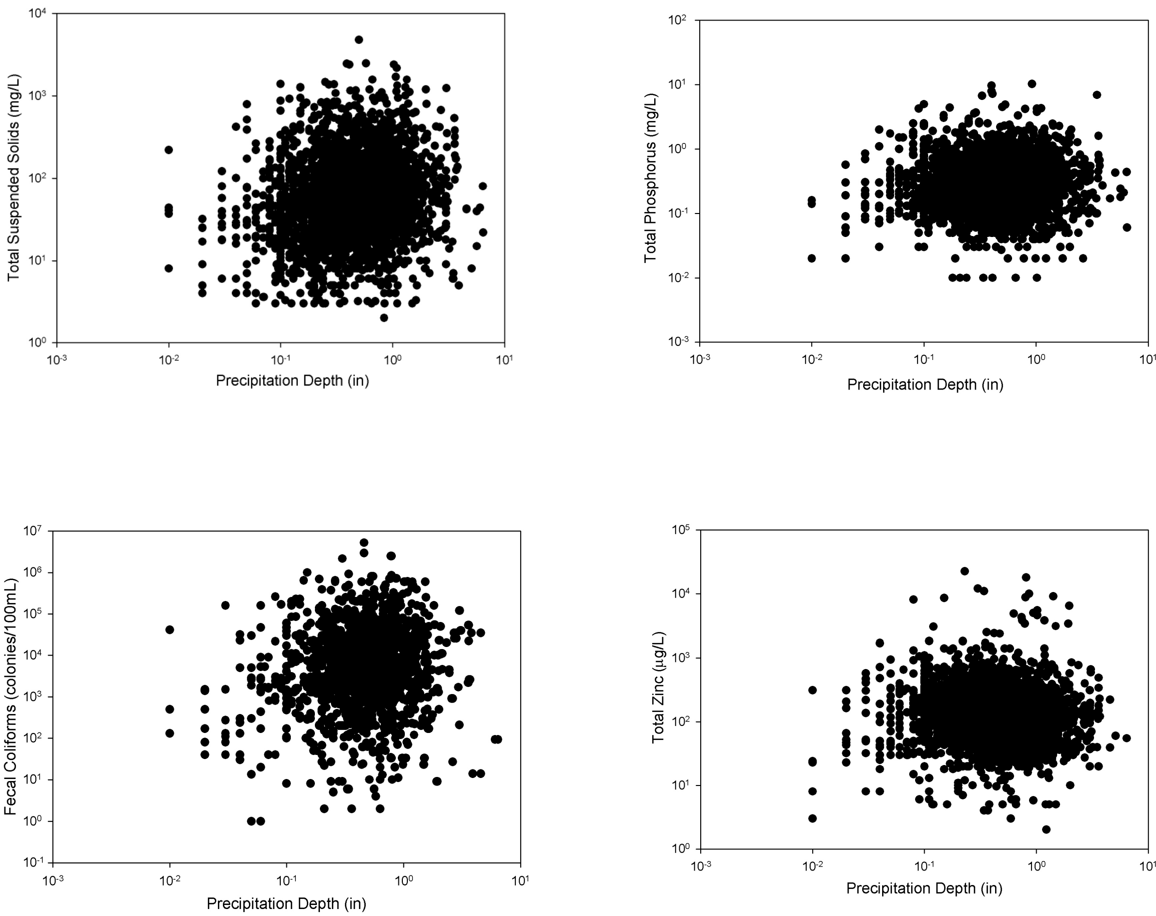


Figure 45. Examples of scatter plots by precipitation depth

Figure 46 shows scatter plots of rainfall and runoff depth for each land use. These should follow a 45 degree line for areas having very large amounts of directly connected impervious areas.

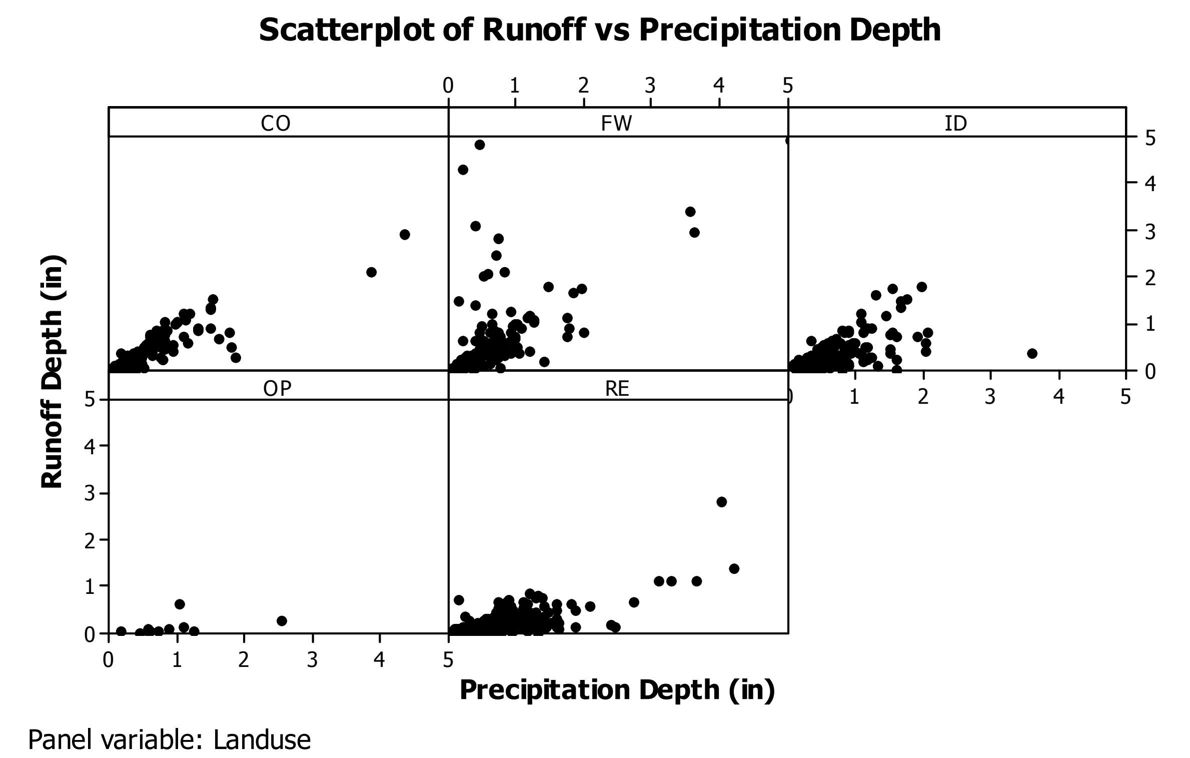


Figure 46. Precipitation depth and runoff depth plotted by land use

These plots show much greater scatter than expected. The freeway plot even indicated larger amounts of runoff than precipitation. This may have occurred due to several reasons: (1) the rainfall was not representative of the drainage area being monitored (especially possible for those sites that relied on off-site rain data); (2) the runoff monitoring was inaccurate (possible when the runoff monitoring relied on stage recording devices and the Manning's equation was applied without local calibration); (3) the drainage area was inaccurately delineated; or (4) when base flows contributed significant amounts of runoff during the event. When reviewing the runoff plots provided in some of the annual reports, significant base flows were observed. It was also apparent that these base flows were not subtracted from the total flows recorded during the rain event. The magnitude of the error would be greater for smaller rain events when the base flows could be much larger than the direct runoff quantity. Base flows commonly occur when a local spring or high groundwater levels enter the storm drainage system. In addition, runoff may still be occurring from a prior large event that ended soon before the current event started (the 3 day antecedent dry period requirement for monitored events was intended to minimize this last cause of base flows).

Antecedent Period without Rain before Monitored Event

The EPA Rain Zones with the longest reported dry interevent periods having data in the NSQD are EPA Rain Zones 6 (southern California) and 7 (Oregon). In these EPA Rain Zones, some antecedent dry periods were reported to be longer than 100 days. Monitored events with the shortest interevent periods of no rains were monitored along the east and south east coasts of the country (EPA Rain Zones 2 and 3). The mean interevent dry period in the western states was about 18 days, while eastern states had mean interevent dry periods of about 5 days. Figure 47 shows box and whisker plots of the number of days having no rain before the monitored event by each EPA Rain Zone.

Samples collected using automatic flow-weighted samplers from watersheds having curbs and gutters and without stormwater controls were used during the following analyses. Only EPA Rain Zone 2 has enough observations to evaluate possible effects of the antecedent dry period on the concentration of stormwater pollutants. Table 40 shows the results from the regression analyses. In residential land uses, 7 out of 12 constituents indicated that antecedent dry period had a significant effect on the median concentrations. All the regression slope coefficients were positive, indicating that as the number of days having no rain increased, the concentrations also increased.

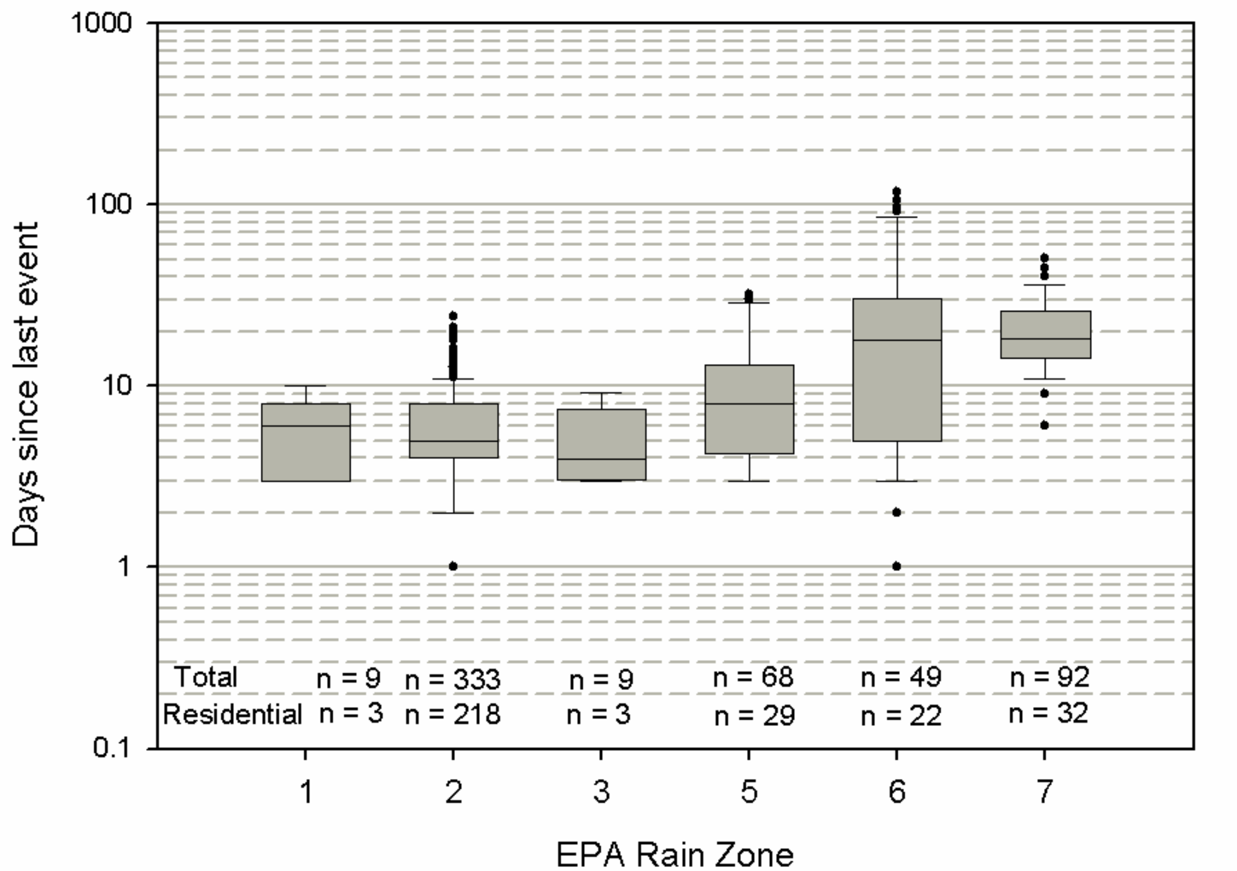


Figure 47. Box and whisker plot of dry days preceding rain event by EPA Rain Zone

Table 40. Regression of Logarithm of Constituent Concentrations by Logarithm of Antecedent Dry Period for Residential Land Use, EPA Rain Zone 2

		Constant		Days since last event			
Constituent	n	Coefficient	p-value	Coefficient	p-value	R^2 adjusted	Significant at 0.05 level?
Oil - Grease mg/L	35	0.737	0	-0.364	0.062	0.074	No
TDS mg/L	208	1.761	0	0.094	0.120	0.007	No
TSS mg/L	214	1.524	0	0.116	0.254	0.001	No
BOD_5 mg/L	211	0.887	0	0.211	***0.004***	0.035	Yes
COD mg/L	206	1.682	0	0.151	***0.032***	0.018	Yes
Ammonia mg/L	204	-0.826	0	0.300	***0.003***	0.039	Yes
NO_3-NO_2 mg/L	208	-0.428	0	0.160	***0.014***	0.024	Yes
TKN mg/L	208	-0.066	0.193	0.232	***0.001***	0.049	Yes
DP mg/L	203	-1.061	0	0.282	***0.002***	0.043	Yes
TP mg/L	214	-0.629	0	0.183	***0.005***	0.031	Yes
Cu µg/L	58	1.082	0	0.025	0.830	0	No
Pb µg/L	53	1.305	0	-0.311	0.277	0.004	No
Zn µg/L	58	1.872	0	-0.058	0.764	0	No

All nutrients (plus organic matter) in residential land uses showed a positive correlation between days since last event and constituent concentration. In all cases, the coefficients of determination (R^2) were smaller than 0.05, indicating that relatively little of the total variation was explained by antecedent dry period. Solids and metals were not affected by the antecedent dry period. Figure 48 shows the regression lines and 95% confidence intervals for four nutrients in residential land uses.

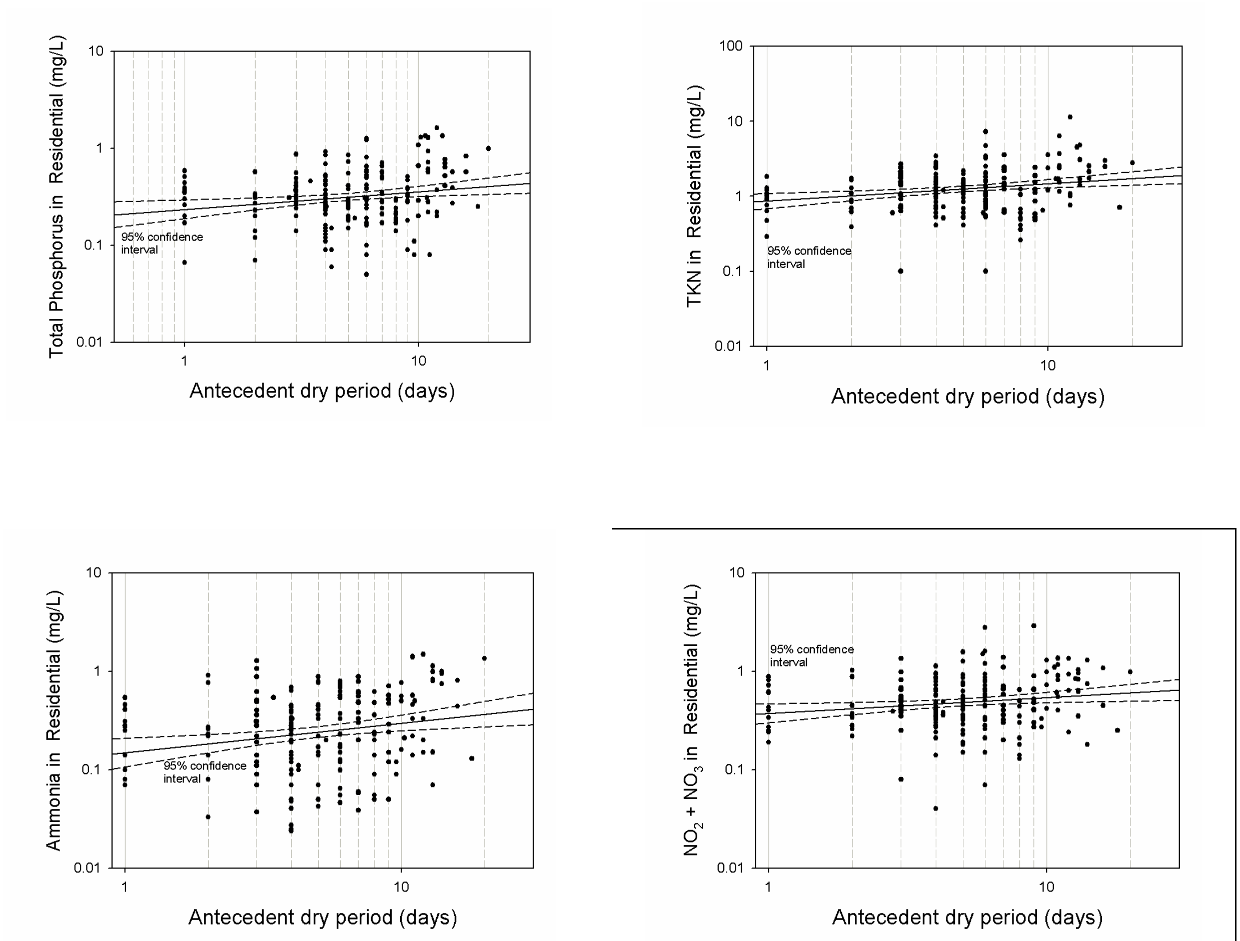


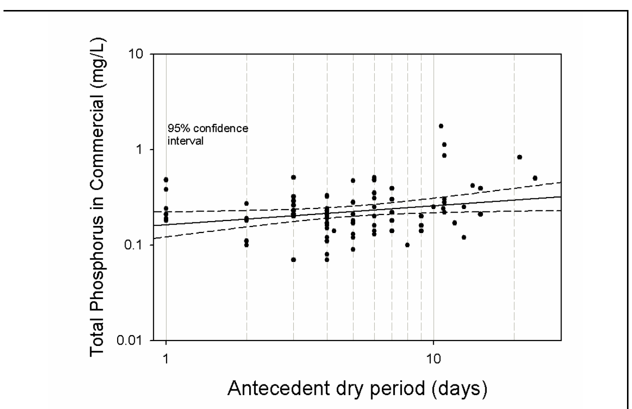
Figure 48. Nutrient concentrations affected by dry periods since last rain in residential land use

Table 41 shows the results from the regression analyses in commercial land uses. Except for nitrates, all the nutrients have positive regressions inside the 95% confidence interval. In commercial land uses, the effects of antecedent dry periods on the median concentrations were less important. Only total phosphorus and total lead had significant regression results. As in the residential case, phosphorus has a positive coefficient with a small coefficient of determination. However, lead decreases with the number of dry days before the storm.

Table 41. Regression of Logarithm of Constituent Concentrations by Logarithm of Antecedent Dry Period for Commercial Land Use, EPA Rain Zone 2

		Constant		Impervious			
Constituent	n	Coefficient	p-value	Coefficient	p-value	R^2 adjusted	Significant at 0.05 level?
Oil - Grease mg/L	25	0.783	0.001	-0.202	0.402	0	No
TDS mg/L	64	1.715	0	0.215	0.169	0.015	No
TSS mg/L	82	1.506	0	0.018	0.872	0	No
BOD_5 mg/L	83	0.971	0	0.149	0.176	0.01	No
COD mg/L	64	1.670	0	0.221	0.093	0.029	No
Ammonia mg/L	64	-0.591	0	0.258	0.175	0.014	No
NO2 mg/L	83	-0.235	0	-0.208	0.176	0.01	No
TKN mg/L	83	-0.006	0.949	0.196	0.109	0.019	No
DP mg/L	61	-1.329	0	0.241	0.160	0.017	No
TP mg/L	83	-0.784	0	0.198	***0.028***	0.047	Yes
Cu μg/L	33	1.081	0	0.959	0.501	0	No
Pb μg/L	33	1.498	0	-1.02	***0.001***	0.261	Yes
Zn μg/L	32	2.21	0	-0.082	0.527	0	No

Figure 49 shows the regression equations for total phosphorus and total lead for data from commercial land uses. The 95% confidence interval of the regression line for total phosphorus can include zero slope lines. This indicates that there is not a strong correlation between antecedent dry period and total phosphorus concentrations. For total lead, the reduction in concentrations with increasing dry periods is more obvious, but not very explicable.



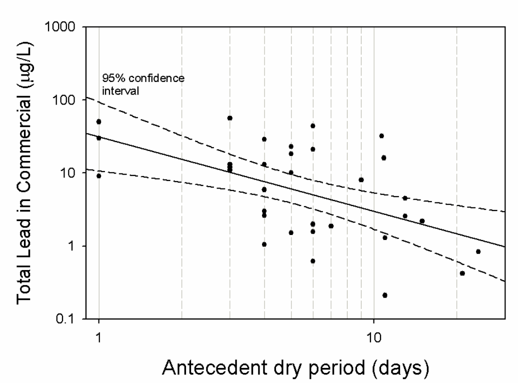


Figure 49. Total phosphorus and total lead concentrations as a function of antecedent dry period in commercial land use areas

The effect of the antecedent dry period on stormwater concentrations at industrial land uses was not significant, except for TSS, as shown on Table 42. Figure 50 is a plot of the TSS concentrations increasing with increasing dry periods.

Table 42. Regression of Logarithm of Constituent Concentrations by Logarithm of Antecedent Dry Period in Industrial Land Use, EPA Rain Zone 2

		Constant		Impervious			
Constituent	n	Coefficient	p-value	Coefficient	p-value	R^2 adjusted	Significant at 0.05 level?
Oil - Grease mg/L	3	0.271	0.773	-0.451	0.700	0	No
TDS mg/L	30	1.651	0	-0.009	0.958	0	No
TSS mg/L	31	1.190	0	0.656	***0.025***	0.134	Yes
BOD_5 mg/L	32	0.780	0	0.201	0.202	0.022	No
COD mg/L	29	1.685	0	0.071	0.622	0	No
Ammonia mg/L	27	-0.487	0.014	-0.084	0.753	0	No
NO2 mg/L	32	-0.154	0.233	-0.124	0.493	0	No
TKN mg/L	32	-0.151	0.215	0.218	0.207	0.021	No
DP mg/L	28	-1.176	0	0.190	0.406	0	No
TP mg/L	32	-0.966	0	0.373	0.11	0.053	No
Cu µg/L	3	1.109	0.124	0.216	0.565	0	No
Pb µg/L	3	0.882	0.197	0.119	0.787	0	No
Zn µg/L	3	2.072	0.056	0.186	0.555	0	No

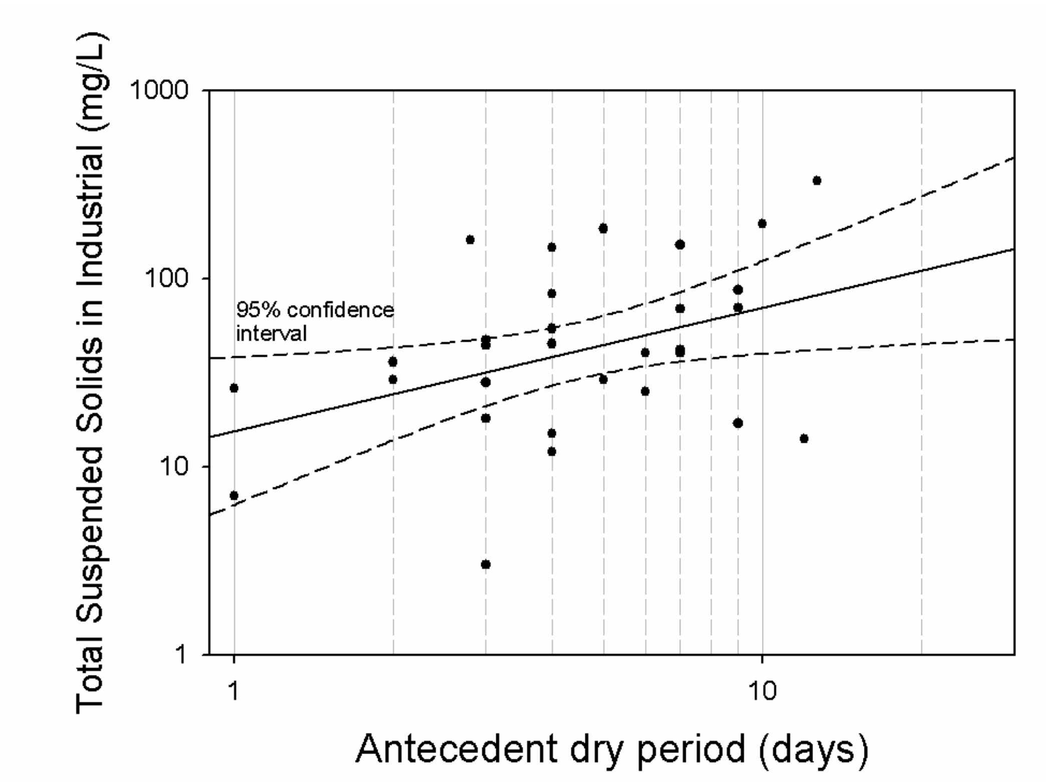


Figure 50. TSS concentrations vs. dry days since rain event in industrial land use areas

Trends in Stormwater Quality with Time

Figure 52 shows a plot of lead concentrations for residential areas only (in EPA Rain Zone 2), for the time period from 1991 to 2002. This plot shows likely decreasing lead concentrations with time. Statistically however, the trend line is not significant due to the large variation in observed concentrations (p = 0.41; there is insufficient data to show that the slope term is significantly different from zero). Likewise the COD concentrations have an apparent downward trend with time, but again, the slope term is not significant (p = 0.12).

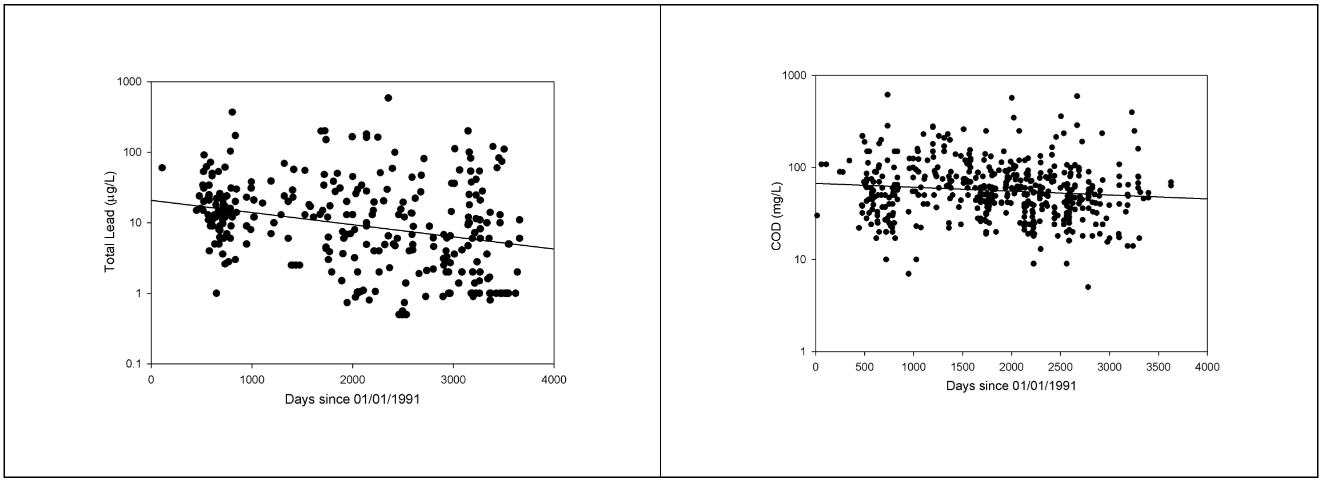


Figure 52. Residential lead and COD concentrations with time (EPA Rain Zone 2 data only)

Except for lead, it is not likely that time between the data collection efforts is the reason why the NURP and NSQD databases have different values.

Summary

Several factors were evaluated using data from the NSQD. Only residential, commercial and industrial land uses in EPA Rain Zone 2 and industrial areas in EPA Rain Zone 3 have enough samples to evaluate factors affecting stormwater concentrations. The effect of each factor cannot be extrapolated to the rest of the country. However they can be used as guidance for communities in other EPA Rain Zones. Additional data from communities that were not included in this first phase of the NSQD database would enable more complete and sensitive analyses. Also, this chapter examined most of these factors in isolation, more as sensitivity analyses and to help identify significant factors. These analyses did not consider factors together and possible interactions.

There is a significant reduction in TSS, nitrite-nitrate, total phosphorus, total copper, and total zinc concentration at sites having wet ponds, the control practice having the largest concentration reductions. No reductions in TKN concentrations were found using wet ponds, but TKN seems to be reduced by dry ponds. Locations with detention storage facilities had smaller reductions of TSS, BOD_5, COD, total lead and total zinc concentrations compared to wet pond sites. Unfortunately, there were few sites in the database having grass swales that could be compared with data from sites having curbs and gutters.

The decision to use automatic or manual sampling methods is not always clear. There were statistical differences found between both methods in residential areas for several constituents. Most communities calculate their EMC values using flow-composited sample analyses. If first flush effects are present, manual sampling may likely miss these more concentrated flows due to delays in arriving at the site to initiate sampling. If the first flush is for a very short duration, time-composited samples may overly emphasize these higher flows. Flow compositing produces more accurate EMC values than time composite analyses. An automatic sampler with flow-weighted samples, in conjunction with a bed load sampler, is likely the most accurate sampling alternative.

There is a certain amount of redundancy (self-correlation) between land use and the percentage of impervious areas, as each land use category generally has a defined narrow range of paved and roof areas. Therefore, it is not possible to test the hypothesis that different levels of impervious (surface coverage) are more important than differences in land use (activities within the area). Residential land uses cover only the lower range of imperviousness, while commercial sites have imperviousness amounts larger than 50%. In order to perform a valid comparison test, the range of imperviousness needs to be similar for both test cases.

Antecedent dry periods before sampling was found to have a significant effect for BOD_5, COD, ammonia, nitrates, TKN, dissolved, and total phosphorus concentrations at residential land use sites. As the number of days increased, there was an increase in the concentrations of the stormwater constituents. This relationship was not observed for freeway sites. This may be associated with the very small drainage areas associated with the freeway sites (drainage areas close to 1 acre), while the drainage areas for residential, commercial and industrial areas ranged between 50 and 100 acres (Figure 2).

No seasonal effects on concentrations were observed, except for bacteria levels that appear to be lower in winter and higher in summer. No effects on concentration were observed according to precipitation depth. Rainfall energy determines erosion and washoff of particulates, but sufficient runoff volume is needed to carry the particulate pollutants to the outfalls. Different travel times from different locations in the drainage areas results in these materials arriving at different times, plus periods of high rainfall intensity (that increase pollutant washoff and movement) occur randomly throughout the storm. The resulting outfall stormwater concentration patterns for a large area having various surfaces is therefore complex and rain depth is just one of the factors involved. The next chapter examines time delivery of pollutants in more detail.

Chapter 6: Comparisons of First 30-minute Samples to 3-hour Composite Samples

Introduction

Sample collection conducted for some of the NPDES MS4 Phase I permits required both a grab and a composite sample for each event. A grab sample was to be taken during the first 30 minutes of discharge, and a flow-weighted composite sample for the entire time of discharge (up to three hours). The initial grab sample was used for the analysis of the "first flush effect," which assumes that more of the pollutants are discharged during the first period of runoff than during later periods. The composite sample was obtained with aliquots collected about every 15 to 20 minutes for at least 3 hours, or until the event ended.

First Flush

First flush refers to an assumed elevated load of pollutants discharged during the beginning of a runoff event. The first flush effect has been observed more often in small catchments than in large catchments (Thompson, *et al*, 1995, cited by WEF and ASCE 1998). In another study, large catchments (>162 Ha, 400 acres) had the highest concentrations observed at the times of flow peak (Soeur, *et al.* 1994; Brown, *et al.* 1995). The presence of a first flush also has been reported to be associated with runoff duration by the City of Austin, TX (Swietlik, *et al*. 1995). Peak pollutant concentrations can occur after the peak discharge, thus some pollutant discharges can be significant for events longer than the time of concentration (Ellis 1986). Adams and Papa (2000), and Deletic (1998) both concluded that the presence of a first flush depends on numerous site and rainfall characteristics.

In this chapter, pollutant characteristics are evaluated using the NSQD database for events that included separate samples collected during both the first 30 minutes and for the entire event (the composite sample), using nonparametric statistical methods. A better analysis of first flush conditions could be performed by using mass discharge curves that relate the total mass discharge as a function of the total runoff volume; however, this procedure requires high resolution flow and concentration information. The NSQD database only contains concentration data from composite samples (and selected first flush samples) and few flow data.

Methodology

A total of 417 storm events having paired first flush and composite samples were available from the NPDES MS4 database. The majority of the events were located in North Carolina (76.2%), but some events were also from Alabama (3.1%), Kentucky (13.9%) and Kansas (6.7%). Table 44 shows the events that were used for this analysis, separated by land use and community. All the events correspond to end-of-pipe samples in separate storm drainage systems.

Table 44. Preliminary Number of Storm Events Selected

State	Community	CO	FW	ID	IS	OP	RE	Total Events	%
AL	Jefferson County	5	2	0	0	0	6	13	3.1
NC	City of Charlotte	8	0	8	0	3	16	35	8.4
NC	City of Fayetteville	18	0	18	18	6	46	106	25.4
NC	City of Greensboro	33	0	33	0	15	33	114	27.3
KY	City of Lexington	12	3	2	0	2	18	37	8.9
KY	City of Louisville	0	0	7	0	0	14	21	5.0
NC	City of Raleigh	18	0	18	0	9	18	63	15.1
KA	City of Wichita	7	0	7	0	0	14	28	6.7
Total Events		101	5	93	18	35	165	417	
%		24.2	1.2	22.3	4.3	8.4	39.6		100

Note: CO (commercial), FW (freeway), ID (industrial), IS (institutional), OP (Open Space) and RE (residential) land uses

The initial task was to select the constituents and land uses that meet the requirements of the statistical comparison tests. Probability plots, box and whiskers plots, concentration vs. precipitation and standard descriptive statistic calculations were performed for 22 constituents for each land use and all areas combined. Nonparametric statistical analyses were performed after these initial analyses. Mann-Whitney and Fligner-Policello tests were most commonly used. Minitab and Systat statistical programs, along with Word and Excel macros, were used during the analysis.

Initial Analyses

One of the conclusions of the NURP program was that most of the constituents in stormwater generally follow a log-normal distribution, especially between the 5th and 95th percentiles (EPA 1983). This characteristic was validated using probability plots during the initial analyses. Results from first flush and composite samples were log-transformed, for different pollutant types, in each land use category.

Figure 53 shows initial statistical results for both phosphorus and COD. Elevated first flush concentrations were evident for COD compared to phosphorus. Probability plots provide useful information about the characteristics of the sample population.

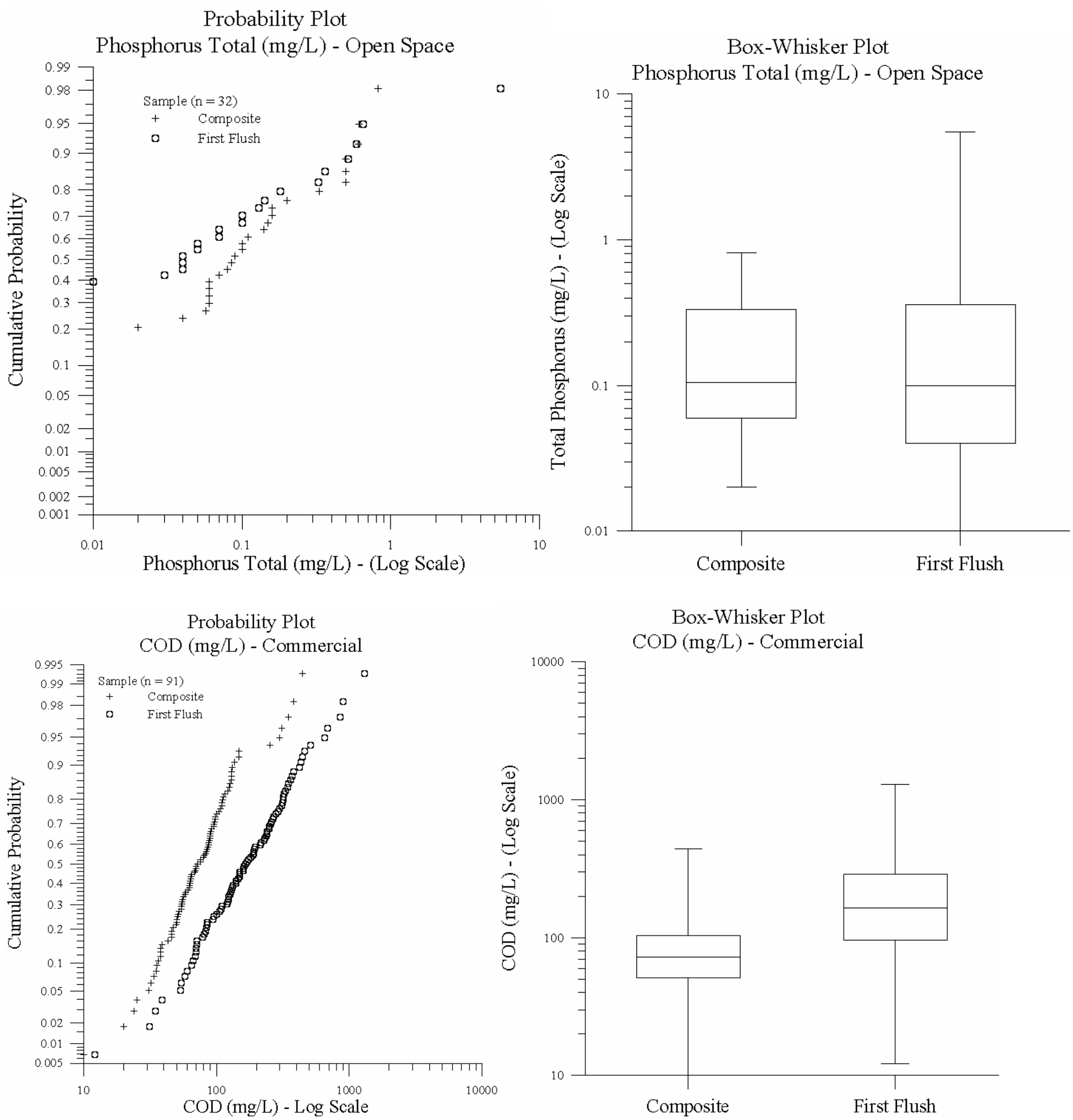


Figure 53. Cumulative probability and box and whiskers plots

Figure 53 is an example for total phosphorus observations from the open space land use. Both sample sets follow a lognormal distribution because most of the points lie on a straight line. The slopes of the lines are different, indicating unequal variances. In this case, about 40% of the first flush samples did not have detected concentrations for phosphorus, while about 20% of the composite samples had non-detected phosphorous concentrations. This plot also indicated that the median concentration of the composite samples is almost twice the median value for the first flush samples.

The next initial analysis used box plots. These plots also represent the distribution of the data, but only show the detectable concentrations. The middle line inside the box represents the median of the data. The top of the box represents the third quartile, and the bottom the first quartile. The whiskers are extended from the 5th to the 95th percentile limits. Values outside these limits are represented with asterisks. The exclusion of the non-detected values changes the median of the data compared to the probability plots. In this example, both of the medians are similar, in contrast with the results of the probability plot. In this example, the variability of the first flush observations is also seen to be larger than the composite data set.

Descriptive statistics for each constituent and land use were calculated to determine if the distributions were symmetrical and if they had the same variance (see Appendix E). This evaluation is needed to select the most appropriate statistical tests. In some conditions, the number of sample pairs was not large enough to allow further analyses. Table 45 shows the results of the initial analysis. Samples having lognormal probability distributions and sufficient data sets were selected for further analyses.

Figure 54 shows the steps that were followed during the nonparametric analysis. The most useful test was the Fligner-Policello test. This test requires independent random samples symmetric about the medians for each data set. The advantage of this test is that it does not require normality or the same variance in each data set (Fligner and Policello 1981). The U statistic and the p-value are shown in the Appendix E for some constituents. Chakraborti (2003) presents a definition and explanation of the Mann-Whitney U test. P-values smaller than 5% (<0.05) indicate that the first flush and composite sample sets have different median concentrations at the 95%, or greater, confidence level.

Table 45. Initial Analyses to Select Data Sets for First Flush Analyses

Constituent	CO	ID	IS	OP	RE	ALL
Turbidity, NTU	Selected	No data	No data	Ned	Selected	Selected
pH, S.U.	Selected	Selected	No data	Ned	Selected	Selected
BOD5, mg/L	Selected	Selected	Box plot FF > Com	Selected	Selected	Selected
COD, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
TSS, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
TDS, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
O&G, mg/L	Selected	Ned	Ned	Ned	Selected	Selected
Fecal Coliform, col/100mL	Selected	Ned	Ned	Ned	Selected	Selected
Fecal Streptococcus, col/100 mL	Selected	Ned	Ned	Ned	Selected	Selected
Ammonia, mg/L	Selected	Selected	Box plot FF > com.	Ned	Selected	Selected
$NO_2 + NO_3$, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
N Total, mg/L	Selected	Selected	Ned	Selected	Selected	Selected
TKN, mg/L	Selected	Selected	Box plot FF > com.	Selected	Selected	Selected
P Total, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
P Dissolved, mg/L	Selected	Selected	Selected	Selected	Selected	Selected
Ortho-P, mg/L	Ned	Selected	Ned	Ned	Selected	Selected
Cadmium Total, μg/L	Selected	Selected	Ned	Selected	Selected	Selected
Chromium Total, μg/L	Selected	Selected	Ned	Selected	Selected	Selected
Copper Total, μg/L	Selected	Selected	Selected	Selected	Selected	Selected
Lead Total, μg/L	Selected	Selected	Selected	Selected	Selected	Selected
Mercury, μg/L	Ned	Ned	Ned	Ned	Ned	Ned
Nickel, μg/L	Selected	Selected	Ned	Ned	Selected	Selected
Zinc, μg/L	Selected	Selected	Selected	Selected	Selected	Selected

* Ned: Not enough data. CO (commercial), FW (freeway), ID (industrial), IS (institutional), OP (Open Space) and RE (residential)

Nonparametric Analyses

If the number of samples is large, and the distributions are normal and have the same variance, a paired Student's t-test is usually a better test to evaluate the hypothesis and support the results of the Fligner-Policello test. To verify that the data distributions are normal, the Anderson-Darling normality test was used (Kottegoda and Rosso 1997). This method uses an empirical cumulative distribution function to check normality. In Appendix E, the p-values of the paired differences are shown. P-values larger than 5% (> 0.05) indicate that the normality requirement was met at the 95% or greater confidence level.

Finally, if the first flush and composite sample distributions are symmetrical (but not necessarily normal), and if they have the same variance, the Mann-Whitney test can be used. If the p-value is larger than 5% (>0.05), the medians of the sample distribution are assumed to be the same, at the 95% or greater confidence level. The preferred test would be the Student's t-test, if the sample characteristics warrant, followed by the Mann-Whitney test and finally the Fligner-Policello test. The selected cases are only for pairs with concentration values above the detection limits. The ratios between the first flush and composite sample median concentrations are also shown. Commercial and residential areas have the highest ratios for most constituents. The smallest ratios were found for open space sites.

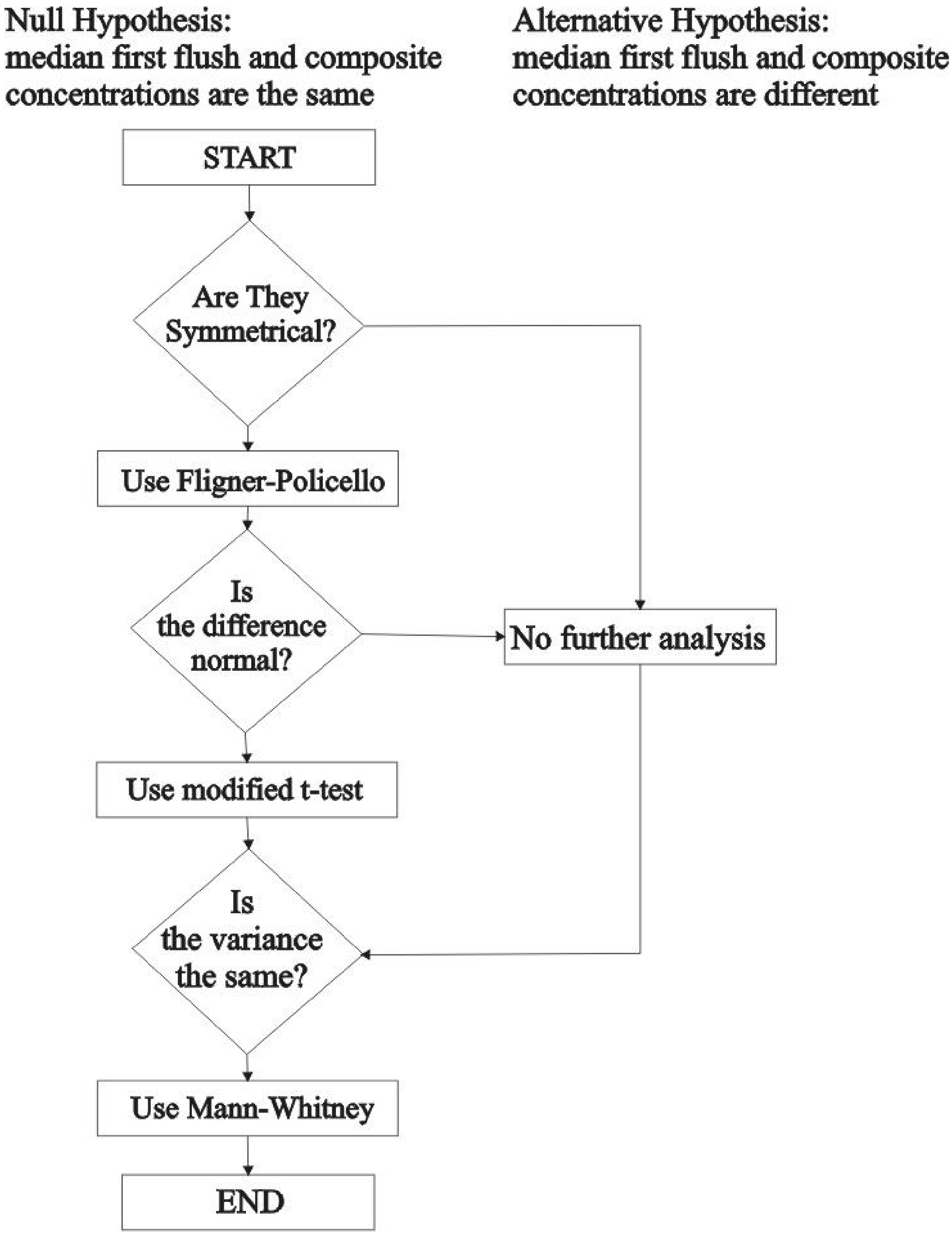


Figure 54. Analysis flow chart

Results

About 83% of the possible paired cases were successfully evaluated. The remaining cases could not be evaluated because the data set did not have enough paired data or they were not symmetrical. Table 46 shows the results of the analysis.

Table 46. Significant First Flushes Ratios (first flush to composite median concentration)

Parameter	Commercial				Industrial				Institutional			
	n	sc	R	ratio	n	sc	R	ratio	n	sc	R	ratio
Turbidity, NTU	11	11	=	1.32			X				X	
pH, S.U.	17	17	=	1.03	16	16	=	1.00			X	
COD, mg/L	91	91	≠	2.29	84	84	≠	1.43	18	18	≠	2.73
TSS, mg/L	90	90	≠	1.85	83	83	=	0.97	18	18	≠	2.12
BOD_5, mg/L	83	83	≠	1.77	80	80	≠	1.58	18	18	≠	1.67
TDS, mg/L	82	82	≠	1.83	82	81	≠	1.32	18	18	≠	2.66
O&G, mg/L	10	10	≠	1.54			X				X	
Fecal Coliform, col/100mL	12	12	=	0.87			X				X	
Fecal Streptococcus, col/100 mL	12	11	=	1.05			X				X	
Ammonia, mg/L	70	52	≠	2.11	40	33	=	1.08	18	16	≠	1.66
$NO_2 + NO_3$, mg/L	84	82	≠	1.73	72	71	≠	1.31	18	18	≠	1.70
N Total, mg/L	19	19	=	1.35	19	16	=	1.79			X	
TKN, mg/L	93	86	≠	1.71	77	76	≠	1.35			X	
P Total, mg/L	89	77	≠	1.44	84	71	=	1.42	17	17	=	1.24
P Dissolved, mg/L	91	69	=	1.23	77	50	=	1.04	18	14	=	1.05
Ortho-P, mg/L			X		6	6	=	1.55			X	
Cadmium Total, μg/L	74	48	≠	2.15	80	41	=	1.00			X	
Chromium Total, μg/L	47	22	≠	1.67	54	25	=	1.36			X	
Copper Total, μg/L	92	82	≠	1.62	84	76	≠	1.24	18	7	=	0.94
Lead Total, μg/L	89	83	≠	1.65	84	71	≠	1.41	18	13	≠	2.28
Nickel, μg/L	47	23	≠	2.40	51	22	=	1.00			X	
Zinc, μg/L	90	90	≠	1.93	83	83	≠	1.54	18	18	≠	2.48
Turbidity, NTU			X		12	12	=	1.24	26	26	=	1.26
pH, S.U.			X		26	26	=	1.01	63	63	=	1.01
COD, mg/L	28	28	=	0.67	140	140	≠	1.63	363	363	≠	1.71
TSS, mg/L	32	32	=	0.95	144	144	≠	1.84	372	372	≠	1.60
BOD5, mg/L	28	28	=	1.07	133	133	≠	1.67	344	344	≠	1.67
TDS, mg/L	31	30	=	1.07	137	133	≠	1.52	354	342	≠	1.55
O&G, mg/L			X				X		18	14	≠	1.60
Fecal Coliform, col/100mL			X		10	9	=	0.98	22	21	=	1.21
Fecal Streptococcus, col/100 mL			X		11	8	=	1.30	26	22	=	1.11
Ammonia, mg/L			X		119	86	≠	1.36	269	190	≠	1.54
$NO_2 + NO_3$, mg/L	30	21	=	0.96	121	118	≠	1.66	324	310	≠	1.50
N Total, mg/L	6	6	=	1.53	31	30	=	0.88	77	73	=	1.22
TKN, mg/L	32	14	=	1.28	131	123	≠	1.65	335	301	≠	1.60
P Total, mg/L	32	20	=	1.05	140	128	≠	1.46	363	313	≠	1.45
P Dissolved, mg/L	32	14	=	0.69	130	105	≠	1.24	350	254	=	1.07
Ortho-P, mg/L			X		14	14	=	0.95	22	22	=	1.30
Cadmium Total, μg/L	30	15	=	1.30	123	33	≠	2.00	325	139	≠	1.62
Chromium Total, μg/L	16	4	=	1.70	86	31	=	1.24	218	82	≠	1.47
Copper Total, μg/L	30	22	=	0.78	144	108	≠	1.33	368	295	≠	1.33
Lead Total, μg/L	31	16	=	0.90	140	93	≠	1.48	364	278	≠	1.50
Nickel, μg/L			X		83	18	=	1.20	213	64	≠	1.50
Zinc, μg/L	21	21	=	1.25	136	136	≠	1.58	350	350	≠	1.59

Note: n = number of total possible events. sc = number of selected events with detected values. R = result. Not enough data (X); not enough evidence to conclude that median values are different (=); median values are different (≠).

The "≠" sign indicates that the medians of the first flush and the composite data set are different. The "=" sign indicates that there is not enough information to reject the null hypothesis at the desired level of confidence (at least at the 95% level). Events without enough data are represented with an "X".

Also, shown on this table are the ratios of the medians of the first flush to the composite data for each constituent and land use combination. Generally, a statistically significant first flush is associated with a median concentration ratio of about 1.4, or greater (the exceptions are where the number of samples in a specific category is much smaller). The largest ratios are about 2.5, indicating that for these conditions, the first flush sample concentrations are about 2.5 times greater than the composite sample concentrations. More of the larger ratios are found for the commercial and institutional land use categories, areas where larger paved areas are likely to be found. The smallest ratios are associated with the residential, industrial, and open spaces land uses, locations where there may be larger areas of unpaved surfaces.

Results indicate that for 55% of the evaluated cases, the median of the first flush data set were different than the composite sample set. In the remaining 45% of the cases, both medians were likely the same, or the concentrations were possibly greater later in the events.

Approximately 70% of the constituents in the commercial land use category had elevated first flush concentrations, about 60% of the constituents in the residential, institutional and the mixed (mostly commercial and residential) land use categories had elevated first flushes, and only 45% of the constituents in the industrial land use category had elevated first flushes. In contrast, no constituents were found to have elevated first flushes in the open space category.

COD, BOD_5, TDS, TKN and Zn all had first flushes in all areas (except for the open space category). In contrast, turbidity, pH, fecal coliform, fecal streptococcus, total N, dissolved and ortho-P never showed a statistically significant first flush in any category. The different findings for TKN and total nitrogen imply that there may be other factors involved in the identification of first flushes besides land use.

Summary

It is expected that peak concentrations generally occur during periods of peak flows (and highest rain energy). On relatively small paved areas, however, it is likely that there will always be a short initial period of relatively high concentrations associated with washing off of the most available material (Pitt 1987). This peak period of high concentrations may be overwhelmed by periods of high rain intensity that may occur later in the event. In addition, in more complex drainage areas, the routing of these short periods of peak concentrations may blend with larger flows and may not be noticeable. A first flush in a separate storm drainage system is therefore most likely to be seen if a rain occurs at relatively constant intensities on a paved area having a simple drainage system.

If the peak flow (and highest rain energy) occurs later in the event, then there likely will not be a noticeable first flush. However, if the rain intensity peak occurs at the beginning of the event, then the effect is exaggerated. Figure 55 shows an example storm in Lexington, KY. Note that in this event there are two periods of elevated peaks, the first occurs one hour after the rain started, the second two hours later. If the concentration remains the same during the entire event, the maximum load will occur during the later periods having the maximum flows (the two peaks), and not during the initial period of the storm. Another factor that needs to be considered is the source of the contaminants and how fast they travel through the watershed. Streets and other impervious areas will contribute flows to the outfall monitoring location before the pervious areas in the drainage area.

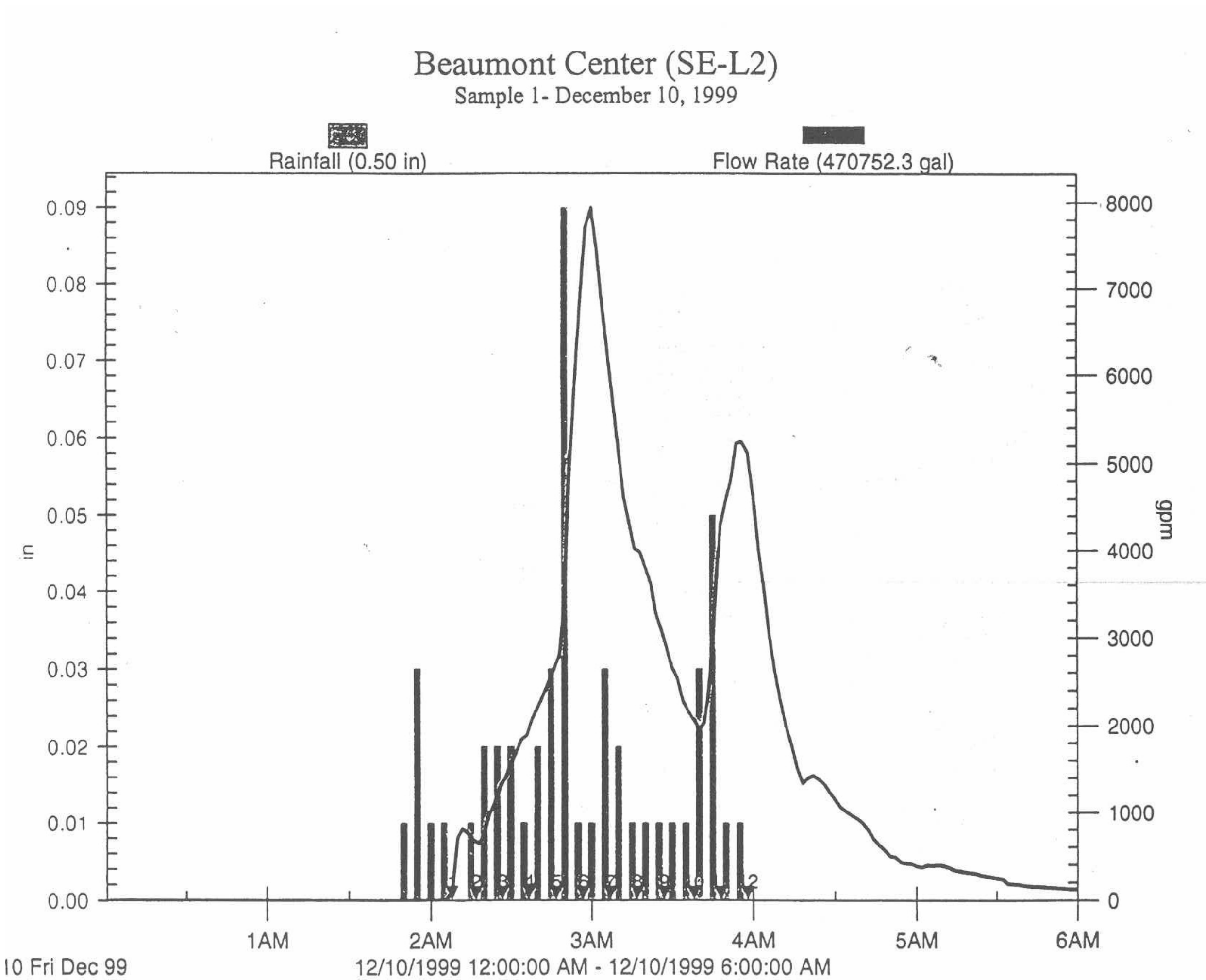


Figure 55. Hydrograph for a storm event (Source: NPDES permit Lexington–KY 2000) (1 in = 25.4 mm, 1 m^3 = 264 gal)

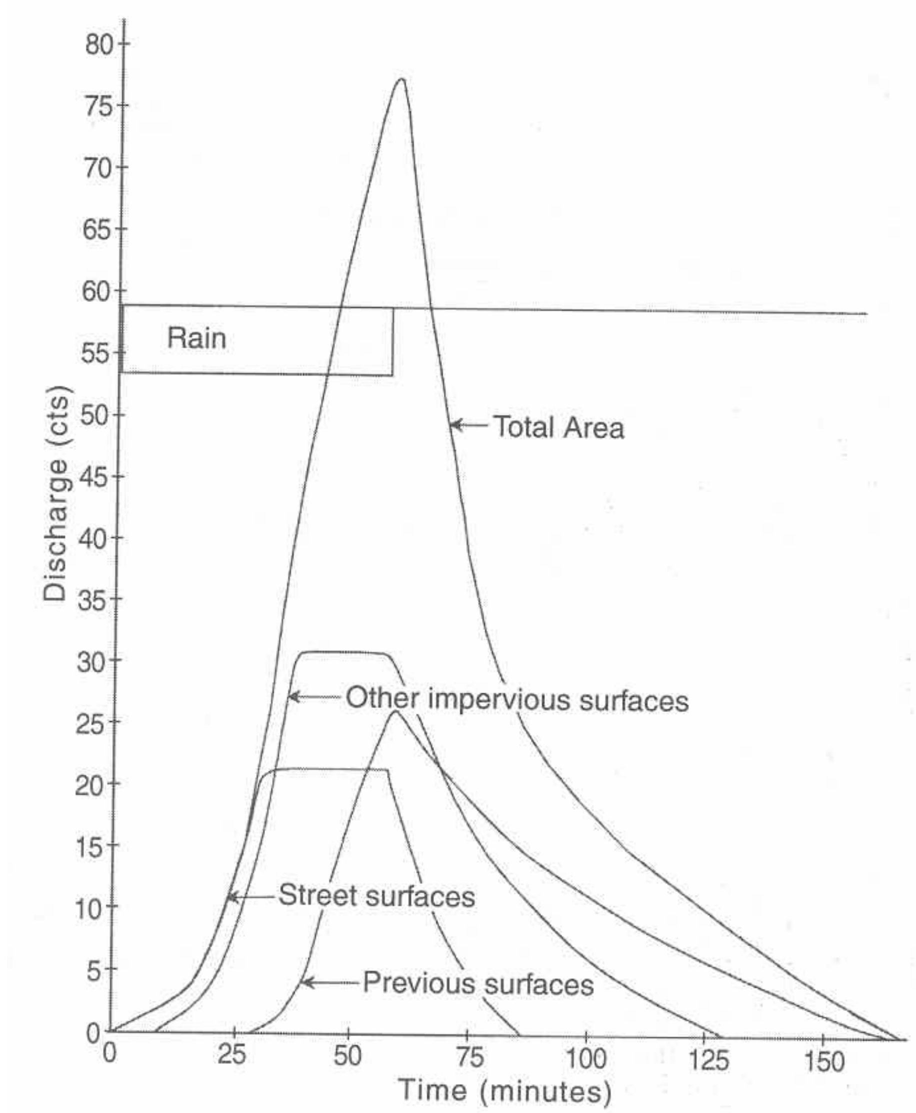


Figure 56. Contributing areas in urban watersheds (Pitt, 1999) (1 m^3/s = 35 cfs)

Figure 56 (Pitt 1999) shows that for an example constant rainfall, the source area flow contribution changes for different rain conditions in an area. If the percentage of impervious surfaces is high, many of the constituents will be discharged faster. This observation agrees with the results observed from the statistical analysis. Commercial areas have a larger frequency of high concentrations at the beginning of the event in contrast to open space areas.

Figure 57 shows that for events (< 12mm, or 0.5 in) in this example medium density residential area, most of the runoff is generated by impervious areas. The average percentage of imperviousness for the monitoring sites was examined. Commercial areas had an average of 83% imperviousness, followed by industrial areas at 70% imperviousness. Institutional and residential land uses were very similar, with 45% and 42% imperviousness respectively. The open space land use category had the smallest imperviousness area, at about 4%. As indicated in Figure 57, larger events can generate more runoff from previous areas than impervious areas. However, it is likely that most of the runoff during the MS4 monitoring activities was associated with the more common small events, and hence, impervious areas were more important.

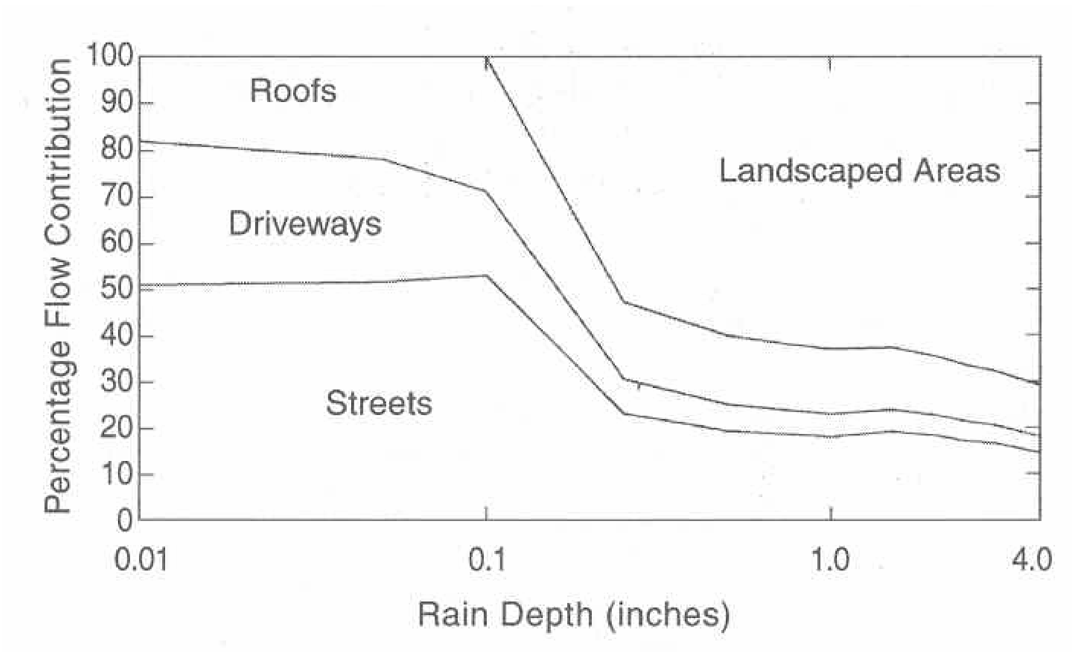


Figure 57. Contributing areas in urban watersheds (Pitt and Voorhees, 1995) (1 in = 25.4 mm)

Probability plots of the precipitation associated with each monitored event for each land use category were prepared to see if there were any significant differences in the ranges of rains observed within each land use category that could have influenced the results. Figure 58 shows that precipitation has the same distribution for almost all the different land uses. The institutional land use category shows a slightly smaller median rain, but this is likely because of the smaller number of events observed in that land use category (18 events). The median precipitation observed during the monitoring at all land uses was about 8 mm (0.3 in), indicating the importance of runoff from the impervious areas.

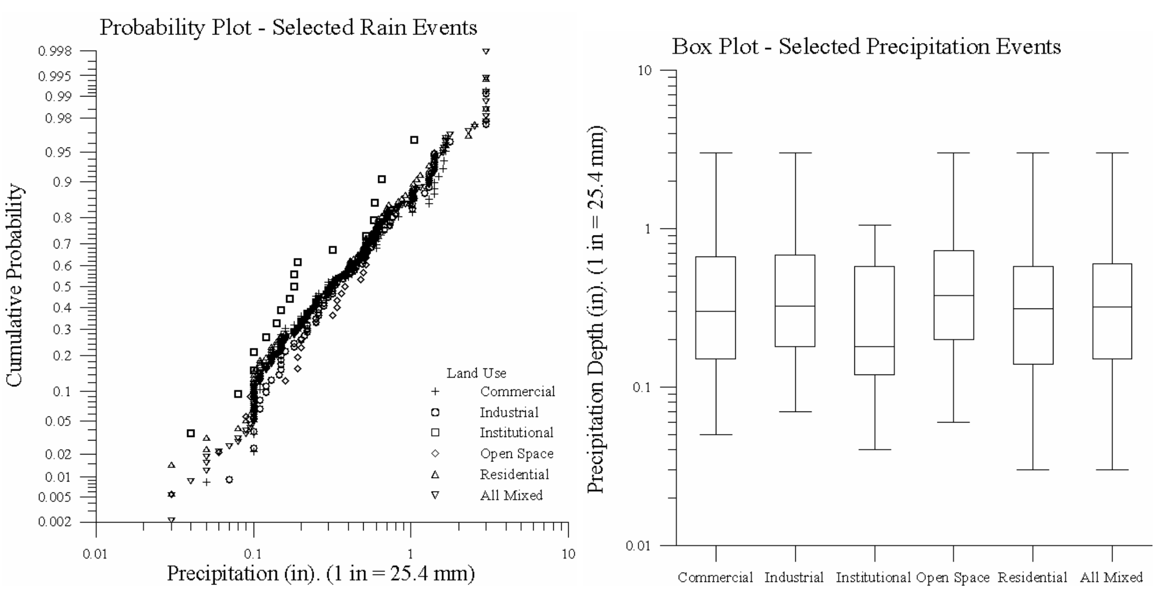


Figure 58. Probability and box and whiskers plot of selected rain events

Finally, another factor that must be considered is the effect of the sampling duration. The guidance provided for monitoring during the Phase I NPDES activities was to collect a sample during the first 30 minutes of the event, and a composite sample only during the first three hours of the event (or the complete event, if shorter than three hours). Figure 59 shows an example case when these conditions can lead to inappropriate conclusions for longer duration events.

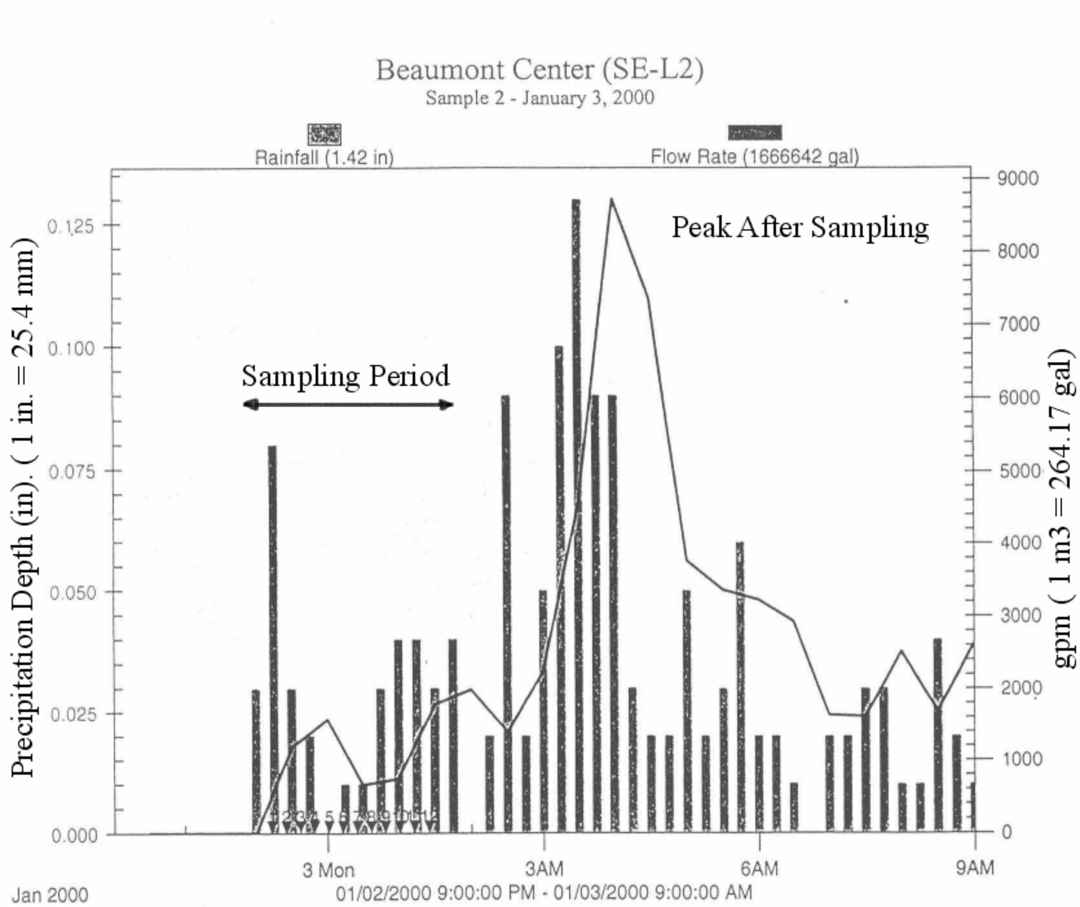


Figure 59. Example of an event with peaks after the sampling period (Source: NPDES permit Lexington-KY, 2000)

The 12 aliquots sampled during the first three hours are shown on the left side of Figure 59. The peak discharge occurred four hours after the event started, as shown on the right side of the figure, and was not represented in the sampling effort. Missing these later storm periods can lead to inappropriate conclusions. It is suggested that for stormwater monitoring, samples should be collected during the complete event and composited before laboratory analyses.

Another sampling example was presented by Roa-Espinosa and Bannerman (1995) who collected samples from five industrial sites using five different monitoring methods. Table 47 shows the ranking of the best methods of sampling based in six criteria. In this table a value between one and five points is assigned to each criterion. Five points indicates that the method is excellent in the specific criterion. Rao-Espinosa and Bannerman concluded that many time-composite subsamples combined for a single composite analysis can provide improved accuracy compared to fewer samples associated with flow-weighted sampling. They also found that time composite subsamples provide better results than samples collected during the first 30 minutes of the event alone.

Table 47. Ranking by Methods of Sampling (Roa-Espinosa, Bannerman, 1995)

Criteria	Flow Composite	Time Discrete	Time Composite	Old Source Sample	New Source Sample	First 30 Minutes
Site Selection	1	1	1	5	5	3
Cost	1	1	3	5	5	5
Technical difficulty	1	1	3	5	5	5
Accuracy	5	5	4	1	5	1
Reproducibility	5	5	5	1	5	1
Representativeness	1	1	3	5	5	1
TOTAL POINTS	14	14	19	22	30	16

Conclusion

A major goal of the present study is to provide guidance to stormwater managers and regulators. Especially important will be the use of this data as an updated benchmark for comparison with locally collected data. In addition, this data may be useful for preliminary calculations when using the "simple method" for predicting mass discharges for unmonitored areas. These data can also be used as guidance when designing local stormwater monitoring programs (Burton and Pitt 2002), especially when determining the needed sampling effort based on expected water quality variations. Additional analyses reported in other chapters expand on these preliminary examples and also investigate other stormwater data and sampling issues.

This investigation of first flush conditions indicated that a first flush effect was not present for all the land use categories, and certainly not for all constituents. Commercial and residential areas were more likely to show this phenomenon, especially if the peak rainfall occurred near the beginning of the event. It is expected that this effect will be more likely to occur in a watershed with a high level of imperviousness, but even so, the data indicated first flushes for less than 50% of the samples for the most impervious areas. This reduced frequency of observed first flushes in these areas most likely to have first flushes is likely associated with the varying rain conditions during the different events, including composite samples that did not represent the complete runoff durations.

Groups of constituents showed different behaviors for different land uses. All the heavy metals evaluated showed higher concentrations at the beginning of the event in the commercial land use category. Similarly, all the nutrients showed higher initial concentrations in residential land use areas, except for total nitrogen and ortho-phosphorus. This phenomenon was not found in the bacteria analyses. None of the land uses showed a higher population of bacteria during the beginning of the event. Conventional constituents showed elevated concentrations in commercial, residential and institutional land uses.

Chapter 7: Effects of Land Use and Geographical Location on Stormwater Quality

Model Building using the NSQD

This chapter describes the methods used to analyze stormwater characterization data in the NSQD in order to determine the best simple method that can be used to calculate the EMC for a site, given the land use, geographical location, and season. These analyses only used those events obtained at single land use sites. This chapter stresses suspended solids analysis as the prototype evaluation procedure that can be used for the other constituents. The later section of this chapter presents results of detailed analyses for other pollutants.

ANOVA Evaluation of Suspended Solids Data

Total suspended solids is one of the most important constituents in stormwater and is commonly used to measure the effectiveness of controls. Unfortunately, there is much controversy concerning TSS monitoring and laboratory analyses. Automatic samplers cannot include bed load and floatable fractions of the solids, and the samplers have reduced efficiency for larger particles (usually larger than about 300 μm). In addition, some laboratories improperly allow the samples to settle before analyses in order to obtain only the suspended portion of the sample, and not the non-filterable fraction as defined by *Standard Methods*. The TSS data in the NSQD were all obtained from outfall monitoring locations, where the amount of particles larger than 300 μm are quite rare, and the laboratories followed proper TSS analytical methods. Analysis of variance (ANVOA) statistical tests were used on natural-log transformed TSS values to identify significant groupings of data, considering both main factors and interactions. The factors examined included land use (residential, commercial, industrial, open space, and freeways), season (spring, summer, fall, and winter) and EPA Rain Zone (the nine EPA rain zones, as shown on Figure 1).

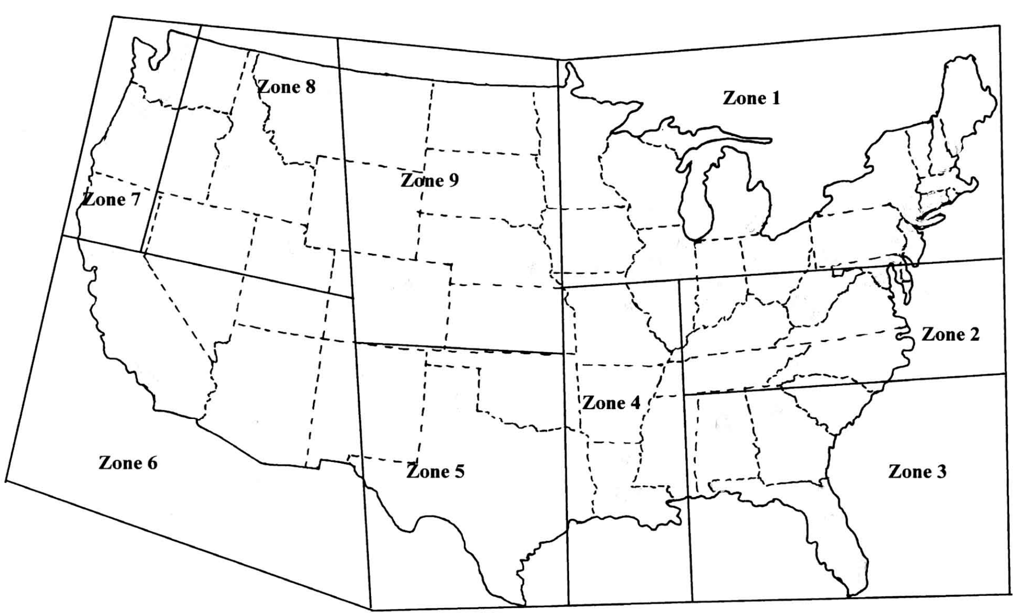


Figure 1. EPA rain zones for the continental US.

Descriptive TSS Statistics

The first step was to calculate simple descriptive statistics for TSS for each of the main factor categories. The TSS concentrations were log transformed (natural log) in order to preserve the normality assumption in the ANOVA analysis. The number of samples, mean, median, maximum, minimum, among other statistics, were calculated in each level of the main factors. Table 1 shows the descriptive statistics for these factors.

Table 1. Descriptive Statistics of the Natural Logarithm (Ln) of TSS mg/L for Single Landuse Categories

Descriptive Statistics: LNTSS by Landuse

Variable	Landuse	N	Mean	Median	TrMean	StDev
LTSS	CO	450	3.8831	3.7377	3.8469	1.1801
	FW	133	4.4644	4.5951	4.4636	1.0680
	ID	423	4.2777	4.3567	4.2842	1.1913
	OP	42	3.945	3.877	3.945	1.717
	RE	977	3.8744	3.8918	3.8650	1.1804

Variable	Landuse	SE Mean	Minimum	Maximum	Q1	Q3
LTSS	CO	0.0556	1.0986	7.7770	3.0910	4.6052
	FW	0.0926	1.0986	8.4764	3.7842	5.0593
	ID	0.0579	1.0986	7.8200	3.4965	5.0752
	OP	0.265	1.099	6.888	2.303	5.426
	RE	0.0378	1.0986	7.8087	3.0910	4.5911

Descriptive Statistics: LNTSS by Season

Variable	Season	N	Mean	Median	TrMean	StDev
LTSS	FA	555	3.8601	3.8501	3.8532	1.1550
	SP	528	4.0990	4.0431	4.0847	1.1968
	SU	400	4.0699	4.0774	4.0670	1.3387
	WI	542	3.9983	3.9512	3.9872	1.1470

Variable	Season	SE Mean	Minimum	Maximum	Q1	Q3
LTSS	FA	0.0490	1.0986	7.7770	3.0910	4.5850
	SP	0.0521	1.0986	7.8200	3.3322	4.8380
	SU	0.0669	1.0986	7.2298	3.0910	5.0876
	WI	0.0493	1.0986	8.4764	3.2958	4.7027

Descriptive Statistics: LNTSS by EPA Rain Zone

Variable	EPA_Rain	N	Mean	Median	TrMean	StDev
LTSS	1	42	3.862	3.761	3.877	1.268
	2	1161	3.7446	3.7612	3.7376	1.1086
	3	120	3.906	3.880	3.898	1.389
	4	218	4.5466	4.4426	4.5320	1.4053
	5	152	4.3056	4.3437	4.3124	1.0898
	6	159	4.6129	4.7005	4.6011	1.0135
	7	141	4.1096	4.2047	4.1142	1.0561
	8	7	4.221	3.970	4.221	0.794
	9	25	5.412	5.587	5.414	0.882

Variable	EPA_Rain	SE Mean	Minimum	Maximum	Q1	Q3
LTSS	1	0.196	1.099	6.447	2.996	4.825
	2	0.0325	1.0986	7.0867	2.9957	4.4543
	3	0.127	1.099	7.030	2.773	4.940
	4	0.0952	1.6094	7.8087	3.4657	5.6204
	5	0.0884	1.0986	7.8200	3.6636	5.1044
	6	0.0804	1.3863	8.4764	4.0431	5.1330
	7	0.0889	1.0986	6.9847	3.4340	4.7664
	8	0.300	3.367	5.858	3.829	4.477
	9	0.176	3.714	7.056	4.684	6.019

There are enough samples to identify if there are any significant differences among the levels and factors, although EOA Rain Zones 1, 8, and 9 and open space have fewer than 50 samples. The range between the minimum and maximum values are similar for all the groups, indicating that there are not any unusual extreme high or low concentration values in the data set. The mean and median values are also close (after the natural-log transformations) indicating data symmetry for each factor level.

During the ANOVA analyses, each factor was identified as a discrete variable. The partial sum of squares was used to identify the effects of the interactions. The results of the ANOVA (using DataDesk 6.1 from MBAWare), including all the interactions are:

Dependent variables

Name	Code
LTSS	LTS

Type of analysis: OLS ANOVA

Factors

Name	Code	Nested in	F/R	Kind
Landuse	Lnd	()	Fix	Disc
Season	Ssn	()	Fix	Disc
RZone	RZn	()	Fix	Disc

Partial (Type 3) Sums of Squares   Design Help

Interactions up to All - way

Add Interaction   Remove Selected Terms   Up   Down

Source	F/R	max df	EMS	F-Denom
Const	-	1	Const	Error
Lnd	F	4	Lnd	Error
Ssn	F	3	Ssn	Error
RZn	F	8	RZn	Error
Lnd*Ssn	F	12	Lnd*Ssn	Error
Lnd*RZn	F	32	Lnd*RZn	Error
Ssn*RZn	F	24	Ssn*RZn	Error
Lnd*Ssn*RZn	F	96	Lnd*Ssn*RZn	Error
Error	R	1846		
Total		2025		

No Modifications

General Results

2026 total cases

ANOVA

Analysis of Variance For **LTSS**
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	32463.3	32463.3	25689	≤ 0.0001
Lnd	4	16.0508	4.01271	3.1754	0.0130
Ssn	3	5.00983	1.66994	1.3215	0.2657
RZn	8	29.529	3.69113	2.9209	0.0030
Lnd*Ssn	12	5.16695	0.430579	0.34073	0.9817
Lnd*RZn	18	32.4181	1.801	1.4252	0.1095
Ssn*RZn	20	31.7578	1.58789	1.2565	0.1982
Lnd*Ssn*RZn	39	48.9396	1.25486	0.99301	0.4830
Error	1921	2427.55	1.26369		
Total	2025	2951.29			

Results for factor Lnd

The probability value for the 3-way interaction term (0.4830) shows that this interaction is not significant in the model. After deleting this three-way interaction, the new ANOVA table is:

Analysis of Variance For **LTSS**
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	32463.3	32463.3	25693	≤ 0.0001
Lnd	4	19.8304	4.9576	3.9236	0.0035
Ssn	3	8.5914	2.8638	2.2665	0.0789
RZn	8	55.0248	6.87809	5.4436	≤ 0.0001
Lnd*Ssn	12	40.5535	3.37946	2.6746	0.0014
Lnd*RZn	18	74.5744	4.14302	3.279	≤ 0.0001
Ssn*RZn	20	40.0732	2.00366	1.5858	0.0477
Error	1960	2476.49	1.26352		
Total	2025	2951.29			

In this case, season and season-rain zone interaction seems not to be significant (probability > 0.05), while all of the remaining factors seem to be important. The mean square error (MSE) is an estimator of the variance in the model. The lower the MSE, the better the model. It was observed that deleting any other source would increase the MSE. The assumption of normality and independence of the residuals for this result was also evaluated as shown in Figure 3.

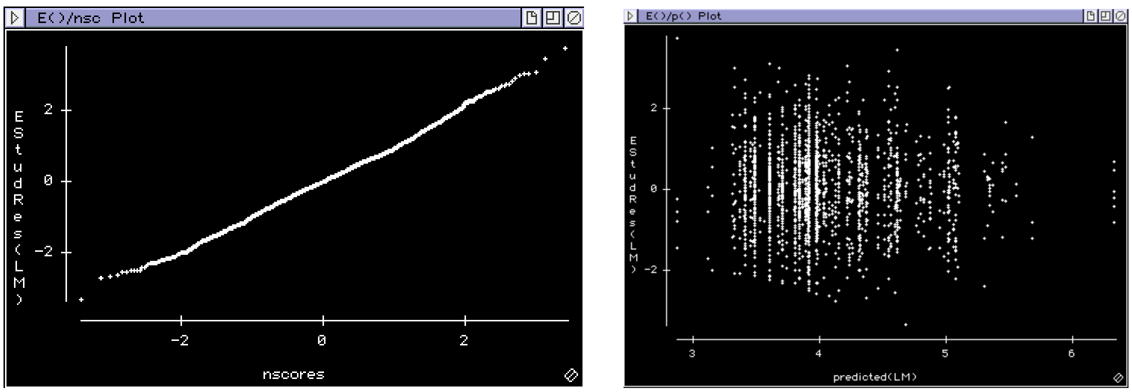


Figure 3. ANOVA results for LN of TSS mg/L using single land uses.

There are not any unusual patterns in the predicted vs. studentized residuals plot. The residuals seem to be independent and normally distributed. The next step is to check if there are any values having large influences or residuals. The potential-residual plot (potential=influence) indicates the data points that have a high influence or residual in the model. A point with an elevated influence indicates that if the point is removed from the dataset, the slope and intercept of the regression line will be affected significantly. DataDesk uses Hadi's influence measure method in preparing the residual-potential plot. The plot identifies unusual observations if they are outside an area described by a hyperbolic trend. Another useful measure is the Cook's distance that considers the influence of each case in all the values. Figure 4 shows the potential residual plot for the natural logarithm of the TSS data.

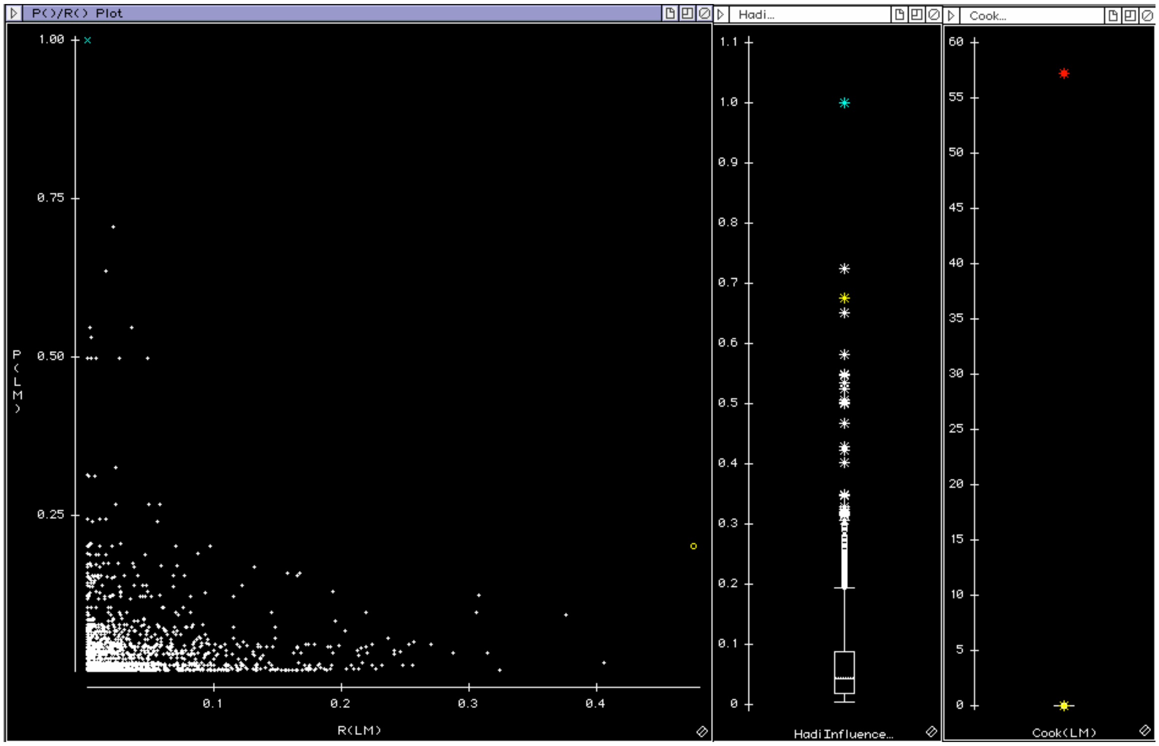


Figure 4. Potential-Residual plot of the natural logarithm of TSS (single land uses).

In this case, the potential residual plot does not indicate any unusual observations in the dataset. All the observations followed a hyperbolic trend; there are not any points outside the area described by this hyperbola. The box plot of the Hadi's influence parameter also does not show any single observation that will influence the whole dataset. The box plot of the Cook's distance indicates a potentially unusual observation. This observation corresponds to a concentration of 46 mg/L in a residential area in EPA Rain Zone 8 during the summer. The unusual characteristic of this observation is that it is the only observation in the database with these characteristics. If this observation is not included in the data plot, the results do not change. The largest influence point is an observation having a concentration of 825 mg/L in an open space area in EPA Rain Zone 2 during the spring. This concentration is not common but it can occur. These data were not deleted from the dataset. Figure 5 shows the box plot of both influence methods.

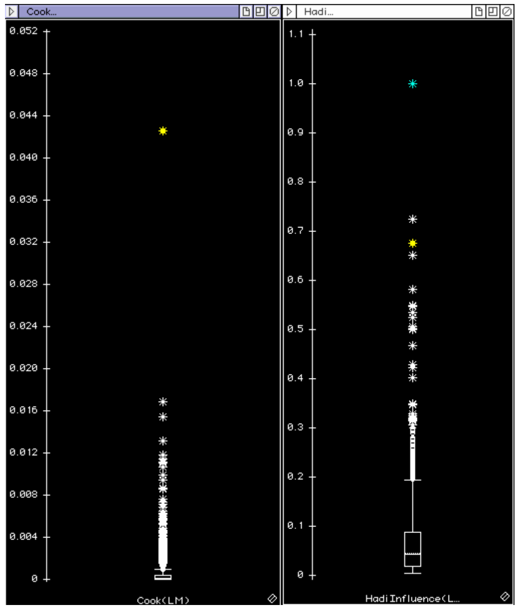


Figure 5. Influence box plot (Cook's distance and Hadi's measure methods).

Because there are no unusual observations, it is possible to evaluate the coefficients for each factor with all the data. A complete examination (all single and multiple interactions) of the coefficients is shown in Table 2.

Table 2. Significant Coefficients for the Complete Factorial Model

LN(TDS mg/L) COEFFICIENTS FOR DIFFERENT LAND USES, SEASONS AND RAIN ZONES					
	Level	**Coefficient**	**std. err.**	**t Ratio**	**prob**
Constant		4.642	0.1742	26.65	< 0.0001
Landuse	**RE**	-0.4871	0.1495	-3.258	0.0011
Season	**None**				
Rain Zone	**2**	-0.8947	0.2155	-4.152	< 0.0001
	4	0.2949	0.1463	2.016	0.0439
	9	0.976	0.4261	2.291	0.0221
Landuse*Season	**SP,OP**	-0.6637	0.3246	-2.045	0.0410
	SU,CO	-0.4617	0.1825	-2.53	0.0115
	SU,OP	0.6597	0.3021	2.183	0.0291
Landuse*Rain Zone	**1,ID**	-2.492	0.7161	-3.479	0.0005
	1,RE	-0.4554	0.2031	-2.242	0.0251
	2,RE	0.4554	0.2031	2.242	0.0251
	5,CO	-1.385	0.6483	-2.136	0.0328
Season*Rain Zone	**2,FA**	-0.3648	0.1386	-2.632	0.0086
	4,SP	0.5739	0.1666	3.444	0.0006
	4,SU	-0.5285	0.1961	-2.695	0.0071
	7,SU	0.6614	0.3321	1.991	0.0466
Landuse*Season*Rain Zone	**None**				

There are 180 possible combinations between the land uses (5), seasons (4) and rain zones (9). The estimated value for any combination is the sum of the coefficients under the conditions of the observation. If the term for a condition being examined is not shown, it was not significant and a zero value is used. Otherwise, the coefficient corresponding to the site condition is used. For example, the following is used to estimate the log value of the TSS for an observation in EPA Rain Zone 4 during spring in a commercial land use. According to Table 2, the expected value is:

Concentration = constant + landuse + season + rain zone + landuse*season + landuse*rain zone + season* rain zone

$$Y = 4.642 + 0 + 0 + 0.2949 + 0 + 0 + 0.5739 = 5.511$$

This corresponds to an expected mean concentration of 247 mg/L. The TSS data in the database for this same group has a mean value of 299 mg/L. This difference is well within the expected error.

After calculating the expected means for each of the 180 possible combinations, a dot plot was created to determine if some groups overlap. For example, it is expected that many of the observations in EPA Rain Zones 1, 3, 4, 6, 7 and 8 will have the same expected TSS concentration values because there were no variations by season for any land uses, except for the residential area. The dot plot of the 180 combinations is shown in Figure 6.

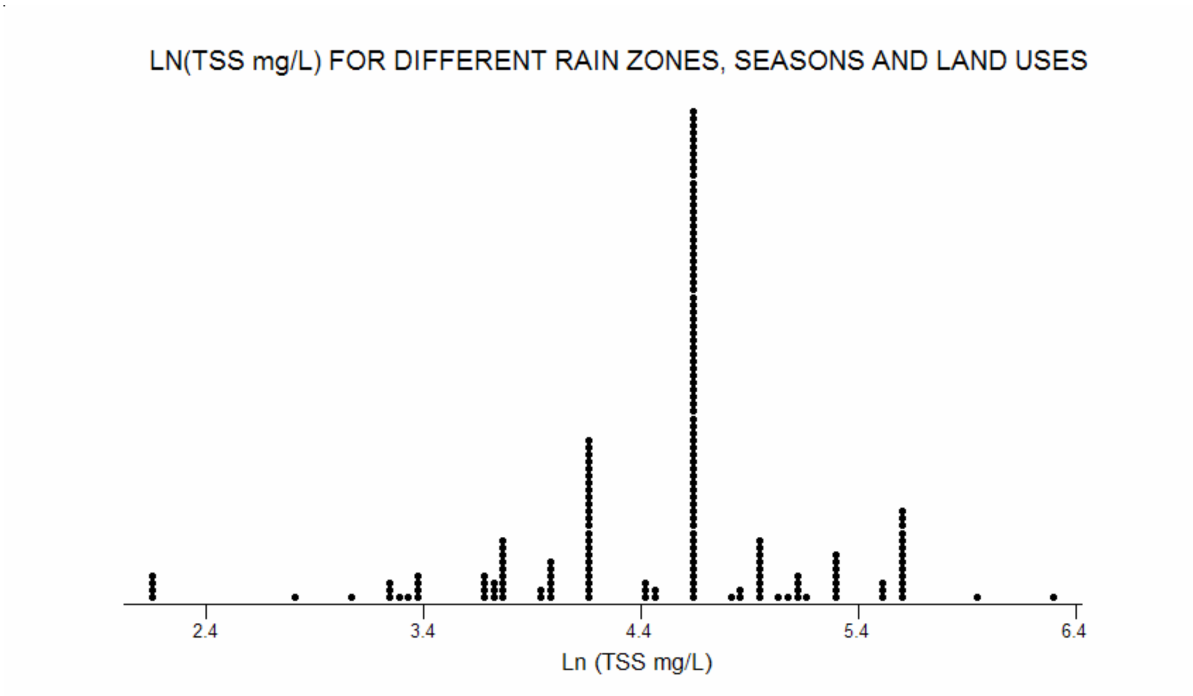


Figure 6. Dot plot of estimated concentrations of Ln TSS.

Figure 6 shows that there are about 17 different groups, at the most. The ANOVA model was reviewed to determine which of the main factors or interactions were the most important. The interaction of Land use*Rain Zone produces by itself the smallest MSE. The new ANOVA table using this interaction is shown in Table 3. The new MSE is 1.29 and is not much larger from the previous MSE using all the significant factors in the model (1.26). Table 4 shows the relevant coefficients using only the reduced model.

Table 3. ANOVA Table using Land Use – Rain Zone Interaction

General Results

2026 total cases

ANOVA

Analysis of Variance For **LTSS**
No Selector

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Const	1	32463.3	32463.3	25078	≤ 0.0001
Lnd*RZn	29	367.524	12.6732	9.7903	≤ 0.0001
Error	1996	2583.76	1.29447		
Total	2025	2951.29			

Results for factor Lnd*RZn

Table 4. Reduced Model

LN (TDS mg/L) COEFFICIENTS FOR DIFFERENT LAND USES, SEASONS AND RAIN ZONES					
	Level	**Coefficient**	**std. err.**	**t Ratio**	**prob**
Constant		4.098	0.07133	57.45	<0.0001
Landuse	**None**				
Season	**None**				
Rain Zone	**None**				
Landuse*Season	**None**				
Landuse*Rain Zone	**1,ID**	-0.8756	0.3235	-2.706	0.0069
	2,CO	-0.3258	0.0954	-3.415	0.0007
	2,FW	1.273	0.331	3.845	0.0001
	2,OP	-0.4027	0.2081	-1.935	0.0531
	2,RE	-0.3944	0.08116	-4.859	<0.0001
	3,CO	-0.5965	0.2429	-2.455	0.0142
	4,CO	0.4018	0.1705	2.357	0.0185
	4,ID	0.5327	0.1611	3.306	0.001
	4,OP	0.8354	0.3235	2.582	0.0099
	4,RE	0.3287	0.1309	2.512	0.0121
	5,CO	-0.7134	0.2528	-2.821	0.0048
	5,ID	0.8913	0.1876	4.751	<0.0001
	6,CO	0.6862	0.3504	1.958	0.0503
	6,FW	0.3606	0.1324	2.723	0.0065
	6,ID	1.229	0.2706	4.54	<0.0001
	6,RE	0.4203	0.217	1.937	0.0529
	7,FW	0.4948	0.2343	2.112	0.0348
	7,ID	0.6974	0.2429	2.87	0.0041
	7,RE	-0.5442	0.1612	-3.376	0.0007
	9,CO	0.9278	0.3859	2.404	0.0163
	9,ID	1.916	0.3859	4.966	<0.0001
Season*Rain Zone	**None**	0			
Landuse*Season*Rain Zone	**None**	0			

All land uses in EPA Rain Zone two (except for freeways) have reduced TSS values when compared with the group average. On the other hand, conditions in EPA Rain Zones 4, 6 and 9 have higher TSS values for the land uses noted. Notice also that industrial and freeway land uses increase the TSS concentrations compared with the other land uses, as expected from the one-way ANOVA tests. Of the 45 possible EPA Rain Zone and land use interactions, 21 have significantly different coefficients and resultant TSS concentrations. All of these possible TSS concentrations, based on this model, are shown in Table 4b.

Table 4b. TSS Concentrations (mg/L) for Different Land Uses and Rain Zones (if values not shown, use 60 mg/L)

	1	2	3	4	5	6	7	8	9
Open space		40		139					
Residential		40		84		92	35		
Commercial		43	33	90	30	120			152
Industrial	25			103	147	206	121		
Freeways		215				86	99		409

Figure 7 shows the groups using the land use*rain zone model. A further reduction in the number of similar groups is not likely possible with this model.

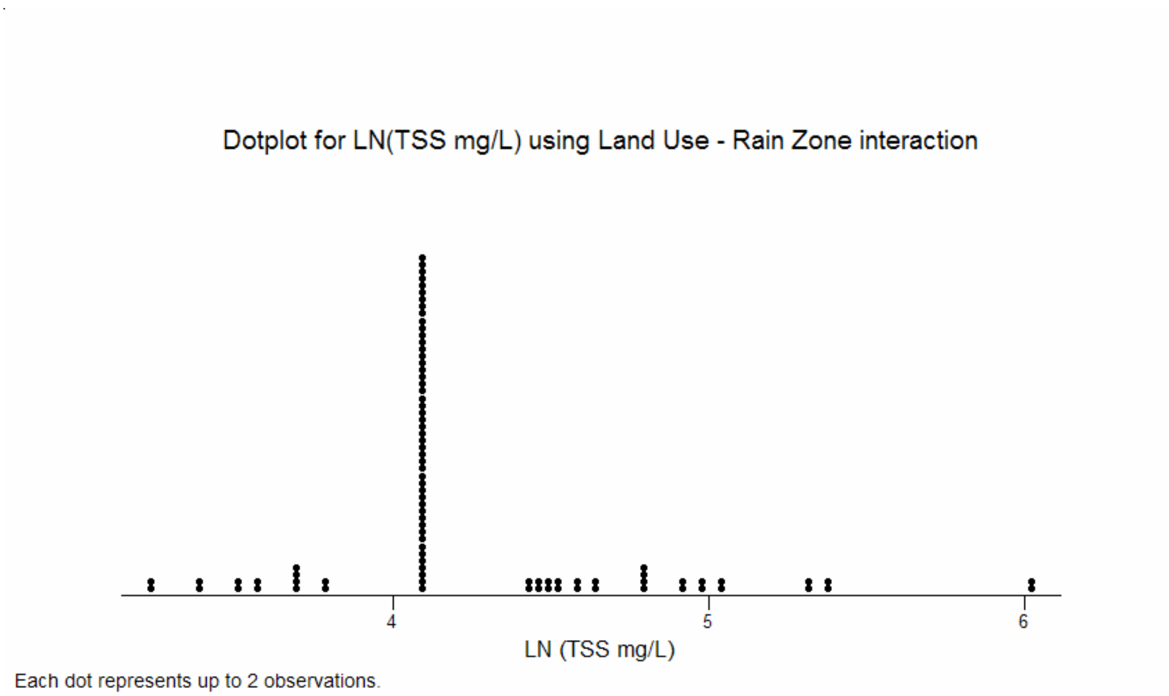


Figure 7. Dot plot using the reduced model

Out of the 45 total land use-rain zone groups, 24 (or 53%) are not affected by significant land use – EPA Rain Zone interaction terms. Seven of the 21 significant groups are smaller than the overall average condition (60 mg/L), while 14 are larger. Only 2 percent of the observations have very large concentrations, they were located in industrial land uses in EPA Rain Zone 9. Figure 8 shows the 5 groups identified with the ANOVA analysis. The variation within the groups is the same as the variation for the whole dataset. The two separate groups located in the upper tail are important. It is not recommended to merge these groups because their concentration differences are very large.

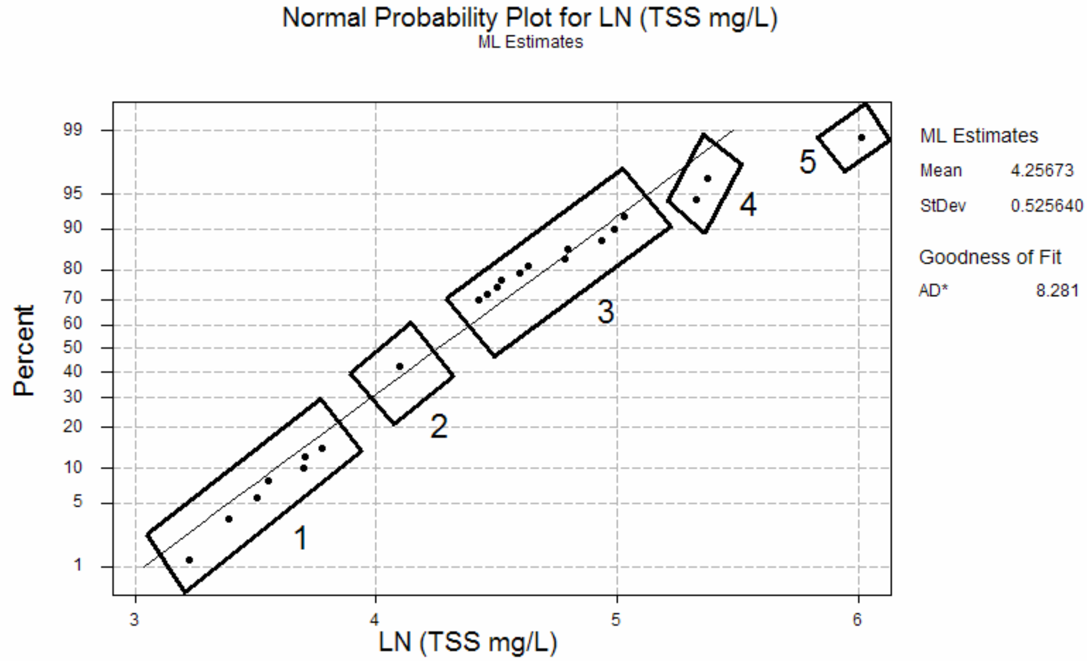


Figure 8. Probability plot using the reduced model (average of the tied points).

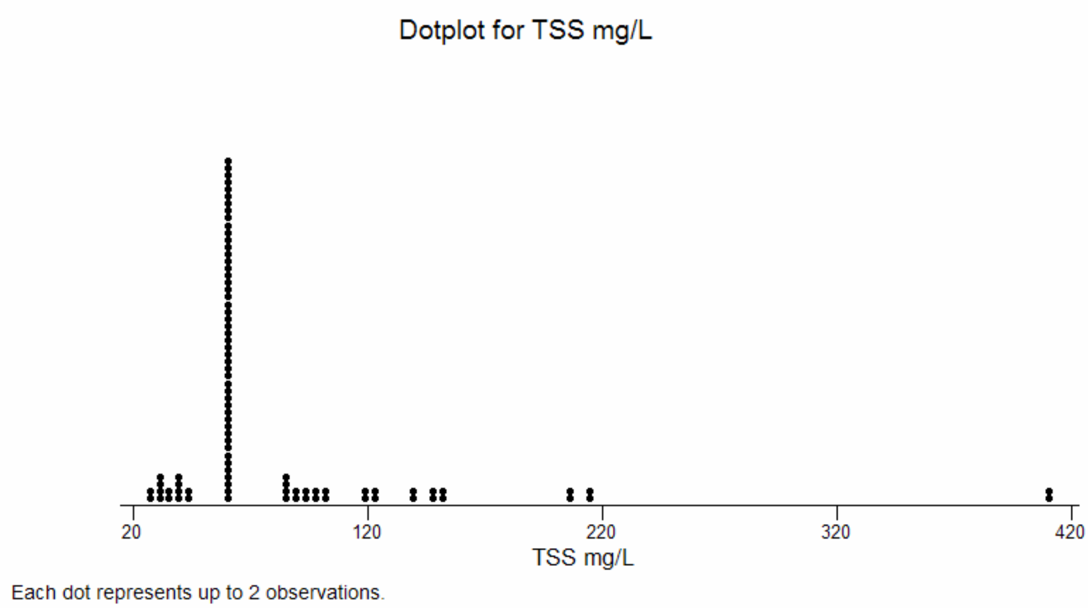


Figure 9. TSS data groups in real space.

There are 2,025 TSS observations for single land uses sites in the NSQD. These observations were classified according to the five groups identified by the above ANOVA model. Figure 11 is a box-whisker plot showing the medians, and 25th, 75th, 5th, and 95th percentiles for each of these groups.

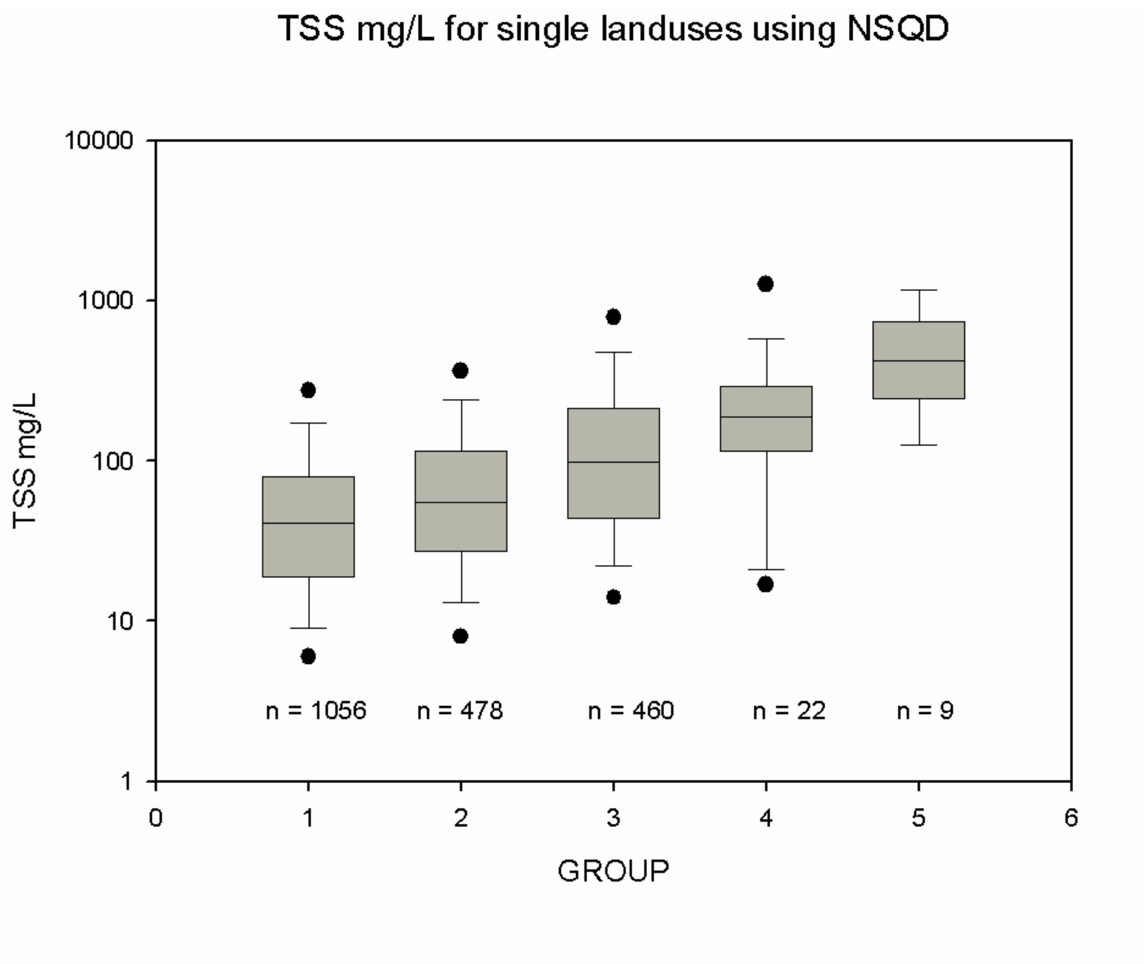


Figure 11. Box plots of five groups also showing 5th and 95th percentiles.

Figure 11 indicates that about half of the TSS single land use data in the NSQD database were in the first group (52%). Most of this data are from residential areas and EPA Rain Zone 2. Twenty-four percent of the observations were not affected by the land use – EPA Rain Zone interaction. Only 1.5% of the data are present in groups 4 and 5. These groups are significantly different than groups 1 and 2. Overall, there are three main levels of TSS concentrations in stormwater: Low (1), Medium (2) and High (3). Other minor categories correspond to groups 4 and 5 and contain the unusually high values.

Table 4c. Five TSS Concentration Categories in NSQD

	Land use*rain zone interactions (Rain Zone: land uses)	**Concentrations (mean ± st. dev. ln mg/L)**	**Range (mean; mean – st. dev. and mean + st. dev., mg/L)**	**Number of single land use TSS observations in category in NSQD**
Low	1: residential 2: open space; residential; commercial 3: commercial 5: commercial 7: residential	3.69±1.12	40 (13 – 123)	1056
Medium	All others not noted elsewhere	4.02±1.11	56 (18 – 169)	478
High	4: residential; commercial; industrial; open space 5: industrial 6: freeways; residential; commercial 7: freeways; industrial 9: commercial	4.60±1.20	99 (30 – 330)	460
Unusually high 1	2: freeways 6: industrial			22
Unusually high 2	9: industrial			9

To evaluate if groups 1 (low) and 2 (medium) are from the same population, a two-sample t test was calculated. The results are as follows:

```
Two-sample T for LTSSG1 vs LTSSG2

           N      Mean     StDev   SE Mean
LTSSG1  1056      3.69      1.12     0.034
LTSSG2   478      4.02      1.11     0.051

Difference = mu LTSSG1 - mu LTSSG2
Estimate for difference:  -0.3370
95% CI for difference: (-0.4577, -0.2162)
T-Test of difference = 0 (vs not =): T-Value = -5.47  P-Value = 0.000  DF = 1532
Both use Pooled StDev = 1.12
```

This test indicates that both groups are from different populations with a p value close to zero. The assumption of equal variances is also valid. The same procedure can be used to evaluate if group 2 (medium) and group 3 (high) are from the same population. The results are as follows:

```
Two-sample T for LTSSG2 vs LTSSG3

          N      Mean     StDev   SE Mean
LTSSG2  478      4.02      1.11     0.051
LTSSG3  460      4.60      1.20     0.056

Difference = mu LTSSG2 - mu LTSSG3
Estimate for difference:  -0.5789
95% CI for difference: (-0.7273, -0.4305)
T-Test of difference = 0 (vs not =): T-Value = -7.66  P-Value = 0.000  DF = 936
Both use Pooled StDev = 1.16
```

The variance of both samples are within 10%. The T statistic and p value corroborates that both distributions are from different populations. A grouped probability plot of the five groups is shown in Figure 12.

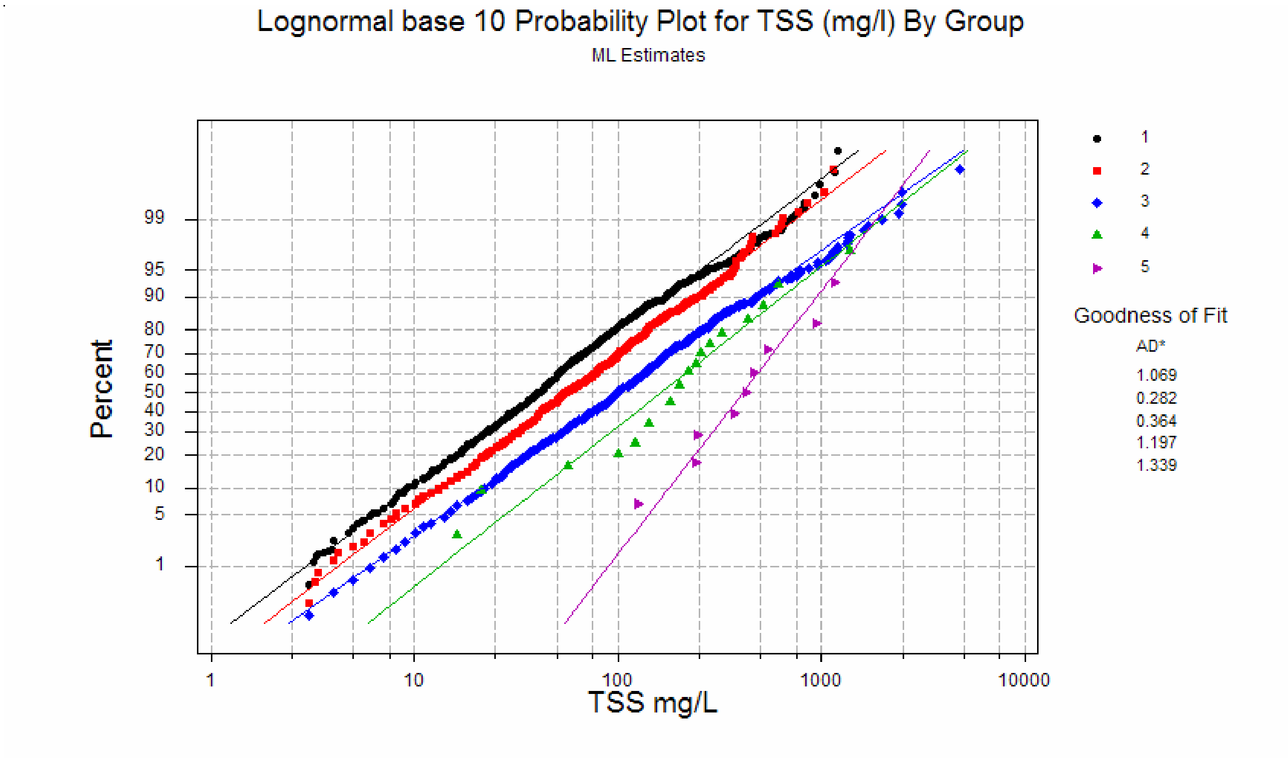


Figure 12. Probability plot of the TSS data in the NSQD using the 5 main sample groups.

The three main groups are clearly defined. Groups 4 and 5 do not have the same numbers of observations as the other groups, so the parameters are not as well described. The upper 5% of the tails in groups 1 and 2 overlap. Group 3 has a slightly larger variance (1.2 vs 1.1) compared with groups 1 and 2. The tails fit the lognormal distribution almost perfectly. The normality test using the Anderson Darling test statistic resulted in a p-value of close to zero for group 1, while for group 2, the p-value was 0.78 and for group 3, the p-value was 0.53. Group 1 fails the normality assumption because of distortion in the upper tail.

Land Use and Geographical Area Effects for All Constituents

This chapter section summarizes the analyses that were conducted to identify significant land use and geographical interactions affecting stormwater concentrations contained in the NSQD. The first step was to select the data for analysis. Only samples collected using flow-weighted automatic samplers were used, in areas not having detention ponds. Also, no sites having only a single monitored event were used.

The second step was to select the following single land uses from the NSQD. The following cross-tabulation summarizes the data counts for samples meeting the above selection criteria in the main three land uses being investigated (CO is for commercial areas, ID is for industrial areas, and RE is for residential areas). The other land uses had many instances of few observations in the EPA Rain Zones.

Tabulated statistics: Landuse, EPA_Rain_Zone

Rows: Landuse Columns: EPA_Rain_Zone

	1	2	3	4	5	6	7	8	9	All
CO	3	123	6	16	42	34	41	0	9	274
ID	3	109	16	17	47	70	33	0	3	298
RE	6	331	18	31	71	38	40	7	7	549
All	12	563	40	64	160	142	114	7	19	1121

EPA Rain Zones 1, 8 and 9 do have not enough data observations in each land use group. Therefore, only EPA Rain Zones 2 thru 7 were included in these analyses. A single land use site corresponds to a watershed with a predominant land uses, where other land uses present in the watershed represent 10% or less of the total area. Therefore, these analyses represent stormwater observations from about the southern half of the country, plus the Pacific Northwest. Data in the NSQD are much sparser in the northeastern states, the upper mid west, the northern Great Plains and western mountain states. The initial NSQD data collection efforts focused on the mid-Atlantic and southern states, while additional data also became available in the southwest and west coast states, allowing at least this partial geographical analysis.

The third step was to estimate the non-detected observations using the Cohen's Maximum likelihood method. The estimation was performed site by site; only samples collected at the same location were used to estimate the censored observations.

For these calculations, the General Linear Model (GLM) was used to identify significant two-way interactions between these land uses and the EPA Rain Zones. The associated Minitab file used is: `RECOID NOSINGLE NOPOND AU FLOW.MPJ.` In all cases, an $\alpha = 5\%$ criterion was high-lighted, although all p values are tabulated for comparison.

Significant Land Use and Geographical Interactions affecting MS4 Stormwater Quality

The following tables summarize the most common stormwater constituents and how they are affected by the interaction of land use and geographical area (residential, commercial, and industrial areas only, and for EPA Rain Zones 2 through 7). The small tables summarize the overall statistics for the constituent. The larger tables summarize a similar summary for each land use/geographical area subset of data. Overall land use summaries are also shown. Only data collected with flow-weighted automatic samplers, with no ponds, are used for these summaries, as described above. In addition, left-censored (non-detected) values were substituted using Cohen's Maximum likelihood method. Calculated p values are located at the top of each cell on this matrix describing the probability that the data in the subset is different from the overall set of data. The grayed-out cells represent conditions where the p-value is greater than 0.05, the usually selected critical value for identifying significant differences. The other cells are therefore usually interpreted as being significantly different from the overall conditions. Some of the cells have no observations and are therefore left blank, except for the zero sample size. Also, some cells are highlighted because they have few sites represented (0, 1, or 2 sites). The data in these grayed-out and highlighted cells should therefore be used with caution. Overall land use summary statistics are also shown. These could be used for those cells indicated in gray, and for those cells that have very few observations, depending on the test statistics comparing the different land uses for each pollutant (see Table XX below). These matrices display the interaction terms for geographical area (represented by EPA Rain Zone) and land use, plus the test statistics for the land uses separately. The detailed tests for statistical significance for the individual factors for each constituent are presented in Appendix F and were calculated using the General Linear Model (GLM) available in Minitab.

Table XX shows the calculated p values using the Tukey simultaneous tests and the General Linear Model for the land use effects alone. This can be used to help select the most appropriate data summary statistics to use for a specific situation, if the land use/geographical interaction data is not appropriate (with not significantly different, or too few data). If the individual cell values are not available, this table indicates that:

- Constituents that should clearly be separated by land use: copper, lead, and zinc

• Constituents that clearly did not have any significant differences for different land use categories, therefore use overall values: pH, temperature (obvious seasonal effects), TDS, and TKN
• Constituents where residential data should be separated from commercial plus industrial area data: TSS (possible) and nitrates plus nitrites
• Constituents where it is not clear; conflicts in p values when comparing different combinations of land uses: hardness, oil and grease, BOD_5, COD, ammonia, total P, and dissolved P

Table XX. Probability of Concentration Differences Between Land Use Categories (General Linear Model and Tukey Simultaneous Tests)*

Constituent	Overall Land Use p	p for Resid. vs. Commercial	p for Resid. vs. Industrial	p for Commercial vs. Industrial	Comment
pH	0.20	n/a	0.20	n/a	use overall values
temperature	0.99	1.00	1.00	1.00	use overall values (obvious seasonal effect)
hardness	***0.008***	0.18	0.24	***0.005***	not clear
oil and grease	***0.010***	***0.01***	0.89	0.06	not clear
TDS	0.065	0.15	0.81	0.06	use overall values
TSS	***<0.0001***	0.08	***<0.0001***	0.36	not clear, or resid. vs. commercial plus industrial if willing to accept slightly higher p
BOD_5	***0.002***	***0.005***	1.00	***0.004***	not clear
COD	***0.036***	***0.03***	0.62	0.45	not clear
ammonia	***0.001***	***0.0005***	0.28	0.09	not clear
nitrates plus nitrites	***<0.0001***	***0.0007***	***0.0006***	1.00	resid. vs. commercial plus industrial
TKN	0.30	0.99	0.35	0.42	use overall values
total phosphorus	***0.003***	***0.008***	1.00	***0.005***	not clear
dissolved P	***0.021***	***0.020***	0.37	1.00	not clear
copper	***<0.0001***	***<0.0001***	***<0.0001***	***<0.0001***	use individual land use values
lead	***<0.0001***	***0.0015***	***<0.0001***	***0.021***	use individual land use values
zinc	***<0.0001***	***<0.0001***	***<0.0001***	***0.0007***	use individual land use values

* the high-lighted p values are <0.05, the usual critical value to identify differences between data categories

When examining the detailed land use and seasonal interactions in the following tables, it is clear that some of the constituents do not have many significant interactions in these factors, or that there are too few observations (or sites) represented in the NSQD. In these cases, the above Table XX can be used to help select either the significant land use value, or the overall value. The constituents that have few, if any clear geographical area/land use interactions include: pH (I6), temperature, hardness, oil and grease (I5 and I7), TDS (C2), ammonia (C7), and dissolved P (R2 and R5). The values in the parentheses are the significant interaction terms (the land use and the EPA Rain Zone).

pH

Total Observations	366
Total Sites	34
Mean	7.5
Median	7.5
Standard Deviaton	0.67
Coefficient of Variation	0.09
Minimum	5.3
Maximum	9.9

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.7469	-	-	0.1546	0.4828	-	Residential
	observ.	14	8	0	71	22	25	140
	sites	3	1		6	3	3	16
	mean	7.36	6.79		7.52	7.34	7.32	7.40
	median	7.43	6.50		7.50	7.40	7.20	7.40
	st. dev.	0.39	0.42		0.58	0.56	0.71	0.60
	COV	0.05	0.06		0.08	0.08	0.10	0.08
	min.	6.70	6.50		6.20	6.30	6.20	6.20
	max.	7.87	7.40		9.90	8.30	9.10	9.90
COMMERCIAL	p	-	-	-	0.8281	0.058	0.6767	Commercial
	observ.	0	0	0	41	27	27	95
	sites				2	2	3	7
	mean				7.70	6.74	7.36	7.33
	median				7.70	6.70	7.40	7.40
	st. dev.				0.44	0.63	0.66	0.69
	COV				0.06	0.09	0.09	0.09
	min.				6.90	5.70	5.30	5.30
	max.				9.10	8.60	8.38	9.10
INDUSTRIAL	p	0.8245	0.2146	-	0.2121	0.0056	0.6767	Industrial
	observ.	18	9	0	46	53	5	131
	sites	3	1		3	3	1	11
	mean	7.32	6.61		7.98	7.68	7.34	7.65
	median	7.31	6.50		8.00	7.70	7.40	7.70
	st. dev.	0.53	0.22		0.50	0.72	0.36	0.69
	COV	0.07	0.03		0.06	0.09	0.05	0.09
	min.	6.40	6.50		6.60	5.90	6.80	5.90
	max.	8.65	7.00		9.00	9.30	7.70	9.30

Temperature (^{o}C)

Total Observations	303
Total Sites	22
Mean	17.8
Median	17.0
Standard Deviaton	6.3
Coefficient of Variation	0.35
Minimum	5.00
Maximum	31.5

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	-	-	-	0.6256	0.5235	0.8447	**Residential**
	observ.	0	0	0	71	20	20	111
	sites				6	2	2	10
	mean				18.10	19.54	11.35	17.14
	median				18.00	17.25	10.35	16.50
	st. dev.				5.19	7.12	4.41	6.08
	COV				0.29	0.36	0.39	0.37
	min.				7.50	11.00	5.00	5.00
	max.				29.00	28.50	23.00	29.00
COMMERCIAL	p	-	-	-	0.0087	0.4802	0.1639	**Commercial**
	observ.	0	0	0	41	26	21	88
	sites				2	1	2	5
	mean				16.23	20.87	12.72	16.76
	median				17.00	20.25	11.90	15.65
	st. dev.				4.76	7.13	4.54	6.23
	COV				0.29	0.34	0.36	0.40
	min.				6.50	11.00	5.00	5.00
	max.				24.00	30.00	21.00	30.00
INDUSTRIAL	p	-	-	-	0.0584	0.9675	0.2401	**Industrial**
	observ.	0	0	0	47	52	5	104
	sites				3	3	1	7
	mean				19.47	20.21	9.90	19.38
	median				20.50	17.25	9.00	17.50
	st. dev.				5.84	6.40	2.70	6.36
	COV				0.30	0.32	0.27	0.36
	min.				7.00	11.50	7.00	7.00
	max.				30.00	31.50	14.00	31.50

Hardness (mg/L)

Total Observations	245
Total Sites	20
Mean	35.0
Median	29.0
Standard Deviaton	37.4
Coefficient of Variation	1.07
Minimum	1.90
Maximum	443

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.2531	-	-	0.2531	-	-	**Residential**
	observ.	8	0	0	64	0	26	98
	sites	1			6		2	9
	mean	53.7			34.6		22.5	32.9
	median	38.8			31.3		19.5	30.1
	st. dev.	38.6			13.4		16.6	19.0
	COV	0.72			0.39		0.74	0.63
	min.	12.70			15.00		5.00	5
	max.	130			78		66	130
COMMERCIAL	p	-	-	-	0.4432	-	-	**Commercial**
	observ.	0	0	0	39	0	26	65
	sites				2		2	4
	mean				31.7		18.5	26.4
	median				28.0		16.0	26.0
	st. dev.				11.6		14.6	14.3
	COV				0.37		0.79	0.55
	min.				14.00		1.90	1.9
	max.				70		62	70
INDUSTRIAL	p	0.2301	-	-	0.7488	0.2737	-	**Industrial**
	observ.	11	0	0	43	8	20	82
	sites	1			3	1	2	7
	mean	110.8			43.8	13.1	21.4	44.3
	median	54.9			37.0	5.0	17.5	33.5
	st. dev.	135.8			23.6	12.9	16.2	58.9
	COV	1.23			0.54	0.98	0.76	1.76
	min.	10.00			18.00	5.00	5.50	5
	max.	443			137	34	76	443

Oil and Grease (mg/L)

Total Observations	609
Total Sites	87
Mean	17.2
Median	3.0
Standard Deviaton	130
Coefficient of Variation	7.48
Minimum	0.10
Maximum	2980

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.7009	-	0.5378	0.7734	0.8825	0.2588	**Residential**
	observ.	92	0	31	68	24	38	253
	sites	28		2	6	3	4	43
	mean	4.41		5.26	74.45	9.74	3.21	23.66
	median	2.50		4.00	2.00	3.00	1.60	2.50
	st. dev.	4.97		6.04	369.19	29.98	5.20	193.11
	COV	1.13		1.15	4.96	3.08	1.62	77.24
	min.	0.20		0.90	0.10	0.20	0.20	0.10
	max.	32		31	2980	150	31	2980
COMMERCIAL	p	0.5325	-	0.8903	0.0043	0.4621	0.131	**Commercial**
	observ.	52	0	14	41	29	37	173
	sites	12		1	2	2	4	21
	mean	5.78		7.24	44.50	7.90	3.54	14.95
	median	3.80		4.00	4.00	4.00	2.60	3.60
	st. dev.	5.96		6.94	92.55	11.84	3.21	48.02
	COV	1.03		0.96	2.08	1.50	0.91	13.34
	min.	0.80		0.90	0.10	0.40	0.30	0.10
	max.	36		25	359	60	18	359
INDUSTRIAL	p	0.7869	-	0.5016	0.0014	0.3238	0.0124	**Industrial**
	observ.	41	0	15	43	55	29	183
	sites	13		1	3	3	3	23
	mean	5.28		3.99	26.90	6.17	4.69	10.43
	median	2.50		4.00	2.00	4.00	4.00	3.00
	st. dev.	5.51		3.64	76.52	5.58	3.44	38.13
	COV	1.04		0.91	2.85	0.91	0.73	12.71
	min.	0.40		0.50	0.10	0.30	0.60	0.10
	max.	22		14	408	24	16	408

Total Dissolved Solids (mg/L)

Total Observations	891
Total Sites	78
Mean	115
Median	68.0
Standard Deviaton	416
Coefficient of Variation	3.63
Minimum	3.00
Maximum	11200

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.6452	0.8115	0.0608	0.5094	0.2516	0.1005	**Residential**
	observ.	268	17	31	64	34	37	451
	sites	19	4	2	6	3	4	38
	mean	87.5	92.2	228.9	77.4	108.9	56.4	95.0
	median	61.5	63.0	173.0	63.0	85.0	58.0	67.0
	st. dev.	124.1	62.3	195.9	35.0	64.6	35.5	117.8
	COV	1.42	0.68	0.86	0.45	0.59	0.63	1.76
	min.	11.00	27.00	59.00	33.00	42.00	3.00	3.00
	max.	1700	211	1096	191	280	175	1700
COMMERCIAL	p	0.0529	0.1439	0.0668	0.1313	0.689	0.8431	**Commercial**
	observ.	82	6	15	39	20	37	199
	sites	6	2	1	2	2	4	17
	mean	124.1	53.3	204.0	54.4	117.6	56.8	101.2
	median	61.0	42.0	140.0	47.0	81.5	53.0	62.0
	st. dev.	421.9	33.5	200.3	22.2	90.0	36.2	280.5
	COV	3.40	0.63	0.98	0.41	0.77	0.64	4.52
	min.	23.00	29.00	64.00	23.00	33.00	4.00	4.00
	max.	3860	120	880	120	383	124	3860
INDUSTRIAL	p	0.0938	0.1045	0.0004	0.3216	0.4684	0.0941	**Industrial**
	observ.	86	16	16	43	56	24	241
	sites	9	3	1	3	4	3	23
	mean	259.2	112.3	126.3	85.6	134.4	76.7	162.5
	median	56.5	101.6	88.0	71.0	127.0	70.5	77.0
	st. dev.	1232.4	61.5	115.9	46.5	67.0	44.1	739.0
	COV	4.76	0.55	0.92	0.54	0.50	0.58	9.60
	min.	4.50	14.00	30.00	27.00	16.00	7.00	4.50
	max.	11200	224	524	238	373	154	11200

Total Suspended Solids (mg/L)

Total Observations	979
Total Sites	103
Mean	133
Median	54.0
Standard Deviaton	258
Coefficient of Variation	1.94
Minimum	0.43
Maximum	2490

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0046	0.5206	0.503	0.0198	0.3101	0.3709	**Residential**
	observ.	309	18	30	64	33	37	491
	sites	32	4	2	6	3	4	51
	mean	64.6	37.3	376.5	102.1	112.4	75.1	91.5
	median	43.0	20.5	188.5	72.0	95.0	49.0	51.0
	st. dev.	79.8	43.3	559.0	105.4	80.2	120.5	176.3
	COV	1.24	1.16	1.48	1.03	0.71	1.60	3.46
	min.	2.55	4.20	11.00	4.00	3.00	7.00	2.55
	max.	823	180	2462	608	350	757	2462
COMMERCIAL	p	0.4554	0.0058	<0.0001	<0.0001	0.0011	0.1118	**Commercial**
	observ.	113	6	15	40	20	37	231
	sites	12	2	1	2	2	4	23
	mean	59.0	76.7	829.2	67.3	111.8	86.4	119.8
	median	35.0	84.5	609.0	29.5	100.5	50.0	47.0
	st. dev.	79.2	38.1	579.1	109.9	109.1	92.3	251.4
	COV	1.34	0.50	0.70	1.63	0.98	1.07	5.35
	min.	6.70	25.00	176.00	2.00	8.00	8.00	2.00
	max.	629	110	2385	640	510	380	2385
INDUSTRIAL	p	0.0929	0.0039	<0.0001	<0.0001	<0.0001	0.0206	**Industrial**
	observ.	101	16	16	43	57	24	257
	sites	15	3	1	3	4	3	29
	mean	68.9	69.1	335.9	244.9	503.9	182.7	222.1
	median	42.0	27.5	176.0	167.0	297.0	114.5	92.0
	st. dev.	66.9	90.2	395.9	387.2	504.6	222.4	354.4
	COV	0.97	1.31	1.18	1.58	1.00	1.22	3.85
	min.	3.00	0.43	21.00	24.00	16.00	16.00	0.43
	max.	330	320	1183	2490	2325	1080	2490

Biochemical Oxygen Demand, 5-day (BOD_5) (mg/L)

Total Observations	901
Total Sites	91
Mean	15.5
Median	9.0
Standard Deviaton	21.5
Coefficient of Variation	1.39
Minimum	0.75
Maximum	270

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0265	0.3419	0.0007	0.0014	0.7423	<0.0001	**Residential**
	observ.	290	16	31	67	15	37	456
	sites	25	4	2	6	3	4	44
	mean	14.7	9.2	19.6	9.7	32.7	7.0	14.0
	median	11.0	8.5	14.0	7.6	22.0	5.0	9.4
	st. dev.	17.0	6.5	15.2	8.6	30.4	7.7	16.3
	COV	1.16	0.71	0.77	0.89	0.93	1.10	1.75
	min.	2.00	2.00	2.00	2.40	10.00	1.34	1.34
	max.	226	28	69	50	130	41	226
COMMERCIAL	p	0.7121	0.0021	0.0403	0.0058	0.9511	<0.0001	**Commercial**
	observ.	114	6	15	40	13	37	225
	sites	9	2	1	2	2	4	20
	mean	15.8	29.2	22.3	6.9	46.0	10.1	15.8
	median	13.0	21.5	17.0	7.3	43.0	6.4	10.1
	st. dev.	13.5	28.8	15.4	3.1	30.3	9.8	16.5
	COV	0.85	0.99	0.69	0.44	0.66	0.97	1.63
	min.	2.93	7.30	4.00	2.50	9.00	0.75	0.75
	max.	96	83	57	17	100	42	100
INDUSTRIAL	p	0.1136	0.0031	<0.0001	0.9638	0.6756	<0.0001	**Industrial**
	observ.	97	16	16	46	21	24	220
	sites	13	3	1	3	4	3	27
	mean	12.5	7.1	6.0	6.4	56.6	43.7	18.0
	median	8.0	7.2	5.5	6.4	21.0	29.0	8.0
	st. dev.	17.7	4.9	3.3	2.3	72.2	38.8	32.4
	COV	1.41	0.69	0.55	0.36	1.27	0.89	4.05
	min.	1.24	1.00	2.00	2.00	1.40	4.00	1.00
	max.	147	19	14	13	270	160	270

Chemical Oxygen Demand (COD) (mg/L)

Total Observations	849
Total Sites	72
Mean	92.9
Median	60.0
Standard Deviaton	96.3
Coefficient of Variation	1.04
Minimum	2.00
Maximum	906

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0451	0.6485	-	<0.0001	0.0813	0.0007	**Residential**
	observ.	265	16	0	69	37	37	424
	sites	19	4		6	3	4	36
	mean	77.4	47.5		80.7	157.5	47.4	81.2
	median	60.0	37.0		60.0	130.0	31.0	59.5
	st. dev.	69.9	34.2		69.2	95.4	54.5	74.5
	COV	0.90	0.72		0.86	0.61	1.15	1.25
	min.	5.00	6.00		10.00	32.00	9.00	5.00
	max.	620	140		480	370	300	620
COMMERCIAL	p	0.7521	0.0396	-	0.0238	0.7645	0.0274	**Commercial**
	observ.	82	6	0	41	33	37	199
	sites	6	2		2	2	4	16
	mean	92.9	93.8		50.7	225.7	58.3	99.8
	median	67.5	69.0		45.0	200.0	41.0	63.0
	st. dev.	91.1	78.8		28.6	127.1	54.8	101.8
	COV	0.98	0.84		0.56	0.56	0.94	1.62
	min.	8.00	23.00		14.09	77.00	8.00	8.00
	max.	635	240		150	582	330	635
INDUSTRIAL	p	0.0344	0.0311	-	0.0117	0.1328	<0.0001	**Industrial**
	observ.	85	16	0	45	56	24	226
	sites	8	3		3	3	3	20
	mean	66.3	42.9		48.9	242.6	104.2	108.9
	median	54.0	35.0		35.0	225.0	84.5	60.5
	st. dev.	51.5	28.8		39.7	165.7	75.8	121.7
	COV	0.78	0.67		0.81	0.68	0.73	2.01
	min.	2.00	4.00		9.55	18.00	18.00	2.00
	max.	340	116		250	906	284	906

Ammonia (NH_3) (mg/L)

Total Observations	588
Total Sites	45
Mean	0.54
Median	0.31
Standard Deviaton	0.73
Coefficient of Variation	1.34
Minimum	0.00
Maximum	7.80

		2	3	4	5	6	7	Overall
RESIDENTIAL	**p**	0.3204	-	-	-	0.6461	0.238	**Residential**
	observ.	262	0	0	0	20	26	308
	sites	19				1	3	23
	mean	0.33				1.10	0.48	0.40
	median	0.25				0.96	0.11	0.26
	st. dev.	0.28				0.72	1.14	0.49
	COV	0.84				0.66	2.40	1.89
	min.	0.02				0.19	0.00	0.00
	max.	1.49				3.40	5.60	5.60
COMMERCIAL	**p**	0.793	-	-	-	0.0128	0.0079	**Commercial**
	observ.	82	0	0	0	23	25	130
	sites	6				1	3	10
	mean	0.56				2.30	0.42	0.84
	median	0.42				1.70	0.14	0.48
	st. dev.	0.49				1.62	0.84	1.09
	COV	0.88				0.70	1.99	2.30
	min.	0.04				0.92	0.01	0.01
	max.	2.51				7.80	4.20	7.80
INDUSTRIAL	**p**	0.2238	-	-	-	0.0011	0.0003	**Industrial**
	observ.	81	0	0	0	51	18	150
	sites	8				2	2	12
	mean	0.34				1.05	0.43	0.59
	median	0.28				0.83	0.30	0.35
	st. dev.	0.28				0.88	0.47	0.66
	COV	0.85				0.83	1.10	1.89
	min.	0.02				0.03	0.03	0.02
	max.	1.60				5.20	1.70	5.20

Nitrites and Nitrates ($NO_2 + NO_3$) (mg/L)

Total Observations	904
Total Sites	75
Mean	0.78
Median	0.60
Standard Deviaton	0.72
Coefficient of Variation	0.92
Minimum	0.02
Maximum	7.30

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	<0.0001	<0.0001	-	<0.0001	<0.0001	-	**Residential**
	observ.	302	7	0	69	38	36	452
	sites	23	1		6	3	4	37
	mean	0.62	1.61		0.82	1.24	0.81	0.73
	median	0.49	1.29		0.66	1.15	0.75	0.60
	st. dev.	0.44	1.31		0.85	0.54	0.66	0.60
	COV	0.71	0.82		1.03	0.44	0.82	1.01
	min.	0.04	0.08		0.11	0.50	0.08	0.04
	max.	2.90	3.16		7.17	2.96	3.50	7.17
COMMERCIAL	p	<0.0001	-	-	0.001	0.0021	<0.0001	**Commercial**
	observ.	114	0	0	40	30	37	221
	sites	9			2	2	4	17
	mean	0.77			0.54	1.26	0.39	0.73
	median	0.55			0.52	1.00	0.28	0.52
	st. dev.	0.88			0.24	0.91	0.44	0.78
	COV	1.14			0.45	0.72	1.11	1.51
	min.	0.02			0.15	0.06	0.08	0.02
	max.	7.30			1.13	3.90	2.60	7.30
INDUSTRIAL	p	0.0656	0.0093	-	0.1978	0.7055	<0.0001	**Industrial**
	observ.	94	9	0	46	58	24	231
	sites	11	1		3	3	3	21
	mean	0.67	0.44		0.68	1.86	0.21	0.92
	median	0.59	0.24		0.67	1.70	0.17	0.68
	st. dev.	0.49	0.53		0.30	1.01	0.18	0.84
	COV	0.73	1.20		0.44	0.54	0.84	1.23
	min.	0.03	0.05		0.07	0.46	0.02	0.02
	max.	2.36	1.66		1.69	4.70	0.70	4.70

Total Kjeldahl Nitrogen (TKN) (mg/L)

Total Observations	936
Total Sites	83
Mean	2.03
Median	1.32
Standard Deviaton	2.30
Coefficient of Variation	1.13
Minimum	0.00
Maximum	36.0

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.7294	0.3106	-	<0.0001	0.3109	<0.0001	**Residential**
	observ.	302	17	0	69	37	37	462
	sites	23	4		6	3	4	40
	mean	1.76	1.19		2.07	3.90	1.34	1.92
	median	1.32	1.00		1.40	3.50	0.70	1.33
	st. dev.	2.39	0.48		1.80	2.37	2.10	2.32
	COV	1.36	0.40		0.87	0.61	1.56	1.75
	min.	0.10	0.56		0.20	0.50	0.00	0.00
	max.	36.00	2.39		10.00	11.00	11.60	36.00
COMMERCIAL	p	0.0944	0.4593	-	0.0687	0.6993	0.2111	**Commercial**
	observ.	114	6	0	40	34	37	231
	sites	9	2		2	2	4	19
	mean	1.94	1.27		1.14	4.41	1.46	2.07
	median	1.44	0.97		0.87	3.65	0.80	1.39
	st. dev.	1.51	0.86		0.83	2.84	1.59	1.96
	COV	0.78	0.67		0.72	0.64	1.09	1.41
	min.	0.05	0.72		0.26	1.04	0.02	0.02
	max.	8.68	3.00		4.00	12.00	8.30	12.00
INDUSTRIAL	p	0.0382	0.0556	-	0.0825	0.5193	<0.0001	**Industrial**
	observ.	94	15	0	46	64	24	243
	sites	11	3		3	4	3	24
	mean	1.31	0.84		1.05	4.66	2.09	2.19
	median	0.98	0.80		0.83	3.85	1.70	1.20
	st. dev.	1.28	0.48		0.80	3.54	1.30	2.54
	COV	0.97	0.57		0.77	0.76	0.62	2.12
	min.	0.16	0.05		0.10	0.25	0.01	0.01
	max.	10.00	1.77		5.00	16.00	5.90	16.00

Total Phosphorus (P) (mg/L)

Total Observations	941
Total Sites	86
Mean	0.43
Median	0.30
Standard Deviaton	0.52
Coefficient of Variation	1.21
Minimum	0.01
Maximum	7.90

		2	3	4	5	6	7	Overall
RESIDENTIAL	**p**	0.003	0.8219	-	<0.0001	0.0038	<0.0001	**Residential**
	observ.	309	18	0	69	38	37	471
	sites	25	4		6	3	4	42
	mean	0.39	0.35		0.49	0.66	0.29	0.41
	median	0.31	0.16		0.43	0.49	0.20	0.33
	st. dev.	0.28	0.62		0.22	0.77	0.38	0.37
	COV	0.72	1.81		0.45	1.16	1.32	1.12
	min.	0.05	0.07		0.19	0.14	0.04	0.04
	max.	1.89	2.78		1.12	4.96	2.20	4.96
COMMERCIAL	**p**	0.022	0.1594	-	0.0002	0.2391	0.6938	**Commercial**
	observ.	114	6	0	40	34	37	231
	sites	9	2		2	2	4	19
	mean	0.30	0.23		0.27	0.53	0.38	0.34
	median	0.22	0.25		0.11	0.48	0.27	0.23
	st. dev.	0.24	0.15		0.66	0.34	0.53	0.41
	COV	0.82	0.62		2.47	0.65	1.40	1.80
	min.	0.06	0.07		0.02	0.16	0.02	0.02
	max.	1.75	0.46		4.27	2.00	3.30	4.27
INDUSTRIAL	**p**	<0.0001	0.1215	-	0.0013	<0.0001	<0.0001	**Industrial**
	observ.	96	15	0	45	59	24	239
	sites	13	3		3	3	3	25
	mean	0.27	0.24		0.25	1.32	0.61	0.56
	median	0.22	0.17		0.18	1.10	0.59	0.29
	st. dev.	0.23	0.24		0.22	1.23	0.32	0.78
	COV	0.85	1.00		0.89	0.93	0.53	2.70
	min.	0.01	0.03		0.05	0.14	0.06	0.01
	max.	1.29	0.90		1.28	7.90	1.40	7.90

Dissolved Phosphorus (dissolved P) (mg/L)

Total Observations	696
Total Sites	56
Mean	0.18
Median	0.12
Standard Deviaton	0.18
Coefficient of Variation	1.02
Minimum	0.01
Maximum	1.60

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0103	0.5914	-	<0.0001	0.0607	0.002	**Residential**
	observ.	255	14	0	69	20	8	366
	sites	17	4		6	1	2	30
	mean	0.19	0.08		0.29	0.28	0.03	0.20
	median	0.14	0.06		0.27	0.21	0.02	0.17
	st. dev.	0.16	0.07		0.15	0.15	0.03	0.16
	COV	0.86	0.92		0.53	0.56	0.97	0.97
	min.	0.01	0.03		0.02	0.11	0.01	0.01
	max.	1.07	0.26		0.84	0.70	0.09	1.07
COMMERCIAL	p	0.1317	0.0324	-	<0.0001	0.0056	0.8748	**Commercial**
	observ.	73	6	0	40	26	7	152
	sites	4	2		2	1	2	11
	mean	0.10	0.10		0.05	0.42	0.04	0.14
	median	0.07	0.06		0.05	0.39	0.03	0.07
	st. dev.	0.09	0.10		0.03	0.33	0.03	0.20
	COV	0.91	0.95		0.67	0.79	0.75	2.84
	min.	0.01	0.03		0.01	0.05	0.01	0.01
	max.	0.36	0.25		0.17	1.60	0.09	1.60
INDUSTRIAL	p	0.4974	0.0618	-	0.0636	0.3767	0.0194	**Industrial**
	observ.	67	11	0	46	52	2	178
	sites	6	3		3	2	1	15
	mean	0.10	0.04		0.09	0.29	0.09	0.15
	median	0.08	0.03		0.06	0.21	0.09	0.09
	st. dev.	0.08	0.03		0.08	0.26	0.07	0.18
	COV	0.75	0.64		0.89	0.90	0.83	1.99
	min.	0.01	0.03		0.02	0.05	0.04	0.01
	max.	0.45	0.10		0.39	1.50	0.14	1.50

Total Copper (Cu) (µg/L)

Total Observations	746
Total Sites	102
Mean	32.4
Median	14.0
Standard Deviaton	72.2
Coefficient of Variation	2.23
Minimum	0.71
Maximum	1360

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0017	0.0006	<0.0001	0.1063	0.0238	0.0401	**Residential**
	observ.	140	17	31	68	38	37	331
	sites	31	4	2	6	3	4	50
	mean	16.1	21.4	33.0	9.4	18.2	12.1	16.4
	median	10.0	5.0	20.0	8.0	9.5	8.0	10.0
	st. dev.	25.6	26.2	27.0	8.1	29.8	14.2	23.5
	COV	1.59	1.22	0.82	0.86	1.64	1.17	2.35
	min.	0.71	5.00	7.00	1.24	1.40	1.63	0.71
	max.	240	102	103	63	180	81	240
COMMERCIAL	p	0.0199	0.3758	0.0022	0.1283	0.0007	0.6792	**Commercial**
	observ.	66	6	15	38	34	37	196
	sites	12	2	1	2	2	4	23
	mean	18.5	9.7	142.8	13.3	17.7	25.7	28.0
	median	17.0	9.5	130.0	8.5	14.0	18.0	14.4
	st. dev.	9.7	4.5	99.1	16.9	14.6	25.2	45.4
	COV	0.52	0.47	0.69	1.27	0.83	0.98	3.15
	min.	1.70	5.00	5.00	2.00	1.50	3.00	1.50
	max.	49	17	384	82	63	130	384
INDUSTRIAL	p	<0.0001	0.0242	0.7863	0.8903	<0.0001	0.1397	**Industrial**
	observ.	54	15	15	44	66	25	219
	sites	15	3	1	3	4	3	29
	mean	19.5	20.0	257.4	21.2	90.6	46.6	60.7
	median	15.0	14.0	118.0	17.5	89.5	31.0	25.0
	st. dev.	17.0	18.3	358.0	13.3	73.4	34.1	117.9
	COV	0.87	0.91	1.39	0.63	0.81	0.73	4.72
	min.	2.20	2.00	10.00	4.00	2.00	11.00	2.00
	max.	87	59	1360	87	340	120	1360

Total Lead (Pb) (µg/L)

Total Observations	725
Total Sites	100
Mean	41.2
Median	16.0
Standard Deviaton	85.8
Coefficient of Variation	2.08
Minimum	0.13
Maximum	1200

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0233	0.1786	0.0013	0.4	0.4616	0.1456	**Residential**
	observ.	135	15	31	68	30	37	316
	sites	30	4	2	6	3	4	49
	mean	17.1	39.2	26.0	16.7	32.9	19.9	20.8
	median	7.5	5.7	14.0	11.5	24.0	11.0	10.5
	st. dev.	37.4	114.1	39.5	15.8	39.5	35.5	41.3
	COV	2.18	2.91	1.52	0.95	1.20	1.78	3.93
	min.	0.46	1.00	1.00	1.68	0.13	0.60	0.13
	max.	368	450	219	89	190	210	450
COMMERCIAL	p	0.2046	0.811	0.1062	0.9323	<0.0001	0.2779	**Commercial**
	observ.	66	6	15	39	34	37	197
	sites	12	2	1	2	2	4	23
	mean	20.8	11.1	96.5	30.4	15.8	42.2	31.3
	median	13.0	10.9	80.0	16.0	10.0	22.0	16.0
	st. dev.	24.4	4.1	69.1	49.6	14.8	58.8	46.1
	COV	1.17	0.37	0.72	1.63	0.93	1.39	2.88
	min.	0.21	5.00	1.00	1.57	3.00	3.00	0.21
	max.	138	16	219	300	77	290	300
INDUSTRIAL	p	0.0017	0.132	0.2569	0.3919	<0.0001	0.7662	**Industrial**
	observ.	48	10	15	45	69	25	212
	sites	14	3	1	3	4	3	28
	mean	22.5	16.2	214.1	37.0	142.7	47.0	80.9
	median	7.5	15.0	72.0	25.0	87.0	30.0	33.0
	st. dev.	30.8	13.9	309.6	51.5	148.0	41.4	135.7
	COV	1.37	0.85	1.45	1.39	1.04	0.88	4.11
	min.	0.56	1.00	2.00	2.23	2.00	4.00	0.56
	max.	130	40	1200	270	620	170	1200

Total Zinc (Zn) (µg/L)

Total Observations	752
Total Sites	102
Mean	207
Median	110
Standard Deviaton	370
Coefficient of Variation	1.79
Minimum	1.11
Maximum	8100

		2	3	4	5	6	7	Overall
RESIDENTIAL	p	0.0378	0.8403	0.2289	0.0036	0.0487	0.0879	**Residential**
	observ.	140	17	31	68	38	37	331
	sites	31	4	2	6	3	4	50
	mean	73	69	199	72	220	83	102
	median	56	54	130	60	155	58	70
	st. dev.	67	66	286	40	249	109	146
	COV	0.93	0.96	1.44	0.55	1.13	1.31	2.08
	min.	1.11	8.17	10.00	16.00	35.00	6.92	1.11
	max.	532	290	1580	230	1500	650	1580
COMMERCIAL	p	0.0005	0.3658	0.1004	0.0091	0.0131	0.0655	**Commercial**
	observ.	66	6	15	39	34	37	197
	sites	12	2	1	2	2	4	23
	mean	199	121	463	98	235	160	196
	median	150	106	410	63	205	92	140
	st. dev.	156	55	297	97	139	173	182
	COV	0.78	0.45	0.64	0.99	0.59	1.08	1.30
	min.	52.00	71.00	46.00	28.00	64.00	39.00	28.00
	max.	892	220	930	560	660	920	930
INDUSTRIAL	p	0.084	0.1843	0.5258	0.9216	0.4551	0.001	**Industrial**
	observ.	54	15	15	45	70	25	224
	sites	15	3	1	3	4	3	29
	mean	204	144	521	225	505	660	370
	median	173	88	320	128	425	332	255
	st. dev.	141	118	488	253	313	1566	599
	COV	0.69	0.82	0.94	1.13	0.62	2.37	2.35
	min.	10.00	25.00	120.00	28.00	70.81	47.00	10.00
	max.	550	420	1590	1400	1400	8100	8100

Chapter 8: Example Application of the National Stormwater Quality Database (TSS and Nutrient Export Calculations for Chesapeake Bay Watersheds)

Overview

This chapter is a demonstration of how the data contained in the NSQD can be used, especially in conjunction with additional urban area flow data, and rural runoff data to estimate the relative contributions of pollutants in a region. This chapter first summarizes the data used, the statistical tests performed, and the results obtained, as part of our effort to identify the most appropriate nonpoint source runoff characteristics for the Chesapeake Bay watershed, the area having most of the collected data in the NSQD.

Data Availability

Two sources of data were used to estimate nonpoint sources of pollution. The first data source corresponding to discharges from urban areas was obtained from the NSQD for the area. The second data source corresponding to discharges from agricultural land uses and forested land cover was obtained from regional data summaries provided by the EPA's Chesapeake Bay Program, "Smart Growth" project group (Office of Policy, Economics, and Innovation).

Urban Data

Data from within the Chesapeake Bay watershed, as contained in the National Stormwater Quality Database (NSQD version 1.1), were used to determine the most appropriate concentrations for urban stormwater nutrients and suspended solids. The NSQD contains information of stormwater discharge concentrations for 19 counties in Virginia and Maryland (Table 48). More than 1,300 events were monitored in these areas representing residential, commercial and industrial land uses. There were no data reported for open space or freeway land uses. The watersheds monitored in Maryland and Virginia ranged from 3.5 and 882 acres and were between 7 and 90% impervious. Reported events used in these analyses were monitored from October 1990, through December 2000.

Table 48. Urban Monitoring Locations in the Chesapeake Bay Watershed Represented in the National Stormwater Quality Database (NSQD, version 1.1)

Virginia	Maryland
Arlington County	Hartford County
Norfolk County	Baltimore County
Virginia Beach County	Baltimore City
Chesapeake County	Carroll County
Portsmouth County	Howard County
Hampton County	Anne Arundel County
Newport News County	Price George's County
Henrico County	Charles County
Chesterfield County	Montgomery County
Fairfax County	

Data for total nitrogen (the sum of total Kjeldahl nitrogen, TKN, and nitrite plus nitrate, NO_2+NO_3), total phosphorus, and total suspended solids (TSS) were evaluated for use in the Chesapeake Bay watershed.

Rural Data

Chesapeake Bay rural water quality information was reported by the USGS in: *Synthesis of Nutrient and Sediment Data for Watersheds within the Chesapeake Bay Drainage Basin* (Langland, *et al.* 1995), prepared in corporation with the EPA. This report describes the comprehensive database of nutrient and sediment data collected from 1972 through 1992 from 1,058 non-tidal monitoring stations in the Chesapeake Bay watershed. Annual discharge loads were calculated at 48 locations for total nitrogen, at 99 locations for total phosphorus, and at 33 locations for suspended sediment. Many of the stations did not have sufficient samples, or flow data to enable load calculations. The fewer locations available for suspended sediment reflect those stations that evaluated suspended sediment, and not suspended solids. Gray, *et al*. (2000) concluded that suspended solids data are not reliable indicators of suspended sediment due to the laboratory processing associated with TSS analyses. The typical pipetting, or pouring, of a subsample for gravimetric analyses typically under predicts the mass associated with sand-sized particles (> 63 micrometers). If cone or churn splitters were used, then the TSS analyses were found to be reasonable. They also found that the results using the two methods are comparable if the mass of these larger particles comprise less than about 25% of the total sample mass. Since no particle size data was available for the TSS samples, they only used information for locations that had total sediment concentrations. Outfall urban runoff samples typically have less than 20% sand, although some early season samples in northern areas where sand is used for traction control may periodically have close to 50% sand, and some source area samples can also have large sand fractions. The TSS values used in the urban component of the analyses, described previously, are expected to be acceptable, as Chesapeake Bay region samples should not be influenced by appreciable winter sand applications, and these are all outfall samples.

Langland, *et al.* (1995) calculated annual nutrient and sediment loads for the selected locations using an unbiased log-linear regression model. This model enabled them to extrapolate the results to annual conditions, and to recognize both base flow conditions (groundwater recharge to the rivers is a major nitrate source, for example) and higher flows associated with surface runoff during storm periods. This analysis also enabled them to consider the potential septic tank and atmospheric deposition contributions to the annual soluble nitrogen loads. Numerous correlation analyses of annual yields of sediment and nutrients with respect to land use, physiographic province, and rock type. They found that river basins having larger percentages of agricultural land had larger nutrient and sediment yields, and that basins that were urbanized had substantially less yields. Table 49 shows the amount of each major land use category in the watershed, and in the portions of the major states within the watershed. In all cases, the land is dominated by forest and agricultural lands, with all urban lands making up about 12% for Maryland and 9% for Virginia portions of the watershed.

Table 49. Land Uses in the Chesapeake Bay Watershed (Landland, *et al.* 1995)

	Percent of Bay Basin	Pennsylvania	Maryland	Virginia
Woody (forest)	53.9	62.5	32.6	52.4
Herbaceous (agriculture)	30.6	31.1	31.3	28.3
High intensity urban	0.6	0.3	1.2	0.6
Low intensity urban	4.0	2.9	6.4	4.5
Woody urban	1.1	0.6	1.9	1.4
Herbaceous urban	1.6	0.8	2.7	2.3
Water	7.2	11.1	20.8	9.6
Exposed	0.3	0.7	0.2	0.1
Herbaceous wetland	0.8	0	2.9	0.8
Total	**100**	**100**	**100**	**100**

Langland, *et al.* (1995) used Kendall's tau test to examine simple linear correlations between annual nutrient and sediment yields and land use, physiographic province, and rock type in the river basins above each station where the annual loads were calculate. They found that land use was the most important variable for predicting nutrient and sediment yield from a river basin. The strongest, most significant and most consistent correlations were between nutrient and sediment yields and agricultural land use. Table 50 shows selected annual yield and land use data for ten of the "load" stations evaluated by the USGS in their Chesapeake Bay report (Langland, *et al.* 1995).

Unfortunately, they did not determine the unit area yields corresponding to separate land uses. They presented these stations as representing the range of land uses for separate locations.

Table 50. Reported Mean Annual Yields and Land Use (Landland, *et al.* 1995)

		Percentage Land Use			Mean Annual Yields (lb/acre/year)	
Basin	Area (mi^2)	Urban	Agriculture	Forest	Total Nitrogen (TKN plus nitrates)	Total Phosphorus
All 127 "load" basins					6.8	0.70
Predominantly Urban Basins						
01571000	11.2	46.0	26.8	28	8.2	0.80
01589300	32.5	54.4	16.8	27	8.1	na
01593500	38.0	54.5	16.8	21	na	0.67
01646000	57.9	50.9	11.4	28	5.9	0.61
01657655	4.0	48.6	22.5	27	na	0.28
Agricultural and Urban Basins						
01586000	56.6	42.4	51.0	3.4	na	0.41
01616000	16.5	43.8	41.9	13	29.7	4.0
Predominantly Agricultural Basins						
01573810	0.38	1.4	91.0	6.7	42.1	6.3
0157608335	1.42	1.1	63.4	26	26.4	4.5
01639500	102	1.1	69.9	29	14.6	na

Summary of Data and Load Calculations

The "simple" model (Schueler 1987) was used to calculate the nonpoint discharges of TSS, total phosphorus, and total nitrogen for Anne Arundel County, Maryland, for the EPA's Chesapeake Bay Program, "Smart Growth" project (Office of Policy, Economics, and Innovation). The simple model was developed by Schueler to enable rapid calculations of pollutant discharges by multiplying the event mean concentration values for a specific land use, the volumetric runoff coefficient for that land use, and the annual rainfall. With appropriate unit conversions, the result can be expressed as the unit area annual discharge for a specific pollutant. When multiplied by the area corresponding to each of the land uses in the area of concern, the total area pollutant discharges can be calculated, and the relative sources of the discharges can be identified. When working with large watersheds, these calculated values are usually much greater than the monitored in-stream values observed at the watershed outlet, because hindered pollutant transport in the stream or river is not considered. However, it is a suitable method to identify the relative pollutant contributions of different land uses in a county, as in this example.

The volumetric runoff coefficients for each land use category were based on analyses of typical land use surface configurations (mostly the impervious area characteristics) and the rain depth was determined from 50 years of rain records from the Baltimore (BWI) airport. The urban area concentration values were obtained by statistical evaluations of the Maryland and Virginia urban area data contained in the National Stormwater Quality Database, as described in the following subsections of this chapter. The urban runoff and concentration factors are assumed to have excellent reliability. However, some of the urban categories were not represented with regional Chesapeake Bay region data, so these factors were obtained from the national averaged values in the database and are labeled with a moderate reliability. The non-urban values are labeled as having poor to very good reliability, depending on the availability of local data. The agriculture values are from regional information summarized by Staver (1995) and Hartigan (1983) and are assumed to be of very good reliability. The forestlands data are from regional Chesapeake Bay regional data collected by the EPA, Office of Policy, Economics, and Innovation (Richards, personal communication) and are assumed to be of moderate reliability. The other land categories and the extraction lands data are unknown and are of poor reliability. Fortunately, as shown on the following summary tables, the best information is associated with the agricultural, forest, and urban categories which are responsible for almost the entire calculated discharges for the county.

The runoff factors are also indicated with varying reliabilities. The urban lands data all have excellent reliability due to the use of calibrated urban data for varying conditions. The agricultural runoff data is of the poorest reliability due to the uncertainties associated with the many agricultural operations that can have dramatic effects on these values. The natural land runoff values are expected to have moderate reliabilities. The USGS (Langland, *et al*. 1995) reported values are not comparable to these discharge values due to a number of reasons, most specifically because they are in-stream values and are affected by sediment and pollutant transport. The USGS report also did not report unit area loadings for specific land uses and the preliminary calculations resulted in unrealistic results that were highly variable. Tables 51, 52 and 53 list the nutrient and suspended solids data applicable for Chesapeake Bay watershed analyses, based on the analyses performed and outlined later in this chapter.

Table 51. Commercial TSS (mg/L), Mean and COV, (a function of season and rain depth)

	<0.1 inches	0.1 to 0.35 inches	0.35 to 1 inch	>1 inch
Spring	18 (0.72)	31 (0.67)	75 (1.5)	no data
Summer	75 (1.5)	18 (0.72)	75 (1.5)	75 (1.5)
Fall	18 (0.72)	75 (1.5)	18 (0.72)	18 (0.72)
Winter	18 (0.72)	75 (1.5)	75 (1.5)	75 (1.5)

Table 51 shows that storm events with precipitation depths larger than 0.35 inches are more likely to discharge higher TSS concentrations in spring, summer and winter than fall. Table 52 shows the expected total nitrogen concentrations in commercial land uses. There is a clear variation among the seasons and precipitation depth. Storm events smaller than 0.1 inch are expected to have higher total nitrogen discharges during the fall and winter than during the summer and spring seasons. For rain events between 0.35 and 1 inch, the highest concentrations were observed during the summer and fall. Table 53 shows the average concentrations and coefficients of variation for TSS, total phosphorus and total nitrogen for residential, commercial and industrial urban land use areas. This table also includes the expected concentrations in agricultural and forested areas.

Table 52. Commercial Total Nitrogen (mg/L), Mean and COV, (a function of season and rain depth)

	<0.1 inch	0.1 to 0.35 inches	0.35 to 1 inch	>1 inch
Spring	2.0 (0.49)	3.2 (0.50)	2.0 (0.49)	no data
Summer	2.0 (0.49)	2.0 (0.49)	3.2 (0.50)	3.2 (0.50)
Fall	3.2 (0.50)	3.2 (0.50)	3.2 (0.50)	2.0 (0.49)
Winter	3.2 (0.50)	2.0 (0.49)	2.0 (0.49)	3.2 (0.50)

The total runoff discharges for the county can be determined based on the calculated total mass discharges for each land use, and the areas for each land use area. Table 54 shows the percentage of total annual runoff volume produced for each land use by season and rain depth range. About 61% of the total annual runoff volume was produced by events having more than 1 inch of rain, followed by rain events in the range 0.36 to 1 inches (31% of the annual runoff volume), rain events in the range of 0.1 to 0.35 inches (7% of the annual runoff volume), and rain events less than 0.1 inch in depth (with 1% of the annual runoff volume).

Table 53. Average Concentrations by Land Use

Land Use	Constituent	Conditions	Average (COV)
Urban - Residential	TSS		
		Summer rains (between 0.1 and 0.35 inches in depth)	143 (0.71)
		All other rains	58 (0.70)
	TP	Sites having <27% impervious cover:	
		Winter rains	0.28 (0.59)
		All other rains	0.41 (0.65)
		Sites having >27% impervious cover:	
		Winter rains (less than 0.1 inches in depth)	0.16 (0.86)
		All other rains	0.30 (0.63)
	TN	Fall rains (less than 0.1 and greater than 1 inch in depth)	1.4 (0.57)
		Winter rains (0.35 and 1 inch in depth)	1.5 (0.30)
		Fall rains (0.35 and 1 inch) and Winter rains (between 0.1 and 0.35 inches in depth)	1.9 (0.51)
		All other rains	2.4 (0.62)
		Spring and summer rains (between 0.1 and 0.35 inches in depth)	2.6 (0.38)
Urban - Commercial	TSS and TN	See tables 7.4 and 7.5	
	TP	Summer rains >1 inch and fall rains between 0.1 and 0.35 inch	0.46 (0.36)
		All other rains	0.23 (0.71)
Urban - Industrial	TSS	Fall, spring, and summer	77 (1.48)
		Winter	81 (0.93)
	TP	Rains less than 0.35 inches	0.29 (0.81)
		Rains greater than 0.35 inches	0.22 (1.05)
	TN	All conditions	2.1 (0.79)
Rural - Agricultural	Sediment		1115 lb/ac/yr (unreliable estimate)
	TN		40 lb/ac/yr
	TP		5.4 lb/ac/yr
Rural - Forest	Sediment		4500 lb/ac/yr (unreliable estimate)
	TN		0 lb/ac/yr
	TP		0 lb/ac/yr

Table 54. Fraction of Annual Flow Associated with Season and Rain Depth Categories (based on 50 years of rain records at Baltimore, BWI)

	rain range	Fraction of rain depth in category	Ultra low density residential	Low density residential	Medium density residential	High density residential	Commercial	Industrial	Institutional	Open urban land	Freeways
Winter	<0.1 inch	0.0078	0.0008	0.0011	0.0013	0.0016	0.0017	0.0014	0.0039	0.0004	0.0022
	0.1 to 0.35	0.0275	0.0109	0.0137	0.0171	0.0193	0.0184	0.0173	0.0190	0.0086	0.0180
	0.36 to 1	0.0917	0.0727	0.0786	0.0855	0.0860	0.0870	0.0874	0.0857	0.0727	0.0846
	>1	0.1025	0.1394	0.1319	0.1230	0.1202	0.1202	0.1212	0.1188	0.1420	0.1220
Spring	<0.1 inch	0.0091	0.0010	0.0013	0.0016	0.0018	0.0020	0.0017	0.0046	0.0005	0.0026
	0.1 to 0.35	0.0387	0.0153	0.0193	0.0240	0.0272	0.0259	0.0243	0.0267	0.0121	0.0253
	0.36 to 1	0.0999	0.0793	0.0857	0.0932	0.0937	0.0948	0.0953	0.0934	0.0792	0.0922
	>1	0.1075	0.1463	0.1383	0.1290	0.1261	0.1261	0.1271	0.1246	0.1489	0.1280
Summer	<0.1 inch	0.0089	0.0010	0.0010	0.0015	0.0018	0.0020	0.0017	0.0045	0.0005	0.0026
	0.1 to 0.35	0.0366	0.0145	0.0183	0.0228	0.0258	0.0245	0.0230	0.0253	0.0115	0.0240
	0.36 to 1	0.0857	0.0680	0.0735	0.0799	0.0804	0.0813	0.0817	0.0801	0.0679	0.0791
	>1	0.1438	0.1955	0.1849	0.1724	0.1686	0.1685	0.1699	0.1665	0.1991	0.1711
Fall	<0.1 inch	0.0064	0.0007	0.0009	0.0011	0.0013	0.0014	0.0012	0.0032	0.0004	0.0018
	0.1 to 0.35	0.0248	0.0098	0.0124	0.0154	0.0174	0.0166	0.0156	0.0171	0.0078	0.0163
	0.36 to 1	0.0691	0.0548	0.0593	0.0645	0.0648	0.0656	0.0659	0.0646	0.0548	0.0638
	>1	0.1391	0.1892	0.1789	0.1668	0.1631	0.1631	0.1644	0.1611	0.1926	0.1655

The flow weighting factors in Table 54 were used with the statistical analyses of the concentration data to obtain calculated long term averaged concentrations for mass loading calculations. Table 55 shows the urban area concentrations developed for Anne Arundel County using the Chesapeake Bay regional data contained in the National Stormwater Quality Database, along with concentrations and runoff quantities for other county land uses.

Table 55. Total Suspended Solids Concentrations and Volumetric Runoff Coefficients for Land Use Categories in Anne Arundel County, Maryland

Land Use Description	# of acres in 2000	TSS (mg/L)	Concentration reliability?	R_V	R_V reliability?
Large lot subdivision (1 unit/ 5- 10 ac)	0	60	excellent	0.09	excellent
Low-density residential (1 unit/ 5 acres to 2 units/acre)	33,337	60	excellent	0.14	excellent
Medium-density residential (2 to 8 units/acre)	33,791	60	excellent	0.23	excellent
High-density residential (8+ units/acre)	6,274	60	excellent	0.34	excellent
Commercial	11,670	58	excellent	0.72	excellent
Industrial	3,249	80	excellent	0.52	excellent
Institutional (schools, churches, military institutions, etc.)	9,813	58	moderate	0.49	excellent
Open urban land	4,139	50	moderate	0.08	excellent
Transportation	1,557	99	moderate	0.41	excellent
Extractive	1,686	350	poor	0.3	moderate
Deciduous forest	43,901	90	moderate	0.08	moderate
Evergreen forest	4,891	90	moderate	0.08	moderate
Mixed forest	56,621	90	moderate	0.08	moderate
Brush	2,565	90	poor	0.08	moderate
Wetlands	1,643	0	poor	0.65	moderate
Beaches	29	0	poor	0.1	moderate
Bare ground	224	1000	poor	0.3	moderate
Row and garden crops	300	357	very good	0.2	poor
Cropland	42,368	357	very good	0.2	poor
Orchards / vineyards / horticulture	63	357	very good	0.15	poor
Pasture	4,690	145	very good	0.08	moderate
Feeding operations	49	145	very good	0.2	poor
Agricultural building, breeding and training facilities	163	145	very good	0.5	poor

Urban land uses produced slightly lower TSS concentrations compared with those observed in forest areas. However the volumetric runoff coefficients for forests are smaller than any other use, except for open urban land, resulting in the likely lowest annual yields. Bare ground, cropland, vineyards, horticulture, row and garden crops and extractive activities have the highest estimated concentrations amongst the land uses examined. Total phosphorus concentrations are presented in Table 56.

In this case, the highest concentrations were assumed for croplands, row and garden crops, orchards, vineyards and horticulture. The lowest concentrations were assumed for forested areas. One order of magnitude separates the minimum and maximum concentrations. This difference can be associated with the use of fertilizers and associated nutrient discharges. For urban areas, industrial and commercial land use areas had lower phosphorus concentrations than residential land use areas. Table 57 shows the average urban area concentrations for long term analyses, based on statistical analyses examining site factors for this regional data. Only phosphorus had different concentrations associated with different site categories that were tested.

Table 56. Total Phosphorus Concentrations for Land Use Categories in Anne Arundel County, Maryland

Land Use Description	# of acres in 2000	TP (mg/L)	Concentration reliability?
Large lot subdivision (1 unit/5- 10 acres)	0	0.38	excellent
Low-density residential (1 unit/ 5 acres to 2 units/acre)	33,337	0.38	excellent
Medium-density residential (2 to 8 units/acre)	33,791	0.3	excellent
High-density residential (8+ units/acre)	6,274	0.3	excellent
Commercial	11,670	0.25	excellent
Industrial	3,249	0.23	excellent
Institutional (schools, churches, military institutions, etc.)	9,813	0.27	moderate
Open urban land	4,139	0.25	moderate
Transportation	1,557	0.25	moderate
Extractive	1,686	0.5	poor
Deciduous forest	43,901	0.1	moderate
Evergreen forest	4,891	0.1	moderate
Mixed forest	56,621	0.1	moderate
Brush	2,565	0.38	poor
Wetlands	1,643	0.38	poor
Beaches	29	0.1	poor
Bare ground	224	0.38	poor
Row and garden crops	300	1.00	very good
Cropland	42,368	1.00	very good
Orchards / vineyards / horticulture	63	1.00	very good
Pasture	4,690	0.38	very good
Feeding operations	49	0.38	very good
Agricultural building, breeding and training facilities	163	0.38	very good

Table 57. Urban Area Stormwater Concentrations

<table>
<tr><th>Land Use</th><th>Constituent</th><th>Conditions</th><th>Average value for long-term analyses (mg/L)</th></tr>
<tr><td rowspan="4">Urban – Residential</td><td>TSS</td><td></td><td>60</td></tr>
<tr><td rowspan="2">TP</td><td>Sites having <27% impervious cover (ultra low and low density areas)</td><td>0.38</td></tr>
<tr><td>Sites having >27% impervious cover (medium and high density areas)</td><td>0.30</td></tr>
<tr><td>TN</td><td></td><td>2.1</td></tr>
<tr><td rowspan="3">Urban – Commercial</td><td>TSS</td><td></td><td>58</td></tr>
<tr><td>TP</td><td></td><td>0.25</td></tr>
<tr><td>TN</td><td></td><td>2.6</td></tr>
<tr><td rowspan="3">Urban – Industrial</td><td>TSS</td><td></td><td>80</td></tr>
<tr><td>TP</td><td></td><td>0.23</td></tr>
<tr><td>TN</td><td></td><td>2.1</td></tr>
</table>

Table 58 shows the summary for total nitrogen. Similar to the total phosphorus case, the largest nitrogen concentrations were predicted for croplands, vineyards, row and garden crops orchards and horticulture activities. The lowest concentrations were observed in open urban land and forested areas. The ratio between largest and smallest concentrations was approximately 2 to 1.

Table 58. Total Nitrogen Calculated Concentrations for Land Use Categories in Anne Arundel County, Maryland

Land Use Description	# of acres in 2000	TN (mg/L)	Concentration reliability?
Large lot subdivision (1 unit/5- 10 acres)	0	2.1	excellent
Low-density residential (1 unit/ 5 acres to 2 units/acre)	33,337	2.1	excellent
Medium-density residential (2 to 8 units/acre)	33,791	2.1	excellent
High-density residential (8+ units/acre)	6,274	2.1	excellent
Commercial	11,670	2.6	excellent
Industrial	3,249	2.1	excellent
Institutional (schools, churches, military institutions, etc.)	9,813	2	moderate
Open urban land	4,139	1.3	moderate
Transportation	1,557	2.3	moderate
Extractive	1,686	1.5	poor
Deciduous forest	43,901	1.5	moderate
Evergreen forest	4,891	1.5	moderate
Mixed forest	56,621	1.5	moderate
Brush	2,565	1.5	poor
Wetlands	1,643	1.5	poor
Beaches	29	1.5	poor
Bare ground	224	1.5	poor
Row and garden crops	300	2.92	very good
Cropland	42,368	2.92	very good
Orchards / vineyards / horticulture	63	2.92	very good
Pasture	4,690	2.2	very good
Feeding operations	49	2.2	very good
Agricultural building, breeding and training facilities	163	2.2	very good

Using the simple model, it is possible to calculate the total annual discharges from these different non point sources. Table 59 shows the total estimated runoff discharged by year, and the total discharges of suspended solids, total nitrogen and total phosphorus for each of the major land use categories. Urban sites produced most of the runoff and total nitrogen, followed by agricultural and forested areas. Half of the total suspended solids were produced by agricultural activities, followed by urban areas (30%), forested areas (12%), and other lands (10%). Urban and agricultural sites combined (in about equal fractions) produced almost 90% of the phosphorus loads. Forested areas only produced about 4% of the total phosphorus annual loads. The remaining phosphorus discharges were produced by other land uses.

Table 59. Discharges by Major Land Use Categories in Anne Arundel County, Maryland

	Runoff Yield		Total Suspended Solids	
	Total Discharge (ft^3/year)	Percent of Total	Total Discharge (kg/year)	Percent of Total
Urban	4.2×10^9	60.6	7.4×10^6	28.7
Agricultural	1.3×10^9	18.7	1.3×10^7	50.0
Forest	1.2×10^9	17.1	3.0×10^6	11.9
Other lands	2.6×10^8	3.7	2.4×10^6	9.4
Total County	7.0×10^9		2.6×10^7	

	Total Phosphorus		Total Nitrogen	
	Total Discharge (kg/year)	Percent of Total	Total Discharge (kg/year)	Percent of Total
Urban	3.6×10^4	46.4	2.6×10^5	60.5
Agricultural	3.6×10^4	45.8	1.1×10^5	25.0
Forest	3.4×10^3	4.3	5.1×10^4	11.9
Other lands	2.8×10^3	3.6	1.1×10^4	2.6
Total County	7.8×10^4		4.3×10^5	

The final values used during for the calculations are summarized in Tables 60 and Table 61. For each of the main land uses, the percentage of impervious areas (indicating the percentage connected and disconnected), the volumetric runoff coefficient and the TSS, total phosphorus, and total nitrogen concentrations are shown. The volumetric runoff coefficients, and curve numbers, were calculated using 50 years of precipitation data from the BWI airport in Baltimore.

Table 60. Urban Land Use Categories Used in Anne Arundel County, Maryland

Description	Note	Average percentage of Impervious areas	TSS (mg/L)	Total P (mg/L)	Total N (mg/L)	Comments
High density residential	Rv = 0.34, 47% pervious, 39.9% dir con imp, and 13.1% dis con impervious	53	60	0.3	2.1	Rv and CN calculated using 50 yrs of BWI rains and concentration factors from MD and VA MS4 data
Medium density residential	Rv = 0.23, 62.3% pervious, 24.2% directly con imp, and 13.5% disconnected impervious	37.8	60	0.3	2.1	Rv and CN using 50 yr BWI rain and concentration factors from MD and VA MS4 data
Low density residential	Rv = 0.14, 79.6% pervious, 14.9% dir con impervious, and 5.5% disconnected imp.	20.4	60	0.38	2.1	1 unit/5 ac to 2 units/ac. Calc Rv and CN using 50 years BWI rains and concentration factors from MD and VA MS4
Ultra low den residential	Rv=0.09, 90.4% pervious, 5.6%directly con imp and 4% discon impervious	9.6	60	0.38	2.1	1 unit/5 to 10 ac, calc 50 yr, concentration factors from MD and VA MS4 data
Freeways and other main roads with paved drainage	Rv = 0.41, 49.5% pervious, 50.5% dir con impervious.	50.5	99	0.25	2.3	Calc using 50 yrs of BWI rains and concentration factors from national MS4 data
Commercial (shopping centers)	Rv=0.72, 8.3% pervious, and 91.7% dir con imp.	91.7	58	0.25	2.6	50 yr of BWI rains and concentration factors from MD and VA MS4 data
Institutional (schools, churches, military, etc.)	Rv=0.49, 36.4% pervious, 61.3%dir con imp, and 2.3%discon imp.	63.6	57.9	0.35	1.57	Calculated from 50 yr BWI rains and concentration factors from national average institutional MS4 data.
Industrial (medium)	Rv=0.52, 16.7% pervious, 62.8% dir con imp, and 20.5% discon con imp.	83.3	80	0.23	2.1	CN calc using 50 yr BWI rain and concentration factors from MD and VA MS4 data.
Open urban area	Rv=0.08, 95.1% pervious and 4.9% dir con impervious.	4.9	70	0.12	1.5	CN calc from 50 yr BWI rains and concentration factors from national average urban open area MS4 data

The land uses having the largest amounts of directly connected impervious surfaces were the commercial, institutional, and industrial land use areas. Urban TSS concentrations ranged between 57 and 99 mg/L, total phosphorus concentrations ranged between 0.12 and 0.40 mg/L, and total nitrogen ranged between 1.5 and 2.6 mg/L. Table 61 shows the summaries for the other land uses.

Table 61. Other Land Use Categories Used in Anne Arundel County Calculations

Description	Note	TSS (mg/L)	Total Phosphorus (mg/L)	Total Nitrogen (mg/L)
Fallow	Straight Row. Concentration factors from prior regional data	107	1.3	4.4
Row Crops	Straight Row, small grain. Concentration factors from prior regional data	357	1	2.92
Row and garden Crops	Straight Row. Concentrations factors from prior regional data	357	1	2.92
Orchards, vineyards, horticultural	Concentration factors from prior regional data	357	1	2.92
Pasture or Range	Concentration factors from prior regional data	145	0.38	2.2
Feeding operations	Continuous forage, poor. Concentration factors from prior regional data	145	0.38	2.2
Woods or Forest Land	Deciduous forest (woods, good). Concentration factors from prior regional data	90	0.1	1.5
Woods or Forest Land	Evergreen forest (woods, good condition). Concentration factors from prior regional data.	90	0.1	1.5
Woods or Forest Land	Mixed forest (woods, good). Concentration factors from prior regional data.	90	0.1	1.5
Farmsteads	Agricultural buildings, breeding and training facilities	163	0.38	2.2
Brush	Herbaceous, fair.	90	0.38	1.5
Extractive		1000	0.38	1.5
Wetlands		0	0.38	1.5
Beaches		0	0.1	1.5
Bare ground		1000	0.38	1.5

The largest TSS concentrations were observed in extractive activities and for bare ground, or exposed soil sites. Land uses where the intensive use of fertilizers is most frequent had the largest total phosphorus and total nitrogen concentrations. The lowest nutrient concentrations were observed in forested areas.

Figure 60 shows the area distributions and the relative contributions for major sources of runoff, total suspended solids, total nitrogen and total phosphorus for sites located in Anne Arundel County in Maryland. Forested and urban land use areas represent almost 80% of the total land uses in the county. About 15% of the area is agricultural and the remaining of 5% is associated with other activities.

Urban land use areas produce almost 65% of the total runoff volume for the county, followed by agricultural and forested areas (about 15% each). As expected, impervious surfaces in urban land use areas were responsible for most of the total discharged runoff volume. Agricultural land uses produce almost half of the total TSS discharges, although they make up only about 15% of the county area. Urban land uses are the second major source of TSS in the county, contributing about 28% of the total annual TSS discharges in the county. Forested areas and other land uses contribute the smallest fractions of the total load, with almost 11% each. Urban and agricultural areas combined produced almost 90% of the total phosphorus load, in about equal percentages. Forested areas and other land uses contribute about 10% of the total countywide phosphorus load. Finally, urban land uses contributed almost 60% of the total nitrogen load for the county, followed by agricultural activities (25%), forested areas (13%) and other land uses (2%).

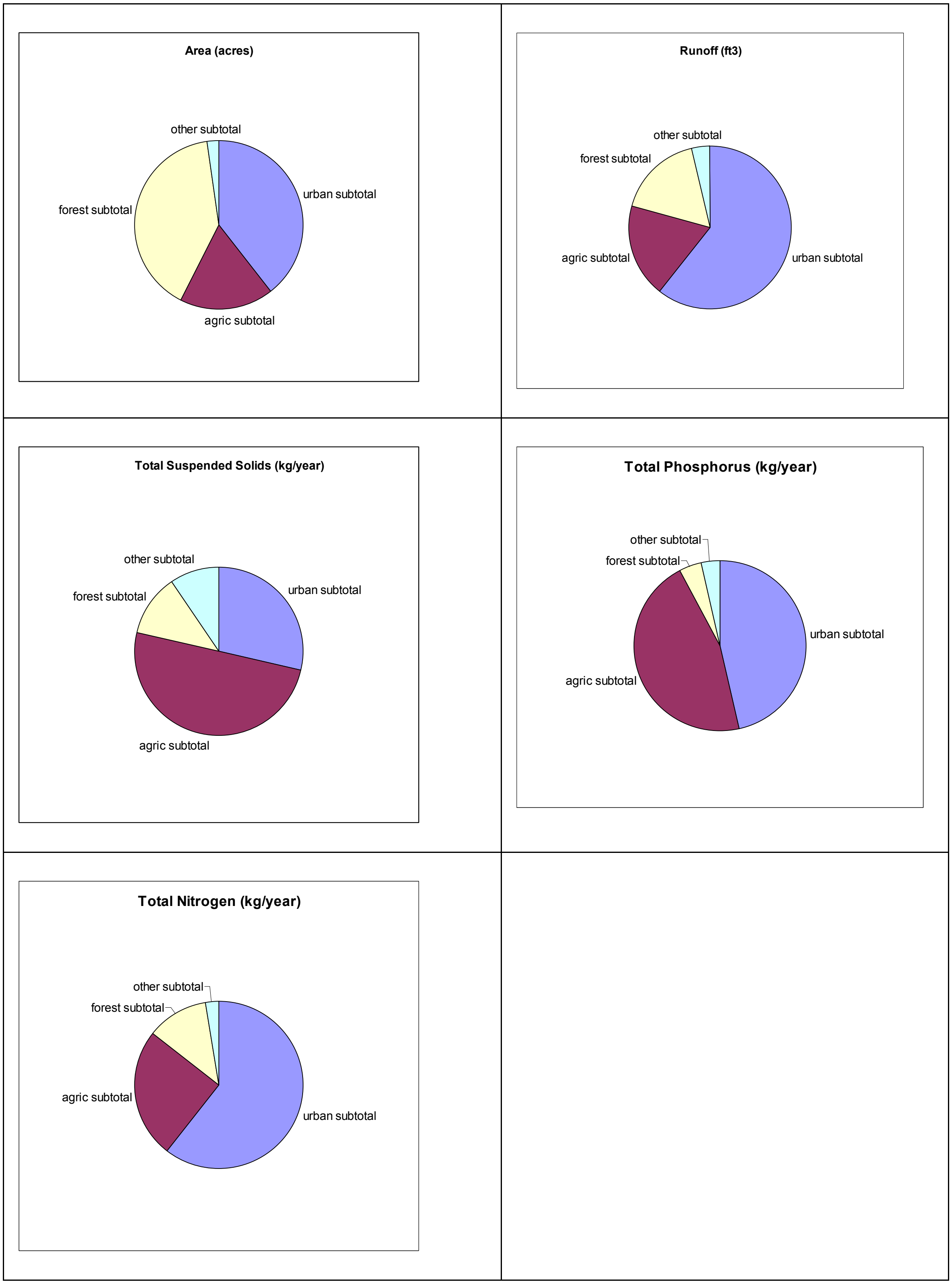


Figure 60. Sources of runoff, TSS and nutrients for different sources in Anne Arundel County

Statistical Analyses Performed

The following discussion describes the statistical analyses performed to identify the different groups in TSS, total nitrogen and total phosphorus by season and precipitations depth for the Chesapeake Bay region data. The objective of the statistical tests was to identify significantly distinct categories of the Chesapeake Bay regional data. Specifically, land use, season, precipitation, percent imperviousness, and watershed drainage area, were considered potentially important factors that would affect the concentration values. In addition, variations of reported concentrations with time were also examined. After appropriate normalization of the data, three-way and two-way ANOVA tests were used to identify the significant factors and interactions between these potentially important factors, while one-way ANOVA tests, along with parametric and nonparametric comparison tests, were used to identify groupings within the range of any one factor. As an example, one-way ANOVA analyses were performed to identify any ranges of percentage of imperviousness that produce different distributions of stormwater constituent concentrations from other ranges, while two-way ANOVA analyses were used to identify any seasonal-total precipitation interactions in the distribution of the stormwater constituent concentrations

Before ANOVA analyses can be conducted, the first step is to examine the data to ensure that it fits a normal probability distribution. If not, the data needs to be transformed. Prior tests (reported in this report) found that most all of the stormwater constituents in the NSQD fit lognormal distributions. In this case, the base 10 logarithm of the original observations adequately followed a normal distribution. Therefore, data from the same population group will fall along the same straight line. Groups in either tail that do not fall on the line can be considered different. This procedure was used in the ANOVA analyses to identify if the concentration values were statistically different for different levels of the factor, or factors, being examined. For example, if the expected values are different for different levels of imperviousness, or different seasons, then those data groupings will not follow the main probability distribution, and the ANOVA test results will indicate a likely significantly different data population. The significant ANOVA coefficients were then used to create a model to predict the concentration values for the different groups. All of the observed conditions within each group will have the same expected concentration value. Once the groups were identified, the mean and standard deviations were calculated from the original observations in the database for each observation in each group, and the data for each group are plotted on probability and box and whisker plots. The following discussion is a detailed description of the tests conducted using the Chesapeake data for total suspended solids.

Residential Area Total Suspended Solids Analyses

The ANOVA tests did not identify any significant groupings for either drainage area, or percentage imperviousness variations. Trends with time since the last rain, and for time since the initiation of the watershed monitoring were also examined, but these analyses did not identify any apparent, or significant, trends for any of the test sites. Initial data evaluations indicated a possible significant variation due to the level of imperviousness in the test watersheds, but when evaluated in conjunction with season and rain depth, these other factors were found to be the only significant factors to describe the variations in TSS concentrations in residential areas. Obviously, the percentage imperviousness values will have a large effect on the amount of runoff volume expected, so the imperviousness will be very important in affecting the mass of pollutants discharged. This is similar to data evaluations for other regions. The Maryland and Virginia data provided a great opportunity to test this hypothesized effect, because there were 13 residential area test watersheds having imperviousness values ranging from 7 to 65% (although most of the data were represented in six watersheds ranging in imperviousness from 20 to 50%). The statistical tests identified two distinct groups of residential TSS data, as represented in the following plots and tables: small summer rains (in the range of 0.1 to 0.35 inches) which had an average TSS concentration of about 143 mg/L, and all other residential conditions which had an average TSS concentration of about 58 mg/L. The following plots and data summaries describe these two data groupings.

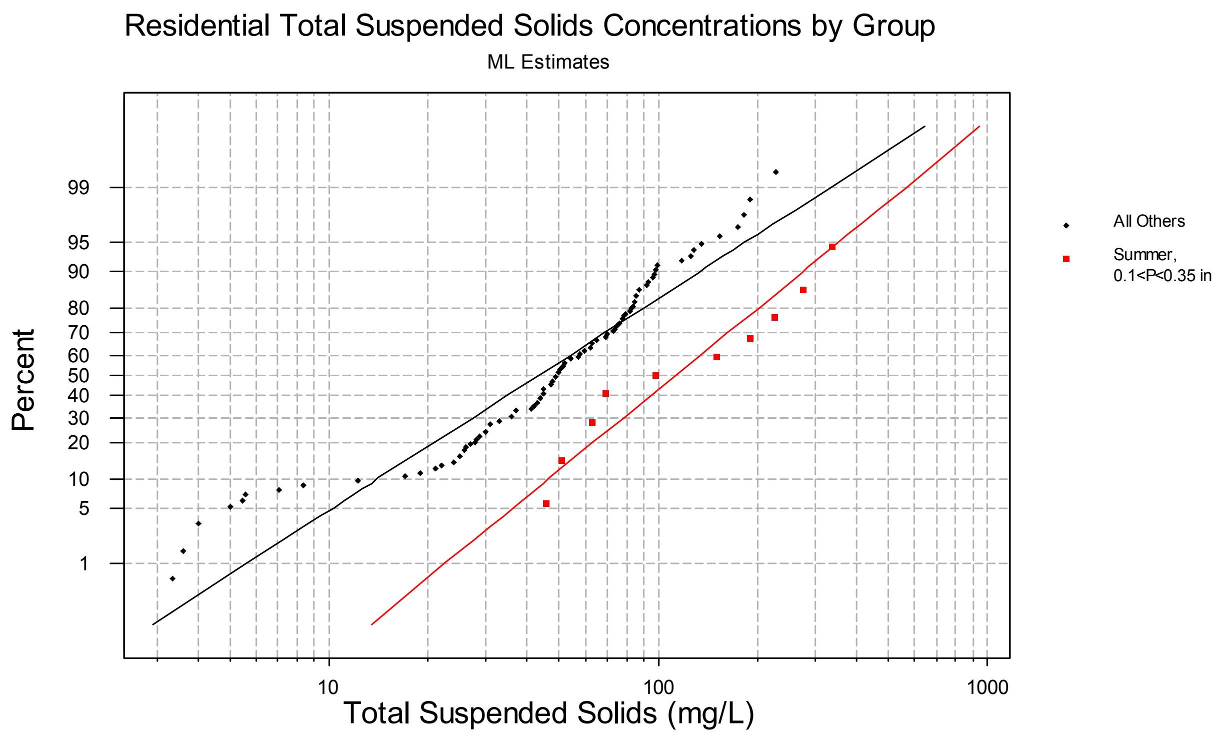


Figure 61. Residential TSS distributions by groups

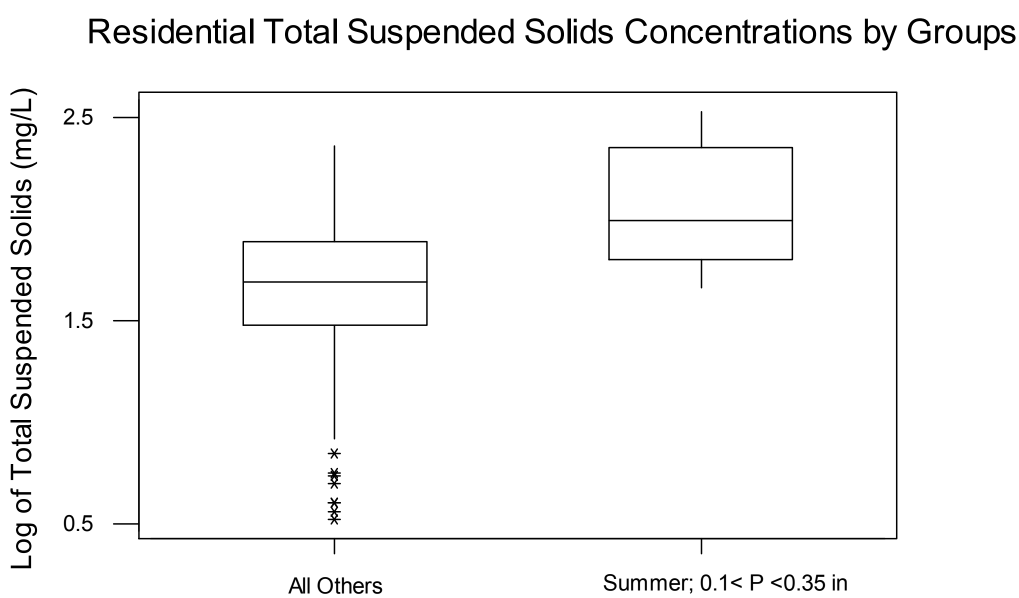


Figure 62. Residential TSS box and whiskers plot distribution by groups

Table 62. Results for Residential TSS (mg/L)

Descriptive statistics of residential TSS (mg/L) by groups:

Groups	N	Mean	Median	StDev	COV
All other rains	111	57.8	49.0	40.5	0.70
Small summer rains	11	143.0	98.0	101.6	0.71

Groups	SE Mean	Minimum	Maximum	Q1	Q3
All other rains	3.85	3.33	229.00	30.0	78.0
Small summer rains	30.6	46.0	337.0	63.0	227.0

Groups

Small summer rains (0.1 to 0.35 inches in depth):
Overall 95% confidence interval of all observed data: 29 to 439 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 75 to 170 mg/L (from fitted probability distribution)

All other rains:
Overall 95% confidence interval of all observed data: 8 to 243 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 37 to 51 mg/L (from fitted probability distribution)

Residential Area Total Phosphorus Data Analyses

The statistical tests of the residential total phosphorus data indicated a significant effect associated with the amount of imperviousness cover in the monitored watersheds. Sites having small amounts of impervious cover (7 to 25%) had significantly higher total phosphorus concentrations than sites having larger amounts of impervious cover (29 to 65%). Winter rains had lower total phosphorus concentrations in each group (all winter rains in the first group, and small winter rains of less than 0.1 inch in the second group). The following plots and data summaries describe these data groupings, separated by the two impervious cover categories.

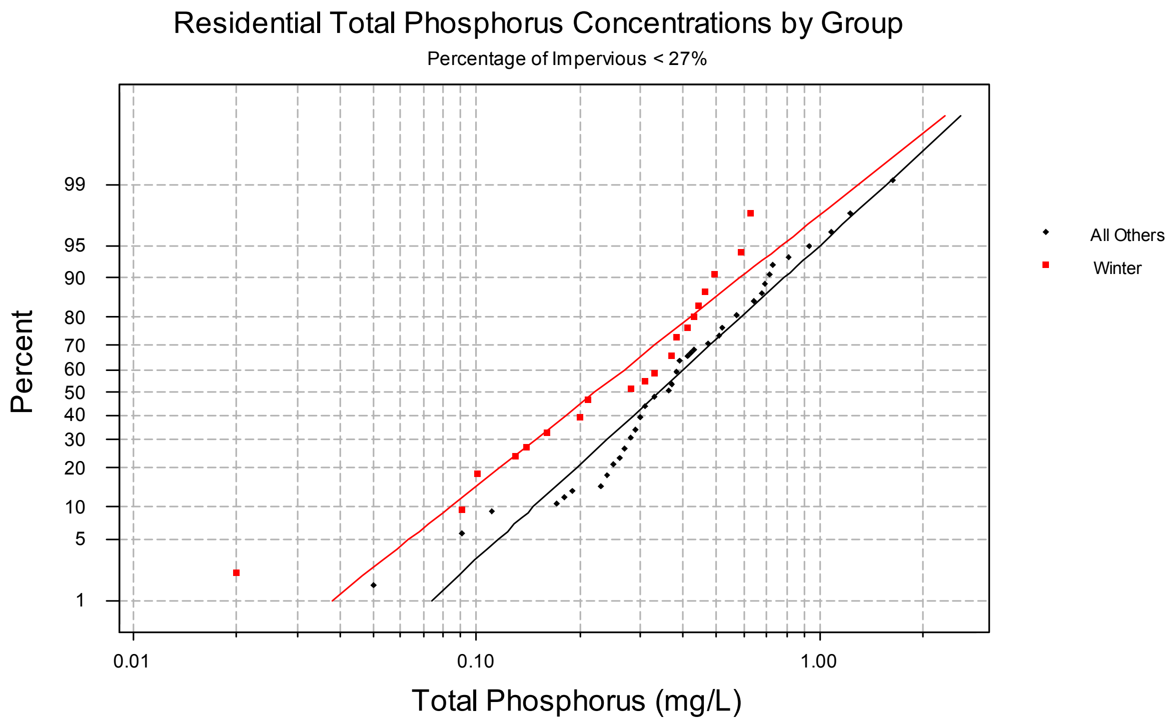


Figure 63. Residential total phosphorus concentrations for sites having < 27% impervious surfaces

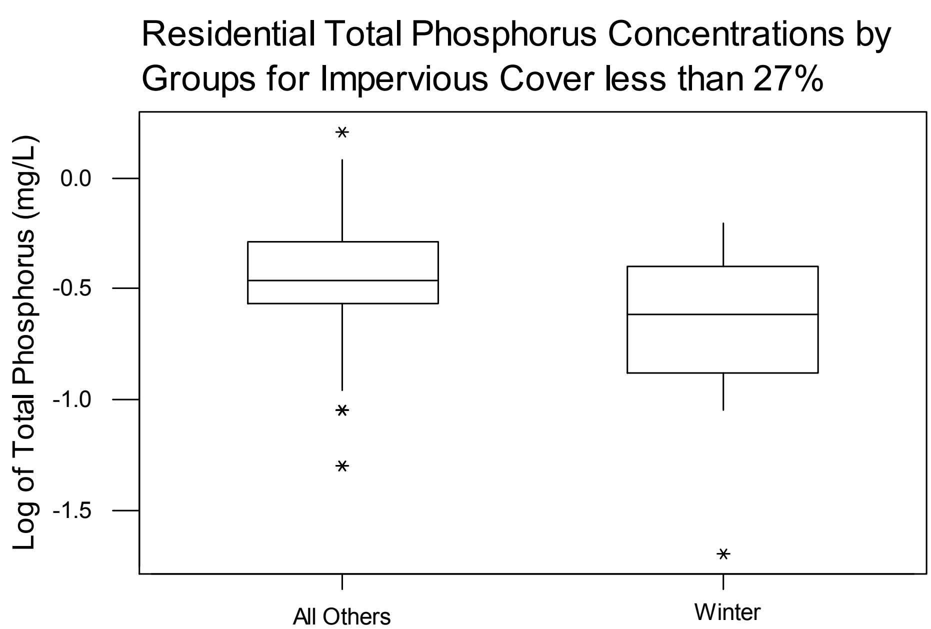


Figure 64. Box and whiskers plot for residential total phosphorus concentrations for sites having < 27% impervious surfaces

Table 63. Results for Residential Total Phosphorus (Impervious < 27%)

Descriptive statistics of residential TSS (mg/L) by groups:					
Season	N	Mean	Median	StDev	COV
All other seasons	72	0.41	0.35	0.27	0.65
Winter	28	0.28	0.25	0.16	0.59
Season	SE Mean	Minimum	Maximum	Q1	Q3
All other seasons	0.031	0.05	1.62	0.27	0.52
Winter	0.031	0.02	0.63	0.13	0.40

Groups

All winter rains:
Overall 95% confidence interval of all observed data: 0.050 to 0.98 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 0.17 to 0.29 mg/L (from fitted probability distribution)

All other seasons:
Overall 95% confidence interval of all observed data: 0.093 to 1.23 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 0.29 to 0.39 mg/L (from fitted probability distribution)

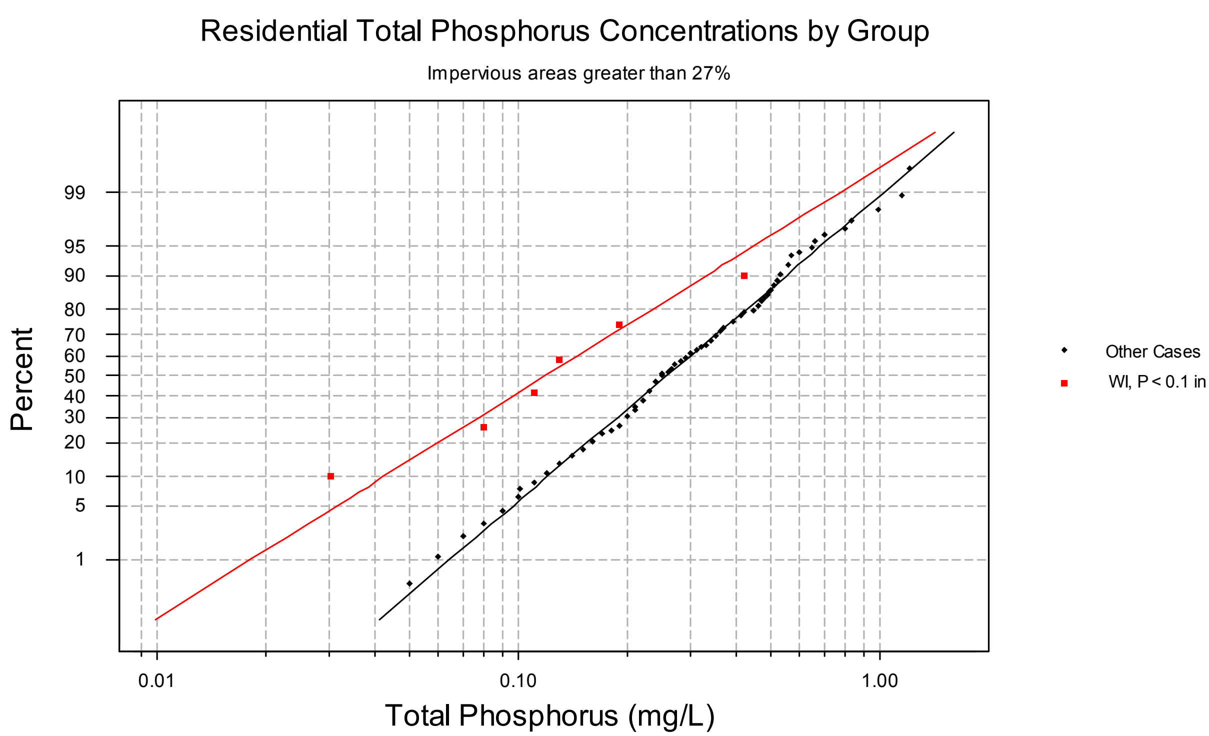


Figure 65. Residential total phosphorus concentrations for sites having > 27% impervious surfaces

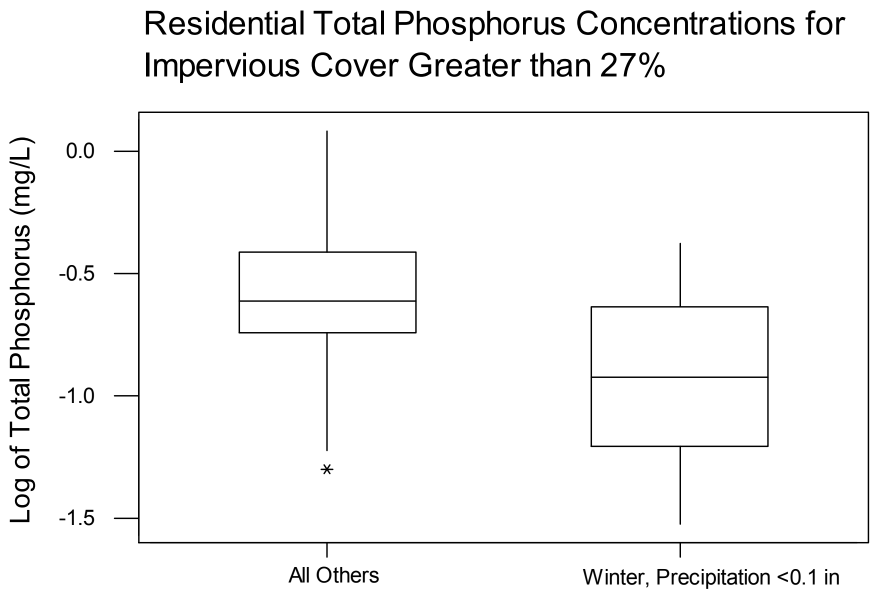


Figure 66. Box and whiskers plot for residential total phosphorus concentrations for sites having > 27% impervious surfaces

Table 64. Results for Residential Total Phosphorus (Impervious > 27%)

Descriptive statistics of residential TSS (mg/L) by groups:

Group	N	Mean	Median	StDev	COV
All others	152	0.30	0.24	0.19	0.63
Small winter rains	6	0.16	0.12	0.14	0.86

Group	SE Mean	Minimum	Maximum	Q1	Q3
All others	0.017	0.05	1.2	0.18	0.39
Small winter rains	0.057	0.03	0.42	0.068	0.25

Groups

Small winter rains <0.1 inch:
Overall 95% confidence interval of all observed data: 0.024 to 0.58 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 0.062 to 0.23 mg/L (from fitted probability distribution)

All other conditions:
Overall 95% confidence interval of all observed data: 0.080 to 0.83 mg/L (from fitted probability distribution)
95% confidence interval of reported median: 0.23 to 0.28 mg/L (from fitted probability distribution)

Residential Area Total Nitrogen Data Analyses

The statistical analysis of the residential total nitrogen data identified several important interactions between season and rain depth. There were no significant factors associated with drainage area, percent imperviousness, or trend with time. Five significant groups were identified for residential total nitrogen concentrations:

1) Fall rains <0.1 and > 1 inch
2) Winter rains between 0.35 and 1 inch
3) Fall for rains between 0.35 and 1 inch, and winter rains between 0.1 and 0.35 inches
4) All other conditions
5) Spring and summer rains between 0.1 and 0.35 inches

The following plots and data summaries describe these five data groupings for residential area total nitrogen concentrations.

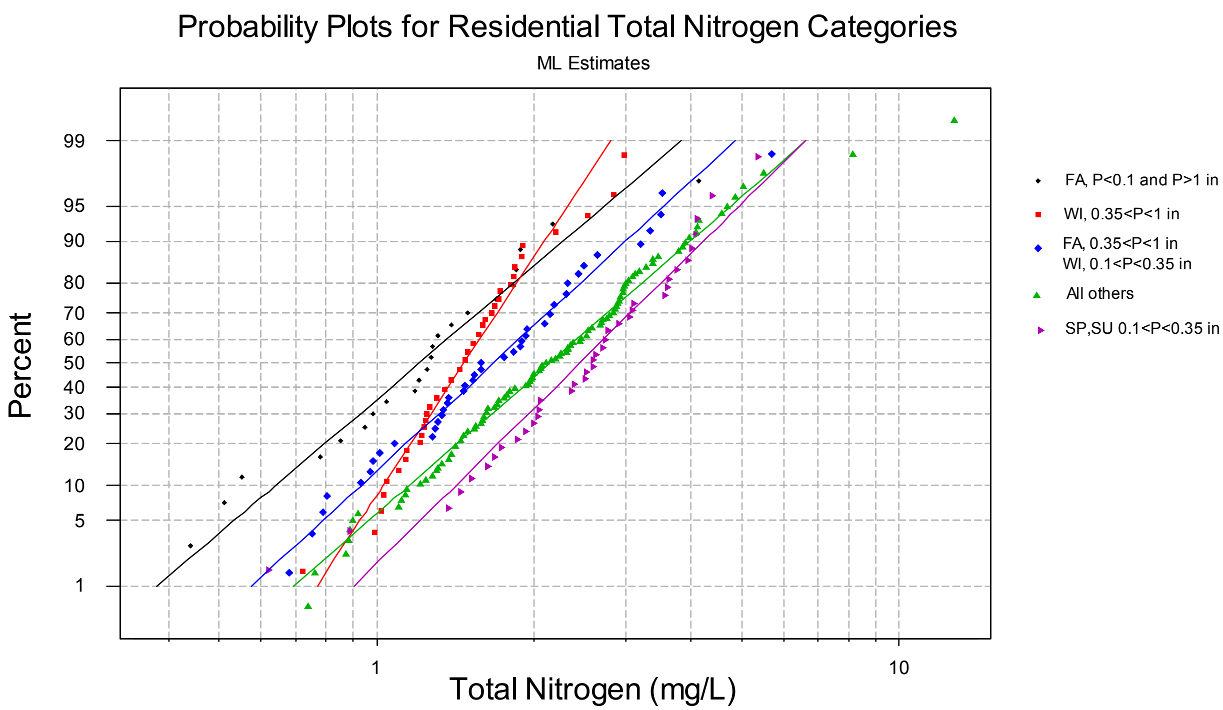


Figure 67. Residential total nitrogen concentration groups

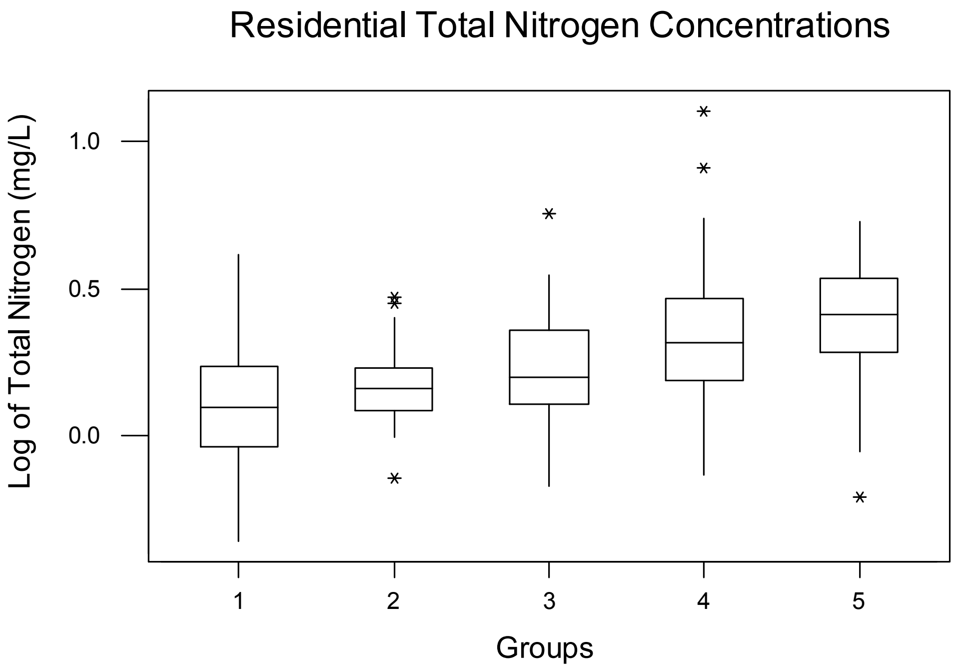


Figure 68. Box and whiskers plot for residential total nitrogen

Table 65. Results for Residential Total Nitrogen

Descriptive statistics of residential TSS (mg/L) by groups:

Groups	N	Mean	Median	StDev	COV
Fall, P<0.1 and P>1	22	1.4	1.3	0.77	0.57
Winter, 0.35<P<1	42	1.5	1.5	0.46	0.30
Fall, 0.35<P<1; WI, 0.1<P<0.35	43	1.9	1.6	0.95	0.51
All other conditions	112	2.4	2.1	1.5	0.62
Sp and Su, 0.1<P<0.35	40	2.6	2.6	1.0	0.38

Groups	SE Mean	Minimum	Maximum	Q1	Q3
Fall, P<0.1 and P>1	0.17	0.44	4.1	0.93	1.7
Winter, 0.35<P<1	0.072	0.72	3.0	1.23	1.7
Fall, 0.35<P<1; WI, 0.1<P<0.35	0.14	0.68	5.7	1.3	2.3
All other conditions	0.14	0.74	13	1.5	2.9
Sp and Su, 0.1<P<0.35	0.16	0.62	5.4	1.9	3.4

Groups
Fall, rains <0.1 inches and rains >1 inch: Overall 95% confidence interval of all observed data: 0.45 to 3.2 mg/L (from fitted probability distribution) 95% confidence interval of reported median: 0.98 to 1.5 mg/L (from fitted probability distribution) Winter, rains between 0.35 and 1 inch: Overall 95% confidence interval of all observed data: 0.85 to 2.5 mg/L (from fitted probability distribution) 95% confidence interval of reported median: 1.3 to 1.6 mg/L (from fitted probability distribution) Fall, rains between 0.35 and 1 inch; and Winter rains between 0.1 and 0.35 inches: Overall 95% confidence interval of all observed data: 0.68 to 4.1 mg/L (from fitted probability distribution) 95% confidence interval of reported median: 1.5 to 1.9 mg/L (from fitted probability distribution) Spring and Summer rains between 0.1 and 0.35 inches: Overall 95% confidence interval of all observed data: 1.1 to 5.7 mg/L (from fitted probability distribution) 95% confidence interval of reported median: 2.1 to 2.8 mg/L (from fitted probability distribution) All other conditions: Overall 95% confidence interval of all observed data: 0.82 to 5.9 mg/L (from fitted probability distribution) 95% confidence interval of reported median: 2.0 to 2.3 mg/L (from fitted probability distribution)

Commercial Area Total Suspended Solids Analyses

The commercial area total solids data appears to be affected by season and rain depth interactions, plus season and rain depth main factors. No affects associated with drainage area, or trends with time or interevent period, were identified.

Commercial Area Total Phosphorus Analyses

The commercial area total phosphorus data appears to be affected by season and rain depth interactions, plus season main factors. No affects associated with drainage area, or trends with time or interevent period, were identified.

Commercial Area Total Nitrogen Analyses

The commercial area total nitrogen data appears to be affected by season and rain depth interactions alone. No affects associated with drainage area, or trends with time or interevent period, were identified.

Industrial Area Total Suspended Solids Analyses

The industrial area total suspended solids data appears to be affected by season main factors alone. No affects associated with rain depth, drainage area, or trends with time or interevent period, were identified.

Industrial Area Total Phosphorus Analyses

The industrial area total phosphorus solids data appears to be affected by rain depth main factors alone. No affects associated with season, drainage area, or trends with time or interevent period, were identified.

Industrial Area Total Nitrogen Analyses

The commercial area total nitrogen solids data does not appear to be affected by any of the factors, or interactions examined. No affects associated with rain depth, season, drainage area, or trends with time or interevent period, were identified.

Summary

In this chapter, the NSQD was used to estimate the expected total suspended solids and nutrient mass discharges from urban, agricultural and forested sources in Anne Arundel County, Maryland, in a typical year. The parameters used in Schueler's simple method are the mean concentrations from each of these sources, the areas associated with each source, the volumetric runoff coefficient for these sources, and the rain depth associated with the period of calculation. The NSQD includes several catchments and more than 1,000 storm events in the Chesapeake Bay area which were used to determine the most appropriate urban area mean stormwater concentrations for residential, commercial and industrial land uses.

The effects associated with different seasons and rain depths on the urban area concentrations of solids and nutrients were also addressed for these regional data in this chapter. ANOVA analyses indicated that there are significant differences in the concentrations according to the seasonal period when the samples were collected, the total precipitation depth, and the interaction between these two factors for some of these pollutants and urban land uses. A stronger influence of the interactions between these factors was observed in residential areas compared with commercial or industrial land use areas.

The data summaries indicated that solids concentrations from forested and urban areas are similar, however the total runoff volume produced in forest areas is very small compared with the urban areas. For this reason, annual mass discharges from forested areas are less than half of the annual mass discharges produced from urban areas, even though the areas for these two main land use categories are similar.

Annual agricultural mass discharges of suspended solids are almost twice those calculated for urban areas. In urban areas, lower TSS concentrations occur, but a much larger fraction of the precipitation is transformed to runoff. Total urban area nitrogen mass discharges are expected to be almost twice the loads discharged from agricultural areas.

Chapter 9: Findings and Conclusions

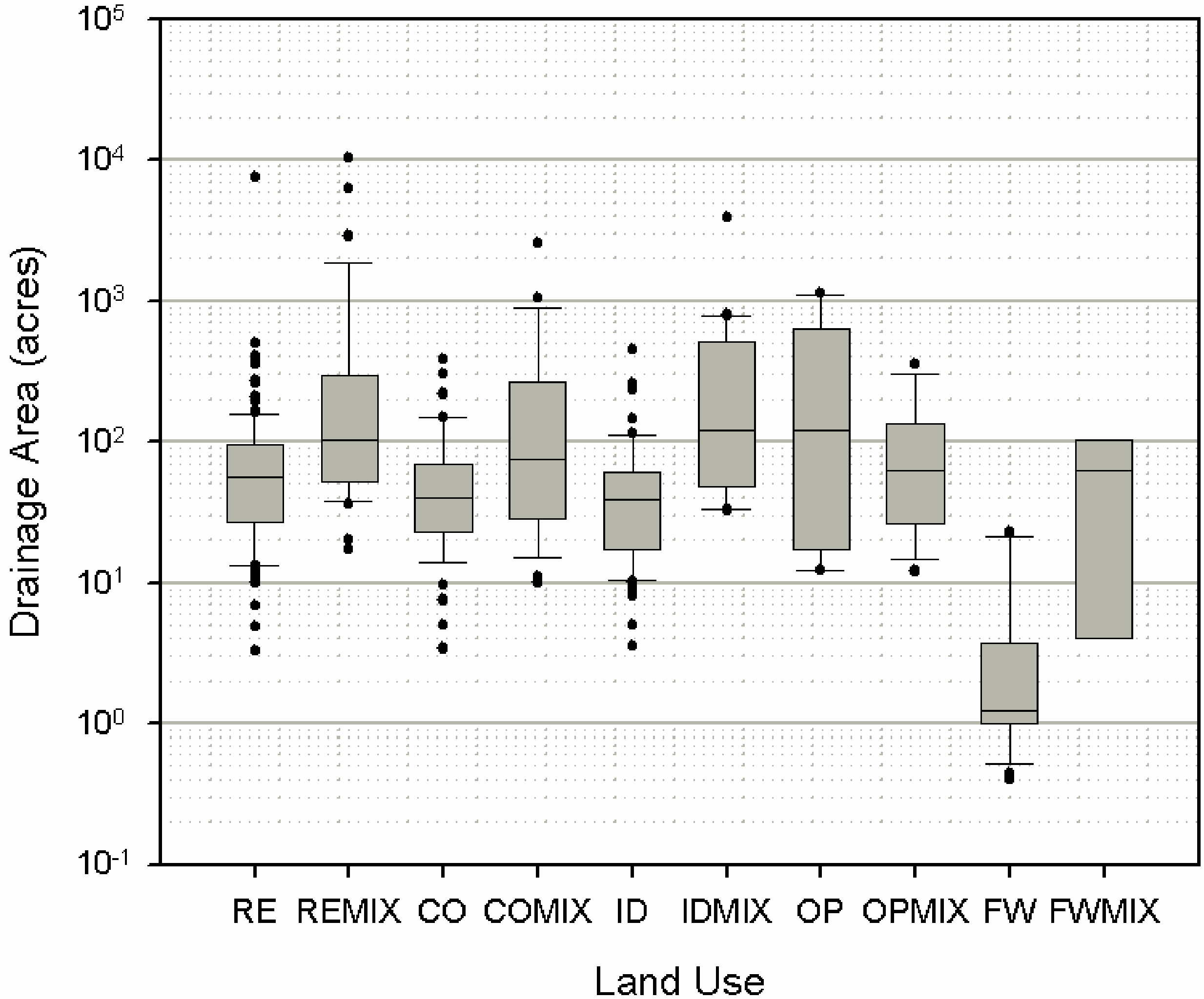
Introduction

The purpose of this report was to examine several commonly accepted assumptions concerning stormwater characteristics (and associated management decisions) by stormwater managers and researchers. These included assumptions relating to the existence of "first flushes;" the effect of the abundance of impervious areas and the length of antecedent dry period on stormwater constituent concentrations; the influences of non-detected observations on stormwater characteristics; among others. These assumptions were evaluated using information contained in the National Stormwater Quality Database (NSQD). More than 3,765 events were monitored at 360 sites throughout the U.S. and the monitored water quality and associated information was included in the first version of the database. Most of the data were collected from residential, commercial and industrial land use areas in the eastern and southern parts of the U.S. (according to the original study design), although most geographical areas are represented.

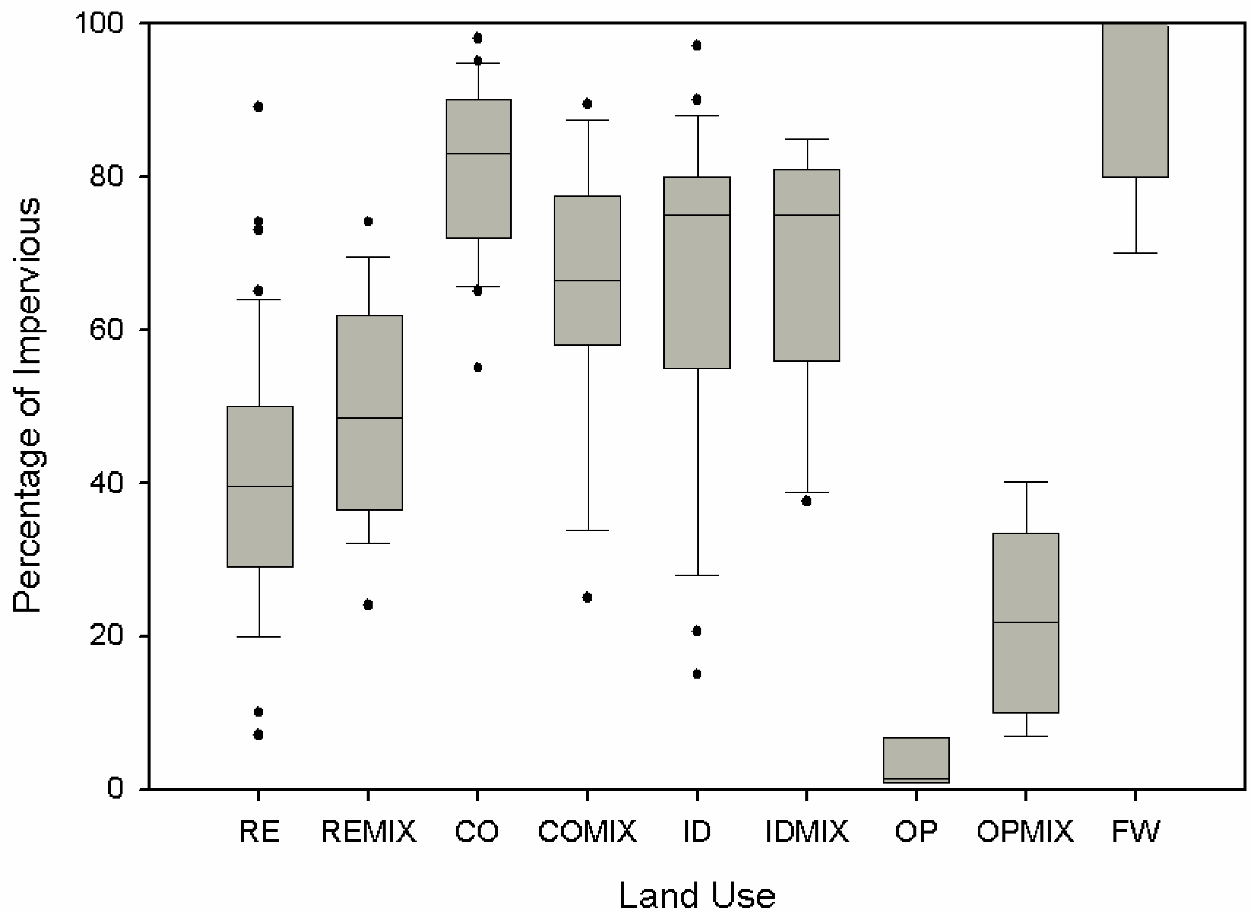
Major Findings, as Reported in Report Chapters

Findings from Chapter 2: The National Stormwater Quality Database (NSQD) Description

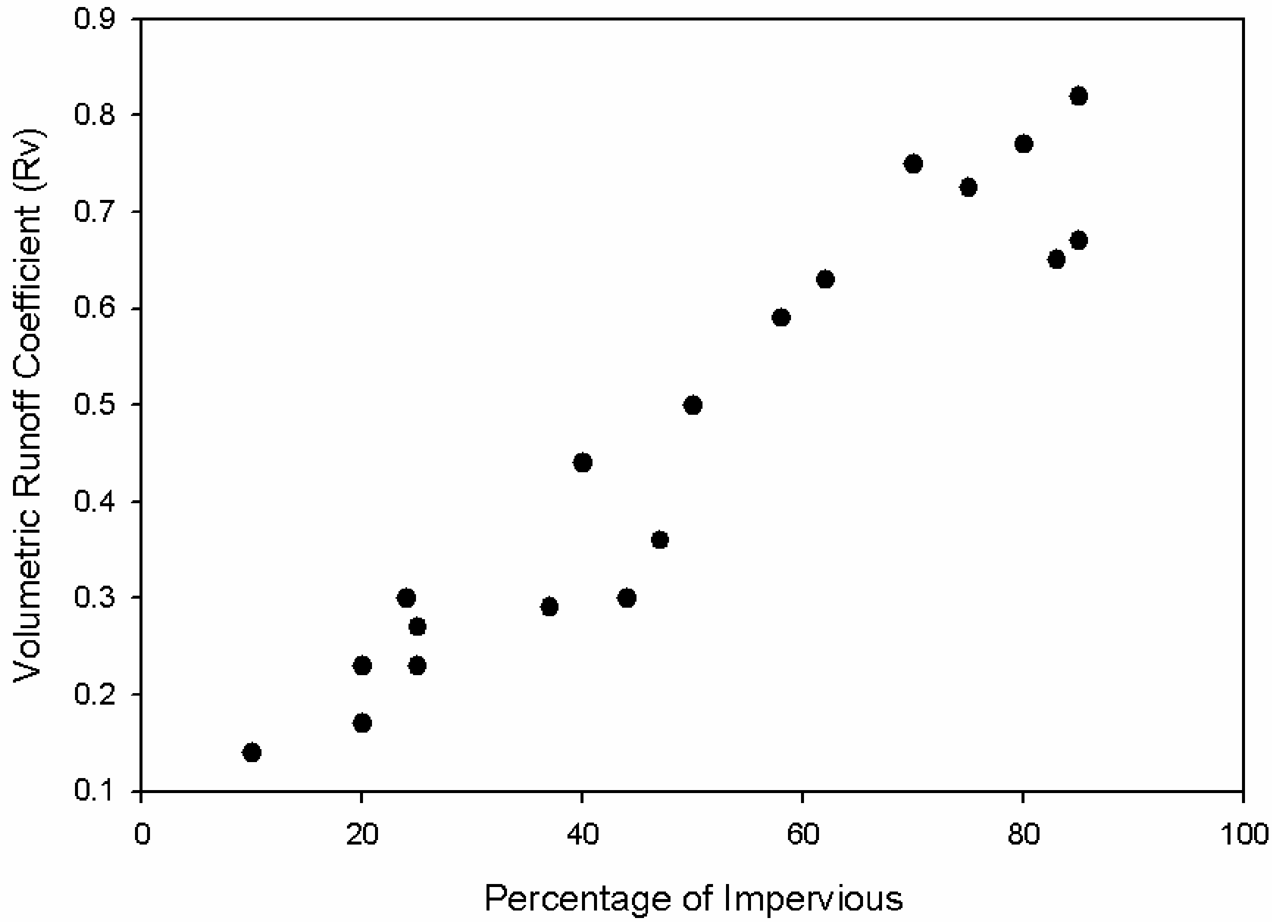
• ***Drainage Areas by Land Use***. Drainage areas for each outfall varied for different land uses, with freeways having the smallest drainage areas and open space having the largest drainage areas. Generally, the median drainage areas ranged from about 40 to 110 acres, excluding the freeway sites which were only about 1.5 acres in size.



• ***Impervious Areas in each Land Use***. The percentage of impervious areas in each drainage area is obviously related to the land uses. Open space, and the open space mixed areas have the lowest fraction of impervious areas (at close to zero and about 20% respectively), while freeways and commercial land uses have the largest fractions of impervious areas (close to 100% and 85%, respectively). Residential areas have about 40% impervious surfaces. The database is not able to distinguish the directly connected vs. the partially connected impervious areas.



• ***Runoff Coefficients and Impervious Cover***. The reported volumetric runoff coefficients were closely related to the percentage of impervious cover. Again, the database cannot separate the directly connected impervious areas from the partially connected areas, so there is some expected variation in this relationship. This relationship significantly affects the mass discharges of pollutants. As noted later in these findings, very few significant relationships were found between the impervious covers and runoff concentrations.



• ***Reported Monitoring Problems***. About 58% of the communities described problems during the monitoring process:

- One of the basic sampling requirements was to collect three samples every year for each of the land use stations. These samples were to be collected at least one month apart during rains having at least 0.1 inch rains, and with at least 72 hours from the previous 0.1-inch storm event. It was also required (when feasible), that the variance in the duration of the event and the total rainfall not exceeded the median rainfall for the area. About 47% of the communities reported problems meeting these requirements. In many areas of the country, it was difficult to have three storm events per year having these characteristics.

- The second most frequent problem, reported by 26% of the communities, concerned backwater tidal influences during sampling, or the outfall became submerged during the event. In other cases, it was observed that there was flow under the pipe (flowing outside of the pipe, in the backfill material, likely groundwater), or sometimes there was not flow at all.

- About 12% of the communities described errors related to malfunctions of the sampling equipment. When reported, the equipment failures were due to incompatibility between the software and the equipment, clogging of the rain gauges, and obstruction in the sampling or bubbler lines. Memory losses in the equipment recording data were also periodically reported. Other reported problems were associated with lighting, false starts of the automatic sampler before the runoff started, and operator error due to misinterpretation of the equipment configuration manual.

• ***Suggested Changes in Monitoring Requirements***:

- Base flows can commonly occur in separate storm drainage systems for a variety of reasons and they may be more important during some seasons than during others. In many cases, they cannot be avoided and should be included in the monitoring program, and their effects need to be recognized as an important flow phase.

- The rain gauges need to be placed close to the monitored watersheds. In the NSQD, about 7% of the events had site precipitation estimated using a rain gauge located at the city airport. About 16% of the events had precipitation depths estimated using their own monitoring network. Some communities had precipitation networks that were used for flood control purposes for the surrounding area. These networks can be considered better than the single airport rain gauge, but should at least be supplemented with a rain gauge located in the monitored watershed. Large watersheds cannot be represented with a single rain gauge at the monitoring station; in those cases the monitoring networks will be a better approach..

- Many of the monitoring stations lacked flow monitoring instrumentation, or did not properly evaluate the flow data. Accurate flow monitoring can be difficult, but it greatly adds to the value of the expensive water quality data.

- The three hour monitoring period that most used may have resulted in some bias in the reported water quality data. For example, it is unlikely that manual samplers were able to initiate sampling near the beginning of the events, unless they were deployed in anticipation of an event later in the day. A more cost-effective and reliable option would be to have semi-permanent monitoring stations located at the monitoring locations and sampling equipment installed in anticipation of a monitored event. Most monitoring agencies operated three to five land use stations at one time. This number of samplers, and flow equipment, could have been deployed in anticipation of an acceptable event and would not need to be continuously installed in the field at all sampling locations.

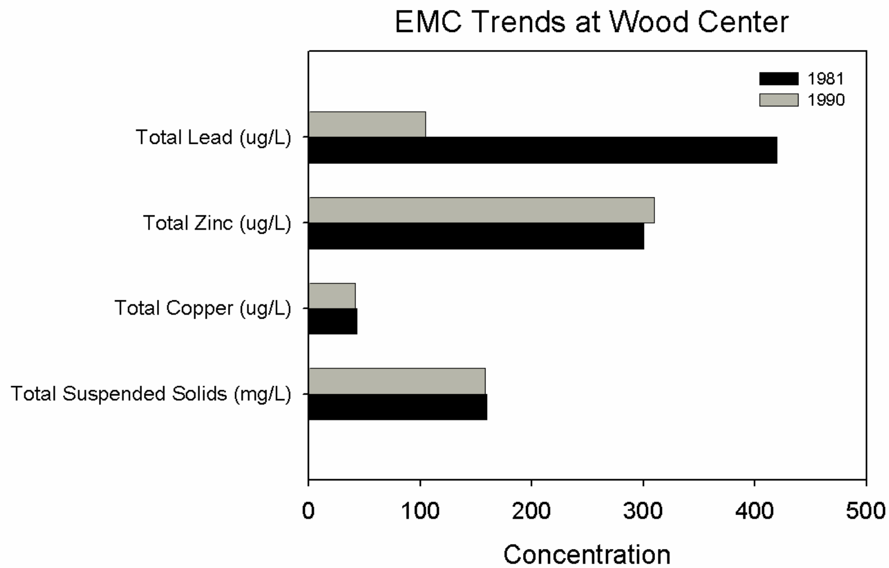
- Some of the site descriptions lacked important information and local personnel sometimes did not have the needed information. This was especially critical for watershed delineations on maps of the area. Also, few of the watershed descriptions adequately described how the impervious areas were connected to the drainage system, one of the most important factors affecting urban hydrologic analyses. In most cases, information concerning local stormwater controls was able to be determined from a variety of sources, but it was not clearly described in the annual reports.

• ***Comparisons of Stormwater Databases***. The NSQD can be compared to the older NURP database:

Comparison of NURP and NSQD Stormwater Databases

Constituent	Units	Source	Event Mean Concentrations		Number of events
			Mean	Median	
Total Suspended Solids	mg/L	NURP	174	113	2000
		NSQD	79	50	3404
Biochemical Oxygen Demand	mg/L	NURP	10	8.4	474
		NSQD	11	8.6	2973
Chemical Oxygen Demand	mg/L	NURP	66	55	1538
		NSQD	71	55	2699
Total Phosphorus	mg/L	NURP	0.34	0.27	1902
		NSQD	0.37	0.29	3162
Dissolved Phosphorus	mg/L	NURP	0.10	0.078	767
		NSQD	0.11	0.078	2093
Total Kjeldahl Nitrogen	mg/L	NURP	1.7	1.4	1601
		NSQD	1.7	1.4	3034
Nitrite and Nitrate	mg/L	NURP	0.84	0.67	1234
		NSQD	0.77	0.61	2983
Copper	µg/L	NURP	67	55	849
		NSQD	18	14	2356
Lead	µg/L	NURP	175	131	1579
		NSQD	24	17	2250
Zinc	µg/L	NURP	176	140	1281
		NSQD	110	88	2888

- The nutrient, COD, and BOD_5 means and medians are very close in both databases, while the suspended solids and metals are much smaller in the NSQD than in the NURP database. As part of their MS4 Phase I application, Denver and Milwaukee (Milwaukee data not yet included in the NSQD) both returned to some of their earlier sampled monitoring stations used during their local NURP projects (EPA 1983). In the time period between the early 1980s (NURP) and the early 1990s (MS4 permit applications), they did not detect any significant differences, except for large decreases in lead concentrations, as shown in the figure below for a Milwaukee site.



Comparison of pollutant concentrations collected during NURP (1981) to MS4 application data (1990) at the same location (personal communication, Roger Bannerman, WI DNR)

- The differences found in both the NURP and the NSQD databases are therefore most likely due to differences in geographical areas emphasized by each database. Half of the events included in the NSQD database were collected in EPA Rain Zone 2 (Maryland, Virginia, North Carolina, Kentucky and Tennessee), while half of the events contained in the NURP database were collected in EPA Rain Zone 1 (Minnesota, Wisconsin, Michigan, Illinois, New York, Massachusetts and New Hampshire). The NSQD best represents the coastal states and the southern states, while NURP best represents the upper Midwest and northeast states.

Findings from Chapter 3: QA/QC Procedures

• ***QA/QC Effort***. QA/QC takes a great deal of time and effort to ensure that the database content is correct and accurate. During this project, about 6 months were spent in collecting the majority of the information from the communities, while more than 9 months were spent in reviewing the accuracy of the data. All data was compared against original information, if at all possible, and all transcribed data was carefully compared to the source data. In addition, the behavior of the data was also carefully reviewed to identify unusual data observations. "Outliers" were not casually eliminated from the dataset unless errors were likely that could not be corrected. Comparisons to associated data and to likely data levels were the most important methods used to identify errors. In addition, unusually high and low observations were all verified.

• ***Non-Detected Analyses***. Left-censored data refers to observations that are reported as below the limits of detection, while right-censored data refers to over-range observations. Unfortunately, many important stormwater measurements (such as for filtered heavy metals) have large fractions of undetected values. These missing data greatly hinder many statistical tests. To estimate the problems associated with censored values, it is important to identify the probability distributions of the data in the dataset and the level of censoring. Most of the constituents in the NSQD follow a lognormal distribution. When the frequencies of the censored observations were lower than 5%, the means, standard deviations and coefficients of variation were almost identical when the censored observations were replaced by half of the detection limit, or estimated using Cohen's Method. As the percentage of non-detected values increases, replacing the censored observation by half of the detection limit instead of estimating them using the Cohen's maximum likelihood method produced lower means and larger standard deviations. Replacing the censored observations by half of the detection limit is not recommended for levels of censoring larger than 15%. The censored observations in the database were replaced using estimated values using Cohen's maximum likelihood method for each site before the statistical tests in this report. Because this method uses the detected observations to estimate the non-detected values, it is not very accurate, and therefore not recommended, when the percentage of censored observations is larger than 40%.

• ***Selection of Analytical Methods***. The best method to eliminate problems associated with left-censored data is to use an appropriate analytical method. By keeping the non-detectable level below 5%, there are many fewer statistical analysis problems and the value of the datasets can be fully realized. The following table summarizes the recommended minimum detection limits for various stormwater constituents to obtain manageable non-detection frequencies (<5%). Some of the open space stormwater measurements (lead, and oil and grease, for example), would likely have greater than 5% non-detects, even with the detection limits shown. The detection limits for filtered heavy metals should also be substantially less than shown on this table.

Suggested Analytical Detection Limits for Stormwater Monitoring Programs to Obtain <5% Non-detects

	Residential, commercial, industrial, freeway	Open Space
Conductivity	20 μS/cm	20 μS/cm
Hardness	10 mg/L	10 mg/L
Oil and grease	0.5 mg/L	0.5 mg/L
TDS	10 mg/L	10 mg/L
TSS	5 mg/L	1 mg/L
BOD_5	2 mg/L	1 mg/L
COD	10 mg/L	5 mg/L
Ammonia	0.05 mg/L	0.01 mg/L
NO_2+NO_3	0.1 mg/L	0.05 mg/L
TKN	0.2 mg/L	0.2 mg/L
Dissolved P	0.02 mg/L	0.01 mg/L
Total P	0.05 mg/L	0.02 mg/L
Total Cu	2 μg/L	2 μg/L
Total Pb	3 μg/L (residential 1 μg/L)	1 μg/L
Total Ni	2 μg/L	1 μg/L
Total Zn	20 μg/L (residential 10 μg/L)	5 μg/L

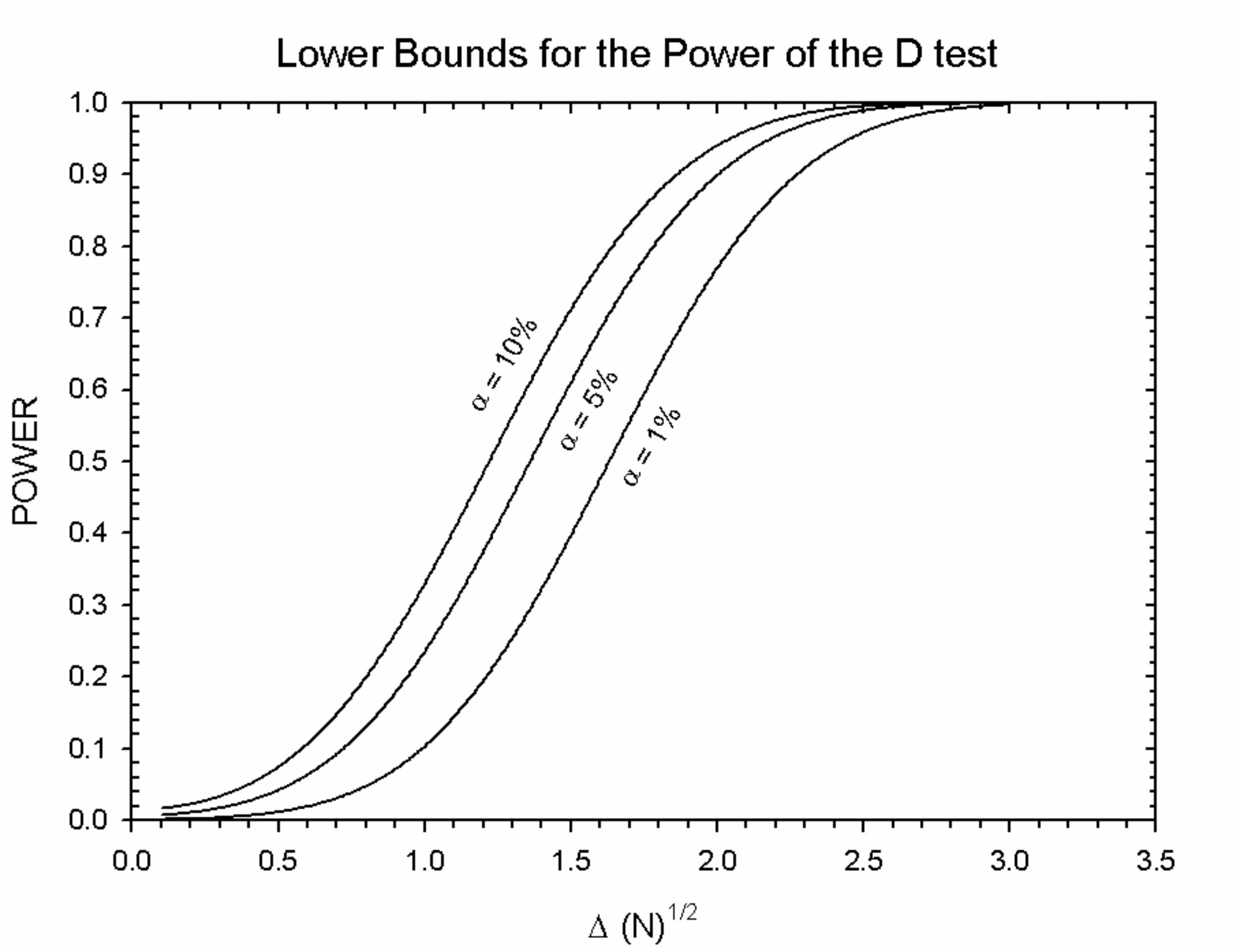
Findings from Chapter 4: Stormwater Quality Descriptions Using the Three Parameter Lognormal Distribution

• ***Statistical Distributions***. Knowing the statistical distributions of stormwater concentrations is a critical step in data analysis. The selection of the correct statistical analyses tools is dependent on the data distribution, and many QA/QC operations depend on examining the distribution behavior. However, much data is needed for accurate determinations of the statistical distributions of the data, especially when examining unusual behavior. The comparison of probability distributions between different data subsets is also a fundamental method to identify important factors affecting data observations.

• ***Log-Normal Statistical Distribution***. Most of the stormwater constituents in the NSQD can be assumed to follow a lognormal distribution with little error. The use of the third parameter does not show a significant improvement in estimating the empirical distribution compared with the 2-parameter lognormal distribution. When the number of samples is very large per category (approximately more than 400 samples) the maximum likelihood and the 2-parameter lognormal distribution better fit the empirical distribution. For large sample sizes, the L-moments method usually unacceptably truncates the distribution in the lower tail. When the sample size is small (<100 samples), the use of the third parameter does not improve the fit with the empirical distribution and the 2-parameter lognormal distribution produces a better fit than the other two methods.

• ***Effects of Data Errors***. Incorrect data observations can have a great effect on the characteristics of the dataset. For example, when only 0.5% of the sample is affected by a factor of a thousand, the coefficient of variation increases almost 12 times more than the correct value. An error of a factor of a thousand occurs periodically, especially for heavy metal values when the concentrations are reported in mg/L units when they are actually in µg/L units.

• ***Data Observations Needed***. Determining the number of data observations needed to compare two datasets with known, and similar distributions, can be readily determined. The following plot shows the power of the D test for 1%, 5%, and 10% levels of confidence of the test for samples size larger than 35. For example, assume that the maximum distance between the alternative cumulative and the estimated cumulative probability distributions is 0.2 (approximately a 20% difference in the concentrations in the datasets), and we want an 80% power (0.8) against the alternative at a 5% level of confidence. To calculate the number of required samples, we read that $\Delta(N)^{0.5}$ is 1.8 for a power of 0.8 and 5% level of confidence. Solving for $N = (1.8/0.2)^2 = 81$ samples. If we want to calculate the number of samples when the difference between the alternative cumulative and the estimated cumulative probability function is 0.05 (a difference of only 5%), with the same power and level of confidence, then 1,296 samples would be required. When the lines are very close together, it is obviously very difficult to statistically show that they are different, and many samples are needed.



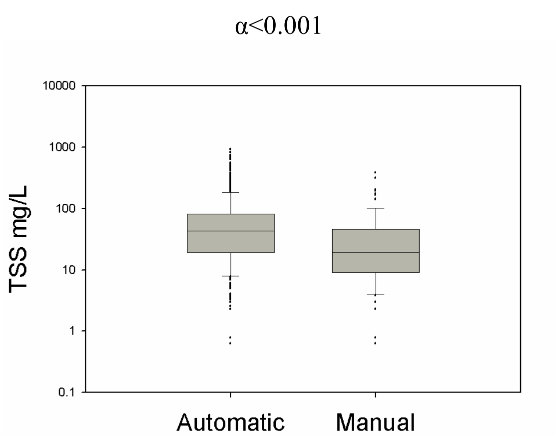
Lower bounds for the power of the D test for α = 1%, 5% and 10% (N>35) (Massey 1951)

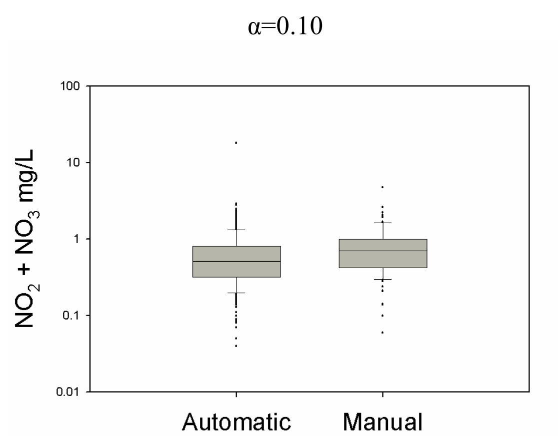
Difference between datasets to be detected (Δ)	Percentage difference between datasets	Number of samples needed, 5% confidence, 80% power [$N = (1.8/\Delta)^2$]
0.05	5	1,300
0.10	10	320
0.25	25	52
0.50	50	13
1.00	100	3

- Obviously, a careful decision has to be made between monitoring budgets and data quality objectives. The sample needs increase dramatically as the difference between datasets become small. Typically, a difference of about 25% (requiring about 50 sample pairs) is a reasonable objective for most stormwater projects. This is especially important when monitoring programs attempt to distinguish test and control conditions associated with stormwater control practices. It is easy to confirm significant differences between influent and effluent conditions at wet detention ponds, as they have relatively high removal rates. Less effective controls are much more difficult to verify, as the sampling program requirements become very expensive.

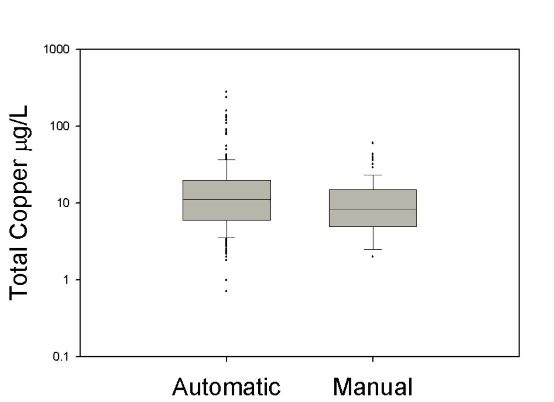
Findings from Chapter 5: Identification of Significant Factors Affecting Stormwater Quality Using the NSQD

• ***Manual vs. Automatic Sampling***. About 80% of the NSQD samples were collected using automatic samplers. It was observed that manual sampling can result in lower TSS concentrations compared to automatic sampling procedures. This may occur, for example, if the manual sampling team arrives after the start of runoff and therefore misses an elevated first flush (if it exists for the site), resulting in reduced event mean concentrations. The following figure contains box and whisker plots comparing resultant sample concentrations when the samples were collected by automatic versus manual sampling methods, for residential land uses in EPA Rain Zone 2.

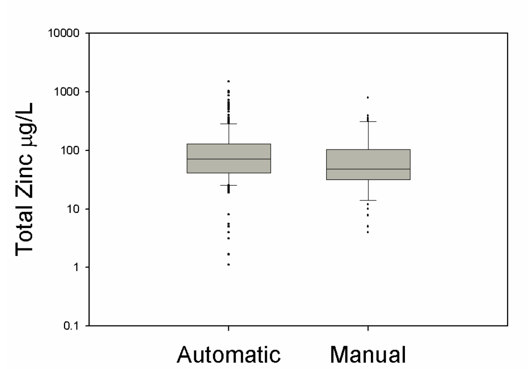




α<0.001



α<0.001



Comparison of reported concentrations in residential land use and EPA Rain Zone 2 for automatic vs. manual sampling methods

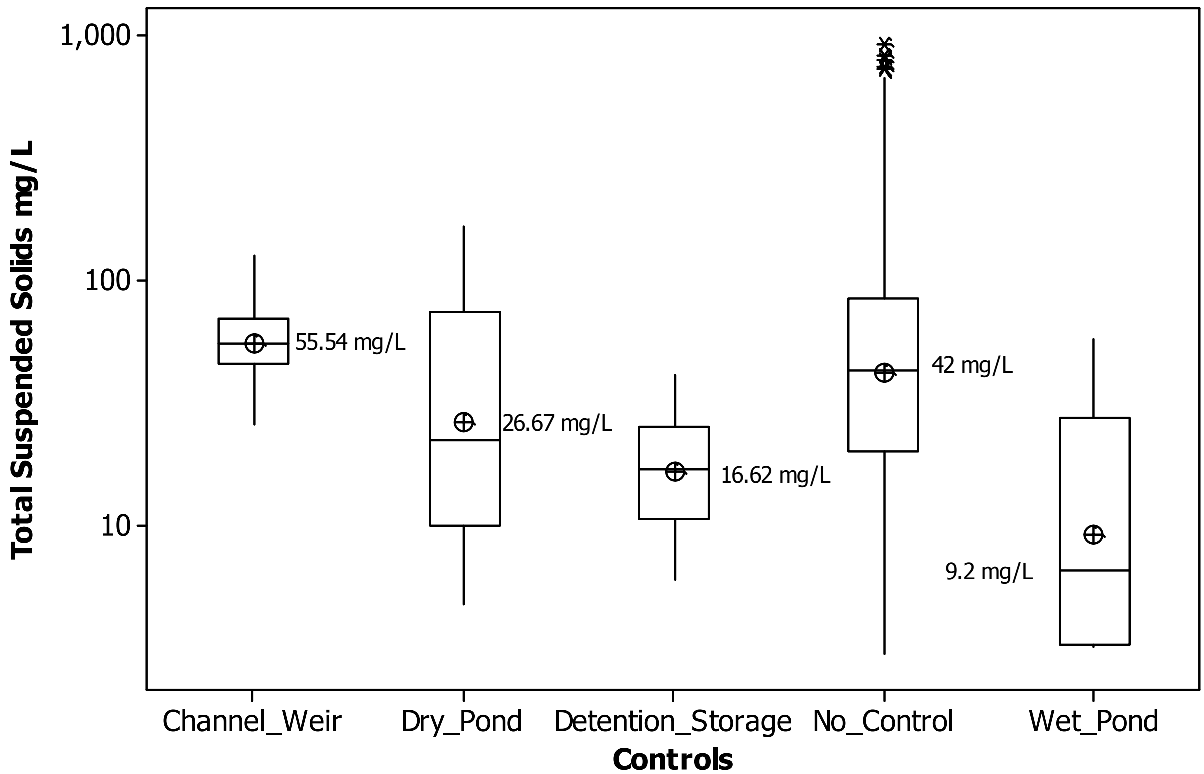
- The decision to use automatic or manual sampling methods is not always clear. There were statistical differences found between both methods in residential areas for several constituents. Most communities calculate their EMC values using flow-composited sample analyses. If first flush effects are present, manual sampling may likely miss these more concentrated flows due to delays in arriving at the site to initiate sampling. If the first flush is for a very short duration, time-composited samples may overly emphasize these higher flows. Flow compositing produces more accurate EMC values than time composite analyses. An automatic sampler with flow-weighted samples, in conjunction with a bed load sampler, is likely the most accurate sampling alternative.

• ***Sample Compositing Methods***. Time and flow-weighted composite options. were also evaluated in residential, commercial, and industrial land uses in EPA Rain Zone 2 and in industrial land uses in EPA Rain Zone 3.

- No significant differences were observed for BOD_5 concentrations using either of the compositing schemes for any of the four categories. A similar result was observed for COD except for commercial land uses in EPA Rain Zone 2, where not enough samples were collected to detect a significant difference. TSS and total lead median concentrations in EPA Rain Zone 2 were two to five times higher in concentration when time-compositing was used instead of flow-compositing.

- Nutrients in EPA Rain Zone 2 collected in residential, commercial and industrial areas showed no significant differences using either compositing method. The only exceptions were for ammonia in residential and commercial land use areas and total phosphorus in residential areas where time-composite samples had higher concentrations. Metals were higher when time-compositing was used in residential and commercial land use areas. No differences were observed in industrial land use areas, except for lead.

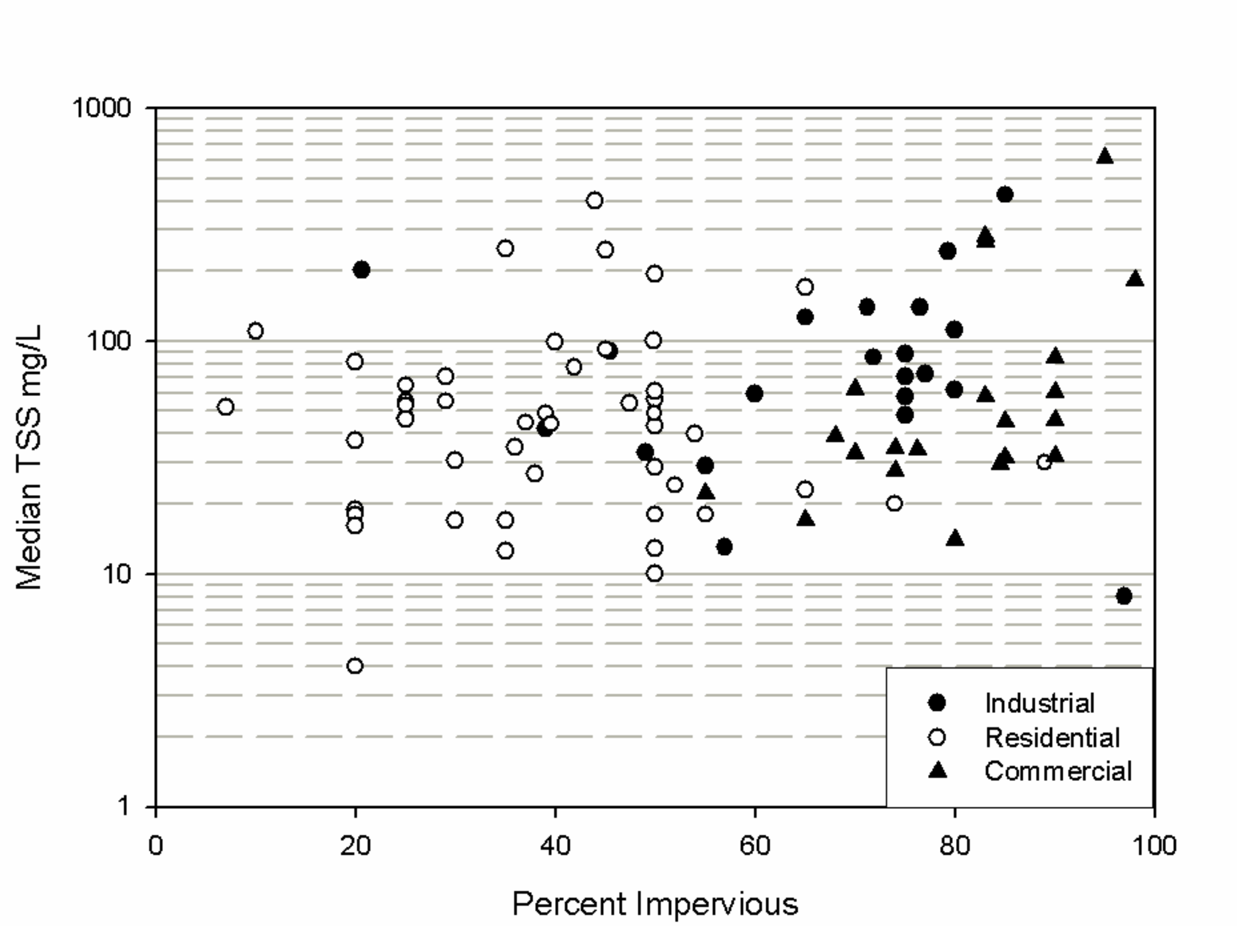
• ***Stormwater Controls***. The following figure shows the observed TSS concentrations in residential areas for EPA Rain Zone 2, for different drainage area stormwater controls (Channel weir: a flow measurement weir in an open channel that forms a small pool; Dry pond: a dry detention pond that drains completely between each storm event; Wet pond: a wet detention pond that retains water between events, forming a small lake or pond; Detention storage: Oversize pipes with small outlet orifices, usually under parking lots).



- There is a significant reduction in TSS, nitrite-nitrate, total phosphorus, total copper, and total zinc concentration at sites having wet detention ponds, the control practice having the largest concentration reductions. No reductions in TKN concentrations were found using wet ponds, but TKN seems to be reduced by dry ponds. Locations with detention storage facilities had smaller reductions of TSS, BOD_5, COD, total lead and total zinc concentrations compared to wet pond sites. Unfortunately, there were few sites in the database having grass swales that could be compared with data from sites having curbs and gutters.

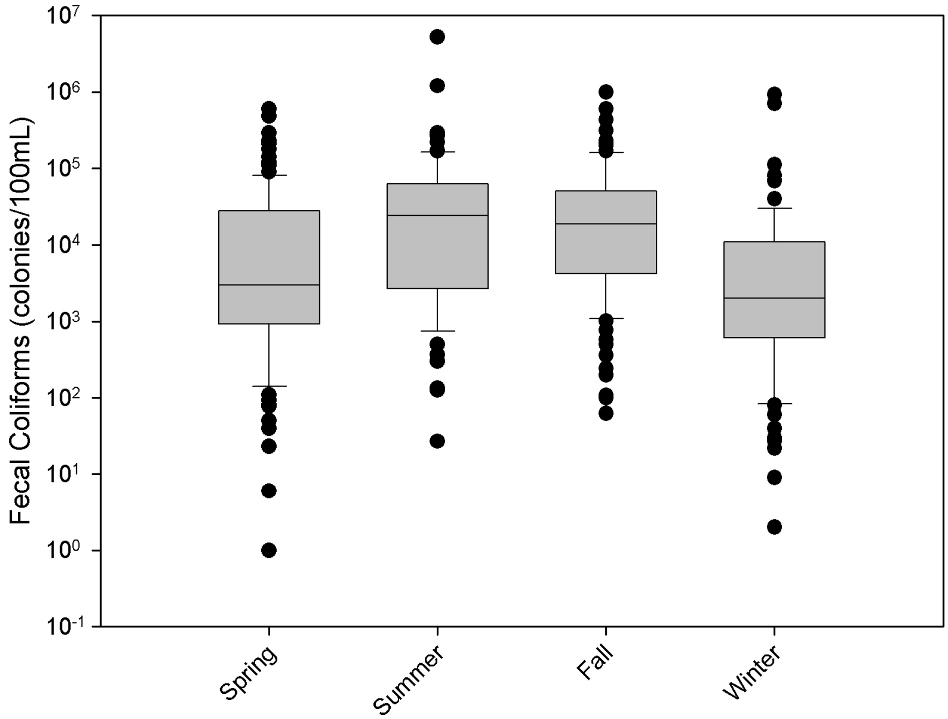
• ***Concentration Effects Associated with Impervious Cover Amounts***. The following plot shows no apparent trend in TSS concentration that can be explained by impervious cover differences. However, it is very likely that a significant and important trend does exist between percent effective imperviousness and the pollutant mass that is discharged. While the relationship between imperviousness and concentration is not clear, the relationship between effective imperviousness and total runoff volume is much stronger (as noted previously) and more obvious as the non-paved areas can infiltrate much water.

- There is a certain amount of redundancy (self-correlation) between land use and the percentage of impervious areas, as each land use category generally has a defined narrow range of paved and roof areas. Therefore, it is not possible to test the hypothesis that different levels of impervious (surface coverage) are more important than differences in land use (activities within the area). Residential land uses cover only the lower range of imperviousness, while commercial sites have imperviousness amounts larger than 50%. In order to perform a valid comparison test, the range of imperviousness needs to be similar for both test cases.

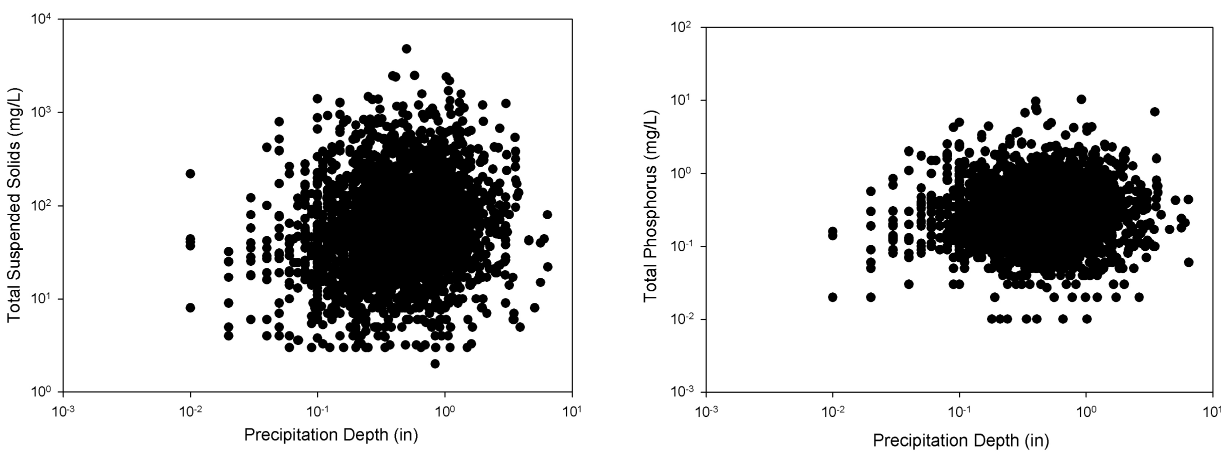


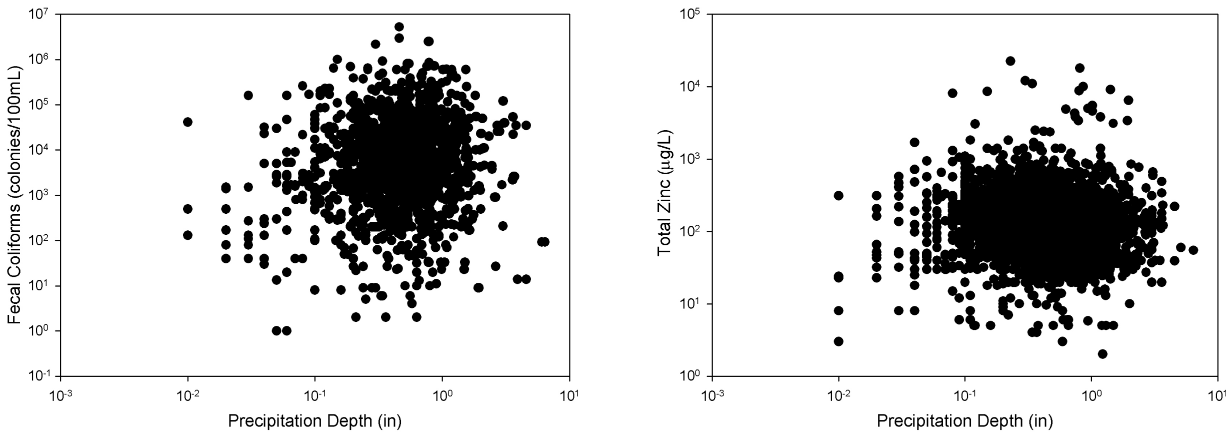
TSS concentrations by impervious cover and single land use

• ***Seasonal Effects***. Another factor that may affect stormwater quality is the season when the sample was obtained. If the few samples collected for a single site were all collected in the same season, the results may not be representative of the whole year. The NPDES sampling protocols were designed to minimize this effect by requiring the three samples per year to be separated by at least 1 month. The few samples still could be collected within a single season, but at least not within the same week. Seasonal variations for residential fecal coliform data are shown in the following figure for all residential areas. The bacteria levels are lowest during the winter season and highest during the summer and fall (a similar conclusion was obtained during the NURP, EPA 1983, data evaluations). The database does not contain any snowmelt data, so all of the data corresponds to rain-related runoff only. No other seasonal trends in stormwater quality were identified.



• ***Rain Depth Effects***. The following figure contains several scatter plots showing concentrations plotted against rain depth. There are no obvious trends of concentration associated with rain depth for the NSQD data..



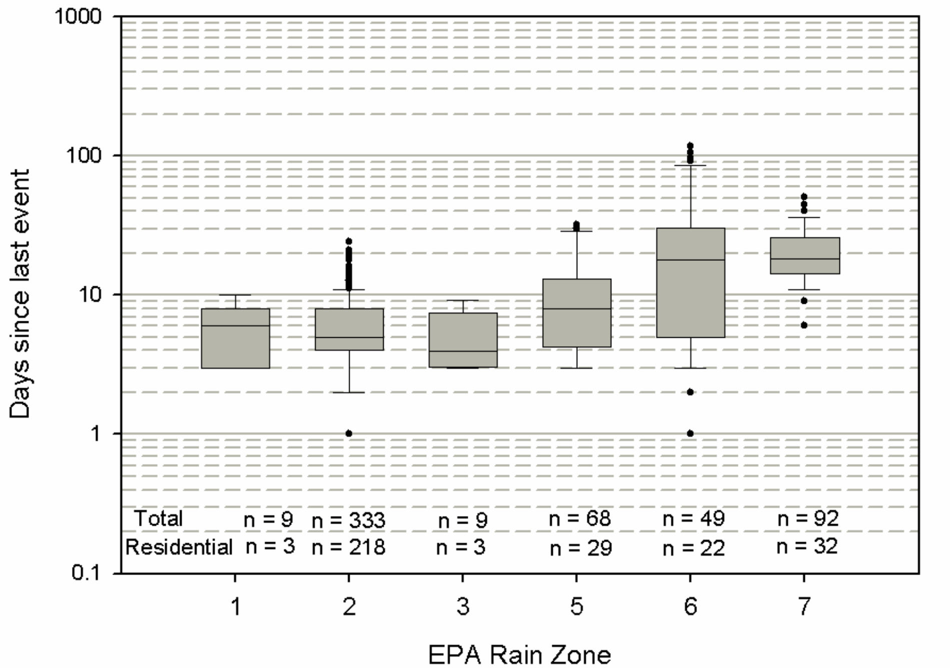


Examples of scatter plots by precipitation depth

- No effects on concentration were observed according to precipitation depth. Rainfall energy determines erosion and washoff of particulates, but sufficient runoff volume is needed to carry the particulate pollutants to the outfalls. Different travel times from different locations in the drainage areas results in these materials arriving at different times, plus periods of high rainfall intensity (that increase pollutant washoff and movement) occur randomly throughout the storm. The resulting outfall stormwater concentration patterns for a large area having various surfaces is therefore complex and rain depth is just one of the factors involved. Chapter 6 examines time delivery of pollutants in more detail.

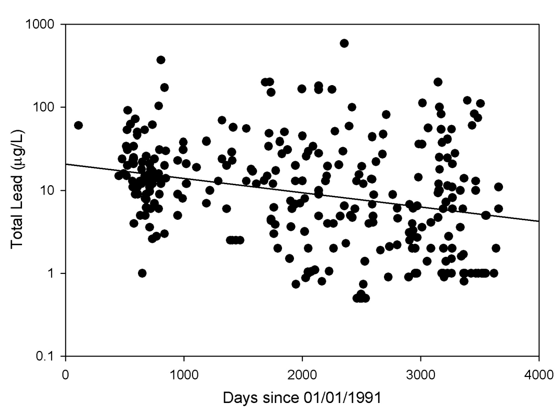
• ***Interevent Period Effects***. The following figure shows box and whisker plots of the number of days having no rain before the monitored events by each EPA Rain Zone. Antecedent dry periods before sampling was found to have a significant effect for BOD_5, COD, ammonia, nitrates, TKN, dissolved, and total phosphorus concentrations at residential land use sites. As the number of days increased, there was an increase in the concentrations of the stormwater constituents. This relationship was not observed for freeway sites.

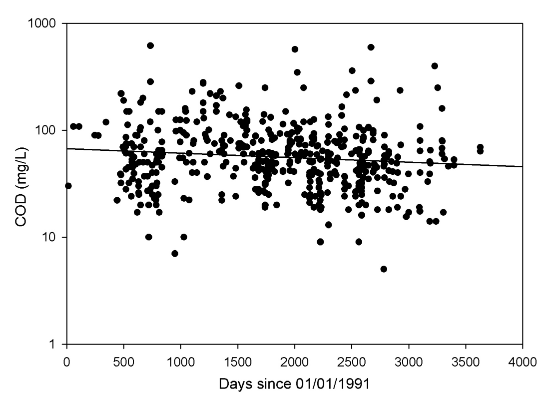
- Only EPA Rain Zone 2 has enough observations to evaluate possible effects of the antecedent dry period on the concentration of stormwater pollutants. In residential land uses, 7 out of 12 constituents indicated that antecedent dry period had a significant effect on the median concentrations. As the number of days having no rain increased, the concentrations also increased.



- Except for nitrates, all the nutrients have positive regressions inside the 95% confidence interval. In commercial land uses, the effects of antecedent dry periods on the median concentrations were less important. Only total phosphorus and total lead had significant regression results. As in the residential case, phosphorus has a positive coefficient with a small coefficient of determination. However, lead decreases with the number of dry days before the storm.

• ***Trends with Time***. The following plots show likely decreasing lead and COD concentrations with time. Statistically however, the trend lines are not significant due to the large variation in observed concentrations.





Findings from Chapter 6: Comparisons of First 30-minute Samples to 3-hour Composite Samples

• ***First-Flush Effects***. Sample collection conducted for some of the NPDES MS4 Phase I permits required both a grab and a composite sample for each event. A grab sample was to be taken during the first 30 minutes of discharge, and a flow-weighted composite sample for the entire time of discharge (up to three hours). The initial grab sample was used for the analysis of the "first flush effect," which assumes that more of the pollutants are discharged during the first period of runoff than during later periods. The composite sample was obtained with aliquots collected about every 15 to 20 minutes for at least 3 hours, or until the event ended.

- About 400 paired sets of 30-minute and 3-hour samples were available for comparisons. The following table shows the results of the analyses.

Significant First Flushes Ratios (first flush to composite median concentration)

Parameter	Commercial				Industrial				Institutional			
	n	sc	R	ratio	n	sc	R	ratio	n	sc	R	ratio
Turbidity, NTU	11	11	=	1.32			X				X	
pH, S.U.	17	17	=	1.03	16	16	=	1.00			X	
COD, mg/L	91	91	≠	2.29	84	84	≠	1.43	18	18	≠	2.73
TSS, mg/L	90	90	≠	1.85	83	83	=	0.97	18	18	≠	2.12
BOD_5, mg/L	83	83	≠	1.77	80	80	≠	1.58	18	18	≠	1.67
TDS, mg/L	82	82	≠	1.83	82	81	≠	1.32	18	18	≠	2.66
O&G, mg/L	10	10	≠	1.54			X				X	
Fecal Coliform, col/100mL	12	12	=	0.87			X				X	
Fecal Streptococcus, col/100 mL	12	11	=	1.05			X				X	
Ammonia, mg/L	70	52	≠	2.11	40	33	=	1.08	18	16	≠	1.66
$NO_2 + NO_3$, mg/L	84	82	≠	1.73	72	71	≠	1.31	18	18	≠	1.70
N Total, mg/L	19	19	=	1.35	19	16	=	1.79			X	
TKN, mg/L	93	86	≠	1.71	77	76	≠	1.35			X	
P Total, mg/L	89	77	≠	1.44	84	71	=	1.42	17	17	=	1.24
P Dissolved, mg/L	91	69	=	1.23	77	50	=	1.04	18	14	=	1.05
Ortho-P, mg/L			X		6	6	=	1.55			X	
Cadmium Total, μg/L	74	48	≠	2.15	80	41	=	1.00			X	
Chromium Total, μg/L	47	22	≠	1.67	54	25	=	1.36			X	
Copper Total, μg/L	92	82	≠	1.62	84	76	≠	1.24	18	7	=	0.94
Lead Total, μg/L	89	83	≠	1.65	84	71	≠	1.41	18	13	≠	2.28
Nickel, μg/L	47	23	≠	2.40	51	22	=	1.00			X	
Zinc, μg/L	90	90	≠	1.93	83	83	≠	1.54	18	18	≠	2.48
Turbidity, NTU			X		12	12	=	1.24	26	26	=	1.26
pH, S.U.			X		26	26	=	1.01	63	63	=	1.01
COD, mg/L	28	28	=	0.67	140	140	≠	1.63	363	363	≠	1.71
TSS, mg/L	32	32	=	0.95	144	144	≠	1.84	372	372	≠	1.60
BOD5, mg/L	28	28	=	1.07	133	133	≠	1.67	344	344	≠	1.67
TDS, mg/L	31	30	=	1.07	137	133	≠	1.52	354	342	≠	1.55
O&G, mg/L			X				X		18	14	≠	1.60
Fecal Coliform, col/100mL			X		10	9	=	0.98	22	21	=	1.21
Fecal Streptococcus, col/100 mL			X		11	8	=	1.30	26	22	=	1.11
Ammonia, mg/L			X		119	86	≠	1.36	269	190	≠	1.54
$NO_2 + NO_3$, mg/L	30	21	=	0.96	121	118	≠	1.66	324	310	≠	1.50
N Total, mg/L	6	6	=	1.53	31	30	=	0.88	77	73	=	1.22
TKN, mg/L	32	14	=	1.28	131	123	≠	1.65	335	301	≠	1.60
P Total, mg/L	32	20	=	1.05	140	128	≠	1.46	363	313	≠	1.45
P Dissolved, mg/L	32	14	=	0.69	130	105	≠	1.24	350	254	=	1.07

Ortho-P, mg/L			X		14	14	=	0.95	22	22	=	1.30
Cadmium Total, µg/L	30	15	=	1.30	123	33	≠	2.00	325	139	≠	1.62
Chromium Total, µg/L	16	4	=	1.70	86	31	=	1.24	218	82	≠	1.47
Copper Total, µg/L	30	22	=	0.78	144	108	≠	1.33	368	295	≠	1.33
Lead Total, µg/L	31	16	=	0.90	140	93	≠	1.48	364	278	≠	1.50
Nickel, µg/L			X		83	18	=	1.20	213	64	≠	1.50
Zinc, µg/L	21	21	=	1.25	136	136	≠	1.58	350	350	≠	1.59

Note: n = number of total possible events. sc = number of selected events with detected values. R = result. Not enough data (X); not enough evidence to conclude that median values are different (=); median values are different (≠).

- Generally, a statistically significant first flush is associated with a median concentration ratio of about 1.4, or greater (the exceptions are where the number of samples in a specific category is much smaller). The largest ratios are about 2.5, indicating that for these conditions, the first flush sample concentrations are about 2.5 times greater than the composite sample concentrations. More of the larger ratios are found for the commercial and institutional land use categories, areas where larger paved areas are likely to be found. The smallest ratios are associated with the residential, industrial, and open spaces land uses, locations where there may be larger areas of unpaved surfaces. Approximately 70% of the constituents in the commercial land use category had elevated first flush concentrations, about 60% of the constituents in the residential, institutional and the mixed (mostly commercial and residential) land use categories had elevated first flushes, and only 45% of the constituents in the industrial land use category had elevated first flushes. In contrast, no constituents were found to have elevated first flushes in the open space category.

- COD, BOD_5, TDS, TKN and Zn all had first flushes in all areas (except for the open space category). In contrast, turbidity, pH, fecal coliform, fecal streptococcus, total N, dissolved and ortho-P never showed a statistically significant first flush in any category.

- This investigation of first flush conditions indicated that a first flush effect was not present for all the land use categories, and certainly not for all constituents. Commercial and residential areas were more likely to show this phenomenon, especially if the peak rainfall occurred near the beginning of the event. It is expected that this effect will be more likely to occur in a watershed with a high level of imperviousness, but even so, the data indicated first flushes for less than 50% of the samples for the most impervious areas. This reduced frequency of observed first flushes in these areas most likely to have first flushes is likely associated with the varying rain conditions during the different events, including composite samples that did not represent the complete runoff durations.

- Groups of constituents showed different behaviors for different land uses. All the heavy metals evaluated showed higher concentrations at the beginning of the event in the commercial land use category. Similarly, all the nutrients showed higher initial concentrations in residential land use areas, except for total nitrogen and ortho-phosphorus. This phenomenon was not found in the bacteria analyses. None of the land uses showed a higher population of bacteria at the beginning of the event. Conventional constituents showed elevated concentrations in commercial, residential and institutional land uses.

Findings from Chapter 7: Effects of Land Use and Geographical Location on Stormwater Quality

• ***ANOVA for land use and geographical location***. All land uses in EPA Rain Zone two (except for freeways) have reduced TSS values when compared with the overall NSQD average. On the other hand, conditions in EPA Rain Zones 4, 6 and 9 have higher TSS values for the land uses noted. Industrial and freeway land uses increase the TSS concentrations compared with the other land uses, as expected from the one-way ANOVA tests. Of the 45 possible EPA Rain Zone and land use interactions, 21 have significantly different coefficients and resultant TSS concentrations. All of these possible TSS concentrations, based on this model, are shown in the following table.

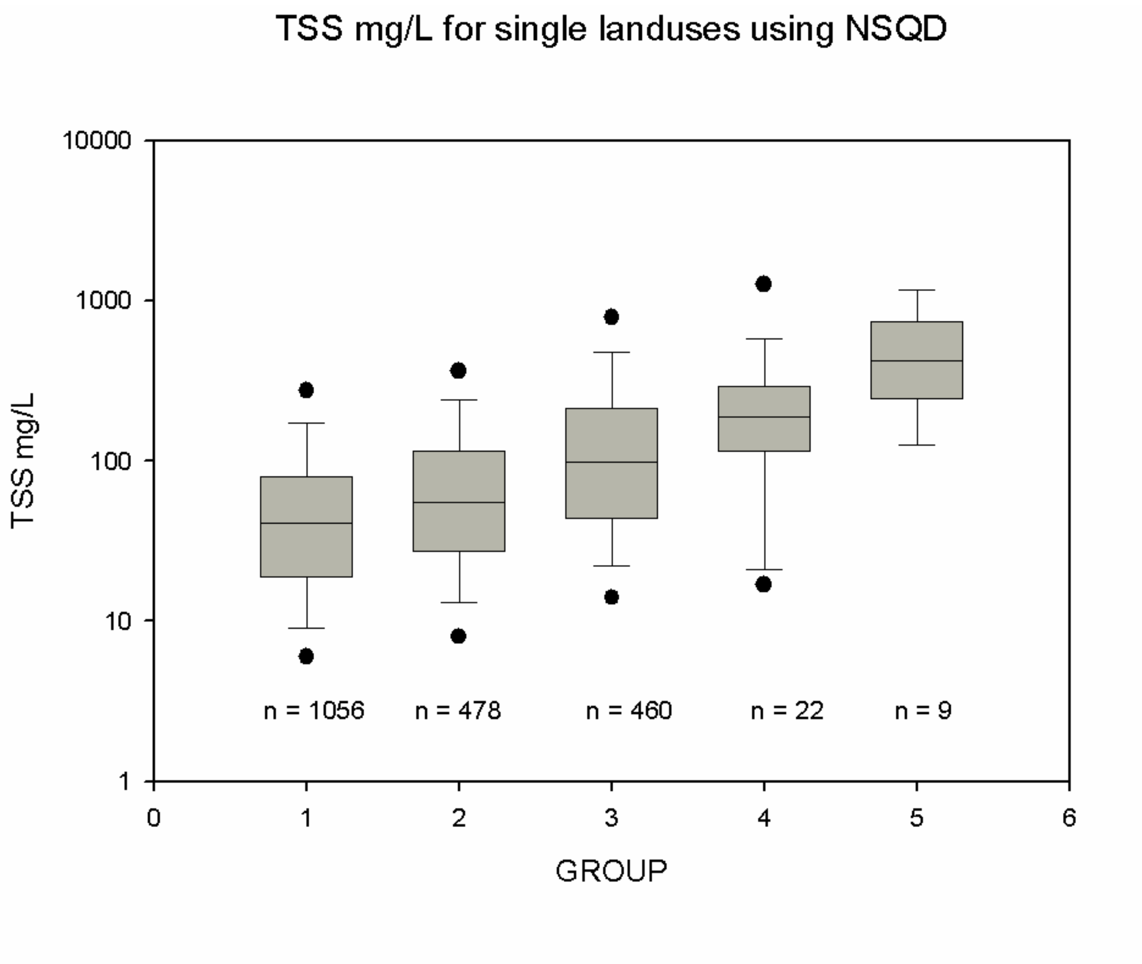
TSS Concentrations (mg/L) for Different Land Uses and Rain Zones (if values not shown, use 60 mg/L)

	1	2	3	4	5	6	7	8	9
Open space		40		139					
Residential		40		84		92	35		
Commercial		43	33	90	30	120			152
Industrial	25			103	147	206	121		
Freeways		215				86	99		409

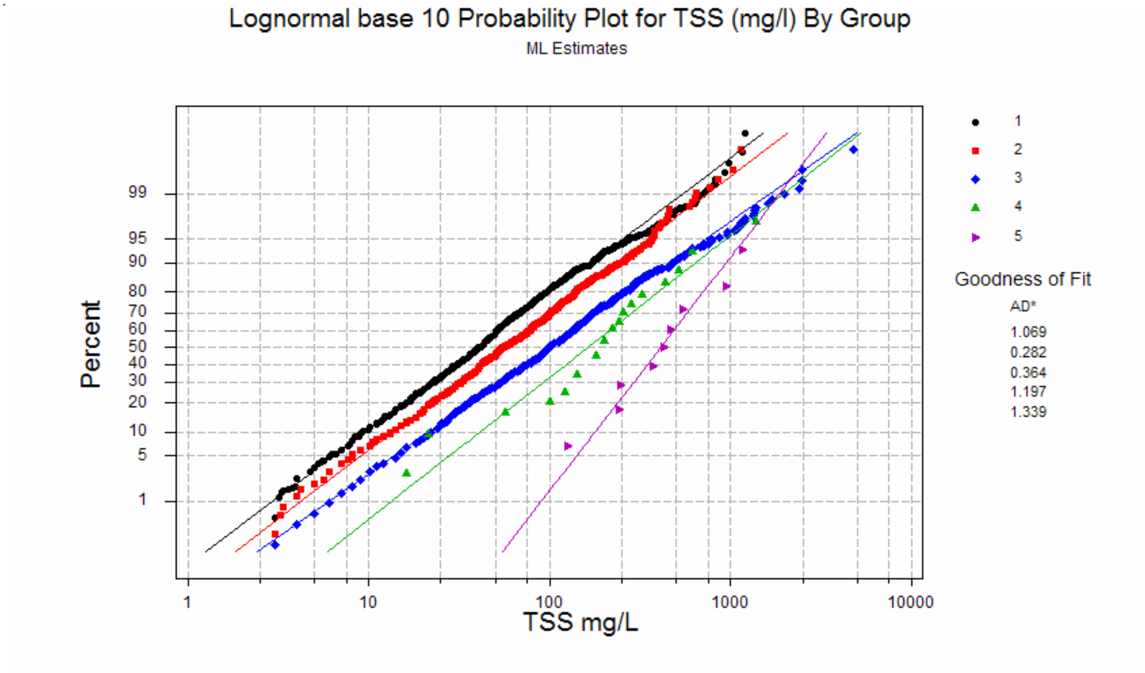
• ***Grouping of TSS Data***. The following table shows the combined groups that had statistically similar TSS concentrations. The figure also indicates that about half of the TSS single land use data in the NSQD database were in the first group (52%). Most of this data are from residential areas and EPA Rain Zone 2. Twenty-four percent of the observations were not affected by the land use – EPA Rain Zone interaction. Only 1.5% of the data are present in groups 4 and 5. These groups are significantly different than groups 1 and 2. Overall, there are three main levels of TSS concentrations in stormwater: Low (1), Medium (2) and High (3). Other minor categories correspond to groups 4 and 5 and contain the unusually high values.

Five TSS Concentration Categories in NSQD

	Land use*rain zone interactions (Rain Zone: land uses)	**Concentrations (mean ± st. dev. in mg/L)**	**Range (mean; mean – st. dev. and mean + st. dev., mg/L)**	**Number of single land use TSS observations in category in NSQD**
Low	1: residential 2: open space; residential; commercial 3: commercial 5: commercial 7: residential	3.69±1.12	40 (13 – 123)	1056
Medium	All others not noted elsewhere	4.02±1.11	56 (18 – 169)	478
High	4: residential; commercial; industrial; open space 5: industrial 6: freeways; residential; commercial 7: freeways; industrial 9: commercial	4.60±1.20	99 (30 – 330)	460
Unusually high 1	2: freeways 6: industrial			22
Unusually high 2	9: industrial			9



Box plots of five groups also showing 5th and 95th percentiles.



Probability plot of the TSS data in the NSQD using the 5 main sample groups.

• ***Land Use and Geographical Area Interactions.*** When examining the detailed land use and seasonal interactions, it is clear that some of the constituents do not have many significant interactions in these factors, or that there are too few observations (or sites) represented in the NSQD. The constituents that have few, if any clear geographical area/land use interactions include: pH (I6), temperature, hardness, oil and grease (I5 and I7), TDS (C2), ammonia (C7), and dissolved P (R2 and R5). The values in the parentheses are the significant interaction terms (the land use and the EPA Rain Zone). If individual land use/geographical interaction cell values are not available, the overall land use, or overall data base summary values should be used:

- Constituents that should clearly be separated by land use: copper, lead, and zinc
- Constituents that clearly did not have any significant differences for different land use categories, therefore use overall values: pH, temperature (obvious seasonal effects), TDS, and TKN
- Constituents where residential data should be separated from commercial plus industrial area data: TSS (possible) and nitrates plus nitrites
- Constituents where it is not clear; conflicts in p values when comparing different combinations of land uses: hardness, oil and grease, BOD_5, COD, ammonia, total P, and dissolved P

Findings from Chapter 8: Example Application of the National Stormwater Quality Database (TSS and Nutrient Export Calculations for Chesapeake Bay Watersheds)

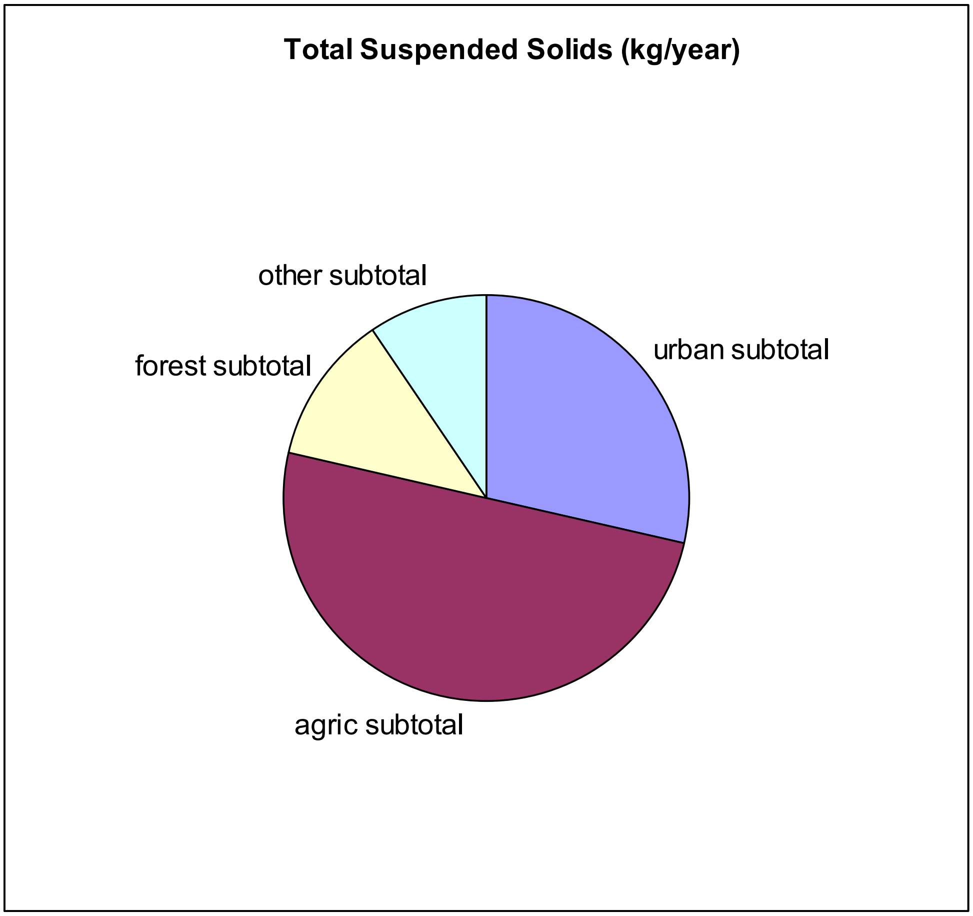
• ***Mass Discharge Calculations***. This chapter demonstrates how the NSQD information can be used to make mass discharge calculations for large drainage areas. This is an example for Maryland's Anne Arundel County, an important tributary of Chesapeake Bay. TSS and nutrient concentrations for the urban land uses in the county were calculated using NSQD data for Maryland and Virginia. Various factors were found to influence these concentrations using ANOVA analyses. Specifically, season, rain depth, and impervious cover were examined for each land use category. The resulting coefficients of variation were all significantly reduced with these categories of data, as shown on the following table.

Average Concentrations by Land Use

Land Use	Constituent	Conditions	Average (COV)
Urban - Residential	TSS		
		Summer rains (between 0.1 and 0.35 inches in depth)	143 (0.71)
		All other rains	58 (0.70)
	TP	Sites having <27% impervious cover:	
		Winter rains	0.28 (0.59)
		All other rains	0.41 (0.65)
		Sites having >27% impervious cover:	
		Winter rains (less than 0.1 inches in depth)	0.16 (0.86)
		All other rains	0.30 (0.63)
	TN	Fall rains (less than 0.1 and greater than 1 inch in depth)	1.4 (0.57)
		Winter rains (0.35 and 1 inch in depth)	1.5 (0.30)
		Fall rains (0.35 and 1 inch) and Winter rains (between 0.1 and 0.35 inches in depth)	1.9 (0.51)
		All other rains	2.4 (0.62)
		Spring and summer rains (between 0.1 and 0.35 inches in depth)	2.6 (0.38)
Urban - Commercial	TSS and TN	See tables 7.4 and 7.5	
	TP	Summer rains >1 inch and fall rains between 0.1 and 0.35 inch	0.46 (0.36)
		All other rains	0.23 (0.71)
Urban - Industrial	TSS	Fall, spring, and summer	77 (1.48)
		Winter	81 (0.93)
	TP	Rains less than 0.35 inches	0.29 (0.81)
		Rains greater than 0.35 inches	0.22 (1.05)
	TN	All conditions	2.1 (0.79)
Rural - Agricultural	Sediment		1115 lb/ac/yr (unreliable estimate)
	TN		40 lb/ac/yr
	TP		5.4 lb/ac/yr
Rural - Forest	Sediment		4500 lb/ac/yr (unreliable estimate)
	TN		0 lb/ac/yr
	TP		0 lb/ac/yr

Total Suspended Solids Concentrations and Volumetric Runoff Coefficients for Land Use Categories in Anne Arundel County, Maryland

Land Use Description	# of acres in 2000	TSS (mg/L)	Concentration reliability?	R_V	R_V reliability?
Large lot subdivision (1 unit/ 5- 10 ac)	0	60	excellent	0.09	excellent
Low-density residential (1 unit/ 5 acres to 2 units/acre)	33,337	60	excellent	0.14	excellent
Medium-density residential (2 to 8 units/acre)	33,791	60	excellent	0.23	excellent
High-density residential (8+ units/acre)	6,274	60	excellent	0.34	excellent
Commercial	11,670	58	excellent	0.72	excellent
Industrial	3,249	80	excellent	0.52	excellent
Institutional (schools, churches, military institutions, etc.)	9,813	58	moderate	0.49	excellent
Open urban land	4,139	50	moderate	0.08	excellent
Transportation	1,557	99	moderate	0.41	excellent
Extractive	1,686	350	poor	0.3	moderate
Deciduous forest	43,901	90	moderate	0.08	moderate
Evergreen forest	4,891	90	moderate	0.08	moderate
Mixed forest	56,621	90	moderate	0.08	moderate
Brush	2,565	90	poor	0.08	moderate
Wetlands	1,643	0	poor	0.65	moderate
Beaches	29	0	poor	0.1	moderate
Bare ground	224	1000	poor	0.3	moderate
Row and garden crops	300	357	very good	0.2	poor
Cropland	42,368	357	very good	0.2	poor
Orchards / vineyards / horticulture	63	357	very good	0.15	poor
Pasture	4,690	145	very good	0.08	moderate
Feeding operations	49	145	very good	0.2	poor
Agricultural building, breeding and training facilities	163	145	very good	0.5	poor



Calculated Sources of TSS for Anne Arundel County, Maryland

Research Hypotheses

The main hypothesis for this research was that commonly accepted assumptions concerning stormwater characteristics are correct and applicable for a wide range of conditions, including different land uses, precipitation depths, seasons, watershed area and geographic locations throughout the U.S. This assumption was evaluated by testing the following hypotheses:

Research Hypothesis 1. Lognormal distributions are robust descriptions of stormwater quality data and a few unusual values have little effect on dataset summary statistical descriptions.

A total of 25 constituents in 5 land uses were evaluated using the NSQD database. In 71% of the cases, lognormal distributions better described the stormwater constituent concentrations compared with gamma and exponential distributions. These last two distributions better represented 10% and 4% of the cases, respectively. In 15% of the cases, lognormal, gamma and exponential distributions did not adequately represent stormwater constituent concentrations. Constituents that mostly were not well described by any of these three distributions included: BOD_5, total arsenic, total cadmium and total copper in residential, commercial and industrial land uses.

Gamma and exponential distributions better described bacteria and nutrient concentrations in open space land use areas.

The use of the 3-parameter lognormal distribution, did not improve the description information compared with the simpler 2-parameter distribution. The 2-parameter lognormal distribution is therefore recommended for those constituents were the use of lognormal distributions produced a better fit of the data.

Unusually elevated values have a significant effect in the mean, median, standard deviation and coefficient of variation of the sample distribution. As an example, when 0.5% of the data are affected by a factor of a thousand (such as may occur when heavy metals are incorrectly expressed with mg/L units when they should be µg/L units), the coefficient of variation will be increased almost 15 times compared to the value when the extreme observations are not present. The effect on the coefficient of variation is larger as the percentage of extreme samples is reduced.

Unusually low values do not have a significant effect on the mean, median, standard deviation, or coefficient of variation, unless the percentage of samples having the low values is higher than 25%.

Research Hypothesis 2. Censored data can be adequately adjusted by substituting half of the detection limit, with little resulting effects on the mean and variance of stormwater datasets.

Replacing non-detected observations by half of the detection limit is appropriate when the percentage of left censored observations (those having concentrations lower than the detection limit) is lower than 15% of the total data set. Replacing the non-detected values with zero will have more extreme effects on these distribution summary values.

Ignoring the non-detected values will result in higher means, medians and standard deviations, and lower coefficients of variation than the true values for the distributions.

The use of the Cohen's maximum likelihood method is recommended to replace the censored observations for those constituents that have lognormal distributions. This is an appropriate method when non-paired statistical tests are to be performed, as the assignment of replacement values for specific tests in not important. However, no replacements are suitable when paired comparison tests are to be conducted, and these tests should only be conducted on the data sets having complete pairs (not using pairs where one or both parts of the pair are below detection). When calculating percentage reductions, or other comparison tests, non-detected effluent concentrations can be used, without substitutions, in calculations to determine the lower limit of removal.

When the number of non-detected observations exceed about 40% of the total number of observations, no substitution method (neither the maximum likelihood method or half of the detection limit) is suitable.

Research Hypothesis 3. Different levels of imperviousness are more important than differences in land use categories when predicting stormwater constituent concentrations.

The use of the impervious area information alone did not reduce the uncertainty about the variability of stormwater constituents. One of the main factors associated with land use concerns the activities that occur in the land use. It is expected that the use of both factors (land use and information about the surface covers in the area, such as the percentage of impervious areas) will reduce the variability of the stormwater concentrations observed, rather than when only one of these factors is considered. However, these tests were only conducted on stormwater concentrations, not on mass discharges. Increases in impervious cover are directly associated with increases in runoff volumes, and therefore in pollutant mass discharges.

When only residential area data from EPA Rain Zone 2 were used, the percentage of impervious areas was found to have a significant effect on the concentration of nitrates. The concentrations of nitrates were reduced as the percentage of impervious cover increased. This is an expected finding; when the impervious areas increase, less landscaping is likely (a major source of nutrient discharges).

No significant relationships were observed between the amount of impervious cover and any stormwater constituent concentration that was examined for industrial and commercial land use areas.

Research Hypothesis 4. Antecedent dry periods have a significant effect on stormwater constituent concentrations.

Antecedent dry periods are not the same for all the EPA Rain Zones in the country. Longer antecedent dry periods occur at west coast sites compared to other locations.

The antecedent dry periods had a positive and significant ($\alpha = 5\%$) effect on the concentration of 7 of 13 constituents examined: nutrients (ammonia, nitrates, TKN, total and dissolved phosphorus), COD and BOD_5 at residential sites located in EPA Rain Zone 2. It was not significant in oil and grease, TDS, TSS, total copper, total lead and total zinc.

Only total phosphorus and total lead concentrations were affected by the antecedent dry period at commercial sites located in EPA Rain Zone 2. Total phosphorus concentrations increased with increasing days before the sampled storm. An opposite relation was observed for total lead at commercial sites.

Only TSS was affected by the antecedent dry period at industrial sites located in EPA Rain Zone 2. A positive relationship was observed, with TSS concentrations increasing as the number of antecedent dry days increased.

Research Hypothesis 5. Outfall samples collected during the "first flush" periods of storms have significantly greater concentrations than total storm composite samples.

The first flush effect was not present for all constituents and all land uses. The phenomenon was most likely to occur in commercial and high density residential land uses, watersheds having high percentages of impervious areas. It was not observed in open space areas, watersheds having low percentages of impervious areas.

TSS, COD, TDS, total copper, total lead, total zinc and TKN had observed flush concentrations that were significantly higher than the composite sample concentrations in those areas where the "first flush" was most likely to occur. pH was the only constituent that did not indicate a first flush effect. Observed elevated first flush concentrations were less than 3 times higher than the corresponding storm composite concentrations.

Recommendations for Future Stormwater Permit Monitoring Activities

• The NSQD is an important tool for the analysis of stormwater discharges at outfalls. About a fourth of the total existing information from the NPDES Phase I program is included in the database. Most of the analyses in this research were performed for residential, commercial and industrial land uses in EPA Rain Zone 2 (the area of emphasis according to the terms of the EPA funded research). Much more data are available from other stormwater permit holders that were not included in this database. Acquiring this additional data for inclusion in the NSQD is a recommended and cost-effective activity and should be accomplished as additional data are also being obtained from on-going monitoring projects.

• The use of automatic samplers, coupled with bedload samplers, is preferred over manual sampling procedures. In addition, flow monitoring and on-site rainfall monitoring needs to be included as part of all stormwater characterization monitoring. The additional information associated with flow and rainfall data will greatly enhance the usefulness of the much more expensive water quality monitoring. Flow monitoring must also be correctly conducted, with adequate verification and correct base-flow subtraction methods applied. A related issue frequently mentioned by the monitoring agencies is the lack of on-site rainfall information for many of the sites. Using regional rainfall data from locations distant from the monitoring location is likely to be a major source of error when rainfall factors are being investigated.

• Many of the stormwater permits also only required monitoring during the first three hours of the rain event. This may have influenced the event mean concentrations if the rain event continued much beyond this time. Flow-weighted composite monitoring should continue for the complete rain duration. Monitoring only three events per year from each monitoring location requires many years before statistically adequate numbers of observations are obtained. In addition, it is much more difficult to ensure that such a small fraction of the total number of annual events is representative. Also, there is minimal value in obtaining continued data from an area after sufficient information is obtained. It is recommended that a more concentrated monitoring program be conducted for a two or three year period, with a total of about 30 events monitored for each site, covering a wide range of rain conditions. Periodic checks can be made in future years, such as repeating concentrated monitored every 10 years, or so (and for only 15 events during the follow-up surveys).

• Finally, better watershed area descriptions, especially accurate drainage area delineations, are needed for all monitored sites. While the data contained in the NSQD is extremely useful, it is believed that future monitoring information obtained as part of the stormwater permit program would be greatly enhanced with these additional considerations.

References

Adams, Barry and Fabian Papa. 2000. Urban Stormwater Management Planning with Analytical Probabilistic Methods. New York, NY: John Wiley and Sons.

Aitchison, John, and James Brown. 1969. The Lognormal Distribution. Cambridge: Cambridge University Press.

American Society of Civil Engineers (ASCE), United States Environmental Protection Agency (EPA). 2000. Determining Urban Stormwater Best Management Practice (BMP) Removal Efficiencies. Task 3.4 Final Data Exploration and Evaluation Report. Office of Water. Washington D.C.

American Society of Civil Engineers (ASCE), United States Environmental Protection Agency (EPA). 2002. Urban Stormwater BMP Performance Monitoring. A Guidance Manual for Meeting the National Stormwater BMP Database Requirements. Office of Water. Washington D.C. EPA.

Arnold, Chester, and James Gibbons. 1996. "Impervious Surface Coverage. The Emergency of a Key Environmental Indicator". Journal of the American Planning Association 62 : 243-58

American Public Health Association (APHA), American Water Works Association (AWWA), and Water Environment Federation (WEF). 1995. Standard Methods for the Examination of Water and Wastewater. 19th edition. EPS Group Inc.

Bailey, Robert. 1993. "Using Automatic Samplers to comply with stormwater regulations". In: Water Engineering and Management 140 : 23 – 25.

Bannerman, Roger, David Owens, Richard Dodds, and Nancy Hornewer. 1993. "Source of Pollutants in Wisconsin Stormwater". Water Science and Technology 28 : 241-59

Bertrand-Krajewski, Jean-Luc, Ghassan Chebbo, and Agnes Saget. 1998. "Distribution of Pollutants Mass vs Volume in Stormwater Discharges and the First Flush Phenomenon". Water Resources 32 : 2341 - 56

Behera, Pradeep K., James Y. Li, and Barry J. Adams. 2000. "Characterization of Urban Runoff Quality: A Toronto Case Study". In: Applied Modeling of Urban Water Systems. Guelph, Ontario: Computational Hydraulics International: (James W. Ed.) CHI.

Berthouex, Paul M., and Linfield C. Brown. 2002. Statistics for Environmental Engineers. Boca Raton, Florida : Lewis Publishers.

Brown, Thomas, William Burd, James Lewis, and George Chang. 1995. "Methods and Procedures in Stormwater Data Collection". In: Stormwater NPDES Related Monitoring Needs. ASCE. 194 - 206

Burton, Allen, and Robert Pitt. 2002. Stormwater Effects Handbook. A Toolbox for Watershed Managers, Scientists, and Engineers. Boca Raton, Florida : Lewis Publishers.

Clark, Shirley, Robert Pitt , and Steven Burian. 2004. Urban wet weather flows – 2003 literature review. Water Environment Research 76

Clark, Shirley, Robert Pitt , and Steven Burian. 2003. Urban wet weather flows – 2002 literature review. Water Environment Research 75

Clark, Shirley, Robert Pitt , and Steven Burian. 2002. Urban wet weather flows – 2001 literature review. Water Environment Research 74

Clark, Shirley, Ronald Rovansek, Leonard Wright, James Heaney, Richard Field, and Robert Pitt. 2001. Urban wet weather flows – 2000 literature review. Water Environment Research 73

Cohen A. 1959. "Simplified Estimators for the Normal Distribution when Samples are Singly Censored or Truncated". Technometrics 1: 217-37

Deletic, Ana. 1998. "The First Flush Load of Urban Surface Runoff". Water Research 32 : 2462-70

Driver, Nancy E., Martha H. Mustard, Bret Rhinesmith, and Robert F. Middleburg. 1985. U.S. Geological Survey Urban Stormwater Database for 22 Metropolitan Areas Throughout the United States. U.S. Geological Survey Open File Report 85-337. Denver, CO: USGS.

Dodson, Roy. 1995. Stormwater Pollution Control, Industry and Construction NPDES compliance. New York, NY: Mc Graw Hill.

Doerfer, John, and Ben Urbonas. Stormwater Quality Characterization in the Denver Metropolitan Area. In: Flood Hazard News 23 (1993): 21 - 23

Ellis, Bryan. 1986. Pollutional Aspects of Urban Runoff. In : Urban Runoff Pollution, NATO ASI Series (G10): 1-38. Berlin, Heidelberg : Spring Verlag Eds.

Fligner Michael, and George E. Policello. 1981. "Robust Rank Procedures for the Behrens-Fisher Problem". Journal of the American Statistical Association 76 : 162 - 68

Gibbons, Jean, and Subhabrata Chakraborti. 2003. Nonparametric Statistical Inference. New York, NY: Marcel Dekker Inc.

Gray, John, Douglas Glysson, Lisa Turcios, and Gregory E. Schwarz. 2000. Comparability of Suspended-Sediment Concentration and Total Suspended Solids Data. Water Resources Investigation Report 00-4191: USGS.

Hollander, Myles, and Douglas Wolfe. 1999. Nonparamentric Statistical Methods. Second Edition. New York, NY : John Wiley and Sons.

Kottegoda, Nathabandu T., and Renzo Rosso. 1998. Statistics, Probability and Reliability for Civil and Environmental Engineers. New York, NY: WCB McGraw-Hill.

Langland, Michael J. Patricia L. Lietman, and Scott Hoffman. 1995. Synthesis of Nutrient and Sediment Data for Watersheds Within the Chesapeake Bay Drainage Basin. U.S. Geological Survey Open File Report 95-4233. Lemoyne, PA: USGS.

Leecaster, Molly, Kenneth C. Schiff, and Liesl Tiefenthaler. 2002. "Assessment of efficient sampling designs for urban stormwater monitoring". Water Research 36: 1556 – 64

Massey, Frank J. Jr. 1950. "A Note on the Power of a Nonparametric Test". Annals of Mathematical Statistics 21 : 440 – 43

Massey, Frank J. Jr. 1951. "The Kolmogorov-Smirnov Test for Goodness of Fit". Journal of the American Statistical Association 46 : 67 – 78

National Council of the Paper Industry for Air and Stream Improvement (NCASI). 1995. A Statistical Method and Computer Program for Estimating the Mean and Variance of Multi-level Left Censored Data Sets. Technical Bulletin No. 703.

Nara, Yukio, Robert Pitt, Rocky Durrans, and Jason Kirby. 2006. The Use of Grass Swales as a Stormwater Control. In: Stormwater and Urban Water System Modeling (James W. Ed.) Guelph, Ontario: Computational Hydraulics International: CHI.

Pitt, Robert, and James McLean. 1986. Toronto Area Watershed Management Strategy Study.Humber River Pilot Watershed Project. Ontario Ministry of the Environment, Toronto, Ontario.

Pitt, Robert. 1986b. The Incorporation of Urban Runoff Controls in the Wisconsin Priority Watershed Program. In: Advanced Topics in Urban Runoff Research (Urbonas B., Roesner L.A. Ed.). Engineering Foundation and ASCE. New York NY: 290 - 313

Pitt, Robert. Small Storm Urban Flow and Particulate Washoff Contributions to Outfall Discharges Ph.D. dissertation. University of Wisconsin. 1987.

Pitt, Robert, and John Voorhees. 1995. Source Loading and Management Model (SLAMM). In: Seminar Watershed Management at the local, County, and State Levels. Center for Environmental Research Information. United States Environmental Protection Agency EPA: EPA/625/R-95/003. Cincinnati, Ohio

Pitt, Robert, Lilburn M., Rocky Durrans, Steve Burian, Sthephan Nix , John Voorhees, and Jeff Martinson. 1999.Guidance Manual for Integrated Wet Weather Flow (WWF) Collection and Treatments Systems for Newly Urbanized Areas. U.S. Environmental Protection Agency.

Pitt, Robert, Alexander Maestre, and Renee Morquecho. 2003. Evaluation of NPDES Phase I Municipal Stormwater Monitoring Data. In: National Conference on Urban Stormwater: Enhancing the Programs at the Local Level. EPA/625/R-03/003

Roa-Espinosa, Aicardo, and Roger Bannerman. 1995. "Monitoring BMP effectiveness at industrial sites". In: Conference on Stormwater NPDES Related Monitoring Needs. ASCE: 467 - 486.

Schueler, Tom, and John Galli. 1995. The Environmental Impacts of Stormwater Ponds. In: Stormwater Runoff and Receiving Systems: Impact, Monitoring and Assessment (CRC Lewis Publishers), Boca Raton, Florida, 177-191

Soeur, Channy, James Hubka , George Chang, and Steve Stecher. 1995. Methods for Assessing Urban Stormwater Pollution. In: Conference on Stormwater NPDES Related Monitoring Needs. ASCE: 558 - 68

Smullen, Jim, and Kelly A. Cave. 2002. National Stormwater Runoff Pollution Database. In: Wet-Weather Flow in the Urban Watershed (Field, R., and D. Sullivan. Ed.). Boca Raton, Florida: Lewis Publishers.

Swietlik, William, William D. Tate, Robert Goo, and Eric Burneson. 1995. Strategies for Using NPDES Storm Water Data. In: Conference on Stormwater NPDES Related Monitoring Needs. ASCE: 244 - 76

United States Environmental Protection Agency (EPA). 1983. Results of the Nationwide Runoff Program. Water Planning Division, PB 84-185552. Washington D.C: EPA

United States Environmental Protection Agency (EPA). 1986. Methodology for Analysis of Detention Basins for Control of Urban Runoff Quality. Office of Water, Nonpoint Source Division. Washington D.C: EPA

Van Buren, Mark, Edgar Watt, and Jiri Marsalek. 1997. "Application of the Lognormal and Normal Distribution to Stormwater Quality Parameters". Water Research

31 (1): 95 – 104

Water Environment Federation (WEF), American Society of Civil Engineers (ASCE). 1998. Urban Runoff Quality Management. WEF Manual Practice Number 23. ASCE Manual and Report on Engineering Practice Number 87.

Appendix A: Sites Included in the Database

The following table shows the number of samples, land use and community for each site, along with the site ID.

Table A1. Site Name and Land Use

State	LOCATION ID	Land use	Jurisdiction	Site ID
AL	ALHUCHIP	ID	City of Huntsville	Chase Industrial Park
AL	ALHUDRAV	RE	City of Huntsville	Drake Avenue
AL	ALHUHURI	RE	City of Huntsville	Hunters Ridge
AL	ALHUMASM	CO	City of Huntsville	Madison Square Mall
AL	ALHUWERP	CO	City of Huntsville	Western Research Park
AL	ALJC004L	CO FW ID RE IS	Jefferson County	C004L
AL	ALJC004R	CO FW ID RE IS	Jefferson County	C004R
AL	ALJCC001	FW ID CO	Jefferson County	C001
AL	ALJCC002	RE OP ID CO	Jefferson County	C002
AL	ALJCC009	RE FW	Jefferson County	C009
AL	ALJCC010	RE FW	Jefferson County	C010
AL	ALJCC012	CO FW RE OP	Jefferson County	C012
AL	ALMOCREO	RE	City of Mobile	Creola
AL	ALMODAPH	CO	City of Mobile	Daphne
AL	ALMOSARA	RE	City of Mobile	Saraland
AL	ALMOSIIV	ID	City of Mobile	Mobile Site IV
AL	ALMOSITV	CO	City of Mobile	Mobile Site V
AL	ALMOSIVI	RE	City of Mobile	Mobile Site VI
AL	ALMOTHEO	ID	City of Mobile	Theodore
AZ	AZMCA001	ID	Maricopa Cnty	48th Street Drain
AZ	AZMCA002	OP	Maricopa Cnty	South Mountain Park
AZ	AZMCA003	ID	Maricopa Cnty	27th Ave at Salt River
AZ	AZMCA004	RE CO	Maricopa Cnty	Aqua Fria at Youngtown
AZ	AZMCA005	CO	Maricopa Cnty	43rd Ave at Peoria
AZ	AZMCA006	RE	Maricopa Cnty	67th Ave Olive Ave at Glendale
AZ	AZTUA001	RE	Tucson	Grant Road and Wilson Avenue
AZ	AZTUA002	RE	Tucson	Greenlee Road

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
AZ	AZTUA003	CO	Tucson	El Con Mall
AZ	AZTUA004	ID	Tucson	17th Street
CA	CAALAL03	ID RE	Alameda County	Woods Street
CA	CAALAL04	RE CO FW	Alameda County	Alice Street and 4th
CA	CAALAL07	CO RE	Alameda County	Cotter Way
CA	CAALAL09	ID	Alameda County	Pacific Street
CA	CAALAL10	ID CO RE	Alameda County	37TH ST 8TH AVE
CA	CACTA001	FW	Caltrans	3 07 Sacramento
CA	CACTA002	FW	Caltrans	4 35 Solano
CA	CACTA003	FW	Caltrans	6 205 Fresno
CA	CACTA004	FW	Caltrans	6 209 Fresno
CA	CACTA005	FW	Caltrans	7 01 Los Angeles
CA	CACTA006	FW	Caltrans	7 127 Los Angeles
CA	CACTA007	FW	Caltrans	7 128 Los Angeles
CA	CACTA008	FW	Caltrans	7 201 Los Angeles
CA	CACTA009	FW	Caltrans	7 202 Los Angeles
CA	CACTA010	FW	Caltrans	7 203 Los Angeles
CA	CACTA011	FW	Caltrans	8 01 Riverside
CA	CACTA012	FW	Caltrans	8 02 San Bernardino
CA	CACTA013	FW	Caltrans	8 03 Riverside
CA	CACTA014	FW	Caltrans	12 01 Orange
CA	CACTA015	FW	Caltrans	12 02 Orange
CO	COCSA001	CO OP RE	Colorado Springs	Sixteenth Hole Valley Hi Golf Course
CO	COCSA002	ID OP	Colorado Springs	Chestnut Street at Douglas Creek
CO	COCSA003	ID CO	Colorado Springs	Beacon Street at Buchanan Street
CO	COCSA004	RE OP	Colorado Springs	Wasatch Street at Cross Lane
CO	COCSA005	OP CO ID	Colorado Springs	Wal-Mart at Eighth Street
CO	CODEA001	CO	Denver Metro	Cherry Creek Storm Drain at Colfax Ave
CO	CODEA002	CO	Denver Metro	Cherry Creek Storm Drain at University Blvd
CO	CODEA003	RE	Denver Metro	North Sanderson Gulch Tributary at Lakewood
CO	CODEA004	ID	Denver Metro	Sand Creek Tributary at 34th and Havana
CO	CODEA005	RE	Denver Metro	Shop Creek at Parker Road
CO	CODEA006	ID	Denver Metro	South Platte River Storm Drain at 54th and Steele
CO	CODEA007	ID	Denver Metro	South Platte River Storm Drain at 7th Ave
CO	CODEA008	CO	Denver Metro	Villa Italia Storm Drain at Lakewood
GA	GAATAT01	ID	City of Atlanta	Ellsworth Industrial Drive
GA	GAATAT02	RE	City of Atlanta	Beverly Road Doncaster Drive
GA	GACLCOSI	ID	Clayton County	Southridge Industrial Park
GA	GACLCOTR	RE	Clayton County	Tara Road

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
GA	GACOC1A2	RE CO	Cobb County	Cobb Long Term 1 Pebble Creek Lot989
GA	GACOC1A3	RE	Cobb County	Cobb Long Term 1 Sewell Mill Creek Roswell Road
GA	GACOCOL2	RE CO	Cobb County	Cobb Long Term 2 Worley Rd Noonday Creek
GA	GADKCOTD	ID CO	Dekalb County	Truman Drive
GA	GAFUCOS1	RE OP	Fulton County	Johns Creek Buice Road
GA	GAFUCOS2	ID OP	Fulton County	Boat Road Blvd Grange Blvd
GA	GAFUCOS3	RE CO	Fulton County	Long Island Creek Northside Drive
ID	IDADA001	CO ID	Ada County Highway District	Koppels Site
ID	IDADA002	RE	Ada County Highway District	Lucky Drive Site
ID	IDADA003	RE FW	Ada County Highway District	Franklin Road Site
ID	IDADA004	CO ID	Ada County Highway District	Production Avenue Site
KA	KATOATWO	RE	City of Topeka	Atwood
KA	KATOBROO	RE	City of Topeka	Brookfield
KA	KATOJACK	CO	City of Topeka	Jackson
KA	KATOSTFE	ID	City of Topeka	Santee
KA	KAWIHUNT	RE	City of Wichita	Huntington
KA	KAWIMCLE	ID	City of Wichita	McLean
KA	KAWISBWY	RE CO	City of Wichita	Broadway
KA	KAWITOWN	CO	City of Wichita	Towne East
KY	KYLOTSR1	RE	City of Louisville	Buechel
KY	KYLOTSR2	ID	City of Louisville	Obannon
KY	KYLOTSR3	RE	City of Louisville	St Matthews
KY	KYLOTSR4	ID	City of Louisville	Okolona
KY	KYLOTSR5	RE CO	City of Louisville	Pleasure Ridge Park
KY	KYLOTSR6	RE CO	City of Louisville	Hurstbourne Acres
KY	KYLXEHL4	OP	City of Lexington	Lakeside golf
KY	KYLXEHL5	OP	City of Lexington	Walnut Hill Chilesburg
KY	KYLXEHL6	FW	City of Lexington	Alumni ManOwar
KY	KYLXEHL7	RE	City of Lexington	Squires Road
KY	KYLXNEL1	RE	City of Lexington	Greenbrier East
KY	KYLXNEL2	RE OP	City of Lexington	Greenbrier
KY	KYLXNEL3	CO	City of Lexington	Eastland
KY	KYLXTBL1	RE	City of Lexington	Mt Vernon
KY	KYLXTBL2	ID	City of Lexington	Leestown
KY	KYLXTBL3	OP	City of Lexington	Viley Road
KY	KYLXWHL1	CO	City of Lexington	Wilhite Drive
MA	MABOA001	RE OP	Boston	Charlestown 29J212
MA	MABOA002	RE	Boston	West Roxbury 13D077 078
MA	MABOA003	CO	Boston	Dorchester 8J102

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
MA	MABOA004	ID	Boston	Brighton 25E037
MA	MABOA005	FW ID	Boston	Hyde Park 2F120
MA	MABOA006	RE	Boston	Mount Vernon 26K099
MA	MABOA007	OP	Boston	Wesley G Ross 6G108
MD	MDAACOMW	ID	Anne Arundel	Midway industrial park MW
MD	MDAACOOD	RE	Anne Arundel County	Odenton OD
MD	MDAACOPP	CO	Anne Arundel	Parole Plaza PP
MD	MDAACORK	RE	Anne Arundel	Rolling Knolls RK
MD	MDAACOSC	CO ID	Anne Arundel	Science Drive SC
MD	MDBACOBC	ID	Baltimore County	Brien Run BC
MD	MDBACOLC	CO RE OP	Baltimore County	Long Quarter Branch LC
MD	MDBACOSC	RE	Baltimore County	Spring Branch SC
MD	MDBACOTC	ID	Baltimore County	Tobasoo creek TC
MD	MDBACOWC	RE	Baltimore County	White Marsh Run WC
MD	MDBCTYBO	ID	Baltimore City	BO
MD	MDBCTYFM	ID	Baltimore City	FM
MD	MDBCTYHA	RE CO	Baltimore City	Hamilton HA
MD	MDBCTYHO	RE	Baltimore City	Home land HO
MD	MDBCTYHR	RE	Baltimore City	Herring Run HR
MD	MDBCTYKO	CO	Baltimore City	Coppers Avenue KO
MD	MDCHCOIP	ID	Charles County	IP
MD	MDCHCOPA	RE	Charles County	PA
MD	MDCHCOPF	RE	Charles County	PF
MD	MDCHCOTG	CO ID	Charles County	TG
MD	MDCLCOBP	CO ID	Carroll County	Route 97 airport industrial BP
MD	MDCLCOCE	RE	Carroll County	Candice estates CE
MD	MDCLCOJS	CO	Carroll County	John street JS
MD	MDCLCOKW	OP RE	Carroll County	Kate Wagner KW
MD	MDCLCOSD	RE ID	Carroll County	Sunset Drive SD
MD	MDHACFBA	XX	Harford County	FBA
MD	MDHACOBP	RE	Harford County	Brentwood Park Woodland Hills
MD	MDHACOCF	CO	Harford County	Constant Friendship CF
MD	MDHACOCS	RE	Harford County	Cool Spring CS
MD	MDHACOGR	RE	Harford County	Green Ridge-II GR
MD	MDHACOIC	ID	Harford County	Greater Harford industrial centre IC
MD	MDHOCODC	CO	Howard County	Dobbin center DC
MD	MDHOCOFM	ID	Howard County	Food market FM
MD	MDHOCOGM	RE	Howard County	Green Moon GM
MD	MDHOCOMH	RE	Howard County	Murray Hill MH
MD	MDHOCOOC	ID	Howard County	Oak land centre OC
MD	MDMOCOBC	CO	Montgomery County	Burtons ville crossing BC
MD	MDMOCOCV	ID	Montgomery County	Coles villeCV

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
MD	MDMOCONV	RE	Montgomery County	Venture V
MD	MDMOCOQA	RE	Montgomery County	Quaint Acres QA
MD	MDMOCOSL	ID	Montgomery County	Southlawn lane SL
MD	MDMOCOWP	CO	Montgomery County	Wheaten plaza WP
MD	MDPGCOS1	CO	Prince Georges County	Brightseat Rd S1
MD	MDPGCOS2	RE	Prince Georges County	Flagstaff Street S2
MD	MDPGCOS3	CO ID	Prince Georges County	Maryland 50 industrial park S3
MD	MDPGCOS4	RE	Prince Georges County	Wayne Place S4
MD	MDPGCOS5	ID	Prince Georges County	John Hanson S5
MD	MDPGCOS6	ID	Prince Georges County	Pennsy Dr N3
MD	MDSHDTDV	OP ID	State Highway	DV
MD	MDSHDTPS	OP ID	State Highway	PS
MN	MNMISD01	RE	City of Minneapolis	E Harriet Pkwy W44 St
MN	MNMISD02	RE	City of Minneapolis	Luella St Orange Ave
MN	MNMISD03	ID	City of Minneapolis	Vandalia st
MN	MNMISD04	RE CO	City of Minneapolis	Charles Ave
MN	MNMISD05	RE CO	City of Minneapolis	E 29 St 31 Ave S
NC	NCCHBREV	ID	City of Charlotte	Brevi1
NC	NCCHHIDD	RE	City of Charlotte	Hiddr2
NC	NCCHHOSK	ID	City of Charlotte	Hoski2
NC	NCCHNANC	RE	City of Charlotte	Nancr1
NC	NCCHROSE	RE ID OP CO	City of Charlotte	Rosem1
NC	NCCHSHEF	OP RE	City of Charlotte	Shefo1
NC	NCCHSIMS	RE	City of Charlotte	Simsr3
NC	NCCHSTAR	CO	City of Charlotte	Starc1
NC	NCCHYARD	CO RE	City of Charlotte	Yardc2
NC	NCFV71ST	IS	City of Fayetteville	71 ST High School 100ft NE Raeford SR1409
NC	NCFVCLEA	RE	City of Fayetteville	3606 Clearwater Drive
NC	NCFVELMS	CO	City of Fayetteville	ELM Street Eutaw Shopping Center
NC	NCFVROSE	RE OP CO	City of Fayetteville	Rose Apartments 225 Tiffany Court

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
NC	NCFVSTRK	OP	City of Fayetteville	Strickland Bridge Road
NC	NCFVTRYO	RE	City of Fayetteville	1740 Tryon Rd
NC	NCFVWINS	ID	City of Fayetteville	Winslow Pine Railroad tracks
NC	NCGRATHE	CO	City of Greensboro	Athena
NC	NCGRCOUN	OP	City of Greensboro	Country Park
NC	NCGRHUST	ID	City of Greensboro	Husbands Street
NC	NCGRMERR	CO	City of Greensboro	Merrit Drive
NC	NCGRRAND	RE	City of Greensboro	Randlem Road
NC	NCGRUNIO	ID CO IS RE	City of Greensboro	Union Street
NC	NCGRWILL	RE	City of Greensboro	Willoughby
NC	NCRASIT1	OP RE	City of Raleigh	I40 400ft east S State Street
NC	NCRASIT2	RE CO	City of Raleigh	Williamson Drive Pineview Street
NC	NCRASIT3	RE CO OP	City of Raleigh	I40 Dandridge Drive Bunche Drive
NC	NCRASIT4	CO RE	City of Raleigh	Williamson Drive Wade Avenue
NC	NCRASIT5	ID OP	City of Raleigh	Pylon Drive 100ft North Hutton Street
NC	NCRASIT6	ID RE	City of Raleigh	South Wilmington Street City Farm Road
NC	NCRASIT7	CO RE ID OP	City of Raleigh	50ft east N West Street Peace Street Dortch Street
OR	ORCCA001	RE	Clackamas County	Bell Station
OR	ORCCA002	RE	Clackamas County	Lake Oswego
OR	ORCCA003	RE	Clackamas County	Milwaukie
OR	ORCCA004	RE	Clackamas County	Oregon City
OR	ORCCA005	CO	Clackamas County	Wilson Road
OR	OREUA001	CO	City of Eugene	C1 Olive Ave
OR	OREUA002	XX	City of Eugene	M1 Bailey Ave
OR	OREUA003	RE	City of Eugene	R1 Coetivy Ave
OR	ORGRA001	ID RE	City of Gresham	E 3 Boeing
OR	ORGRA002	RE CO	City of Gresham	I 13 Riverview St
OR	ORGRA003	RE	City of Gresham	K 4 Fairview Park
OR	ORGRA004	CO RE	City of Gresham	M 16
OR	ORODA001	FW	ODOT	Ashland
OR	ORODA002	FW	ODOT	Astoria

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
OR	ORODA003	FW	ODOT	Eugene
OR	ORODA004	FW	ODOT	Neskowin
OR	ORODA005	FW	ODOT	Portland
OR	ORPOA001	CO	City of Portland	C 1 Jantzen Beach
OR	ORPOA002	CO RE	City of Portland	C 2 Salmon Street
OR	ORPOA003	ID	City of Portland	I 1 Yeon Ave 35th Ave
OR	ORPOA004	ID	City of Portland	I 2 Swan Island
OR	ORPOA005	RE CO	City of Portland	M 1 Columbia Slough
OR	ORPOA006	RE	City of Portland	R 2 Sandy Boulevard
OR	ORPOA007	FW	City of Portland	T 1
OR	ORSAA001	CO RE	City of Salem	Commercial
OR	ORSAA002	CO	City of Salem	Cottage
OR	ORSAA003	ID	City of Salem	Edgewater
OR	ORSAA004	RE	City of Salem	Redleaf
PA	PAPH0864	RE	Philadelphia	Cresheim Creek
PA	PAPH0891	RE	Philadelphia	Tacony Creek
PA	PAPH1014	RE	Philadelphia	Byberry Creek
PA	PAPH1051	RE CO	Philadelphia	Wooden Bridge Run
PA	PAPH1182	OP RE	Philadelphia	North Byberry Creek
TN	TNKXTYAP	ID OP RE	City of Knoxville	Acker Place
TN	TNKXTYFC	RE FW	City of Knoxville	First Creek
TN	TNKXTYGV	RE OP	City of Knoxville	Gallaher View
TN	TNKXTYTC	OP ID IS RE	City of Knoxville	Third Creek
TN	TNKXTYWE	CO RE IS OP	City of Knoxville	Wellington Drive
TN	TNMET207	OP	City of Memphis	207 Walnut Grove
TN	TNMET211	ID	City of Memphis	211 Warford
TN	TNMET231	RE	City of Memphis	231 Raleigh Lagrange
TN	TNMET260	CO RE	City of Memphis	260 Austin Peay
TN	TNMET410	RE	City of Memphis	410 Whitehaven
TX	TXARA001	CO	City of Arlington	The Parks mall AC603
TX	TXARA002	RE	City of Arlington	R Legacy PK AR602
TX	TXARA003	RE	City of Arlington	Trib to W FK Tri AR601
TX	TXARA004	RE	City of Arlington	Trib To Johnson Creek AI604
TX	TXDAA001	ID	City of Dallas	Joes Cr 138
TX	TXDAA002	ID	City of Dallas	Bastille St 325
TX	TXDAA003	RE ID CO	City of Dallas	Knights Branch 34
TX	TXDAA004	RE	City of Dallas	White Rock Creek 86
TX	TXDAA005	RE	City of Dallas	Ash Creek 55
TX	TXDAA006	RE OP	City of Dallas	Newton Creek 189
TX	TXDCA001	OP FW	TXDOT Dallas	Mountain Creek DH902
TX	TXDCA002	OP FW	TXDOT Dallas	Bachman Branch DH901
TX	TXFWA001	OP RE	City of Fort Worth	Clear FK Trin R TRI STG1
TX	TXFWA002	OP IS	City of Fort Worth	Pylon St PY1
TX	TXFWA003	CO RE	City of Fort Worth	West Fk Trinity R BEL1

Table A1.Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
TX	TXFWA004	ID	City of Fort Worth	Dry Branch CRA1
TX	TXFWA005	RE CO OP	City of Fort Worth	Estrn Hills HS EH1
TX	TXGAA001	RE CO ID OP	City of Garland	Mills Branch Tributary GM404
TX	TXGAA002	ID OP	City of Garland	Trib to Duck Creek GI401
TX	TXGAA003	RE CO	City of Garland	Sleepy Hollow St GR402
TX	TXGAA004	CO RE	City of Garland	I635 Outfall at CE GC603
TX	TXHCA001	RE	Harris County	Overbluff
TX	TXHCA002	RE OP	Harris County	Cypress Trace Station
TX	TXHCA003	CO	Harris County	Steeplechase
TX	TXHCA004	ID	Harris County	Bayport
TX	TXHCA005	CO	Harris County	WillowBrook Mall
TX	TXHCA006	RE	Harris County	Little Cypress Creek
TX	TXHCA007	OP	Harris County	Hadden Road
TX	TXHOA001	OP	City of Houston	Briar Forest
TX	TXHOA002	ID	City of Houston	Eleventh Street
TX	TXHOA003	RE	City of Houston	Lazybrook
TX	TXHOA004	CO	City of Houston	Memorial City Mall
TX	TXHOA005	RE	City of Houston	Tanglewilde
TX	TXIRA001	RE	City of Irving	Bear Cr IR501
TX	TXIRA002	ID RE OP	City of Irving	Cottonwood Branch Trib IM504
TX	TXIRA003	ID CO	City of Irving	Hereford Rd II503
TX	TXIRA004	ID CO	City of Irving	Trib to ELM FK II502
TX	TXMEA001	CO FW	City of Mesquite	South mesquite I635 MC801
TX	TXMEA002	RE	City of Mesquite	South Mesquite South Parkway MC802
TX	TXMEA003	RE	City of Mesquite	South Mesquite Bruton Road MC803
TX	TXPLA001	RE CO	City of Plano	Rowlett Cr PR701
TX	TXPLA002	OP FW	City of Plano	Beck Brach PU704
TX	TXPLA003	CO OP	City of Plano	Spring Creek PC702
TX	TXPLA004	CO ID RE	City of Plano	Spring Creek PI703
TX	TXTCA001	FW OP ID	TXDOT Tarrant County	Deer Creek TH904
VA	VAARLCV2	RE	Arlington	Colonial Village CV2
VA	VAARLLP1	RE	Arlington	Little Pimmet LP1
VA	VAARLRS3	CO	Arlington	Randolph Street RS3
VA	VAARLTC4	ID	Arlington	Trades Center TC4
VA	VACHCCC4	CO	Chesterfield County	CoverLeaf Mall CC4
VA	VACHCCC5	RE	Chesterfield County	Buck Rub Drive CC5
VA	VACHCN1A	RE	Chesterfield County	Gates bluff 1A
VA	VACHCN2A	RE	Chesterfield County	Helmsley road 2A
VA	VACHCOF1	ID	Chesterfield County	unnamed OF1

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
VA	VACHCOF2	OP RE	Chesterfield County	Oak river drive OF2
VA	VACHCOF3	RE	Chesterfield County	Kings mill road OF3
VA	VACHCOF4	RE	Chesterfield County	OF4
VA	VACHCOF5	RE	Chesterfield County	Laurel oak road OF5
VA	VACPTC1A	RE	Chesapeake	Briarfield Drive C1A
VA	VACPTSF2	RE	Chesapeake	Woodards Mill SF2
VA	VACPTYC1	RE	Chesapeake	Etheridge rd Mt Pleasant Rd C1
VA	VACPTYC2	RE OP	Chesapeake	Hunningdon Lakes C2
VA	VACPTYC3	RE	Chesapeake	Horse Run Ditch C3
VA	VACPTYC4	CO	Chesapeake	Woodford Square Along Battlefield Blvd C4
VA	VACPTYC5	ID	Chesapeake	Cavalier Industrial Park C5
VA	VACPTYO1	ID	Chesapeake	Paramount Avenue O1
VA	VAFFCOF1	RE	Fairfax County	Apple Ridge Road
VA	VAFFCOF2	RE CO ID	Fairfax County	Sunset Hills Road
VA	VAFFCOF3	RE	Fairfax County	Onley Road
VA	VAFFCOF4	CO	Fairfax County	Green Look Place
VA	VAFFCOF5	RE	Fairfax County	Oakton Terrace Road
VA	VAFFCOF6	CO	Fairfax County	Fairview Park Drive
VA	VAFFCOF7	RE	Fairfax County	Lakeview Drive
VA	VAFFCOF8	RE	Fairfax County	Pumphrey Drive
VA	VAFFCOF9	RE	Fairfax County	Rock Ridge Road
VA	VAFFOF10	ID	Fairfax County	Boston Boulevard
VA	VAFFOF11	ID	Fairfax County	Prosperity Avenue
VA	VAHAHMS2	ID	Hampton	Copeland Industrial Park HMS2
VA	VAHAHMS5	RE OP	Hampton	Grays Landing HMS5
VA	VAHATYH1	CO	Hampton	Commerce Drive H1
VA	VAHATYH2	ID	Hampton	Mingee Drive H2
VA	VAHATYH3	RE	Hampton	Hampton Club H3
VA	VAHATYH4	RE	Hampton	Bay Avenue H4
VA	VAHATYH5	RE	Hampton	Willow Oaks Boulevard H5
VA	VAHCCOC1	CO	Henrico County	Dickens Place C1
VA	VAHCCOC2	CO	Henrico County	Carousel Lane C2
VA	VAHCCON1	ID	Henrico County	Tomlyn Street N1
VA	VAHCCON2	ID	Henrico County	Impala Drive and Galaxy Road N2
VA	VAHCCOR1	RE	Henrico County	Prestwick Circle R1
VA	VAHCCOR2	RE	Henrico County	Westbury Drive R2
VA	VANFTMS5	CO	Norfolk	Village avenue MS5
VA	VANFTMS6	RE	Norfolk	Robin hood road MS6
VA	VANFTMS8	CO	Norfolk	North Hampton MS8
VA	VANFTMS9	CO	Norfolk	Bay side road MS9
VA	VANFTYN1	CO ID	Norfolk	Armistead Avenue N1

Table A1. Site Name and Land Use – Continued

State	LOCATION ID	Land use	Jurisdiction	Site ID
VA	VANFTYN2	RE	Norfolk	Modoc Avenue N2
VA	VANFTYN3	RE	Norfolk	Little creek road N3
VA	VANFTYN4	CO	Norfolk	Military circle N4
VA	VANFTYN5	RE	Norfolk	Sewel's point N5
VA	VANNTMF1	RE	Newport News	Marshall Avenue MF1
VA	VANNTMF4	RE	Newport News	Chesapeake Bay Apartments MF4
VA	VANNTNN1	RE	Newport News	Glendale Road NN1
VA	VANNTNN2	RE OP	Newport News	Shields Road NN2
VA	VANNTNN3	CO	Newport News	Patrick Henry Mall NN3
VA	VANNTNN4	CO ID	Newport News	Oyster Point Park Jefferson Ave NN4
VA	VANNTNN5	CO RE ID	Newport News	Oyster Point Park Thimble Shoals Blvd NN5
VA	VANNTSF4	RE	Newport News	Central Parkway SF4
VA	VANNTSF6	RE	Newport News	Jefferson Avenue SF6
VA	VANNTYI2	ID OP	Newport News	City Line Rd I2
VA	VAPMTYP1	CO	Portsmouth	Cradock Shopping center P1
VA	VAPMTYP2	RE	Portsmouth	West park homes P2
VA	VAPMTYP3	RE CO	Portsmouth	Church land shopping center P3
VA	VAPMTYP4	RE	Portsmouth	Edgefield apartmentsP4
VA	VAPMTYP5	RE	Portsmouth	South Hampton P5
VA	VAVBTYA1	OP ID	Virginia Beach	Morris Neck Road A1
VA	VAVBTYI1	ID	Virginia Beach	Airport Industrial Park I1
VA	VAVBTYM2	RE CO OP ID	Virginia Beach	Ketlam Road M2
VA	VAVBTYR1	RE	Virginia Beach	Homestead Drive R1
VA	VAVBTYV1	RE	Virginia Beach	Bow creek V1
VA	VAVBTYV2	RE	Virginia Beach	Salem Road V2
VA	VAVBTYV3	CO OP	Virginia Beach	Haygood V3
VA	VAVBTYV4	ID	Virginia Beach	Viking Drive V4
VA	VAVBTYV5	RE OP	Virginia Beach	Holland road V5

Table A2. Site Characteristics

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Conveyance	Control	First Sample	Last Sample	Number of Samples
ALHUCHIP	19.5	3			GS		08/27/92	09/12/00	8
ALHUDRAV	20	3			GS		09/19/01	09/19/01	1
ALHUHURI	78.3	3			CG		08/27/92	09/19/01	9
ALHUMASM	87	3			CG		08/27/92	09/19/01	9
ALHUWERP	130	3			CG		08/27/92	09/19/01	9
ALJC004L	2564	3			CG		09/19/01	01/19/02	2
ALJC004R	1047	3			CG		09/20/01	01/19/02	2
ALJCC001	336	3			CG		11/27/01	01/19/02	2
ALJCC002	750	3			CG		11/27/01	03/20/02	2
ALJCC009	112	3			CG		08/31/01	03/09/02	2
ALJCC010	167	3			CG		08/31/01	03/09/02	2
ALJCC012	244	3			CG		12/17/01	12/17/01	1
ALMOCREO	74	3			GS		02/10/93	04/15/93	3
ALMODAPH	14	3					02/16/93	04/20/93	3
ALMOSARA	64	3			GS		01/24/93	04/04/93	3
ALMOSIIV	450	3					02/11/93	04/15/93	3
ALMOSITV	304	3			CG		01/24/93	04/04/93	3
ALMOSIVI	194	3			CG		01/24/93	04/04/93	3
ALMOTHEO	27	3			GS		01/24/93	03/30/93	3
AZMCA001	39	6	80				11/10/91	07/22/98	27
AZMCA002	1120	6	1		GS		01/12/92	02/07/92	2
AZMCA003	45	6	15				12/10/91	07/22/98	27
AZMCA004	81	6	33		CG		10/27/91	08/22/92	6
AZMCA005	3.4	6	94		CG		12/04/92	08/07/98	26
AZMCA006	17.8	6	60		CG		03/07/94	09/11/98	20
AZTUA001	103	6			CG		07/25/96	12/04/01	13
AZTUA002	48.3	6			CG		08/26/96	12/04/01	12
AZTUA003	29	6			CG		08/14/96	12/04/01	11
AZTUA004	83	6			CG		09/24/96	12/11/01	11
CAALAL03	168	6			CG		02/15/90	03/25/93	20
CAALAL04	20	6			CG		03/02/90	02/27/91	5
CAALAL07	78	6			CG		01/13/90	03/17/91	5
CAALAL09	260	6			CG		01/13/90	03/17/91	9
CAALAL10	144	6			CG		03/02/90	03/19/91	8
CACTA001	0.69	6	95		CG		01/23/01	03/10/02	14
CACTA002	1.61	6	100		CG		10/28/00	03/06/02	16
CACTA003	1.85	6	70		CG		01/23/01	03/23/02	10
CACTA004	0.44	6	70		CG		01/23/01	03/23/02	11
CACTA005	0.99	6	100		CG		11/26/97	03/25/99	8
CACTA006	0.99	6	100		CG		01/25/99	03/20/99	3
CACTA007	0.99	6	100		CG		01/25/99	03/25/99	4

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
CACTA008	3.16	6	80		CG		01/17/00	01/27/02	19
CACTA009	4.18	6	80		CG		11/20/99	03/17/02	24
CACTA010	0.96	6	80		CG		11/08/99	03/17/02	26
CACTA011	0.4	6	100		CG		11/10/97	02/09/99	4
CACTA012	0.99	6	100		CG		11/13/97	02/09/99	4
CACTA013	1.41	6	100		CG		11/10/97	02/09/99	3
CACTA014	0.99	6	100		CG		11/13/97	01/25/99	4
CACTA015	0.99	6	100		CG		11/13/97	01/25/99	4
COCSA001	80	9	58.1		CG		06/03/92	11/21/92	7
COCSA002	105.6	9	37.5		CG		05/31/92	07/29/02	7
COCSA003	110.72	9	55.9		CG		06/05/92	11/21/92	7
COCSA004	209.28	9	34.2		CG		05/26/92	11/21/92	7
COCSA005	31.36	9	40.1		CG		06/10/92	12/06/92	7
CODEA001	150	9	83		CG		06/05/92	07/12/92	3
CODEA002	55	9	83		CG		04/14/92	07/12/92	3
CODEA003	269	9	20		CG		03/22/92	08/23/92	4
CODEA004	498	9	85		CG		05/21/92	07/10/92	3
CODEA005	495	9	44		CG		06/06/92	08/23/92	3
CODEA006	636	9	85		CG		06/08/92	07/10/92	3
CODEA007	56	9	85		CG		03/28/92	07/02/92	3
CODEA008	146	9	83		CG		03/28/92	05/31/92	3
GAATAT01	28	3			CG		03/08/95	02/16/97	10
GAATAT02	95	3			CG		10/04/95	02/16/97	9
GACLCOSI	18	3			CG		11/29/95	03/19/00	20
GACLCOTR	125	3			CG		05/01/95	03/16/00	24
GACOC1A2	63.6	3					01/27/96	02/28/00	17
GACOC1A3	7590.4	3					08/24/00	03/19/01	6
GACOCOL2	2947	3					01/19/95	03/12/01	22
GADKCOTD	115	3			CG	WP	12/13/93	06/06/00	25
GAFUCOS1	10339	3			GS		11/10/94	04/25/01	22
GAFUCOS2	3915	3			CG		10/30/94	04/25/01	19
GAFUCOS3	6257	3			GS		01/06/95	04/25/01	22
IDADA001	10.9	8			CG		08/11/99	04/19/01	7
IDADA002	105	8			CG		04/29/99	04/11/01	7
IDADA003	17	8			CG		04/29/99	07/30/01	9
IDADA004	18	8					04/11/01	04/11/01	1
KATOATWO	38	4	55		CG		04/27/98	09/13/02	15
KATOBROO	18.5	4	25		CG		04/27/98	09/13/02	16
KATOJACK	218	4	65		CG		04/27/98	09/13/02	16
KATOSTFE	39.5	4	75		CG		04/27/98	08/16/02	17
KAWIHUNT	36	4	50				02/09/98	10/05/01	16

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
KAWIMCLE	30	4	65				02/09/98	10/05/01	16
KAWISBWY	250	4	60				02/09/98	10/05/01	16
KAWITOWN	40	4	90				02/09/98	10/05/01	16
KYLOTSR1	96.64	2	39.6		CG		01/15/91	10/05/91	3
KYLOTSR2	108.16	2	20.6		GS		02/05/91	10/05/91	3
KYLOTSR3	134.4	2	35		CG		03/01/91	12/12/91	3
KYLOTSR4	43.52	2	45.5		GS		03/12/91	04/15/92	4
KYLOTSR5	84.48	2	68.9		CG		03/27/91	05/12/92	5
KYLOTSR6	180.48	2	63.5		CG		04/04/91	12/12/91	3
KYLXEHL4	13	2	10		GS		10/07/98	10/07/98	1
KYLXEHL5	550	2			GS		07/30/98	07/30/98	1
KYLXEHL6	1.3	2			CG		10/24/97	01/05/98	3
KYLXEHL7	4.8	2			CG		10/24/97	01/05/98	3
KYLXNEL1	32	2			GS		10/07/98	10/07/98	1
KYLXNEL2	580	2			CG	WP	07/30/98	10/07/98	2
KYLXNEL3	73	2			CG		06/03/92	09/27/96	12
KYLXTBL1	71	2			CG		06/03/92	09/27/96	12
KYLXTBL2	94	2			GS		06/30/92	09/27/96	12
KYLXTBL3	205	2			GS		06/19/92	09/21/96	5
KYLXWHL1	38	2			CG		06/03/92	09/27/96	13
MABOA001	40.4	1	74		CG		04/11/92	08/14/92	5
MABOA002	86.7	1	52		CG		04/17/92	06/24/92	3
MABOA003	5	1	55		CG		04/11/92	06/24/92	3
MABOA004	32	1	97		CG		04/11/92	06/24/92	3
MABOA005	102.7	1	38		CG		04/17/92	06/24/92	3
MABOA006	3.3	1	74		CG		06/02/01	07/17/01	3
MABOA007	12.2	1			GS		09/25/01	09/25/01	1
MDAACOMW	5	2	94	E	CG		07/31/92	09/25/92	3
MDAACOOD	28	2	41	E	CG		08/11/92	10/09/92	3
MDAACOPP	25	2	85	E	CG		08/11/92	11/14/00	26
MDAACORK	12	2	41	E	CG		08/28/92	10/30/92	3
MDAACOSC	26	2	41	E	CG		08/11/92	10/09/92	3
MDBACOBC	25.3	2	60				12/15/93	03/08/94	3
MDBACOLC	225	2	70				10/20/93	01/15/98	19
MDBACOSC	83.5	2	30				12/15/93	06/19/98	26
MDBACOTC	144.06	2					01/12/94	04/07/94	3
MDBACOWC	73	2	7				01/12/94	03/21/94	3
MDBCTYBO	48.43	2					04/16/93	03/21/94	3
MDBCTYFM	45.96	2					05/30/92	11/04/93	3
MDBCTYHA	104.4	2	32	E			05/17/95	12/14/00	66
MDBCTYHO	354.09	2					06/05/92	09/25/92	3

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
MDBCTYHR	38.8	2	54				05/30/92	10/20/93	3
MDBCTYKO	54.36	2					04/26/93	03/21/94	3
MDCHCOIP	11	2			CG		10/04/95	01/02/96	3
MDCHCOPA	10	2			CG		09/22/95	01/24/96	3
MDCHCOPF	50	2			GS		01/02/96	04/30/96	3
MDCHCOTG	10	2			CG		10/14/95	10/14/95	1
MDCLCOBP	15	2					07/21/94	12/10/94	3
MDCLCOCE	22.35	2	26	E			12/10/93	11/21/94	3
MDCLCOJS	20	2	91	E			12/21/93	11/01/94	3
MDCLCOKW	66.2	2	11	E			03/21/94	09/22/94	3
MDCLCOSD	36	2	49	E			02/21/94	11/10/94	3
MDHACFBA		2					11/05/93	11/10/94	3
MDHACOBP	69.7	2	16	E		WP	02/17/99	12/16/00	18
MDHACOCF	14.4	2					01/05/93	04/10/93	2
MDHACOCS	51	2					04/21/93	09/22/94	3
MDHACOGR	80	2					04/26/93	04/10/94	2
MDHACOIC	10	2					08/06/93	09/22/94	2
MDHOCODC	7.5	2	90		CG	WP	12/15/93	04/13/94	3
MDHOCOFM	3.5	2	77		CG	GS	12/15/93	11/01/94	3
MDHOCOGM	29.5	2	38		CG	WP	12/10/93	11/01/94	3
MDHOCOMH	19	2	65		CG	WP	12/10/93	04/13/94	3
MDHOCOOC	11.7	2	49		CG	WP	11/17/93	03/21/94	3
MDMOCOBC	14.2	2	83	E	CG		05/25/94	09/22/95	3
MDMOCOCV	11.5	2	55	E		WP	08/13/96	09/25/00	37
MDMOCONV	75.4	2	57	E	CG		05/04/94	03/08/95	3
MDMOCOQA	34.2	2	45	E	CG		05/04/94	10/27/95	3
MDMOCOSL	81	2	92	E	CG	OT	09/22/94	09/22/95	3
MDMOCOWP	70	2	96	E	CG	OT	05/25/94	10/27/95	3
MDPGCOS1	19.7	2	47	E	CG		08/11/92	01/22/97	26
MDPGCOS2	57.3	2	45	E	CG		08/11/92	09/25/00	63
MDPGCOS3	34.4	2	96	E	CG		08/11/92	03/04/93	3
MDPGCOS4	102.5	2	33	E	CG		08/11/92	03/04/93	3
MDPGCOS5	41.3	2	83	E	CG		08/11/92	03/04/93	3
MDPGCOS6	42.4	2			GS		10/23/94	08/20/97	28
MDSHDTDV	4	2			CG		06/14/99	06/06/00	8
MDSHDTPS	20	2			GS		02/11/98	06/21/00	13
MNMISD01	143	1			CG		05/06/01	10/13/01	10
MNMISD02	95	1			CG		05/06/01	10/13/01	9
MNMISD03	80	1			CG		05/20/01	11/12/01	10
MNMISD04	63	1			CG		05/06/01	11/12/01	9
MNMISD05	100	1			CG		05/20/01	11/12/01	10

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
NCCHBREV	15.1	2	75		CG		05/13/92	03/03/93	4
NCCHHIDD	20	2	30		CG		05/13/92	12/10/93	5
NCCHHOSK	17.4	2	71.83				05/13/92	03/03/93	4
NCCHNANC	10.9	2	20		GS		08/27/92	02/21/94	4
NCCHROSE	78.87	2	42.55		CG		05/13/92	08/12/92	3
NCCHSHEF	42.5	2	20.68				06/04/92	02/21/94	3
NCCHSIMS	6.8	2	50		CG		05/13/92	03/03/93	4
NCCHSTAR	14.1	2	70		GS		05/13/92	03/03/93	4
NCCHYARD	88.6	2	68.21		CG		05/13/92	03/03/93	4
NCFV71ST	36	2	45		CG		01/21/93	06/15/99	18
NCFVCLEA	12	2	20				01/04/93	04/01/99	14
NCFVELMS	40	2	90		CG		01/04/93	04/01/99	18
NCFVROSE	39.27	2	50		CG		12/17/92	06/15/99	14
NCFVSTRK	85	2	1				02/07/93	06/16/96	6
NCFVTRYO	25	2	50				01/21/93	04/01/99	18
NCFVWINS	12	2	75				01/04/93	06/15/99	18
NCGRATHE	23	2	90		CG		07/06/95	04/15/99	17
NCGRCOUN	18.5	2	2		GS		06/19/95	04/01/99	15
NCGRHUST	13	2	75		CG		06/01/95	05/14/99	16
NCGRMERR	21	2	74		CG		05/10/95	04/15/99	16
NCGRRAND	26	2	50		CG		06/01/95	06/15/99	17
NCGRUNIO	33	2	75		CG		06/01/95	01/23/99	17
NCGRWILL	13	2	20				05/19/95	05/14/99	16
NCRASIT1	21	2					05/19/93	03/16/00	9
NCRASIT2	42	2					05/19/93	03/16/00	9
NCRASIT3	110	2					05/19/93	03/16/00	9
NCRASIT4	30	2					05/19/93	03/16/00	9
NCRASIT5	32	2					05/19/93	03/16/00	9
NCRASIT6	58	2					05/19/93	03/16/00	9
NCRASIT7	467	2					05/19/93	03/16/00	9
ORCCA001	15	7			CG		11/18/92	11/18/92	1
ORCCA002	120	7			CG		11/18/92	10/14/94	6
ORCCA003	165	7			CG		03/01/93	10/14/94	5
ORCCA004	50	7			CG		11/18/92	10/14/94	6
ORCCA005	41	7			CG		11/18/92	10/14/94	6
OREUA001	380	7			CG		09/23/92	05/21/96	16
OREUA002	886	7			CG		09/23/92	05/21/96	15
OREUA003	377	7			CG		09/23/92	05/21/96	15
ORGRA001	292	7			CG		03/02/93	04/11/96	6
ORGRA002	789	7			CG		03/02/93	04/11/96	6
ORGRA003	73	7			CG		03/02/93	04/11/96	6

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
ORGRA004	64	7			CG		03/02/93	04/11/96	6
ORODA001	22.4	7			CG		12/04/95	12/04/95	1
ORODA002	1.2	7			CG		10/02/95	10/02/95	1
ORODA003	18.2	7			CG		01/07/95	01/18/96	5
ORODA004	3.6	7			CG		10/12/95	10/12/95	1
ORODA005	23.1	7			CG		01/07/95	01/18/96	6
ORPOA001	35	7			CG		05/07/91	10/25/95	13
ORPOA002	75	7			CG		05/07/91	03/03/96	16
ORPOA003	46	7			CG		05/07/91	03/03/96	14
ORPOA004	49	7			CG		08/09/91	10/25/95	13
ORPOA005	91	7			CG		05/07/91	03/03/96	14
ORPOA006	85	7			CG		05/07/91	03/03/96	13
ORPOA007	10	7			CG		05/07/91	03/03/96	14
ORSAA001	31	7			CG		01/07/95	01/14/96	6
ORSAA002	40	7			CG		01/07/95	01/14/96	6
ORSAA003	35	7			CG		01/07/95	01/14/96	6
ORSAA004	72	7			CG		01/07/95	01/14/96	6
PAPH0864	22	2	84	E	CG		09/10/92	09/25/92	2
PAPH0891	35	2	83	E	CG		09/22/92	10/09/92	2
PAPH1014	22	2	82	E	CG		09/10/92	10/09/92	3
PAPH1051	223	2	87	E	CG		09/22/92	10/09/92	2
PAPH1182	31	2	57	E	CG		09/10/92	10/09/92	3
TNKXTYAP	582.4	2	44		GS	WP	03/27/91	06/30/01	63
TNKXTYFC	2880	2	40			WP	03/06/92	06/07/01	47
TNKXTYGV	224	2	37		GS		08/14/91	08/25/99	39
TNKXTYTC	352	2	34				02/13/92	04/11/00	54
TNKXTYWE	364.8	2	60				04/08/91	05/03/00	51
TNMET207	157	2			GS	WP	06/21/00	04/23/01	5
TNMET211	45	2			CG		01/11/00	04/23/01	4
TNMET231	26	2			CG		01/11/00	04/23/01	4
TNMET260	294	2			CG		07/20/00	05/17/01	4
TNMET410	154	2			CG		06/21/00	04/23/01	4
TXARA001	38.8	5	76.2		CG		10/28/92	03/08/01	22
TXARA002	160.6	5	47.4		CG		10/28/92	03/08/01	21
TXARA003	77	5	89		CG		10/29/92	04/14/93	7
TXARA004	85.5	5	80.9			WP	12/09/92	03/28/93	7
TXDAA001	9	5	80		CG		03/03/92	09/21/92	7
TXDAA002	49.5	5	80		CG		03/03/92	03/08/01	19
TXDAA003	486.7	5			CG		12/02/97	05/04/01	21
TXDAA004	59.1	5	84.5		CG		02/22/92	04/11/01	20
TXDAA005	71.3	5	50		CG		02/12/92	09/21/92	7

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
TXDAA006	38.9	5	44.9		CG		02/22/92	03/24/01	20
TXDCA001	115.36	5	10				09/03/97	03/08/01	17
TXDCA002	12.05	5	33				01/11/98	03/08/01	9
TXFWA001	61.7	5	21.9				02/22/92	08/12/92	7
TXFWA002	151.6	5	27.7	* I	CG		02/03/92	03/08/01	21
TXFWA003	136	5	66.5		CG		03/09/92	10/28/92	7
TXFWA004	73.7	5	79.3				03/24/92	03/08/01	21
TXFWA005	150.8	5	61.4				04/17/92	03/24/01	23
TXGAA001	268	5			CG		12/02/97	03/27/01	23
TXGAA002	33.9	5	67.3				06/20/92	03/27/01	22
TXGAA003	67.3	5	55.4				09/01/92	01/23/93	7
TXGAA004	36.2	5	84.6		CG		09/01/92	01/23/93	7
TXHCA001	560	4			GS		01/29/99	03/27/01	8
TXHCA002	95	4	65				07/20/92	04/16/01	14
TXHCA003	32	4			CG		02/11/99	06/22/01	8
TXHCA004	99	4	71.25				04/07/93	03/08/01	14
TXHCA005	81	4	95		CG		06/30/92	04/07/93	6
TXHCA006	401	4	45				06/30/92	04/29/93	6
TXHCA007	872	4				WP	07/15/92	11/09/93	6
TXHOA001	44	4	5.7		GS		06/30/92	11/19/92	7
TXHOA002	232	4	76.5				06/22/92	05/31/01	16
TXHOA003	65	4	45		GS		06/22/92	03/27/01	14
TXHOA004	24	4	98		CG		07/19/92	11/22/99	12
TXHOA005	38	4	65		CG		06/22/92	07/19/01	16
TXIRA001	65.3	5	41.9		CG		09/03/92	03/24/01	22
TXIRA002	127.7	5			CG		03/18/99	05/28/01	22
TXIRA003	43.4	5	77.3			WP	08/24/92	01/09/93	7
TXIRA004	43.9	5	77.8				09/21/92	01/28/93	7
TXMEA001	45.9	5	89.4		CG		02/24/93	03/24/01	22
TXMEA002	45.4	5	49.8		CG		03/11/93	06/25/93	7
TXMEA003	46.2	5	49.9		CG	WP	02/10/93	05/23/93	7
TXPLA001	51.4	5	54.3		CG		12/09/92	04/14/93	7
TXPLA002	73.5	5					11/09/98	04/11/01	22
TXPLA003	22.7	5	73.5		CG		12/09/92	05/04/01	25
TXPLA004	49	5	81.6				01/09/93	06/09/93	7
TXTCA001	63.13	5	27				02/06/97	03/24/01	15
VAARLCV2	24.7	2	35			DS	02/11/98	01/19/01	9
VAARLLP1	38.7	2	35				10/20/99	03/04/01	8
VAARLRS3	14	2	74			DS	09/21/99	01/19/01	8
VAARLTC4	36	2	39				02/03/98	06/01/01	13
VACHCCC4	60	2	80		CG		08/12/96	12/10/01	13

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Conveyance	Control	First Sample	Last Sample	Number of Samples
VACHCCC5	10	2	50		CG		08/12/96	12/10/01	13
VACHCN1A	10	2	10		CG		08/19/99	01/08/01	4
VACHCN2A	60	2	20		CG		08/19/99	01/08/01	4
VACHCOF1	22.5	2					04/16/93	10/26/93	3
VACHCOF2	19.05	2	10		CG		04/16/93	10/08/98	8
VACHCOF3	13.5	2	20		CG		04/16/93	02/01/99	11
VACHCOF4	38.5	2					04/16/93	12/15/93	3
VACHCOF5	55.6	2	50		CG		04/16/93	12/10/01	16
VACPTC1A	130	2	25	E	CG GS		11/01/97	06/15/99	8
VACPTSF2	91	2	10	E	GS		04/16/93	10/26/93	3
VACPTYC1	57	2	25	E	CG		02/26/93	12/05/96	7
VACPTYC2	188	2	25	E	CG	WP	02/26/93	01/24/99	15
VACPTYC3	32	2	50	E	CG		02/26/93	02/18/99	15
VACPTYC4	28	2	85	E	CG		04/16/93	02/02/99	14
VACPTYC5	16	2	57	E	CG		03/27/93	01/15/99	15
VACPTYO1	14	2	50	E	CG		02/26/93	04/16/93	3
VAFFCOF1	32.3	2			CG		03/18/92	08/11/92	3
VAFFCOF2	20.1	2	50	E	CG	DP	07/03/92	08/01/00	14
VAFFCOF3	63.9	2			CG		06/18/92	09/02/92	3
VAFFCOF4	108.8	2	70	E	CG	WP	04/21/92	09/03/00	13
VAFFCOF5	39.7	2			CG	DP	04/16/92	09/22/92	3
VAFFCOF6	213.4	2	21	E	CG	WP	07/12/92	11/10/00	14
VAFFCOF7	49.9	2	25	E	CG		06/24/92	11/29/00	15
VAFFCOF8	57.5	2			CG		04/21/92	09/02/92	3
VAFFCOF9	63.8	2	50	E	CG	WP	07/21/92	09/02/00	13
VAFFOF10	82	2			CG		04/21/92	08/11/92	3
VAFFOF11	37.9	2	66	E	CG		06/26/97	11/29/00	11
VAHAHMS2	793	2	67	E	CG		11/26/92	01/21/93	3
VAHAHMS5	53	2	28	E			11/12/92	02/12/93	3
VAHATYH1	115	2	80	E	CG		11/12/92	05/14/99	18
VAHATYH2	47	2	70	E	CG		11/26/92	04/24/99	19
VAHATYH3	18	2	40	E	CG		11/12/92	06/20/99	17
VAHATYH4	134	2	25	E	CG		11/12/92	04/24/99	17
VAHATYH5	35	2	25	E	CG		11/12/92	04/24/99	17
VAHCCOC1	65	2	89	E	CG		11/13/92	12/20/92	2
VAHCCOC2	70	2	87	E	CG		10/30/92	01/05/93	3
VAHCCON1	75	2	89	E	CG		12/18/92	01/22/93	2
VAHCCON2	23	2	89	E	CG		11/22/92	01/22/93	3
VAHCCOR1	40	2	61	E	CG		11/03/92	01/05/93	3
VAHCCOR2	70	2	57	E	CG		11/03/92	01/05/93	3
VANFTMS5	56	2			CG		04/22/92	07/27/92	3

Table A2. Site Characteristics – Continued

LOCATION ID	Area (acres)	EPA Rain Zone	% Impervious	Q	Convey ance	Control	First Sample	Last Sample	Number of Samples
VANFTMS6	68	2					05/05/92	07/27/92	3
VANFTMS8	65	2					04/22/92	07/27/92	3
VANFTMS9	40	2					05/05/92	08/27/92	3
VANFTYN1	43	2	47		CG		04/22/92	02/12/00	28
VANFTYN2	97	2	25		CG		05/30/92	12/14/99	30
VANFTYN3	27	2	37		CG		04/22/92	12/14/99	28
VANFTYN4	43	2	70		CG		04/22/92	12/14/99	28
VANFTYN5	39	2	25		CG		06/09/92	02/18/00	28
VANNTMF1	39	2	50		CG		10/04/92	01/21/93	3
VANNTMF4	12	2	73		CG		12/28/92	03/13/93	3
VANNTNN1	75	2	40		OT		10/31/92	04/02/99	12
VANNTNN2	397	2	24		CG	DP	12/10/92	04/16/99	15
VANNTNN3	24	2	85		CG	WP	10/04/92	04/16/99	15
VANNTNN4	294	2	58		CG	WP	10/04/92	04/16/99	16
VANNTNN5	83	2	62		OT		12/28/92	04/02/99	11
VANNTSF4	111	2	30		GS		12/10/92	02/26/93	3
VANNTSF6	207	2	37		GS		10/04/92	03/03/93	4
VANNTYI2	49	2	73		GS		10/04/92	01/21/93	3
VAPMTYP1	27.2	2	68		CG		01/16/93	05/14/99	18
VAPMTYP2	101.1	2	36		CG		02/26/93	06/20/99	17
VAPMTYP3	46	2			CG		01/16/93	05/14/99	17
VAPMTYP4	35.3	2	39		CG		12/20/92	06/20/99	17
VAPMTYP5	53.5	2	14	E	CG		12/20/92	05/23/99	17
VAVBTYA1	225	2	7				07/01/92	10/30/92	5
VAVBTYI1	8	2	90				06/09/92	10/04/92	3
VAVBTYM2	310	2	35				10/04/90	10/30/92	4
VAVBTYR1	49	2	25				05/07/92	09/19/92	5
VAVBTYV1	63	2	29		OT		03/26/92	02/28/99	27
VAVBTYV2	260	2	29		OT	WP	05/07/92	02/18/99	30
VAVBTYV3	25	2	25		CG		04/12/92	02/28/99	33
VAVBTYV4	29	2	55		CG		04/12/92	03/14/99	30
VAVBTYV5	882	2	47		OT	WP	05/07/92	03/14/99	28

Appendix B: Modified Values in the Database

Description

The following table indicates the values that were modified in the database. The column "Order" corresponds to the row number in the table. The column "Problem" indicates the reason why the value was deleted or modified. In the case that the information available can solve the problem, the action was described in the column "action". The last column indicates the community where the event was located.

Table B1. Modified Values in the NSQD

Order	Constituent	Original value	Problem	Action	Location_ID
1373	TSS	10100	High	Delete	Acker Place
890	TSS	53000	High	Delete	Philadelphia
1909	TP	80.1	High - Ortho very Low	Delete	Louisville
1707	TP	35	High	Delete	Lexington
1629	TP	15.4	High	Delete	
1907	Ortho P	60.1	High	Delete	Louisville
3135	Dis Zn / Tot Zn		High Ratio	Delete Total	Boston
3118	TDS	17900	High TDS	Deleted	Boston
2893	Dis Cu / Tot Cu		High Ratio	Delete Dissolved Copper	Portland
2883	Dis Cu / Tot Cu		High Ratio	Delete Dissolved Copper	Ada County
561	Dis P / Tot P		Wrong Dissolved Values	corrected	Portsmouth
562	Dis P / Tot P		Wrong Dissolved Values	corrected	Portsmouth
563	Dis P / Tot P		Wrong Dissolved Values	corrected	Portsmouth

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
221	Dis P / Tot P		Wrong Dissolved Values	Corrected	Hampton
223	Dis P / Tot P		Wrong Dissolved Values	Corrected	Hampton
1707	Dis P / Tot P		High Values	Deleted	Lexington
1999	Dis P / Tot P		High Dissolved Value	Deleted	Cobb
2301	Dis P / Tot P		Low Total	Delete both	Fayetteville
2268	Dis P / Tot P		Low Total	Delete both	Fayetteville
2293	Dis P / Tot P		Low Total	Delete both	Fayetteville
4315	Dis P / Tot P		Low Total	Delete both	Raleigh
4306	Dis P / Tot P		Low Total	Delete both	Raleigh
4351	Dis P / Tot P		Low Total	Delete both	Raleigh
4342	Dis P / Tot P		Low Total	Delete both	Raleigh
4055	Dis P / Tot P		Low Total	Delete both	Greensboro
4070	Dis P / Tot P		Inverted	Corrected	Greensboro
4197	Dis P / Tot P		Low Total	Delete both	Greensboro
4249	Dis P / Tot P		Low Total	Delete both	Greensboro
4085	Dis P / Tot P		Low Total	Delete both	Greensboro
4217	Dis P / Tot P		Low Total	Delete both	Greensboro
4068	Dis P / Tot P		Wrong values	Corrected	Greensboro
4038	Dis P / Tot P		Low Total	Delete both	Greensboro
4134	Dis P / Tot P		Low Total	Delete both	Greensboro
4233	Dis P / Tot P		Low Total	Delete both	Greensboro
4149	Dis P / Tot P		Low Total	Delete both	Greensboro
3698	Dis P / Tot P		Low Total	Delete both	Greensboro
4024	Dis P / Tot P		Low Total	Delete both	Greensboro
2150	Dis P / Tot P		Low Total	Delete both	Fulton
1449	Dis P / Tot P		Low Total	Delete both	Knoxville
1617	Dis P / Tot P		Low Total	Delete both	Knoxville
1596	Dis P / Tot P		Low Total	Delete both	Knoxville
1616	Dis P / Tot P		Low Total	Delete both	Knoxville
1460	Dis P / Tot P		Low Total	Delete both	Knoxville
1707	TKN	290	High	Deleted	Lexington
1000	TKN	250	High	Deleted	Baltimore City
4149	TKN	147	High	Deleted	Greensboro
2699	TKN	120	High	Deleted	Maricopa
3136	NO2 NO3	1690	High	Deleted	Atlanta
3281	NO2 NO3	50	High	Deleted	Houston 06/30/92
3331	NO2 NO3	48	High	Deleted	Houston 06/30/92
3289	NO2 NO3	32.1	High	Deleted	Houston 06/30/92
3305	NO2 NO3	28	High	Deleted	Houston 06/30/92
48	COD BOD	< 5	COD low	Deleted	Chesterfield 02/03/98
737	COD	5050		Deleted	Bow creek V1

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
97	BOD		Weird Value <30	Deleted	Chesterfield 09/27/99
1895	BOD	610		Deleted	Pleasure_Ridge_Park
2676	COD BOD	4300		Deleted	Maricopa
3299	COD BOD	1260		Retyped	Houston
4087	BOD	545		Deleted	Husbands_Street
4343	COD BOD			Deleted	Williamson_Drive_Wade_Avenue
4399	COD BOD			Deleted	50ft_east_N_West_Street_Peace_Street_Dortch_Street
96	Cadmium		Weird Values Dissolved Higher	Deleted	Chesterfield 08/19/99
97	Cooper		Weird Values Dissolved Higher	Deleted	Chesterfield 09/27/99
85	Cooper		Weird Values Dissolved Higher	Deleted	Chesterfield 09/27/99
110	Lead		Weird Values Dissolved Higher	Deleted	Chesterfield 09/27/99
76	Zinc		Weird Values Dissolved Higher	Deleted	Chesterfield 09/27/99
110	Zinc		Weird Values Dissolved Higher	Deleted	Chesterfield 09/27/99
3288	Antimony		Elevated values for the same set of samples	Deleted	Houston
3304	Antimony		Elevated values for the same set of samples	Deleted	Houston
3330	Antimony		Elevated values for the same set of samples	Deleted	Houston
3281	Antimony		Elevated values for the same set of samples	Deleted	Houston
3289	Antimony		Elevated values for the same set of samples	Deleted	Houston
3305	Antimony		Elevated values for the same set of samples	Deleted	Houston
3331	Antimony		Elevated values for the same set of samples	Deleted	Houston
3276	Oil & Grease		Detection Limit is different	Replace as a detected value	Harris County
2836	Conductivity		Elevated Value. Two samples	Use mean value	Colorado Springs
4077	Turbidity		NT in cell	Move to qualifier	Greensboro
446	TDS <46		Wrong Qualifier	Delete Qualifier..	Norfolk N2
2128	TDS	0.065	Factor of a thousand.	Change value from 0.065 to 65	Fulton County
2257	TDS		Wrong value	corrected (32 mg/L)	Fayetteville
3136	TDS		Value not clear in hardcopy	Delete value <31 mg/L	Atlanta

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
2649	TDS		TDS>TS	Delete pair	Maricopa
2658	TDS		TDS>TS	Delete pair	Maricopa
2660	TDS		TDS>TS	Delete pair	Maricopa
2664	TDS		TDS>TS	Delete pair	Maricopa
2665	TDS		TDS>TS	Delete pair	Maricopa
2672	TDS		TDS>TS	Delete pair	Maricopa
2686	TDS		TDS>TS	Delete pair	Maricopa
2699	TDS		TDS>TS	Delete pair	Maricopa
2704	TDS		TDS>TS	Delete pair	Maricopa
2706	TDS		TDS>TS	Delete pair	Maricopa
2709	TDS		TDS>TS	Delete pair	Maricopa
2713	TDS		TDS>TS	Delete pair	Maricopa
2714	TDS		TDS>TS	Delete pair	Maricopa
2715	TDS		TDS>TS	Delete pair	Maricopa
2716	TDS		TDS>TS	Delete pair	Maricopa
2717	TDS		TDS>TS	Delete pair	Maricopa
2719	TDS		TDS>TS	Delete pair	Maricopa
2720	TDS		TDS>TS	Delete pair	Maricopa
2721	TDS		TDS>TS	Delete pair	Maricopa
2725	TDS		TDS>TS	Delete pair	Maricopa
2726	TDS		TDS>TS	Delete pair	Maricopa
2729	TDS		TDS>TS	Delete pair	Maricopa
2730	TDS		TDS>TS	Delete pair	Maricopa
2734	TDS		TDS>TS	Delete pair	Maricopa
2740	TDS		TDS>TS	Delete pair	Maricopa
2742	TDS		TDS>TS	Delete pair	Maricopa
2750	TDS		TDS>TS	Delete pair	Maricopa
2751	TDS		TDS>TS	Delete pair	Maricopa
3321	TSS		Wrong Detection Limit	Change for value	Houston
3154	TSS		Wrong Detection Limit	Change for <1	Atlanta
1117	TSS		Wrong Detection Limit	Deleted	Montgomery County
2543	TSS		First Flush was compared with TDS	Change for 160	Los Angeles
32	COD		Weird Detection Limit	Delete Qualifier..	Arlington
5	COD		Weird Detection Limit	Delete Qualifier. Quantification limit = 5mg/L	Arlington
16	COD		Weird Detection Limit	Delete Qualifier. Quantification limit = 5mg/L	Arlington
859	Fecal Streptococcus		Atypical Growth	Delete value >6000	Fairfax
1401	Ammonia		Typo in detection limit	Change <2 to <0.2 mg/L	Knoxville

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
2299	Ammonia		Detection Limit is different	Change for <0.05 (other sites)	Fayetteville
2886	Ammonia		Different Detection Limit	Change <0.14 to <0.20	Portland
2899	Ammonia		Different Detection Limit	Change <0.14 to <0.20	Portland
2915	Ammonia		Different Detection Limit	Change <0.14 to <0.20	Portland
2942	Ammonia		Different Detection Limit	Change <0.14 to <0.20	Portland
2956	Ammonia		Different Detection Limit	Change <0.14 to <0.20	Portland
1999	TKN		Detection Limit is Weird	Delete <0.6	Cobb County
2097	TKN		Typo in Detection Limit	Delete <0.0 for <0.2	Fulton County
2257	TKN		Rows seem to be wrong	Change 46 for 1.2	Fayetteville
2257	TDS		Rows seem to be wrong	Change 1.2 for 46	Fayetteville
2300	Nitrogen Total Organic		Value = 0, Grab = 1.46	Delete value	Fayetteville
2336	Nitrogen Total Organic		Value = 0, Grab = 1.14	Delete value	Fayetteville
13	Phosphorus Dissolved		Low Value	Change 0.009 by 0.09	Arlington
1488	Phosphorus Dissolved		Values lower than DL	Change by <0.02	Knoxville
1527	Phosphorus Dissolved		Values lower than DL	Change by 0.02	Knoxville
1580	Phosphorus Dissolved		Values lower than DL	Change by 0.02	Knoxville
4079	Beryllium		Detection limit	Change <0.6 by <0.06	Greensboro
4245	Cadmium		Detection limit	Change <0.4 by <0.04	Greensboro
2150	Cadmium		Wrong Columns	Copy 16000 in TotCol 230 Fec	Fulton
2128	Cadmium		Wrong Columns	Cd, Tot col and Fec Col in correct columns	Fulton
1107	Cadmium		LD in cell	Replace by detection limit	Howard County
1110	Cadmium		LD in cell	Replace by detection limit	Howard County
2864	Cadmium		Detection limit	Replace <2.5 by <0.5	Ada County
1107	Chromium		LD in cell	Replace by detection limit	Howard County

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
1110	Chromium		LD in cell	Replace by detection limit	Howard County
2935-3419	Cyanide		Factor of a thousand.	Multiply by 1000	Texas
18	Cyanide		Wrong Detection Limit	Changed	Arlington
29	Cyanide		Wrong Detection Limit	Changed	Arlington
35	Cyanide		Wrong Detection Limit	Changed	Arlington
2460	Conductivity	1	The value is the detection limit	Deleted	CAALA001
2871	Conductivity	2.5	Conductivity was collected in grab sample	Deleted	IDADA002
2662	DO	11.6	Evaluated by temperature	Deleted	AZMCA001
3016	DO	10.2	Evaluated by temperature	Deleted	ORODA005
3076	DO	15	Evaluated by temperature	Deleted	OREUA001
3077	DO	12.2	Evaluated by temperature	Deleted	OREUA001
3078	DO	17	Evaluated by temperature	Deleted	OREUA001
3092	DO	12.1	Evaluated by temperature	Deleted	OREUA002
3093	DO	18.4	Evaluated by temperature	Deleted	OREUA002
3097	DO	14	Evaluated by temperature	Deleted	OREUA002
3107	DO	11.5	Evaluated by temperature	Deleted	OREUA003
3108	DO	19.2	Evaluated by temperature	Deleted	OREUA003
3115	DO	16.3	Evaluated by temperature	Deleted	MABOA001
3120	DO	21.8	Evaluated by temperature	Deleted	MABOA002
3122	DO	10.2	Evaluated by temperature	Deleted	MABOA002
3126	DO	15.4	Evaluated by temperature	Deleted	MABOA004
3129	DO	14.6	Evaluated by temperature	Deleted	MABOA005
3020	HARDNESS	<1	Weight of evidence compared with conductivity and TDS	Deleted	ORODA005
4065	TSS	66	Turbidity high but TSS low, checked with other parameters	TSS to 660	NCGRHUST

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
2619	Turbidity	5.2	Grab samples higher than composite	Calculate time composite	CACTA010
2620	Turbidity	4.3	Grab samples higher than composite	Calculate time composite	CACTA010
1350 - 1351	Various		Values don't seem correct. The community changed lab after this samples	Delete event	GACLCOSI
65	TDS	1	Low value	Use <5	VACHCOF5
875	TDS	2	Low value	Use <5	VAFFOF10
1611	TDS	2	Low value	Use <5	TNKXTYWE
1645	TDS	5406	Elevated value without support	Deleted	TNMET410
2102	TDS	4000	Elevated value without support	Deleted	GAFUCOS1
2122	TDS	4100	Elevated value without support	Deleted	GAFUCOS2
2144	TDS	4200	Elevated value without support	Deleted	GAFUCOS3
2128 , 2150	Various		Values don't seem correct. The community changed lab after this samples	Delete event	GAFUCOS2,GAFUCOS3
2155	TDS	<	Missing detection limit	Deleted	GAFUCOS3
2699	TDS	1290	Elevated value without support	Deleted	AZMCA003
2965	TDS	3	Low Value	Use <5	ORPOA006
2942	TDS	4	Low Value	Use <5	ORPOA005
3691	TDS	1	Low Value	Use <5	TXIRA002
3772	TDS	1	Low Value	Use <5	TXIRA002
3119	TDS	17900	Elevated value, but other samples support it.	Keep with ?	MABOA001
16	TSS	2	Low Value	Use <5	VAARLCV2
19	TSS	1	Low Value	Use <5	VAARLRS3
48	TSS	2	Low Value	Use <5	VACHCOF2
65	TSS	1	Low Value	Use <5	VACHCOF5
76	TSS	2	Low Value	Use <5	VACHCOF5
266	TSS	2	Low Value	Use <5	VAHATYH3
498	TSS	2	Low Value	Use <5	VANFTYN4
812	TSS	1	Low Value	Use <5	VAFFCOF3
842	TSS	2.5	Low Value	Use <5	VAFFCOF6
935	TSS	2	Low Value	Use <5	MDAACORK
965	TSS	1.8	Low Value	Use <5	MDBACOSC
1160	TSS	2.87	Low Value	Use <5	MDMOCOCV
1441	TSS	1	Low Value	Use <5	TNKXTYFC
1482	TSS	1	Low Value	Use <5	TNKXTYGV

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
2014	TSS	2	Low Value	Use <5	GACOC1A2
2028	TSS	2.2	Low Value	Use <5	GACOCOL2
2143	TSS	1	Low Value	Use <5	GAFUCOS3
2235	TSS	2	Low Value	Use <5	NCFVTRYO
3143	TSS	2.99	Low Value	Use <5	GAATAT02
3149	TSS	1.83	Low Value	Use <5	GAATAT01
3150	TSS	1.98	Low Value	Use <5	GAATAT01
3265	TSS	2	Low Value	Use <5	TXHCA005
3451	TSS	0.5	Low Value	Use <5	TXDAA004
3453	TSS	0.5	Low Value	Use <5	TXDAA004
3552	TSS	0.5	Low Value	Use <5	TXFWA003
3647	TSS	0.5	Low Value	Use <5	TXGAA003
3775	TSS	2	Low Value	Use <5	TXPLA002
3776	TSS	2	Low Value	Use <5	TXPLA002
3781	TSS	1	Low Value	Use <5	TXPLA002
1314	BOD	0.73	Low Value	Use <1	MDSHDTPS
1317	BOD	0.41	Low Value	Use <1	MDSHDTPS
1322	BOD	0.91	Low Value	Use <1	MDSHDTPS
3868	BOD	0.7	Low Value	Use <1	NCCHSHEF
32	COD	<150	Unusual Detection Limit	Deleted	VAARLTC4
2250	COD	1500	Unusual elevated value, no evidence	Deleted	NCFVTRYO
3479	COD	1300	Unusual elevated value, no evidence	Deleted	TXDAA006
1897	Ammonia	60.3	Unusual elevated value, no evidence	Deleted	KYLOTSR5
1907	Ammonia	60.5	Unusual elevated value, no evidence	Deleted	KYLOTSR6
1909	Ammonia	30.4	Unusual elevated value, no evidence	Deleted	KYLOTSR6
2699	Ammonia	64	Unusual elevated value, no evidence	Deleted	AZMCA003
8	NO2 NO3	13	Unusual elevated value, no evidence	Deleted	VAARLLP1
1314	NO2 NO3	7.05	Unusual elevated value, no evidence	Deleted	MDSHDTPS
2011	NO2 NO3	6.3	Unusual elevated value, no evidence	Deleted	GACOC1A2
2030	NO2 NO3	>0.2	Unusual Detection Limit	Deleted	GACOCOL2
2140	NO2 NO3	9.3	Unusual elevated value, no evidence	Deleted	GAFUCOS3
2966	NO2 NO3	6.5	Unusual elevated value, no evidence	Deleted	ORPOA006
1905	TN	0.39	TN < NH3	Deleted both	KYLOTSR6
1907	TN	0.9	TN < NH3	Deleted both	KYLOTSR6
1909	TN	3	TN < NH3	Deleted both	KYLOTSR6

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
1600	TN	1.5	TN < NH3	Deleted both	TNKXTYWE
4324	TN	3.42	TN < TKN	Deleted both	NCRASIT3
4387	TN	6.65	TN < TKN	Deleted both	NCRASIT6
3281	TN	50.2	Unusual elevated value, no evidence	Deleted	TXHOA001
3289	TN	33	Unusual elevated value, no evidence	Deleted	TXHOA002
3305	TN	28.9	Unusual elevated value, no evidence	Deleted	TXHOA003
3331	TN	49.7	Unusual elevated value, no evidence	Deleted	TXHOA005
1907	Nitrogen Nitrite	40	Unusual elevated value, no evidence	Deleted	KYLOTSR6
1907	Phosphate Ortho	60.1	Unusual elevated value, no evidence	Deleted	KYLOTSR6
2978	Phosphate Ortho	0.8	Unusual elevated value, no evidence	Deleted	ORPOA007
3419	Phosphorus Total	0.005	Low value	Deleted	TXDAA002
4073	Antimony	0.02	Grab sample	Deleted	NCGRHUST
2006	Cadmium Total	40.16	Unusual elevated value, no evidence	Deleted	GACOC1A2
2007	Cadmium Total	42.7	Unusual elevated value, no evidence	Deleted	GACOC1A2
2036	Cadmium Total	122	Unusual elevated value, no evidence	Deleted	GACOCOL2
2128	Cadmium Total	16000	Unusual elevated value, no evidence, seems to be wrong columns	Deleted	GAFUCOS2
1131	Chromium Total	120	Unusual elevated value, no evidence	Deleted	MDMOCOWP
797	Copper Total	396	Unusual elevated value, no evidence	Deleted	VAFFCOF2
889 - 900	Mercury Total	0		Deleted	PAPH
1790	Nickel Total	200	Unusual elevated value, Detection limit 20	Change by 20	KYLXWHL1
3299	Nickel Total	325	Unusual elevated value, no evidence	Deleted	TXHOA002
3321	Nickel Total	720	Unusual elevated value, no evidence	Deleted	TXHOA004
3504	Nickel Total	0.013	Low value	Deleted	TXDCA001
3515	Nickel Total	0.01	Low value	Deleted	TXDCA002
2456	Selenium Total	0.3	Low value	Deleted	CAALA001
2457	Selenium Total	0.4	Low value	Deleted	CAALA001
2458	Selenium Total	0.068	Low value	Deleted	CAALA001
2459	Selenium Total	0.2	Low value	Deleted	CAALA001
2460	Selenium Total	0.059	Low value	Deleted	CAALA001
2461	Selenium Total	0.13	Low value	Deleted	CAALA001

Table B1. Modified Values in the NSQD – Continued

Order	Constituent	Original value	Problem	Action	Location_ID
2462	Selenium Total	0.095	Low value	Deleted	CAALA001
1082	Silver Total	290	Unusual elevated value, no evidence	Deleted	MDCLCOJS
1262	Silver Total	90	Unusual elevated value, no evidence	Deleted	MDPGCOS4
2006	Zinc Total	0.11	Low value	Deleted	GACOC1A2
2007	Zinc Total	0.19	Low value	Deleted	GACOC1A2
3704	Zinc Total	1	Low value	Deleted	TXIRA002
3777	Zinc Total	1	Low value	Deleted	TXPLA002
3514	Runoff	2.296	High Value	Deleted	TXDCA002
3515	Runoff	0.909	High Value	Deleted	TXDCA002
3201	Runoff	3.25	High Value	Deleted	KATOBROO
1364	Runoff	1.318	High Value	Deleted	GACLCOSI
2401	Runoff	1.73	High Value	Deleted	NCFVWINS

Appendix C: Methods to Estimate Non-Detected Values in Stormwater Datasets

Introduction

A few large stormwater quality databases have been prepared in the past 20 years (EPA 1983; Smullen 2002, for example). The data collected generally shows that there are important variabilities in stormwater pollutant concentrations for different land uses. Other factors that some researchers have found to be important include: imperviousness, slope, and size of the watershed. However, these databases include numerous instances where the laboratory results are reported to be "below detection." Statistical analyses can be greatly affected by these uncertain values, depending on their number and percentage of occurrence. There are several schemes that have generally been used to overcome the problems associated with these non-detected values.

The NSQD database has collected data representing more than 3,700 storm events in the U.S., including information about the location of the monitoring station, watershed characteristics, hydrology, and chemical constituents. Each community has the flexibility to choose the equipment and analytical methods to detect the constituents in the stormwater. Chemical constituents in this database had been preliminary analyzed for different land uses (Pitt, *et al*. 2003). It has been observed while preparing the NSQD database, that different methods and procedures had been used for the analyses of the samples. The use of different methods generates different detection limits in the database for the same constituent.

Datasets containing values below the detection limits (censored data) complicate the statistical analyses, even including the basic calculations of the means and variance. Most of the time, "left-censored" data are of concern (observations below the detection limit). However, there are situations where "right-censored" data may occur, especially for bacteria analyses, when the observations are greater than the upper limit of the dilution. Three main approaches to the analysis of censored data can be found in the literature: substitution, statistical estimation, and graphical methods. In this chapter, these methods will be presented using different data sets.

Analysis of Multiple Censored Data

Estimation methods for single censored data have been widely discussed in the literature. However, in the case of multiple censored data (datasets affected by several different detection limits), the situation is not the same. Helsel and Cohn (1988) continued the previous work of Guilliom and Helsel, but for multiple censored data.

Eight methods were studied in the multiple censored cases:

1) ZE: Censored data are assumed to equal zero.

2) DL: Censored data are assumed to equal the detection limit.

3) HA: Censored data are assumed to equal half the detection limit.

4) LR: Entire data set is log transformed and is assumed to be normally distributed. Censored data is estimated using least squares regression.

5) MR: Plotting positions are calculated using equations given by Hirsch and Stedinger (1987).

6) LM: Concentrations are assumed log normally distributed with parameters using the Cohen method. The mean and standard deviation of the untransformed values were estimated using the equations given by Aitchison and Brown (1969).

7) MM: This method uses the maximum likelihood method, but for the case of multiple censored data. Cohen (1976)

8) AM: Adjusted maximum likelihood procedure of Cohn (1988). The AM method is the same as the MM but makes a first order correction in the bias.

When the LR and LM methods were used, all the points below the highest of the censoring thresholds were treated as less than that censoring level. This will simplify the problem as a single censored occurrence. In the last three methods listed above, it is assumed that the data is log-normally distributed.

The results indicate that the MM and MR methods are improvements compared to the results obtained with the single threshold assumption. The MR, MM and AM methods were also compared. A higher RMSE (root mean squared error) for the moments was estimated by the MM method. The AM method present lower errors than the plotting position method (MR), but it is less robust for distributions different than log-normal. The substitution methods present a higher error than the MR or the AM methods.

One of the main problems using these methods was to assume that that the data is lognormal. There is no certainty that water quality follows this distribution. For that reason, robust methods are considered very important in water quality analysis. When data depart from the lognormal distribution, the RMSE of the mean and standard deviation values, when using the MM and AM methods, can be larger than 1000%. Helsel and Cohn (1988) indicate that in water quality data the lognormal distribution and the gamma with a coefficient of variation of two are very common. The MR model present better results when the distribution is not known.

They also evaluate the plotting position using the Weibull, Blom, and Hazel equations. There really is not an effect in the results when any of these equations are used.

If the distribution is unknown, the MR method should be chosen. If there is certainty that the distribution is lognormal, the AM method is recommended. The previous methods were evaluated with copper observations in commercial areas during the fall. Table C1 shows the original observations, and Table C2 show the log-transformed observations.

Table C1. Copper Observations in Commercial Areas

2	2	3	5	5	5	5	5.2	5.4	5.5	6	6
6	6.5	6.5	7	8	8	8.1	8.4	9	9	9	10
10	10	10	10	10	10	10	10	10	11	11	13
13	13.4	14	14	14	14	14	14	14.4	14.5	15	15
15	17	17	17	17	17.1	18	19	19	20	20	20
20	20	20	20	20	20	21	21	21	22	22	22.4
23	24	24	26	26.6	26.9	29	29	30	30	33	33.7
36	37	37	40	41.3	42	50	50	50.5	50.7	59.4	60
60	61	62	70	100	130	130	175	<10	<10	<10	<10
<10	<10	<20									

Table C2. Copper Observations in Commercial Areas – (Log Values)

0.3	0.3	0.5	0.7	0.7	0.7	0.7	0.7	0.73	0.74	0.78	0.78
0.8	0.81	0.8	0.9	0.9	0.9	0.9	0.9	0.95	0.95	0.95	1
1	1	1	1	1	1	1	1	1	1.04	1.04	1.11
1.1	1.13	1.2	1.2	1.15	1.15	1.2	1.2	1.16	1.16	1.18	1.18
1.2	1.23	1.2	1.2	1.23	1.23	1.3	1.3	1.28	1.3	1.3	1.3
1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.32	1.34	1.34	1.35
1.4	1.38	1.4	1.4	1.43	1.43	1.5	1.5	1.48	1.48	1.52	1.53
1.6	1.57	1.6	1.6	1.62	1.62	1.7	1.7	1.7	1.7	1.77	1.78
1.8	1.79	1.8	1.9	2	2.11	2.1	2.2	<1	<1	<1	<1
<1	<1	<1.30									

Figure C1 shows the probability plots when censored data was deleted, replaced by the detection limit, replaced by half of the detection limit, and estimating the values below the highest detection limit using the LR method. The results indicated that there is a bias in the mean value when the censored data is deleted or replaced by the detection limit. When data is replaced by half of the detection limit, or is estimated by the LR methods, the results are very similar. Notice that in the LR method, all the values below the highest detection limit are considered censored. This assumption changes the level of censoring from 6% to 58%, but even at this level of censoring, the results are very close to those obtained with the substitution methods. Because the transformed data seems to follows a log-normal distribution, it is possible to estimate the moments using only the upper side of the line.

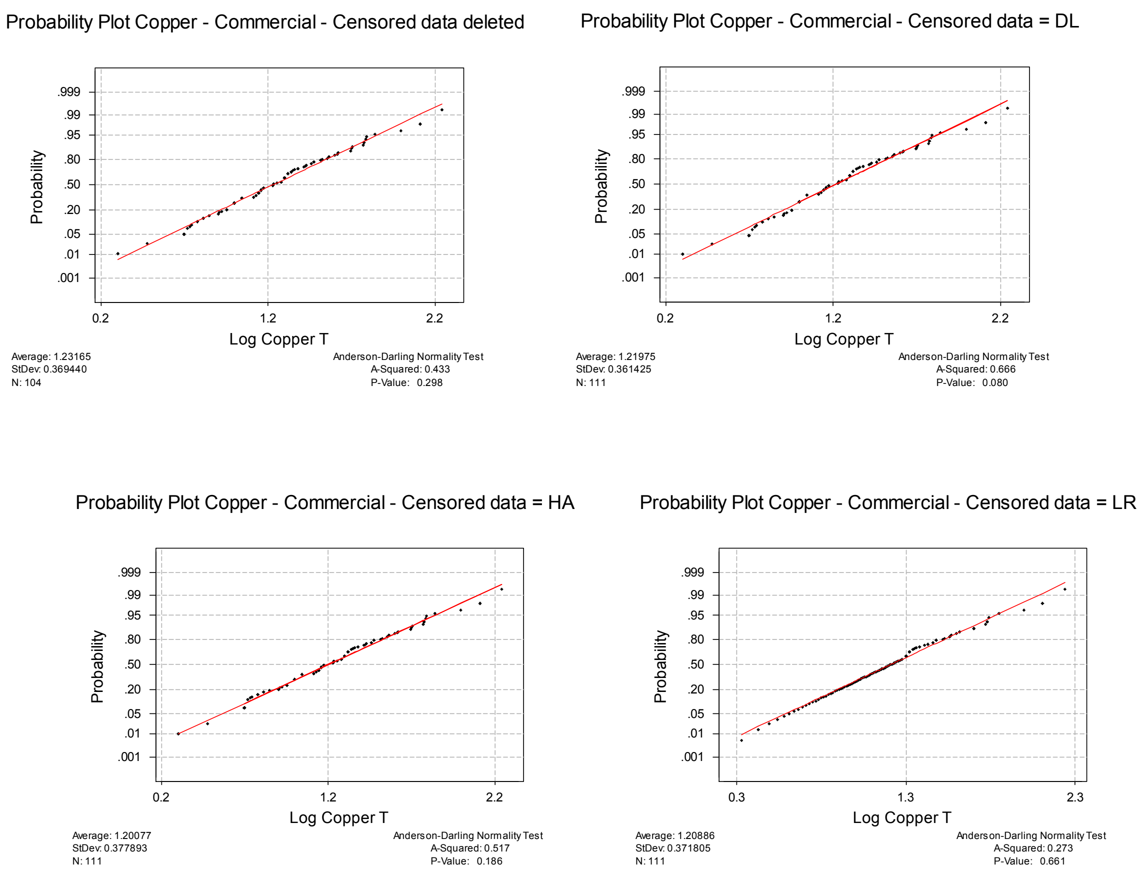


Figure C1. Probability plot for different replacements of the censored data for copper

The MR method was proposed by Hirsch and Stedinger (1987). They define a variable Aj as the number of uncensored data greater than the detection limit j and below the next detection limit. In the copper data, there are 47 uncensored observations above the highest detection limit (A3=47, DL=20). There are 34 observations between the detection limits of 20 and 10 μg/L, (A2=34, DL=10). Finally, there are 23 uncensored observations above the minimum detection limit, zero (A1=23, DL=0). This assumption must be made in the case that the smallest value in the dataset is not a censored value. The parameter Bj is defined as the number of censored and uncensored observations below the j detection limit. In the example case, there are 64 observations below the second detection limit (B3=64), 29 observations below the first detection limit (B2=29), and zero observations below the detection limit zero (B1=0). The method uses the probability of exceeding the jth detection limit pe,j to calculate the probability position of each observation.

(C.1) $$p_{e,j} = p_{e,j+1} + \frac{A_j}{A_j + B_j}\left(1 - p_{e,j+1}\right)$$

The calculations are easier going from higher to lower values. In the copper example, there are three detection limits; by definition, the probability of exceeding a fourth detection limit is zero. The probability of exceeding the third, second and first detection limits are 0.423, 0.735 and 1, respectively.

The Weibull formula was used to calculate the plotting position of the censored data in the range between the probabilities of exceeding boundaries.

(C.2) $$p(i) = (1 - p_{e,j}) + (p_{e,j} - p_{e,j+1})\frac{i}{A_j + 1}$$

This formula indicates that the values observed between the j and j+1 range are distributed according to the Weibull formula. The plotting position for the censored data follows the same concept; distribute the censored data between the limits using the Weibull formula. For the censored observations, the plotting positions can be calculated as:

(C.3) $$pc(i) = (1 - p_{e,j})\frac{i}{C_j + 1}$$

The formula calculates the position of the ith censored observation, among the C tied observations, in the jth detection limit. The probability plot is shown in Figure C2.

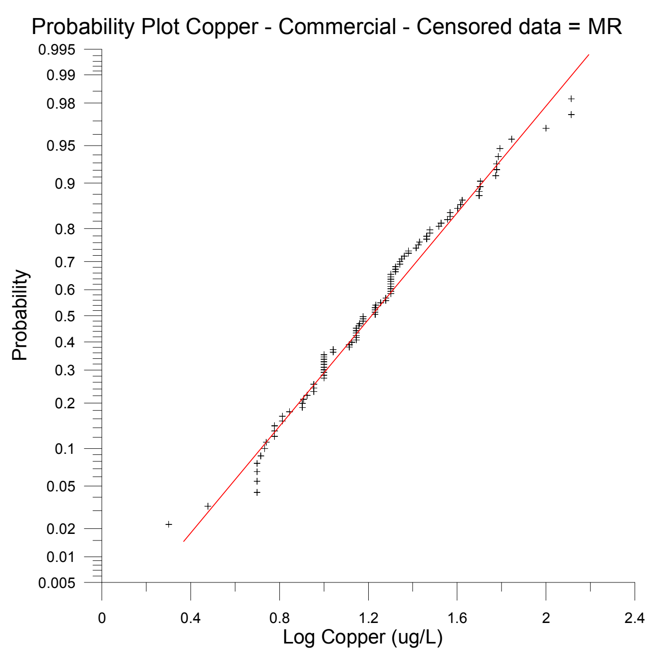


Figure C2. Probability plot using the MR method.

The MM method presented by Cohen (1976) uses the maximum likelihood method for the three parameter lognormal distribution.

(C.4) $$f(x;\mu,\sigma,\gamma)=\frac{1}{\sigma\sqrt{2\pi}(x-\gamma)}\exp\left[-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2}\right]$$

$$\gamma<x<\infty$$

In this distribution, X is a random variable lognormal. The method assumes that there are k censored observations and n noncensored observations. For a censored level Tj, the transformed yj will be:

(C.5) $$y_i=\ln(T_j-\gamma)$$

and,

(C.6) $$\xi_j=\frac{y_i-\mu}{\sigma}$$

Cohen created a new variable Zj that is used to solve the maximum likelihood estimators.

(C.7) $$Z_j=Z(\xi_j)=\frac{\varphi(\xi_j)}{1-\Phi(\xi_j)}$$

Where:
$\phi(\xi_j)$ = normal density function N(0,1)
$\Phi(\xi_j)$ = normal cumulative density function N(0,1)

Three simultaneous equations can be solved to estimate the parameters μ, σ, and γ.

(C.8) $$\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]+\sigma\sum_{j=1}^{k}Z_j=0$$

(C.9) $$\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2+\sigma^2\left[\sum_{j=1}^{k}\xi_jZ_j-n\right]=0$$

Cohn suggested assuming a γ value and solve for μ and σ from the first two equations. After that, the γ parameter can be recalculated using the third equation. In some cases, the parameter g does not converge. In that case the following approximation must be used.

(C.10) $$\gamma=x_k-\exp(\mu+\sigma\xi_k)$$

Where x_k is the k^{th} order statistic in the sample.

The AM method is the best alternative in the case that the distribution is lognormal. In any other case, it was found that elevated bias and rmse are obtained in the mean and the variance (>1000%) if the distribution is different.

The National Council of the Paper Industry for Air and Stream Improvement had created seven technical bulletins about statistical methods used with environmental data sets (NCASI 1995). One of the reports presents a decision tree to select the appropriate statistical method and a description of the Cohen's multilevel MLE procedure. This method was recommended after compare it with other methods, such as replacement/deletion, D-log procedure, regression of normal order statistics balancing techniques, and graphical techniques.

The diagram indicates that in some cases of multiple detection limits, the problem can be solved using single censoring point (SCP) methods, for example when all the non-detected values are smaller than the detected values. In other situations, the simplification cannot be done and multiple censoring point (MCP) methods must be used. The Cohen's maximum likelihood method obtained the mean and variance estimates from the logarithm of the likelihood function and obtaining the partial derivate in respect to the mean and the variance. The following equation defines the log likelihood function:

(C.11)
$$L(S) = -n\ln(\sigma) - \frac{1}{2}\sum_{1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2 + \sum_{1}^{k} r_i \ln[F_i] - n\ln[\sqrt{2\pi}]$$

Where:

- S = sample set containing a total of N censored observations and fully quantified values
- x_i = i^{th} fully quantified value
- μ = population mean
- σ = population standard deviation
- k = number of censored levels
- r_i = number of censored values at each censored level i
- n = number of noncensored observations
- $F_i = F(\xi_i)$ = area under standard normal curve at f
- $\xi_i = (T_i - \mu)/\sigma$, standard normal variate for the i^{th} censoring level
- $\phi(t) = (2\pi)^{-1} \exp[-t^2/2]$, ordinate value of normal variate , $f/(\xi_i)$
- T_i = the limit of detection of the i^{th} level of censoring.

The derivates are:

(C.12)
$$\frac{\partial L}{\partial \mu} = \frac{n}{\sigma}\left[\frac{(\bar{x}-\mu)}{\sigma} - \sum_{1}^{k}\frac{r_i}{n}Z_i\right] = 0$$

(C.13) $$\frac{\partial L}{\partial \sigma} = \frac{n}{\sigma}\left[\frac{s^2 + (\bar{x} - \mu)^2}{\sigma^2} - 1 - \sum_{1}^{k} \frac{r_i}{n} \xi_i Z_i\right] = 0$$

Z_i = the hazard function ϕ_i/F_i.

s^2 = Sample variance.

NCASI (1995) includes the program source code for using Cohen's method, in FORTRAN and SAS. The procedure and the code presented in the technical bulletin No. 703 were used to estimate the censored observations for this NSQD research.

Appendix D: Unusual Sites Identified Using Xbar Plots

This appendix describes sites having unusual stormwater concentrations for all land uses, besides the residential areas that were described in Chapter 3.

Evaluation of the Methods Selected to Estimate Non-Detected Observations

Three methods were used to estimate appropriate substitution values for the non-detected observations: delete them ("ignore"), replace them by half of the detection limit ("HD") or estimate them using the Cohen's maximum likelihood method (an extrapolation of the probability plot of the data) ("estimate"), as presented in the preceding appendix. The following discusses the analyses for each constituent for each land use category.

Hardness

Total hardness was detected in all samples, except in industrial land use areas where less than 2% of the samples were not detected. Changes in the average, median, standard deviation and coefficient of variation were not significant if the non-detected values were ignored, estimated, or replaced by half of the detection limit. Table D1 shows that there are no important differences in the industrial land descriptions using any of these three methods.

Table D1. Summary Statistics for Estimated Observations for Total Hardness (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	250	250	250	139	139	139	138	138	138
% Detected	100.00			100.00			96.38		
Minimum	3.00	3.00	3.00	1.90	1.90	1.90	5.50	5.00	5.00
Maximum	401.00	401.00	401.00	356.00	356.00	356.00	888.00	888.00	888.00
Average	43.32	43.32	43.32	62.03	62.03	62.03	68.83	66.52	66.52
Median	32.00	32.00	32.00	38.90	38.90	38.90	39.00	38.50	38.50
Standard Dev.	44.87	44.87	44.87	65.17	65.17	65.17	104.55	103.32	103.32
Coeff. of Var.	1.04	1.04	1.04	1.05	1.05	1.05	1.52	1.55	1.55

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	8	8	8	127	127	127
% Detected	100.00			100.00		
Minimum	11.00	11.00	11.00	5.00	5.00	5.00
Maximum	270.00	270.00	270.00	1000.00	1000.00	1000.00
Average	145.25	145.25	145.25	57.19	57.19	57.19
Median	150.00	150.00	150.00	34.00	34.00	34.00
Standard Dev.	85.12	85.12	85.12	105.95	105.95	105.95
Coeff. of Var.	0.59	0.59	0.59	1.85	1.85	1.85

Figure A1 shows probability plots for industrial land use hardness values. The plot indicates that the mean value is smaller when the non-detected values are either estimated or replaced by half of the detection limit. The lower 40% of the distribution is displaced to the left. All the non-detected values were observed at 10 mg/L. The upper 60% of the distribution is not affected by the non-detected values.

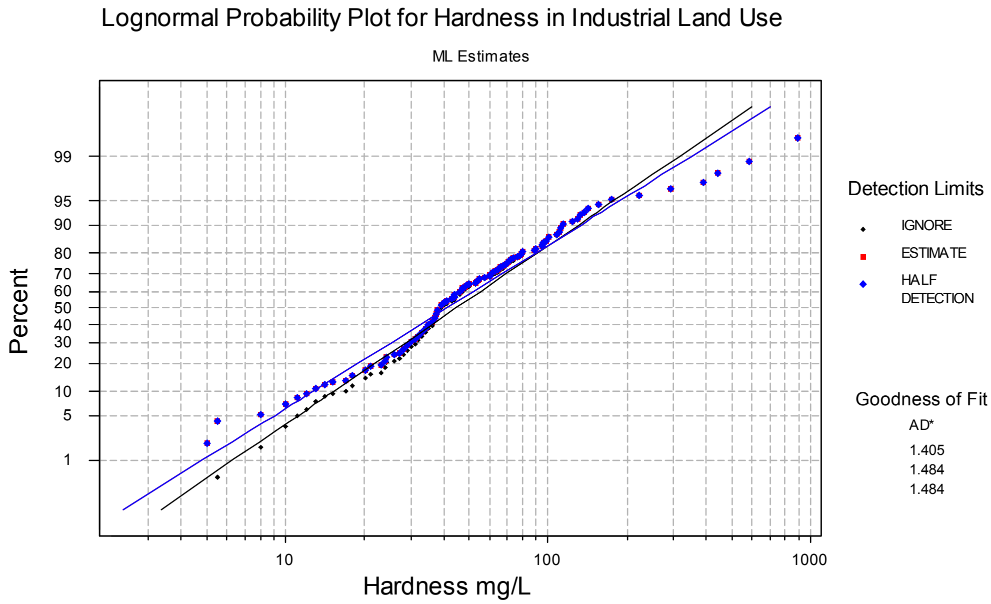


Figure D1. Estimated hardness distributions in industrial land use areas

Oil and Grease

Oil and grease had censored data for 37% and 72% of the observations. Table D2 shows the differences in the descriptive statistics using the three methods. The greatest change occurred in the coefficient of variation values for freeway sites. The mean oil and grease values increased in a range of 30% to 60% when the censored observations were ignored. The difference was below 4% when the censored observations were replaced using Cohen's maximum likelihood method, or replaced by half of the detection limit.

Table D2. Summary Statistics for Estimated Observations for Oil and Grease (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	533	533	533	308	308	308	327	327	327
% Detected	57.79			70.78			65.14		
Minimum	0.20	0.02	0.20	0.80	0.03	0.25	0.50	0.00	0.25
Maximum	2980	2980	2980	359	359	359	11000	11000	11000
Average	22.85	13.87	13.89	12.63	9.42	9.39	62.87	41.40	41.39
Median	3.85	2.50	2.50	4.70	3.00	3.00	5.00	2.50	2.60
Standard Dev.	175.53	133.76	133.76	39.75	33.80	33.81	753.77	608.56	608.56
Coeff. of Var.	7.68	9.65	9.63	3.15	3.59	3.60	11.99	14.70	14.70

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	19	19	19	60	60	60
% Detected	36.84			71.67		
Minimum	0.50	0.50	0.50	3.00	0.50	0.25
Maximum	3.70	3.70	3.70	30.00	30.00	30.00
Average	1.53	1.09	1.09	8.49	6.57	6.45
Median	1.30	0.50	0.50	8.00	4.65	4.65
Standard Dev.	1.07	0.93	0.93	5.28	5.42	5.52
Coeff. of Var.	0.70	0.85	0.85	0.62	0.83	0.86

The probability plot in residential land use areas indicates that the lower tail is better described with the Cohen estimated method (Figure D2). The upper tail was the same for the estimated and the half detection limit method. About 40% of the non-detected values were at the < 1 mg/L level, and another 40% were at the < 5 mg/L level. The estimated values better describe the lower tail, however there was no significant differences in the means, standard deviations and coefficients of variation. This case is very important because the level of censoring was large (42.2%). Ignoring the non-detected values increased the mean value by more than 64% and the standard deviation by more than 30%, and reduces the coefficient of variation in 20%.

The analyses for commercial land use data resulted in a similar trend as observed for the residential land use areas (Figure D3). There is a better description of the lower tail, but the mean, standard deviation and coefficient of variation values are almost the same if the censored data are replaced by half of the detection limit, or if they are estimated. The most frequent reported level of non-detected values was < 5 mg/L, followed by < 1 mg/L. The average was increased by 34%, and the standard deviation by 18%, when the censored data was ignored, and the coefficient of variation was reduced about 12% when the non-detected values were ignored.

Figure D4 shows the probability plot for oil and grease data at industrial land use areas and illustrates the case when an unusual value was present in the dataset. The maximum observation was larger by a factor of 2,200 compared with the median value of the distribution. This generates a coefficient of variation of 12 when the censored data are ignored, or 14.7 in the case when they are estimated or replaced by half of the detection limit.

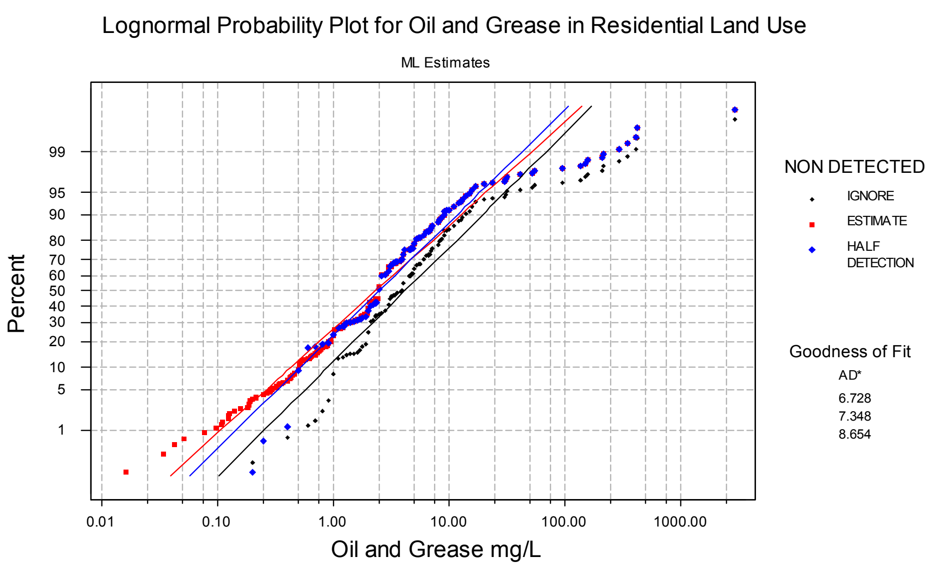


Figure D2. Estimated oil and grease distributions in residential land use areas

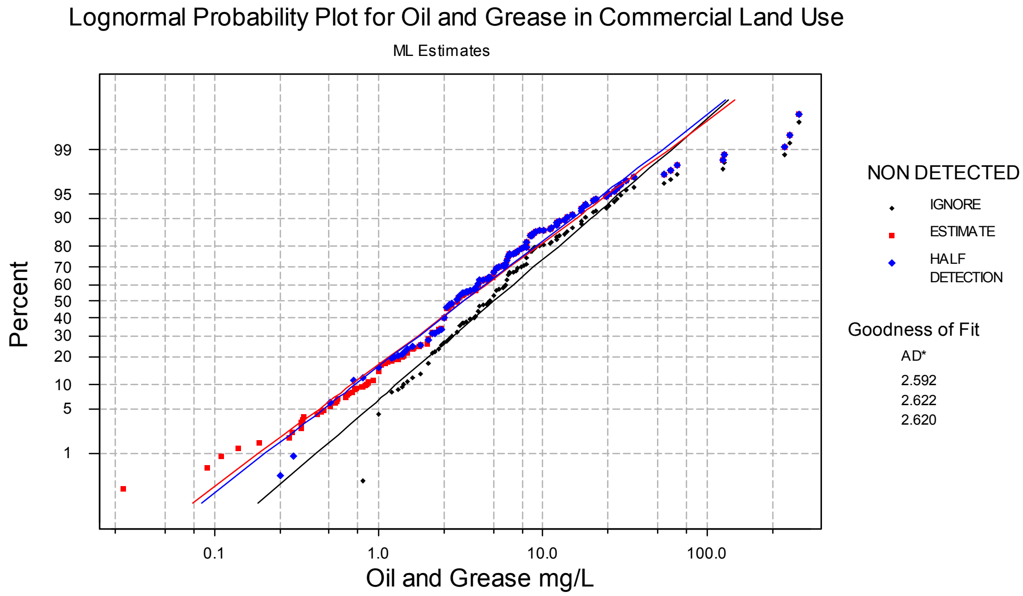


Figure D3. Estimated oil and grease distributions in commercial land use areas

The percentage of detected values for oil and grease in open space areas was very low (only 7 of 19 observations were detected) (Figure D5). Almost all of the non-detected values were at < 1 mg/L. It was not possible to use the Cohen's maximum likelihood method in this case because the percentage of non-detected values was too high. Ignoring the non-detected values will increase the mean value by almost 40% compared when the non-detected values were replaced with half of the detection limit.

The probability plot for freeway oil and grease values indicate that estimating or replacing the censored observations for half of the detection limit does not cause a significant difference in the coefficient of variation (Figure D6). The coefficient of variation was 3% larger when half of the detection limit was used instead of Cohen's method. A different situation occurs when the non-detected values were ignored. In this case, the coefficient of variation was reduced by 30% compared with the estimated method.

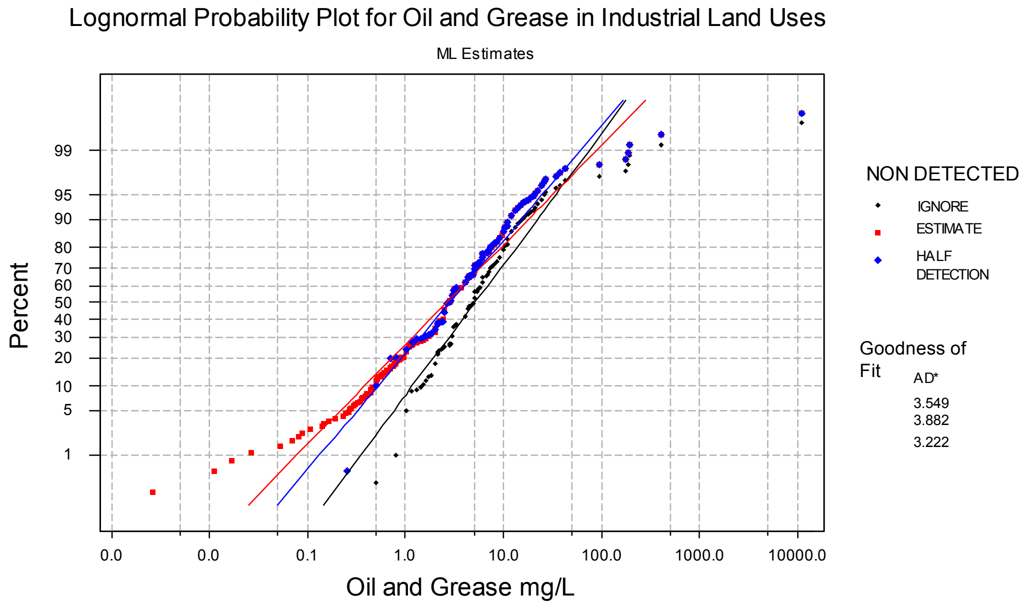


Figure D4. Estimated oil and grease distributions in industrial land use areas

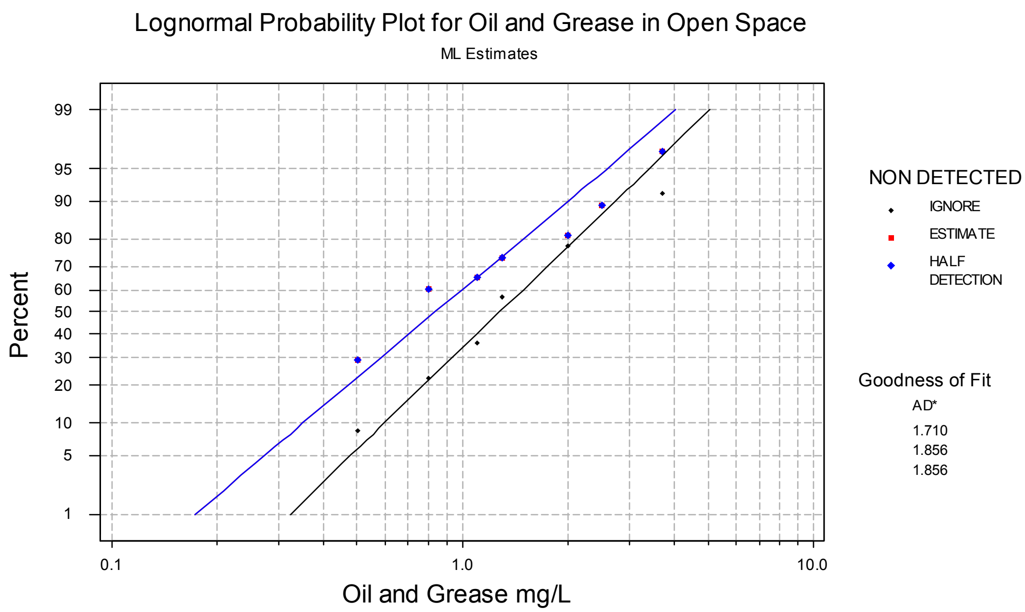


Figure D5. Estimated oil and grease distributions in industrial land use areas

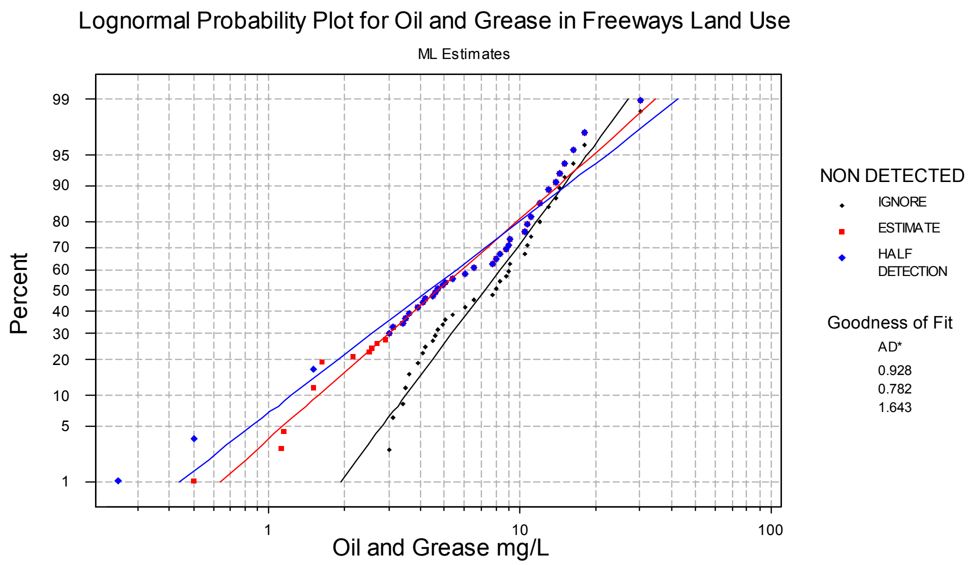


Figure D6. Estimated oil and grease distributions in freeway land use areas

Total Dissolved Solids (TDS)

In all the land use categories, the percentages of non-detected TDS values were very low. The lowest percentage was observed in open space areas, with 2% not detected. No important differences were observed in the means, standard deviations and coefficients of variation when the non-detected values were ignored, estimated using with the Cohen method, or substituting with half the detection limit. Descriptive statistics for each of the three methods are shown in Table D3.

Table D3. Summary Statistics for Estimated Observations for TDS (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	861	861	861	399	399	399	412	412	412
% Detected	99.19			99.50			99.51		
Minimum	3.00	3.00	0.50	4.00	4.00	4.00	4.50	1.78	2.50
Maximum	1700	1700	1700	3860	3860	3860	11200	11200	11200
Average	96.26	95.54	95.50	109.94	109.44	109.42	161.99	161.23	161.22
Median	72.00	70.50	70.50	74.00	74.00	74.00	91.00	89.50	89.50
Standard Dev.	102.45	102.35	102.38	208.76	208.36	208.37	582.40	581.09	581.09
Coeff. of Var.	1.06	1.07	1.07	1.90	1.90	1.90	3.60	3.60	3.60

Land use	OPEN SPACE Ignore	Estimate	HD	FREEWAY Ignore	Estimate	HD
Observations	45	45	45	97	97	97
% Detected	97.78			98.97		
Minimum	32.00	10.79	2.50	12.00	5.85	0.50
Maximum	542	542	542	470	470	470
Average	151.41	148.28	148.10	95.31	94.39	94.34
Median	124.50	119.00	119.00	77.50	77.00	77.00
Standard Dev.	109.83	110.58	110.82	76.38	76.52	76.59
Coeff. of Var.	0.73	0.75	0.75	0.80	0.81	0.81

Figure D7 shows the probability plots for residential land use TDS concentrations. The plot indicates that using half of the detection limit lowers values compared to the Cohen's maximum likelihood method. The upper 95% of the distributions are identical for the three cases. The probability plots don't indicate significant differences among the three methods for the remaining land uses. For example, Figure D8 shows the probability plots for commercial areas. The three lines overlap, except for a small fraction in the lower tail of the distribution.

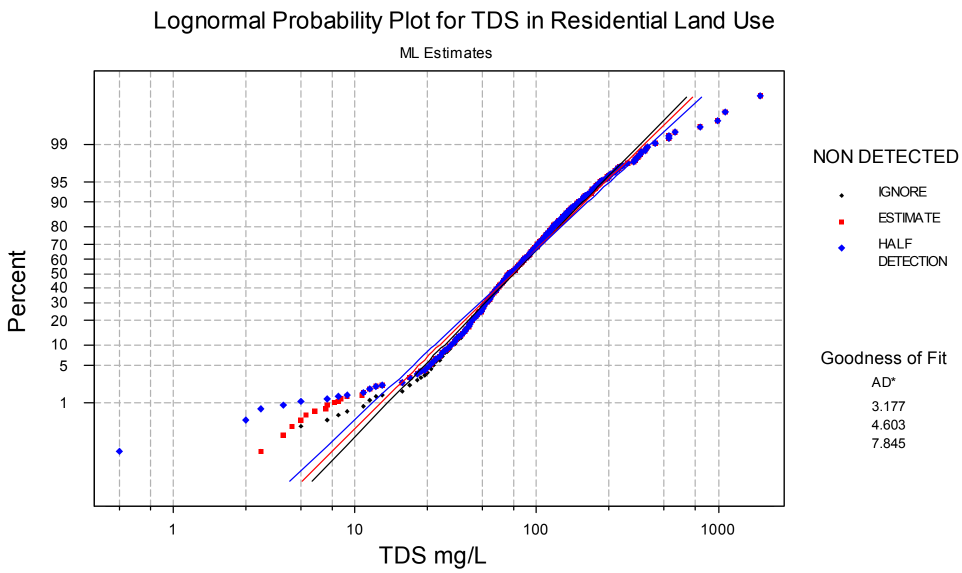


Figure D7. Estimated TDS distributions in residential land use areas

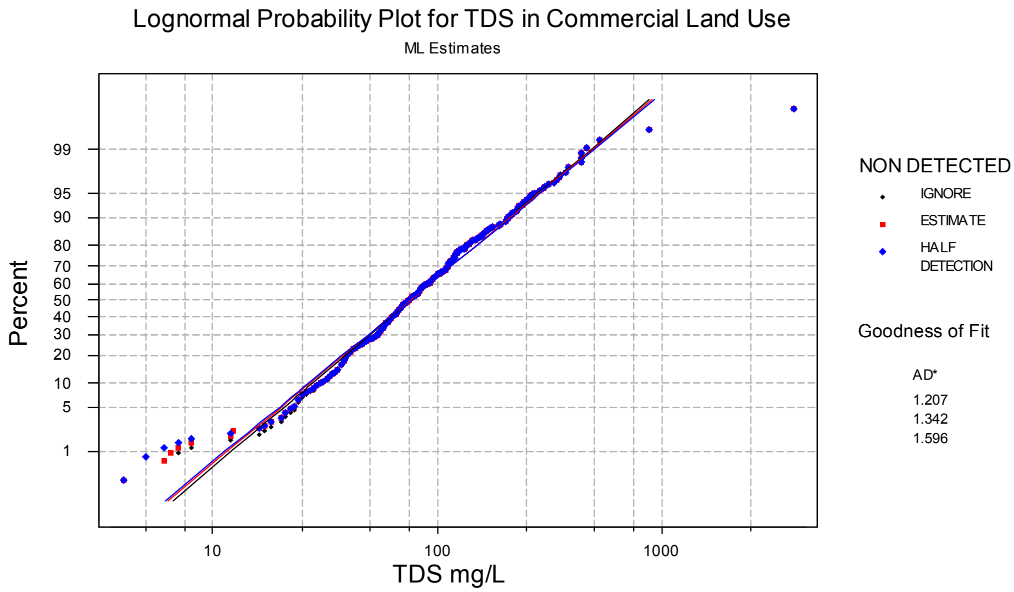


Figure D8. Estimated TDS distributions in commercial land use

Total Suspended Solids (TSS)

The results for TSS were similar to above described results for TDS, the maximum level of non-detected values was observed in open space areas, where about 5% of the observations were censored. Table D4 indicates that there are not any relevant differences in means, standard deviations or coefficients of variation for any of the three methods.

Table D4. Summary Statistics for Estimated Observations for TSS (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	991	990	991	458	457	458	427	426	427
% Detected	98.59			98.25			99.06		
Minimum	3.00	0.63	0.25	3.00	1.56	0.25	3.00	0.43	0.50
Maximum	2462	2462	2462	2385	2385	2385	2490	2490	2490
Average	99.84	98.53	98.46	110.06	108.45	108.18	142.44	141.36	141.12
Median	49.00	48.00	48.00	42.00	41.00	41.00	78.00	76.36	76.00
Standard Dev.	179.12	178.29	178.22	218.51	217.22	217.05	218.76	218.35	218.15
Coeff. of Var.	1.79	1.81	1.81	1.99	2.00	2.01	1.54	1.54	1.55

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	44	44	44	134	134	134
% Detected	95.45			99.25		
Minimum	3.00	1.22	0.50	3.00	3.00	0.50
Maximum	980	980	980	4800	4800	4800
Average	176.88	168.98	168.91	173.39	172.13	172.10
Median	48.50	39.00	39.00	99.00	98.50	98.50
Standard Dev.	263.04	259.44	259.49	448.85	447.39	447.41
Coeff. of Var.	1.49	1.54	1.54	2.59	2.60	2.60

The probability plots indicate that the lower values were better estimated using half of the detection limit, rather than the Cohen's method. This indicate that with large numbers observations and small percentages of non-detected values, replacing the missing data by half of the detection limit will produce similar means compared to those obtained when using the maximum likelihood method. Figure A9 shows the probability plot for TSS concentrations for residential land use areas. The three curves overlap, indicating than the three methods will produce practically the same result.

The probability plot for open space has the lower number of observations among the five land uses. In this case, the pattern observed in the three methods was almost the same. The coefficient of variation increases only 3% when the censored data was estimated with the Cohen method, or replaced by half of the detection limit.

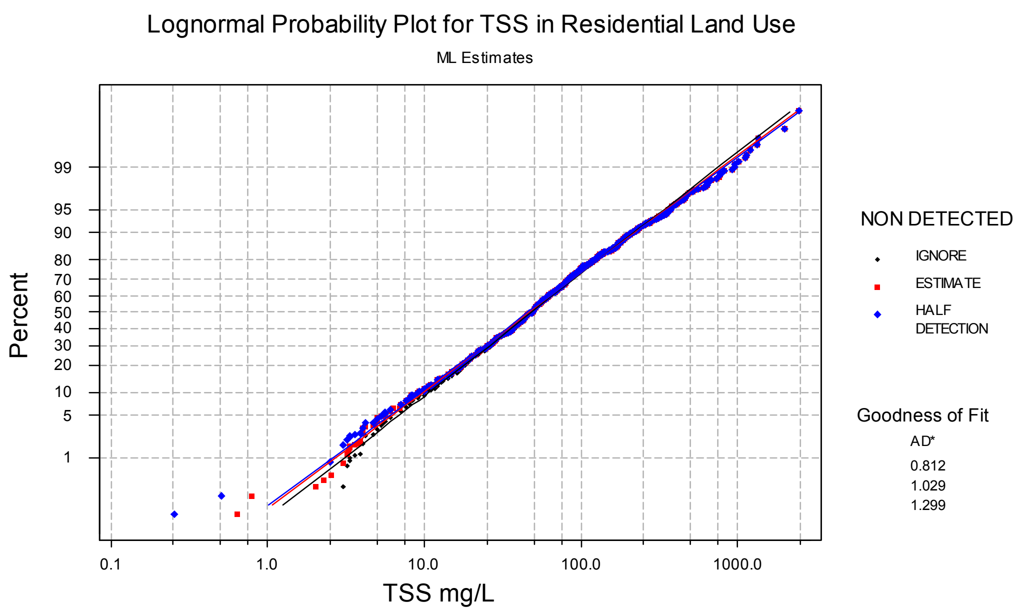


Figure D9. Estimated TSS distributions in residential land use areas

Biochemical Oxygen Demand (BOD_5)

The percentage of non-detected values for BOD_5 was higher in open space and freeway areas compared with the other land uses (Table D5). The lowest concentrations were observed in open space areas with a median BOD_5 value of 4 mg/L. Freeways, commercial and residential land use areas have similar concentrations, with 15 mg/L average BOD_5 values. The highest BOD_5 concentration was observed at an industrial land use site, however a single unusual BOD_5 observation of 6,920 mg/L had a large effect on the mean, standard deviation and coefficient of variation values.

The lognormal probability plot for industrial land use areas showed one unusual BOD_5 observation. This BOD_5 concentration was 35 times larger than the second highest observation. This unusual value increased the standard deviation almost 18 times compared with the other land uses. Figure D10 shows the probability plot for BOD_5 concentrations at industrial land use areas.

Table A5. Summary Statistics for Estimated Observations for BOD_5 (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	941	941	941	432	432	432	406	406	406
% Detected	97.56			97.45			95.32		
Minimum	1.00	1.00	0.50	2.00	0.75	0.50	1.00	0.55	0.50
Maximum	350	350	350	150	220	150	6920	6920	6920
Average	15.05	14.97	14.84	18.16	18.58	18.14	35.92	34.65	34.47
Median	9.00	9.00	9.00	11.00	11.00	11.00	9.00	9.00	9.00
Standard Dev.	22.25	22.34	22.11	20.25	22.59	20.63	351.89	343.62	343.61
Coeff. of Var.	1.48	1.49	1.49	1.12	1.22	1.14	9.80	9.92	9.97

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	44	43	44	26	26	26
% Detected	86.36			84.62		
Minimum	1.00	0.62	0.50	2.0	1.5	1.5
Maximum	20	20	20	89	89	89
Average	6.25	5.68	5.74	14.86	13.06	12.88
Median	5.40	4.00	4.00	8.0	6.5	6.5
Standard Dev.	4.30	4.34	4.38	18.68	17.67	17.76
Coeff. of Var.	0.69	0.76	0.76	1.26	1.35	1.38

Open space and freeway areas had the largest level of non-detected BOD_5 values. The mean value for open space areas increased by 10% when the censored data were ignored. No significance difference was observed for the variance values (Figure D11). Estimating the non-detected value using Cohen's method, or replacing the non-detected values by half of the detection limit results in almost the same means, standard deviations and coefficients of variation values.

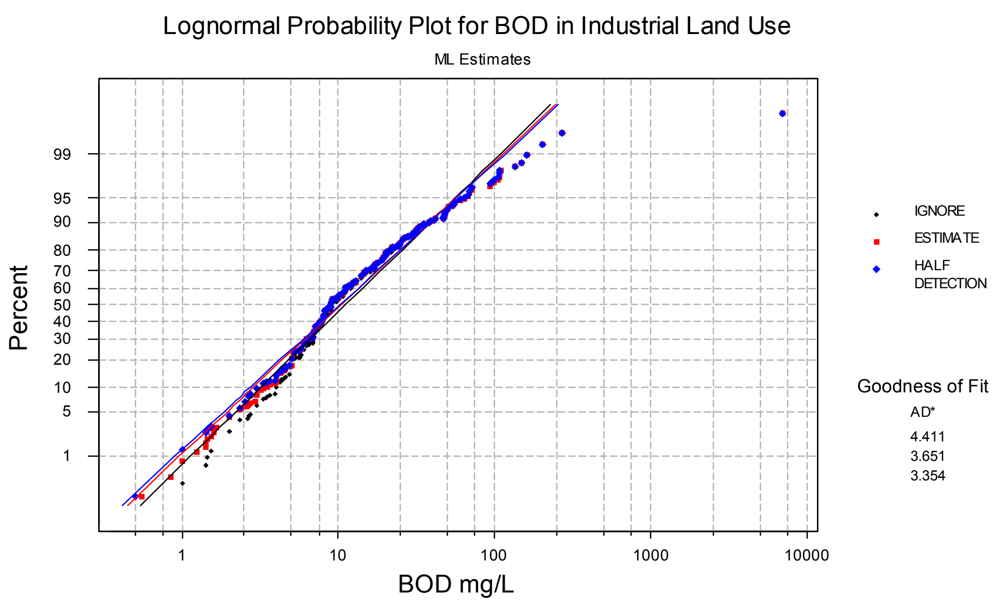


Figure A10. Estimated BOD_5 distributions in industrial land use areas

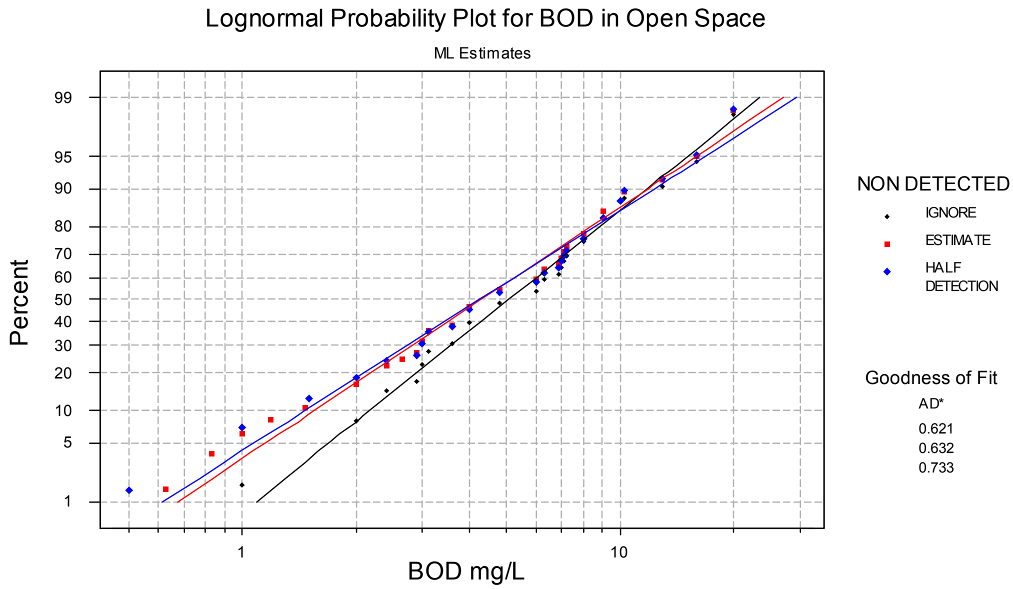


Figure D11. Estimated BOD_5 distributions for open space land use areas

Chemical Oxygen Demand (COD)

Differences in the means, averages and coefficients of variation for COD concentrations between the different methods for replacing the censored data were not important, except for the open space land use area where the level of non-detected observations was high (close to 25%) (Table D6). In the remaining land use areas, the frequency of non-detected values was smaller than 2%.

Table D6. Summary Statistics for Estimated Observations for COD (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	796	796	796	373	373	373	361	361	361
% Detected	98.87			98.39			98.89		
Minimum	5.00	1.74	0.50	4.00	1.96	0.50	2.00	2.00	2.00
Maximum	620	620	620	635	635	635	1260	1260	1260
Average	74.34	73.55	73.52	94.11	92.70	92.63	103.23	102.26	102.17
Median	55.00	53.60	53.60	60.00	59.00	59.00	60.00	59.00	59.00
Standard Dev.	69.12	69.12	69.15	94.39	94.28	94.34	127.35	126.97	127.03
Coeff. of Var.	0.93	0.94	0.94	1.00	1.02	1.02	1.23	1.24	1.24

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	44	44	44	67	67	67
% Detected	75.00			98.51		
Minimum	8.00	3.70	5.00	2.44	2.44	2.44
Maximum	476	476	476	1012.82	1012.82	1012.82
Average	51.47	40.93	40.76	140.99	139.10	138.96
Median	42.10	24.85	24.85	100.00	100.00	100.00
Standard Dev.	79.11	70.73	70.78	148.89	148.56	148.69
Coeff. of Var.	1.54	1.73	1.74	1.06	1.07	1.07

One characteristic of the COD probability plot is that the lower tail does not follow the trend showed by the rest of the distribution. Figure D12 shows an example COD distribution for residential land use areas. This effect is increased when the censored data is estimated or replaced by half of the detection limit.

The mean value in open space land use areas was increased by 25% when the censored data was ignored (Figure D13). In contrast, the coefficient of variation was reduced by almost 12 % when the non-detected values were ignored. No significant differences can be observed when the censored data was estimated using Cohen's method or replaced with half of the detection limit.

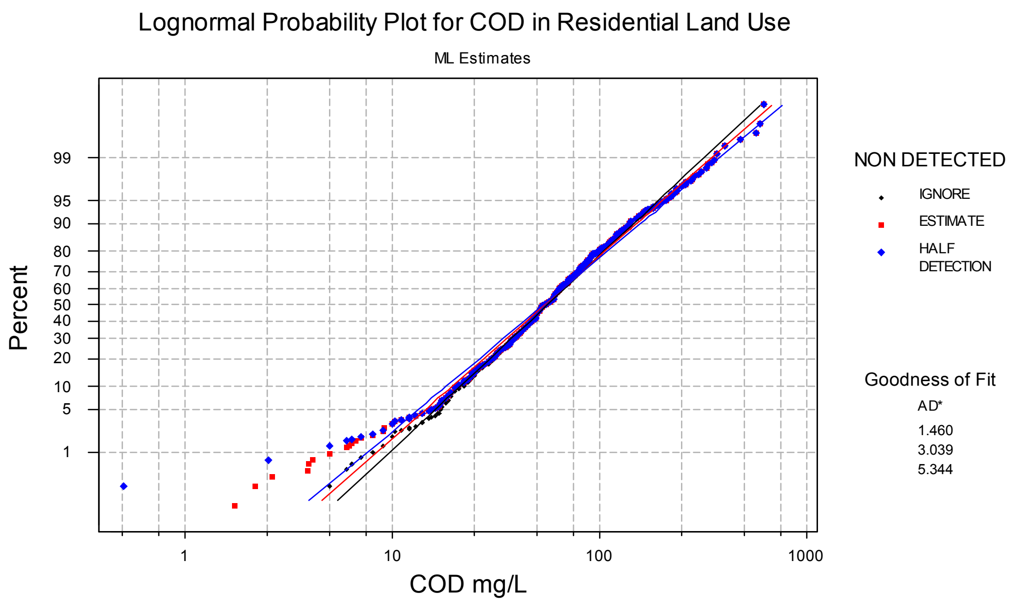


Figure D12. Estimated COD distributions in residential land use areas

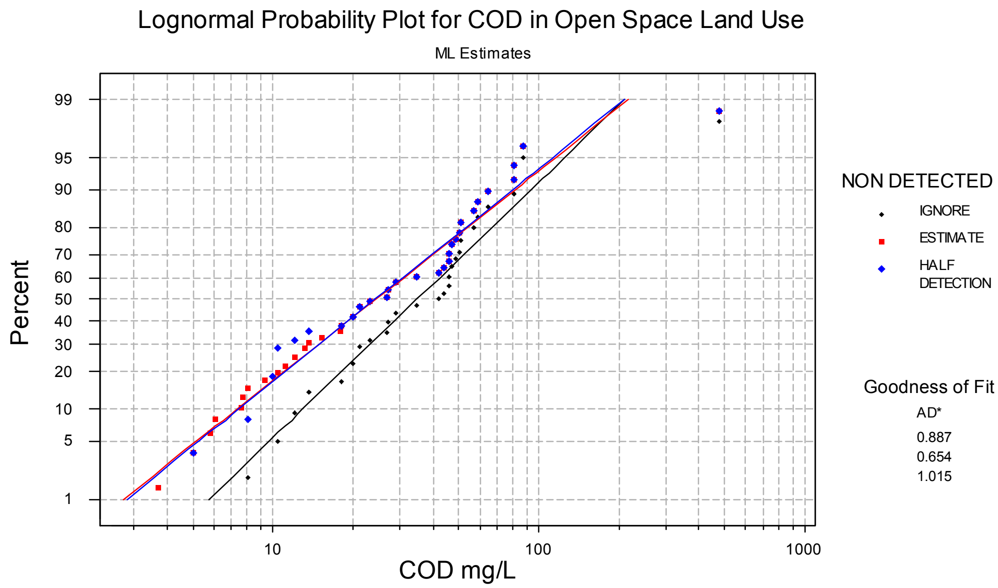


Figure D13. Estimated COD distributions in open space land use areas

Ammonia (NH_3)

Ammonia had one of the largest levels of censored observations of the common stormwater constituents examined in detail. The percentage of non-detected observations was about 20%, except for open space areas where it is more than 80%. The highest ammonia concentrations were observed at the freeway sites. Ignoring the censored observations increased the mean values by about 15%, while ignoring the non-detected values increased the coefficients of variation by almost 15%.

Table A7. Summary Statistics for Estimated Observations for Ammonia (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	595	595	595	299	299	299	253	252	253
% Detected	81.51			83.28			83.40		
Minimum	0.01	0.00	0.01	0.02	0.00	0.01	0.03	0.00	0.01
Maximum	5.60	5.60	5.60	7.80	7.80	7.80	9.84	9.84	9.84
Average	0.47	0.40	0.39	0.85	0.73	0.73	0.78	0.68	0.68
Median	0.32	0.27	0.25	0.50	0.41	0.40	0.47	0.38	0.36
Standard Dev.	0.51	0.48	0.48	1.02	0.97	0.97	0.96	0.91	0.91
Coeff. of Var.	1.09	1.20	1.22	1.20	1.32	1.33	1.23	1.35	1.35

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	32	32	32	79	79	79
% Detected	18.75			87.34		
Minimum	0.07	0.02	0.01	0.08	0.08	0.08
Maximum	1.80	1.80	1.80	11.87	11.87	11.87
Average	0.64	0.27	0.26	1.73	1.53	1.52
Median	0.18	0.25	0.25	1.07	0.90	0.90
Standard Dev.	0.79	0.38	0.38	2.24	2.16	2.16
Coeff. of Var.	1.24	1.43	1.44	1.30	1.41	1.42

The probability plots showed that replacing the non-detected values by half of the detection limit resulted in lower values than if the Cohen's method was used. The Anderson Darling statistic for normality increased when the censored data was estimated, indicating a better fit to a normal distribution. Figure D14 shows the probability plot for ammonia for commercial land use areas. In open space areas, the estimated values don't seem to fit the log normal distribution (Figure D15). Estimating the censored observations using Cohen's method when more than 80% of the observations were below the detection limit is certainly not recommended.

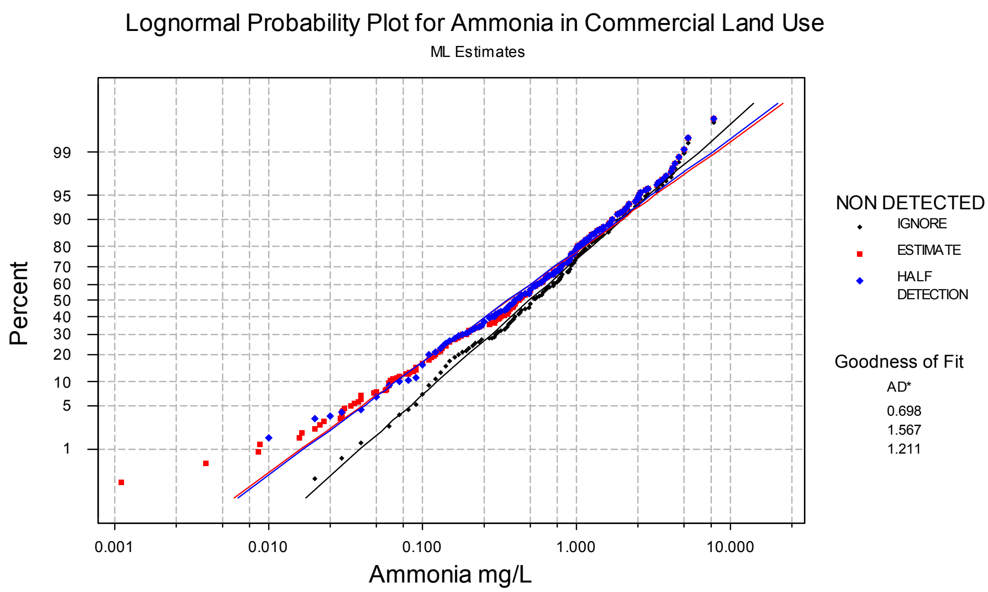


Figure D14. Estimated ammonia distributions in commercial land use areas

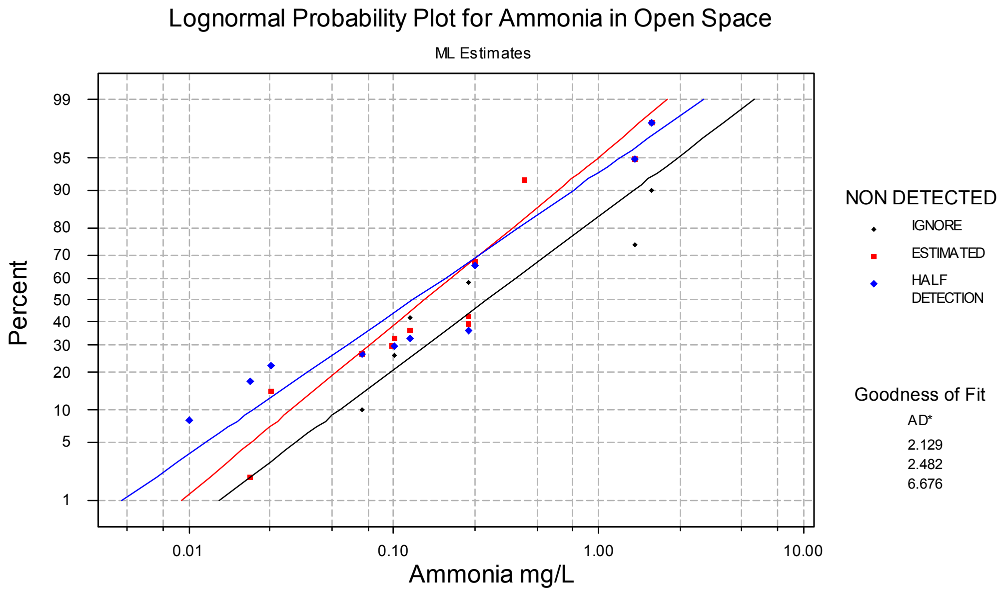


Figure D15. Estimated Ammonia distributions in open space land use areas

Nitrite and Nitrate ($NO_2 + NO_3$)

The percentages of non-detected values was smaller than 5% in all the land uses for nitrites plus nitrates, except for open space areas where the level of censored values was higher than 15%. There were no significant differences in the means, standard deviations and coefficients of variation, except for the open space data set, when the alternative substitution methods were used.

Table D8. Summary Statistics for Estimated Observations for $NO_2 + NO_3$ (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	927	927	927	425	425	425	417	417	417
% Detected	97.41			98.12			96.16		
Minimum	0.01	0.01	0.01	0.03	0.02	0.01	0.02	0.02	0.01
Maximum	18.00	18.00	18.00	8.21	8.21	8.21	8.40	8.40	8.40
Average	0.76	0.75	0.74	0.86	0.85	0.85	0.98	0.95	0.94
Median	0.59	0.58	0.58	0.61	0.60	0.60	0.73	0.72	0.70
Standard Dev.	0.87	0.86	0.86	0.91	0.91	0.91	0.87	0.86	0.86
Coeff. of Var.	1.14	1.15	1.16	1.06	1.08	1.08	0.89	0.91	0.91

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	44	44	44	25	25	25
% Detected	84.09			96.00		
Minimum	0.09	0.02	0.05	0.10	0.10	0.10
Maximum	3.33	3.33	3.33	3.00	3.00	3.00
Average	0.99	0.84	0.84	0.51	0.50	0.50
Median	0.59	0.50	0.50	0.28	0.28	0.26
Standard Dev.	0.88	0.88	0.88	0.63	0.62	0.62
Coeff. of Var.	0.89	1.04	1.04	1.23	1.23	1.25

The probability plots for residential, commercial and industrial land use areas show a different trend for the lower tail of the distribution up to the 10th percentile for the different methods. The departures from normality are more evident in the case when the censored observations are replaced by half of the detection limit (Figure D16). In open space areas, when the censored data was estimated or replaced, the coefficient of variation increased almost 17% due the elevated level of censoring (Figure D17). There were no observed differences in the means, standard deviations and coefficients of variation when the censored values were replaced by half of the detection limit or estimated using Cohen's method.

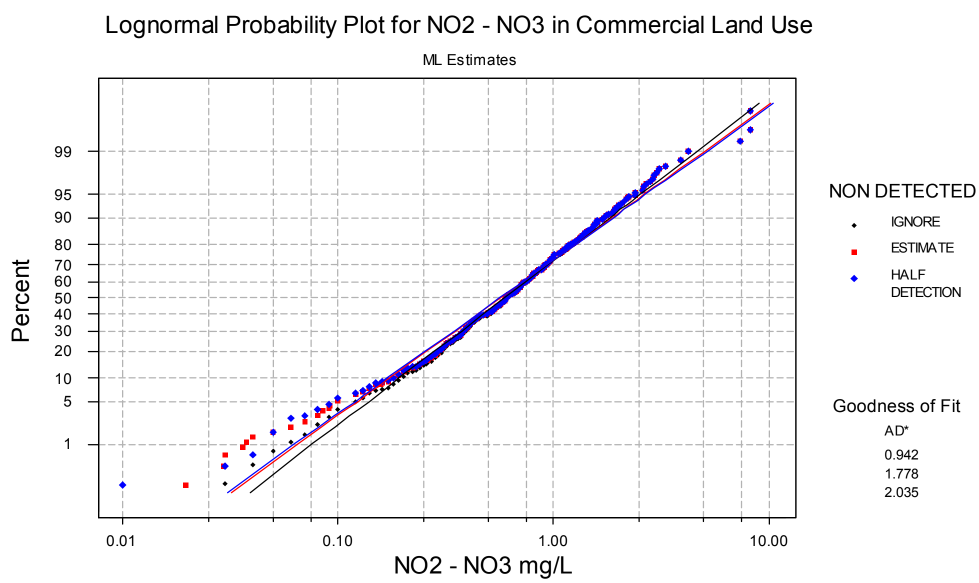


Figure D16. Estimated nitrate - nitrite distributions in commercial land use areas

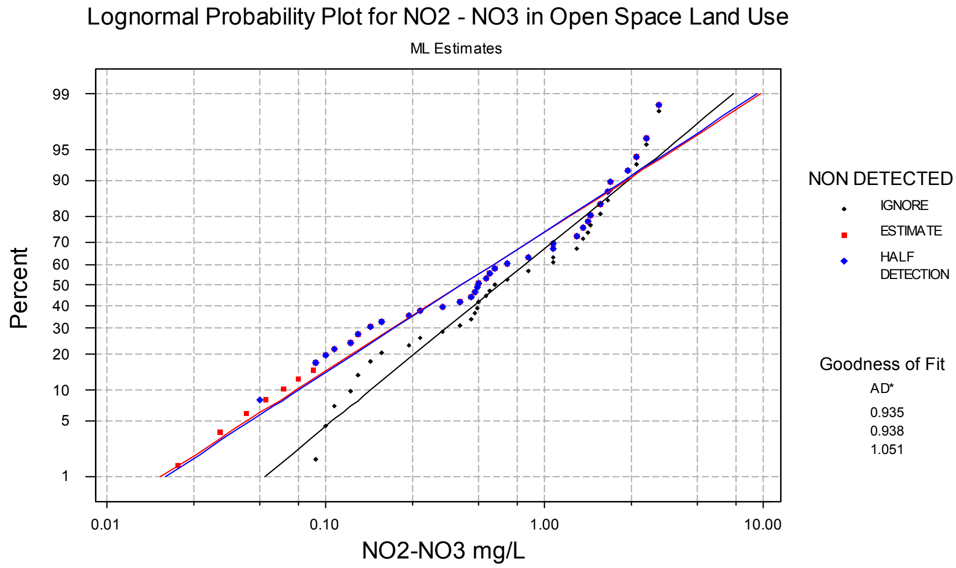


Figure D17. Estimated nitrate - nitrite distributions in open space land use areas

Total Kjeldahl Nitrogen (TKN)

The level of censoring for TKN was smaller than 4% for all land use areas except for open space areas. The highest TKN concentrations were observed in freeway areas, and the lowest TKN concentrations were observed in open space areas (Table D9). Large changes in the coefficient of variation were observed in open space areas when using Cohen's method (an increase of 15%) and when replacing the censored values by half of the detection limit (increases of 22%).

Table D9. Summary Statistics for Estimated Observations for TKN (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	957	957	957	449	449	449	439	439	439
% Detected	96.76			97.33			95.90		
Minimum	0.05	0.00	0.01	0.05	0.02	0.01	0.05	0.01	0.01
Maximum	36.00	36.00	36.00	15.00	15.00	15.00	25.00	25.00	25.00
Average	1.96	1.91	1.90	2.23	2.18	2.17	2.23	2.17	2.16
Median	1.43	1.40	1.40	1.59	1.55	1.55	1.40	1.37	1.37
Standard Dev.	2.05	2.04	2.04	2.08	2.07	2.08	2.56	2.53	2.54
Coeff. of Var.	1.05	1.07	1.07	0.93	0.95	0.96	1.15	1.17	1.18

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	45	45	45	125	125	125
% Detected	71.11			96.80		
Minimum	0.20	0.20	0.20	0.20	0.19	0.05
Maximum	4.70	4.70	4.70	36.15	36.15	36.15
Average	1.35	1.08	1.03	3.29	3.20	3.19
Median	0.74	0.50	0.50	2.00	1.93	1.93
Standard Dev.	1.20	1.10	1.13	4.49	4.44	4.45
Coeff. of Var.	0.89	1.02	1.09	1.37	1.39	1.39

The lognormal probability plot follows a straight line, except for the lower tail up to the 5th percentile (Figure D18). The effect on the Anderson Darling statistic is increased when the censored data is estimated. The effect is higher when the non-detected values are replaced by half of the detection limit, instead of being estimated using Cohen's maximum likelihood estimator. In open space areas when the level of censoring is elevated and the number of observations is low, the Cohen's estimated method did not follow a lognormal distribution. In Figure A19, two groups seem to exist, but it is important to mention that more than 44% of the total TKN observations were lower than 0.5 mg/L. All the censored values in this land use were located at 0.5 mg/L TKN.

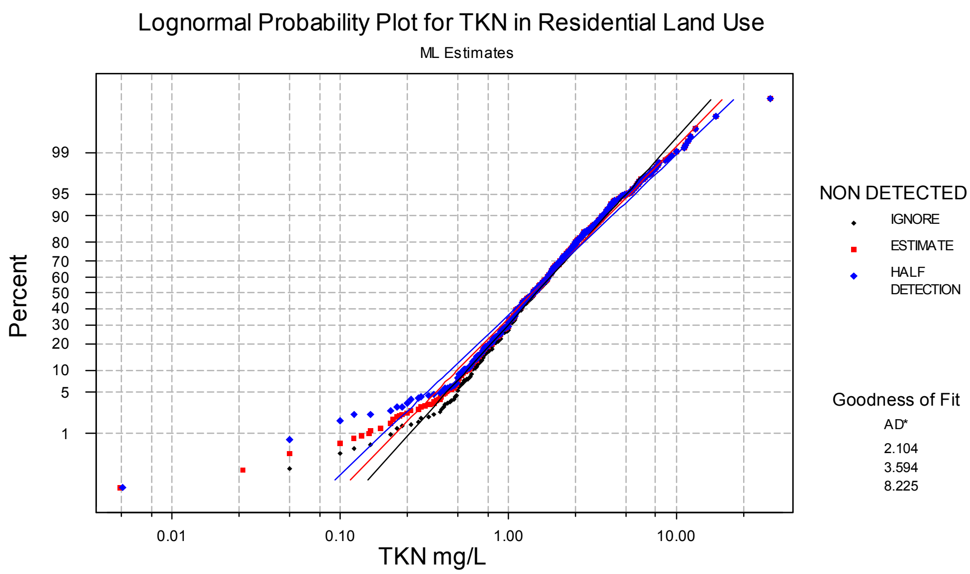


Figure D18. Estimated TKN distributions in residential land use areas

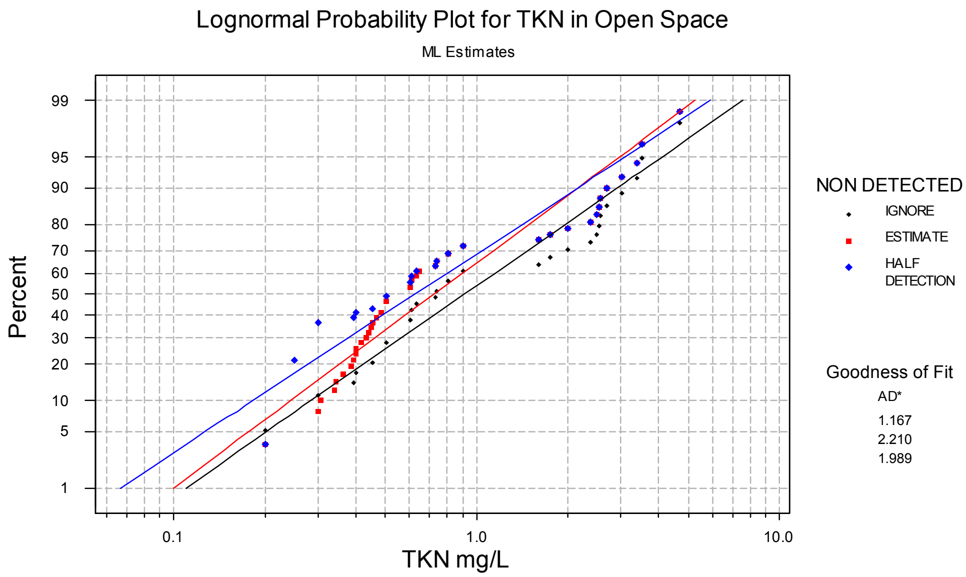


Figure D19. Estimated TKN distributions in open space land use areas

Total Phosphorus

Total phosphorus has low level of censored observations (less than 5%) at all land use areas, except for open space (where it is close to 15%) (Table D10). Variations in the coefficient of variation were not significant, except in open space areas where ignoring the censored observations reduces the coefficient of variation by almost 7%.

Table D10. Summary Statistics for Estimated Observations for Total Phosphorus (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	963	963	963	446	446	446	434	434	434
% Detected	96.88			95.74			95.85		
Minimum	0.01	0.00	0.01	0.02	0.01	0.01	0.02	0.01	0.01
Maximum	6.90	6.90	6.90	3.35	3.35	3.35	7.90	7.90	7.90
Average	0.42	0.41	0.41	0.35	0.34	0.34	0.46	0.45	0.45
Median	0.30	0.30	0.30	0.22	0.22	0.22	0.26	0.25	0.25
Standard Dev.	0.47	0.47	0.47	0.40	0.39	0.39	0.64	0.63	0.63
Coeff. of Var.	1.13	1.14	1.14	1.16	1.16	1.16	1.39	1.41	1.40

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	46	46	46	128	128	128
% Detected	84.78			99.22		
Minimum	0.02	0.01	0.01	0.06	0.05	0.02
Maximum	15.40	15.40	15.40	7.19	7.19	7.19
Average	0.68	0.59	0.60	0.43	0.43	0.43
Median	0.31	0.22	0.25	0.25	0.25	0.25
Standard Dev.	2.43	2.24	2.24	0.76	0.76	0.76
Coeff. of Var.	3.54	3.77	3.74	1.76	1.77	1.77

When the censored data is ignored, the observations followed a lognormal distribution. However, if the non-detected values are replaced by half of the detection limit or estimated using the Cohen method, the lower tail has lower values than expected.

There is an unusual observation 20 times higher than the second highest observation for the open space data (Figure D20). The most frequent non-detected observation was <0.5 mg/L. Replacing the censored observations by half of the detection limit produces values smaller than those estimated by Cohen's method. In the freeway plot, it was observed that the higher observations are higher than the lognormal trend. The upper 20th percentile has a different slope than the remaining observations shown on the distribution.

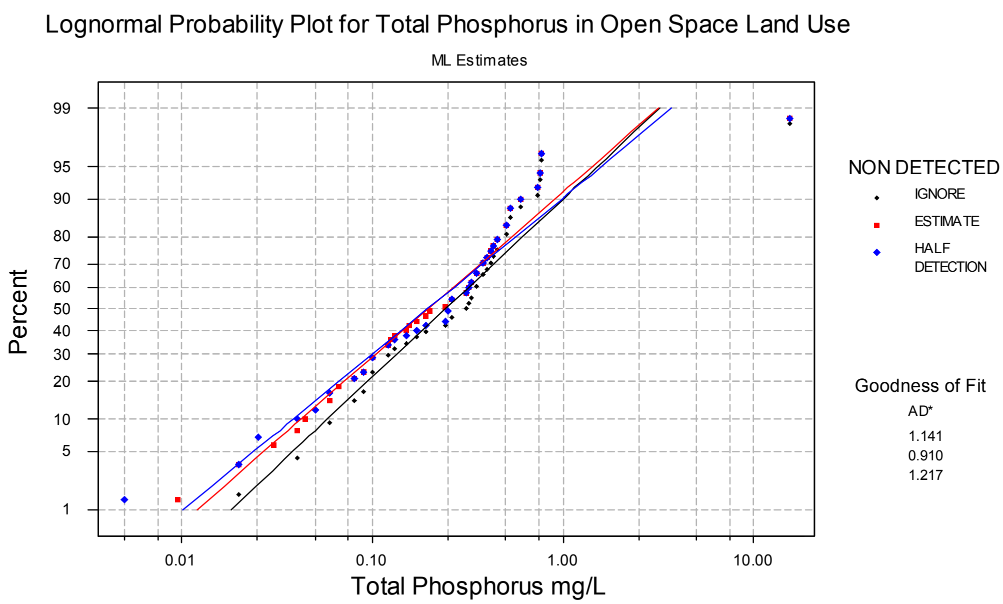


Figure D20. Estimated total phosphorus distributions in open space land use areas

Dissolved Phosphorus

Dissolved phosphorus has a large amount of non-detected values in all the land use areas (about 13 to 20%), except for freeways where only 5% of the observations were censored. In general, ignoring the non-detected values increased the means and standard deviations and reduced the coefficients of variation. Table D11 shows the descriptive statistics for dissolved phosphorus.

Table D11. Summary Statistics for Estimated Observations Dissolved Phosphorus (mg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	738	738	738	323	323	323	325	325	325
% Detected	84.15			81.11			87.38		
Minimum	0.009	0.001	0.005	0.01	0.00	0.01	0.003	0.003	0.003
Maximum	1.69	1.69	1.69	1.60	1.60	1.60	1.60	1.60	1.60
Average	0.23	0.20	0.20	0.21	0.18	0.19	0.17	0.16	0.16
Median	0.17	0.14	0.14	0.11	0.09	0.09	0.11	0.10	0.10
Standard Dev.	0.21	0.21	0.21	0.27	0.25	0.25	0.20	0.19	0.19
Coeff. of Var.	0.94	1.04	1.05	1.24	1.35	1.34	1.18	1.23	1.23

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	44	44	44	22	22	22
% Detected	79.55			95.45		
Minimum	0.010	0.003	0.005	0.06	0.01	0.01
Maximum	0.52	0.52	0.52	6.97	6.97	6.97
Average	0.18	0.16	0.17	0.78	0.75	0.75
Median	0.13	0.09	0.14	0.20	0.20	0.20
Standard Dev.	0.16	0.15	0.15	1.66	1.63	1.63
Coeff. of Var.	0.89	0.95	0.87	2.13	2.18	2.18

As in the previous cases, ignoring the censored observations results in larger mean values. There were no observed practical differences between the maximum likelihood method and replacing the non-detected values with half of the detection limit (Figure D21). , Dissolved phosphorus had the lowest level of censoring at freeway sites. The probability plot indicates that the distribution is heavy in the tails; the slope between the 20th and 60th percentiles is higher than in the tails (Figure D22).

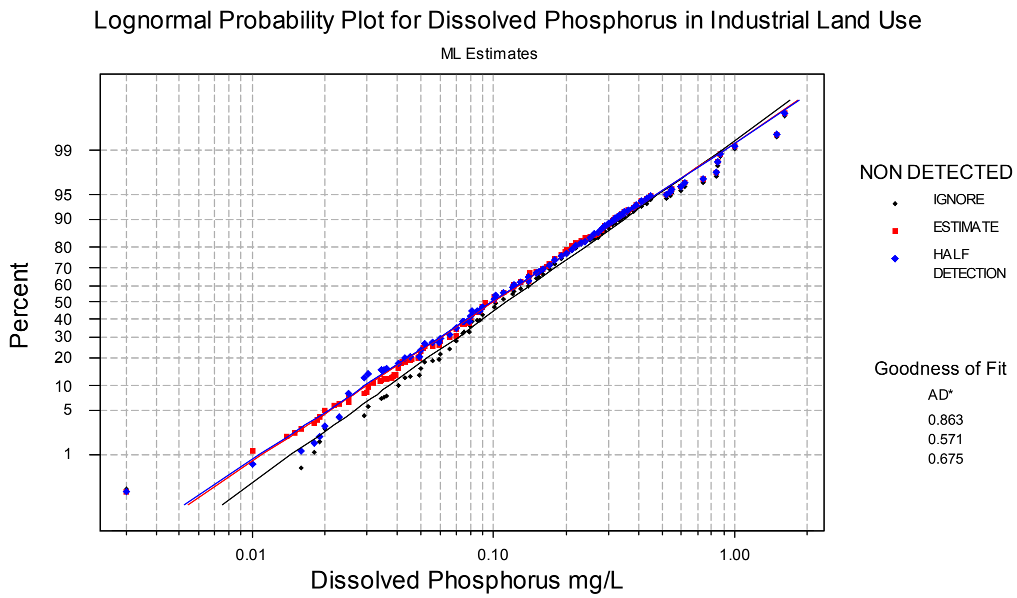


Figure D21. Estimated dissolved phosphorus distributions in industrial land use areas

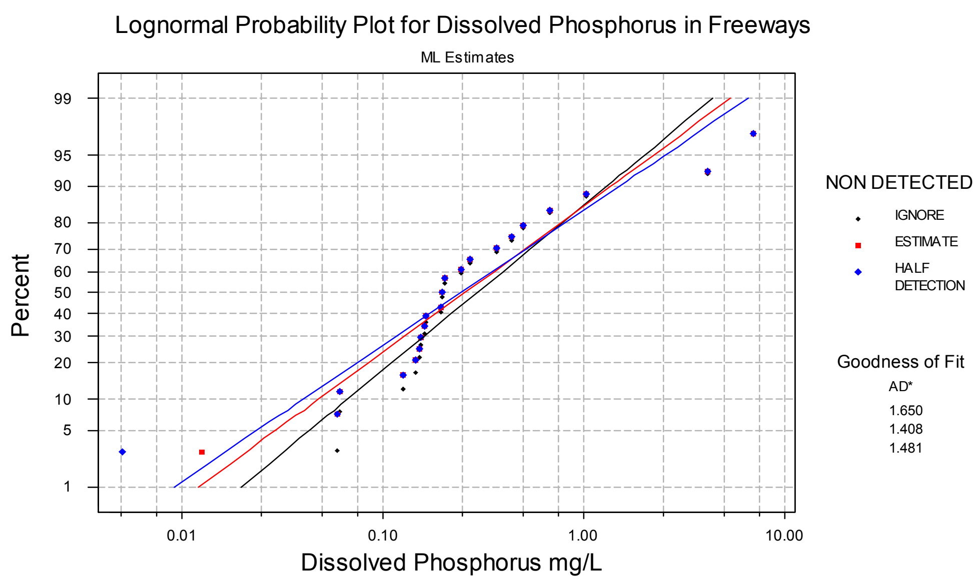


Figure D22. Estimated dissolved phosphorus distributions in freeways land use areas

Total Cooper (Cu)

The levels of censoring for copper vary from 1 to 15% among the different land uses. When the non-detected values are estimated or replaced by half of the detection limit, the coefficients of variation increased between 1% and 6%, in addition there is a reduction in the means and standard deviations. Table D12 shows the descriptive statistics for each method by land use.

Table D12. Summary Statistics for Estimated Observations for Total Cooper (μg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	799	799	799	387	387	387	415	415	415
% Detected	83.60			92.76			89.64		
Minimum	1.00	0.25	0.23	1.50	1.50	1.00	1.97	1.77	1.00
Maximum	590	590	590	384	384	384	1360	1360	1360
Average	21.06	18.54	18.51	29.02	27.47	27.30	47.00	43.37	42.98
Median	12.00	10.00	10.00	17.00	15.60	15.00	21.88	20.00	20.00
Standard Dev.	38.51	35.70	35.69	42.92	41.73	41.79	93.81	89.47	89.60
Coeff. of Var.	1.83	1.93	1.93	1.48	1.52	1.53	2.00	2.06	2.08

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	39	39	39	97	97	97
% Detected	74.36			98.97		
Minimum	2.00	2.00	2.00	5.00	5.00	5.00
Maximum	210	210	210	244	244	244
Average	19.15	15.79	15.65	48.29	47.86	47.85
Median	10.00	5.30	5.00	34.70	33.40	33.40
Standard Dev.	38.97	33.98	34.00	45.91	45.87	45.89
Coeff. of Var.	2.04	2.15	2.17	0.95	0.96	0.96

The lognormal probability plots for residential and commercial land use areas indicate that the upper 5th percentile of the copper concentrations have higher values than expected if the distribution was lognormal. This observation is important because the upper tail of the distribution has an important effect in the mean and standard deviation values of the dataset.

In open space areas, replacing the non-detected values by the Cohen's method or replacing the non-detected values by half of the detection limit, reduce the means and standard deviations of the distribution by 18% and 13%, respectively. The probability plot for freeway areas is almost a perfect lognormal trend. In this case, the level of non-detected values was only 1%, and the difference in the coefficients of variations was also 1%.

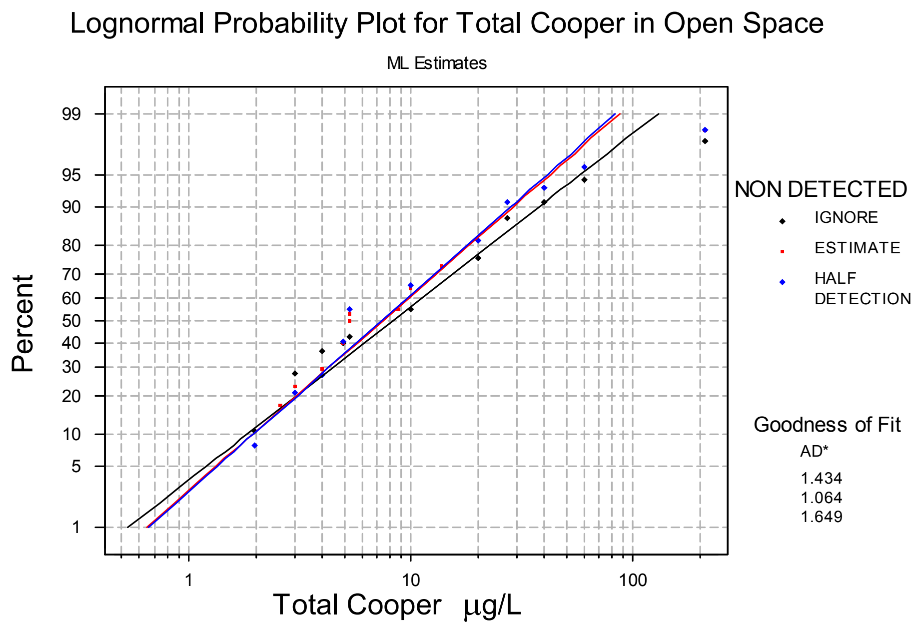


Figure D23. Estimated total cooper distributions in open space land use areas

Total Lead

The level of non-detected values for lead varied from 0 to 58%. All the observations at the freeway sites indicate a presence of lead, in addition to the highest concentration among the land uses. Open land use areas had the highest level of non-detected lead values. There was about a 10% reduction in the coefficient of variation when the censored data were ignored. Table D13 shows the descriptive statistics for each method.

The probability plots indicate that when replacing the censored data by half of the detection limit, the values are smaller than when using Cohen's method (Figure D24). Estimating the censored values reduces the Anderson Darling statistic. In open space areas, most of the censored values were observed at < 40 mg/L, < 50 mg/L and < 100 mg/L. In all land use areas, almost 80% of the lead observations were smaller than 50 mg/L. In open space areas, the estimated means, standard deviations and coefficients of variation are dubious because most of the censored observations were located in the upper part of the distribution (the frequency of non-detectable observations was quite high, at about 58%).

Table D13. Summary Statistics for Estimated Observations for Total Lead (µg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	788	723	788	377	355	377	411	377	411
% Detected	71.32			85.41			76.40		
Minimum	0.50	0.03	0.10	1.00	0.21	0.35	1.00	0.21	0.50
Maximum	585	585	585	689.07	689.07	689.07	1200	1200	1200
Average	26.00	21.03	22.08	37.42	34.27	33.84	70.10	59.52	57.49
Median	12.00	8.20	10.00	18.00	17.00	17.00	25.00	20.00	20.00
Standard Dev.	48.98	44.21	43.17	59.53	57.56	56.07	128.57	119.79	115.57
Coeff. of Var.	1.88	2.10	1.96	1.59	1.68	1.66	1.83	2.01	2.01

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	45	29	45	107	107	107
% Detected	42.22			100.00		
Minimum	0.20	0.08	0.10	1.60	1.60	1.60
Maximum	150	150	150	450	450	450
Average	28.39	19.21	23.98	48.77	48.77	48.77
Median	10.00	3.16	10.00	25.00	25.00	25.00
Standard Dev.	47.36	40.10	33.70	70.74	70.74	70.74
Coeff. of Var.	1.67	2.09	1.41	1.45	1.45	1.45

Total Zinc

The percentage of non-detected zinc values was smaller than 4%, except for open space areas where it was close to 30% (Table D14). No important changes in the coefficient of variations were observed, except for open space areas where ignoring the censored values reduced the coefficients of variation by 13%.

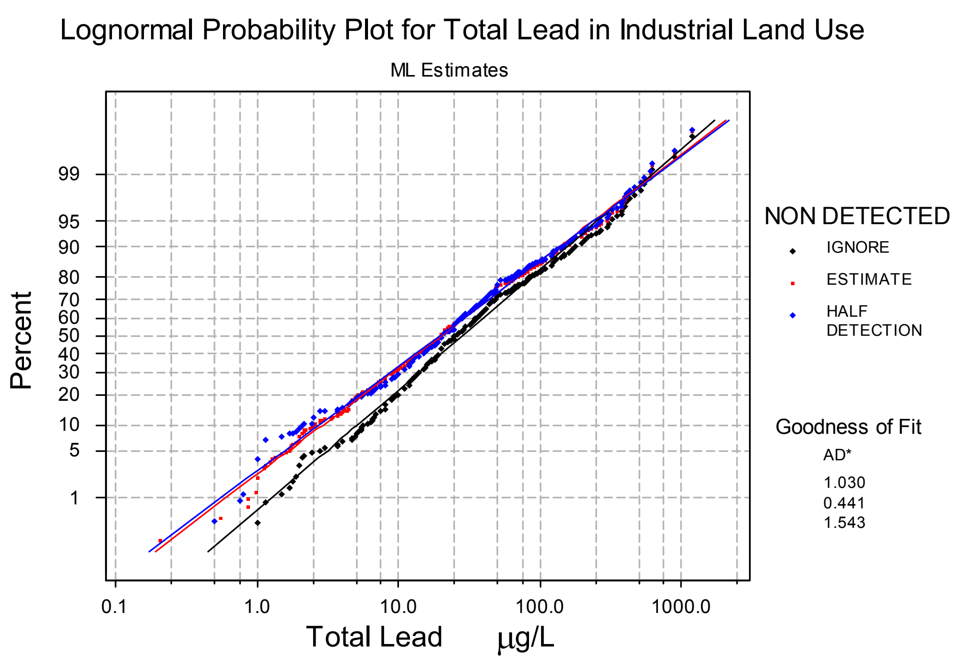


Figure D24. Estimated total lead distributions in industrial land use areas

Table D14. Summary Statistics for Estimated Observations for Total Zinc (μg/L)

	RESIDENTIAL			COMMERCIAL			INDUSTRIAL		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	810	810	810	392	392	392	432	432	432
% Detected	96.42			98.98			98.61		
Minimum	3.00	0.48	0.30	5.00	5.00	5.00	5.77	3.05	2.00
Maximum	1580	1580	1580	3050	3050	3050	8100	8100	8100
Average	116.70	113.53	113.23	225.32	224.06	223.55	318.25	315.02	314.34
Median	73.00	70.00	70.00	150.00	150.00	150.00	209.50	204.50	201.00
Standard Dev.	151.81	150.25	150.24	275.81	274.74	274.96	474.36	471.89	472.21
Coeff. of Var.	1.30	1.32	1.33	1.22	1.23	1.23	1.49	1.50	1.50

	OPEN SPACE			FREEWAY		
Land use	**Ignore**	**Estimate**	**HD**	**Ignore**	**Estimate**	**HD**
Observations	45	45	45	93	93	93
% Detected	71.11			96.77		
Minimum	5.00	2.00	2.50	6.00	6.00	2.50
Maximum	390	390	390	1829	1829	1829
Average	72.44	55.90	55.62	279.43	271.63	271.52
Median	40.00	20.00	20.00	200.00	194.49	194.49
Standard Dev.	96.88	85.85	85.99	281.16	279.87	279.98
Coeff. of Var.	1.34	1.54	1.55	1.01	1.03	1.03

The probability plot indicates that in the lower tail, replacing the non-detected observations by half of the detection limit will create smaller values than when estimating them using Cohen's method (Figure D25). In open space areas, if the censored data are estimated using Cohen's method, there is a reduction in the mean and variance of the dataset of 23% and 12%, respectively, however the coefficients of variation increased by 15% (Figure D26).

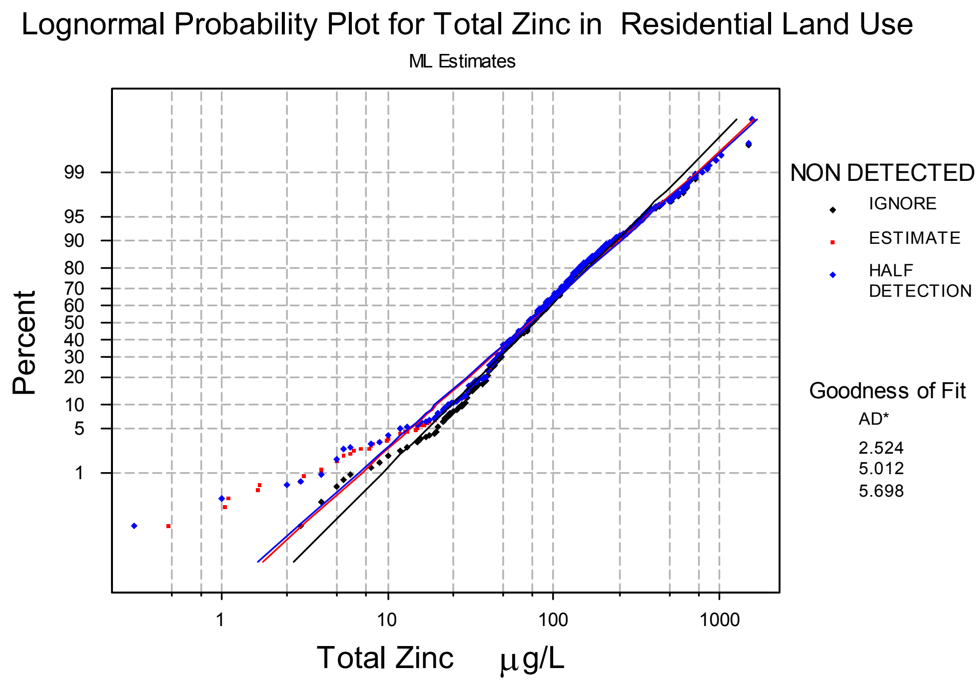


Figure D25. Estimated total zinc distributions in residential land use areas

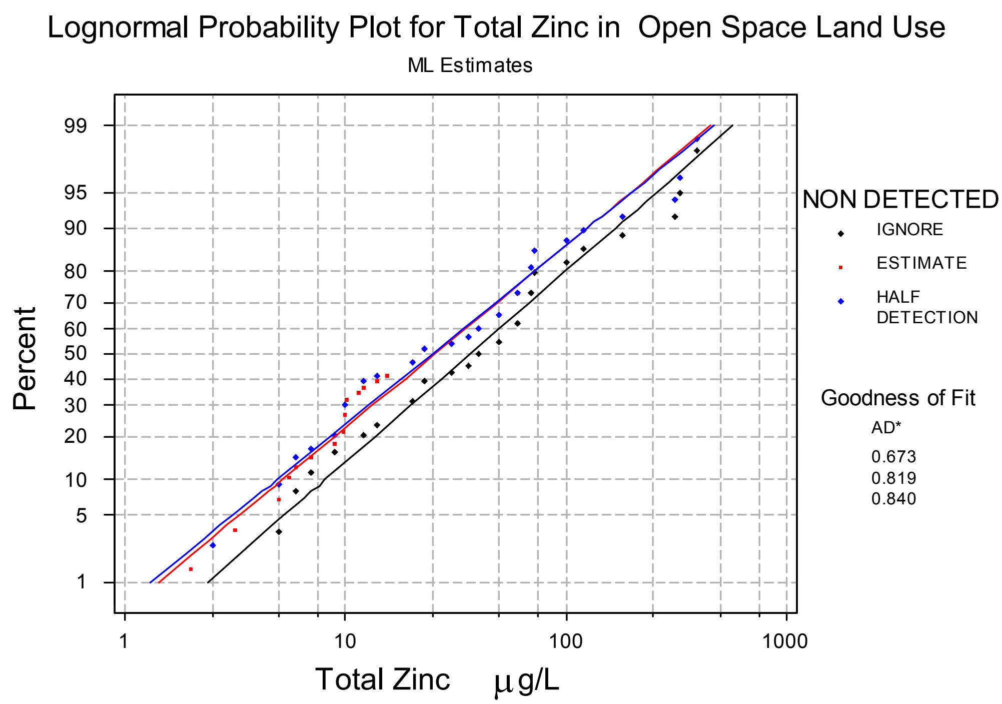


Figure D26. Estimated total zinc distributions in open space land use areas

Sites with Unusual TSS Concentrations for Different Land Uses

This section presents the continuation of the example presented in Chapter 3, where sites having unusual conditions were identified and examined more carefully to try to understand the reasons for these values. A similar procedure was followed in this appendix for the commercial, industrial and mixed land use areas to complement the Chapter 3 analyses, which were conducted for residential areas only.

Residential and Mixed Residential Locations

The box and whisker plot (Figure D27 shows TSS concentrations by rain zone and location) indicates that there is only one site that seems to have a different TSS concentration probability distribution compared to the remaining sites in this group. The site of interest is located in a residential-commercial area in Wooden Bridge Run, Philadelphia (PAPH1051), and has much lower concentrations that the other sites. Only two samples were collected at this site, and both were below 15 mg/L. The few samples available reduce the significance of this observation, however.

The results from the Xbar S chart analyses for mixed residential land uses are presented in Table D15. These analyses consider the numbers of samples and the variability of the data from each site, compared to the complete data set in the category being examined.

Table D15. Sites failing Xbar and S Chart Tests in Mixed Residential Land Use Areas

EPA Rain Zone	Sites Failing Xbar Chart Test	Sites Failing S Chart Test
ALL	9COCSA004(H) 2NCFVROSE(L) 7ORPOA005(H) 2TNKXTYGV(H) 5TXFWA005(H) 2VAVBTYV5(L)	GAFUCOS3(H)
1	None	None
2	TNKXTYGV(H) VAVBTYV5(L)	None
3	None	None
4	None	None
5	None	None
6	None	None
7	None	None
8	None	None
9	None	None

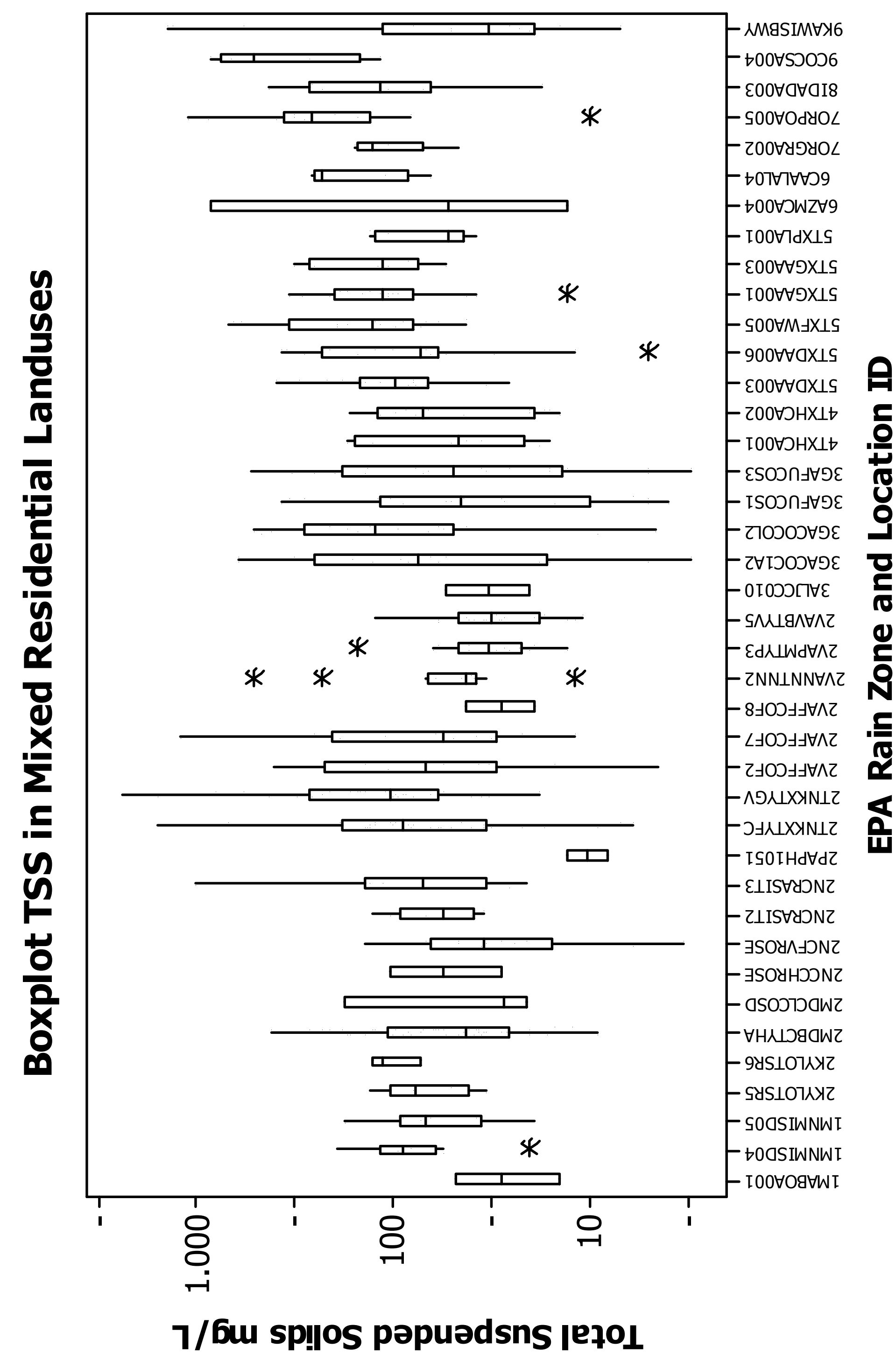


Figure D27. Box and whiskers plots for TSS concentrations at mixed residential land use areas

When 40 mixed residential sites were examined, only six sites were designated as being "out of control." These sites had unusual concentrations that were outside a band described by three standard deviations from the mean values. Two sites with means (log) below the lower control limit were located in EPA Rain Zone 2 (North Carolina and Virginia). The site located at Long Island Creek in Fulton County, Georgia, has the largest standard deviation among the mixed residential sites examined. However, the S chart indicates that this site is in control compared to other sites in EPA Rain Zone 3.

There are 21 sites located in mixed residential land use areas and in EPA Rain Zone 2, 17 sites have more than one observation each. Two sites, one above the upper control limit and one below the control limit, were observed in the Xbar chart. The site with the high median (log) value is located in Gallaher view, Knoxville, Tennessee (TNKXTYGV, 38 observations, median TSS = 105 mg/L). This site information included construction activity in the north part of the watershed, and a self-storage business, north and east of Cedar Hills apartments. The site located in Holland road, Virginia Beach (VAVBTYV5, 26 observations, median TSS = 32 mg/L) has wet ponds in the watershed that seem to control high concentrations, but the average value is the same as the other mixed residential sites.

The ANOVA analyses indicate that there is at least one EPA Rain Zone with TSS concentrations different than the other EPA Rain Zone with a p-value smaller than 1%. The Dunnett's comparison test at a family error rate of 5% indicates that EPA Rain Zones 5 and 7 have higher concentrations than those observed in EPA Rain Zone 2.
In summary, at a family error rate of 5%, higher concentrations occurred in EPA Rain Zones 5 (six sites, median TSS = 108 mg/L) and 7 (two sites, TSS = 175 mg/L) compared with EPA Rain Zone 2 (21 sites, TSS = 59 mg/L). The Kurskal-Wallis test indicates that there is a significant difference in the TSS median concentrations (with a p-value close to zero). Site TNKXTYGV has higher characteristics than the other residential mixed sites, most likely due to the noted construction activity close to the outfall location.

Commercial and Mixed Commercial Locations

Box plots Xbar and S charts and ANOVA tests were used for commercial land use data. Figure D28 identifies a site with high TSS concentration in EPA Rain Zone 4 (KATOJACK, 15 observations, median TSS = 603 mg/L). In general, it seems that sites in EPA Rain Zone 7 and 9 have higher concentrations than the other EPA Rain Zones. No other trend or variation among EPA Rain Zone was identified from the box plot.

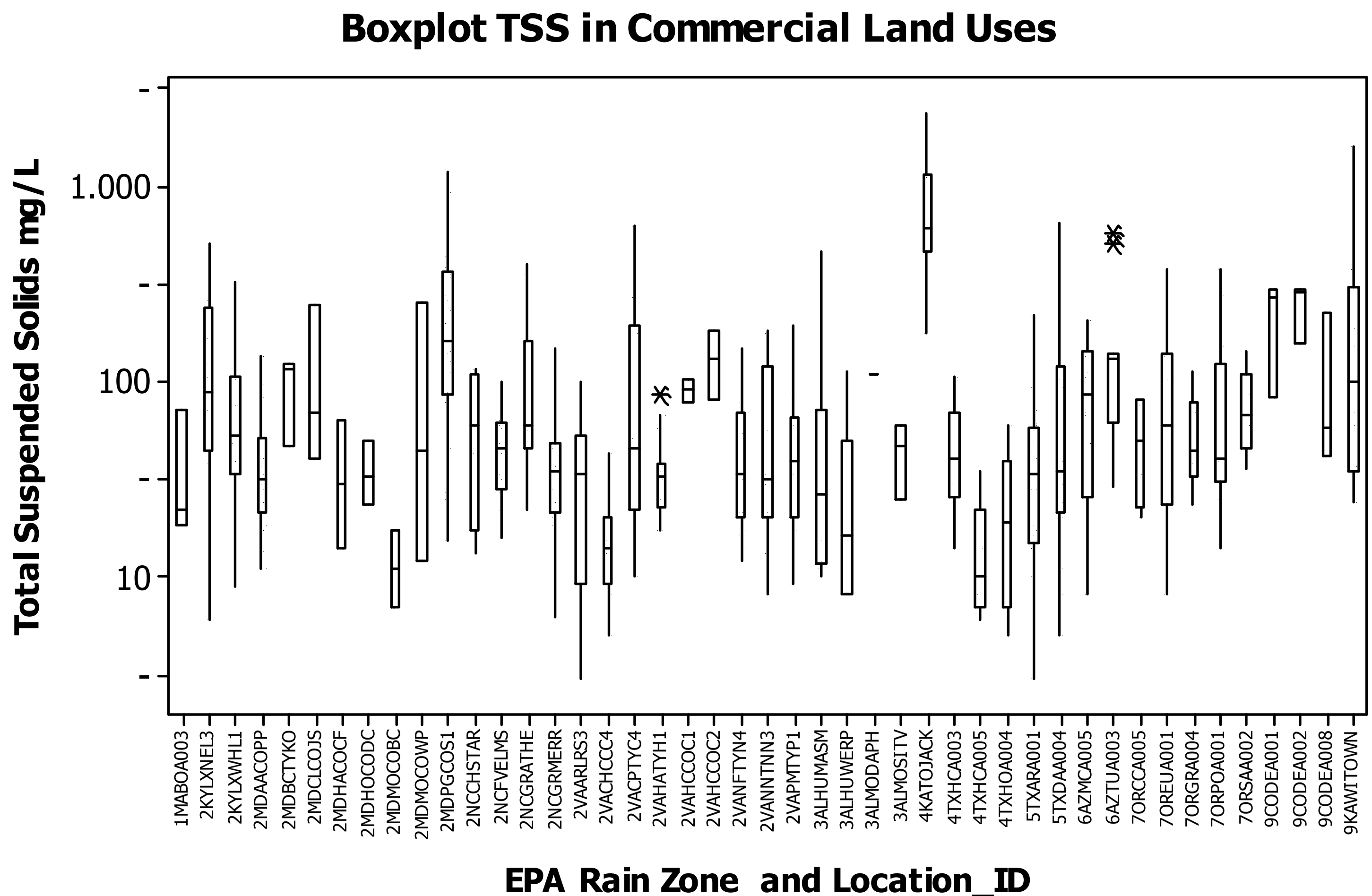


Figure D28. Box and whiskers plots for TSS concentrations at commercial land use areas

The second approach was to identify unusual sites by EPA Rain Zone using Xbar S charts. The results for commercial land uses are presented in Table D16.

Table D16. Sites failing Xbar and S Chart Tests in Commercial Land Use Areas

EPA Rain Zone	Sites Failing Xbar Chart Test	Sites Failing S Chart Test
ALL	2MDPGCOS1(H) 2VACHCCC4(L) 3ALHUWERP(L) 4KATOJACK(H) 4TXHCA005(L) 4TXHOA004(L) 9KAWITOWN(H)	None
1	None	None
2	2MDPGCOS1(H) 2VACHCCC4(L)	None
3	None	None
4	4KATOJACK(H)	None
5	None	None
6	None	None
7	None	None
9	None	None

The Xbar S plot did not indicate any trend by geographical region for the 45 sites. Sites with low concentrations were observed in EPA Rain Zones 2, 3 and 4. There were three sites identified with concentrations above the control limit, one in EPA Rain Zone 2, another in EPA Rain Zone 9, and a site identified by the box plot located in EPA Rain Zone 4.

In EPA Rain Zone 2, two sites were found outside the control limits. MDPGCOS1 is located in a shopping center in Arena Plaza, Price Georges County, Maryland. 26 samples were collected at this location. The median TSS concentration for this site is 158 mg/L. No reason was given for the high observed TSS concentrations. The second site is located at Clover Leaf Mall in Chesterfield County, Virginia (VACHCCC4, 12 observations, 60 acres, median TSS = 14 mg/L). There is no clear reason that explains the low concentrations found at this location. No sites outside the control limits were found in other EPA Rain Zones except for EPA Rain Zone 4. This outfall is located in Jackson Street in Topeka, Kansas. The high TSS concentrations may have been affected by tracking of sediment from a sand quarry close to the watershed. There were collected 16 samples collected between April 1998 and Septembers 2002.

The ANOVA test indicated that there was a significant difference among EPA Rain Zones (P-value = 0). The Dunnett's comparison test, with a family error of 5%, indicates that TSS concentrations compared with EPA Rain Zone 2 (median TSS = 48 mg/L) are larger in EPA Rain Zones 4 (median TSS= 82 mg/L) and 9 (median TSS = 128 mg/L). The median TSS concentrations at the remaining EPA Rain Zones are not statistically different than those observed in EPA Rain Zone 2.

There are 24 sites located in mixed commercial land use areas with more than one observation. EPA Rain Zone 2 has the largest number of sites (10 sites), followed by EPA Rain Zone 5 (5 sites). Figure D29 shows the box plots for mixed commercial land uses by EPA Rain Zone.

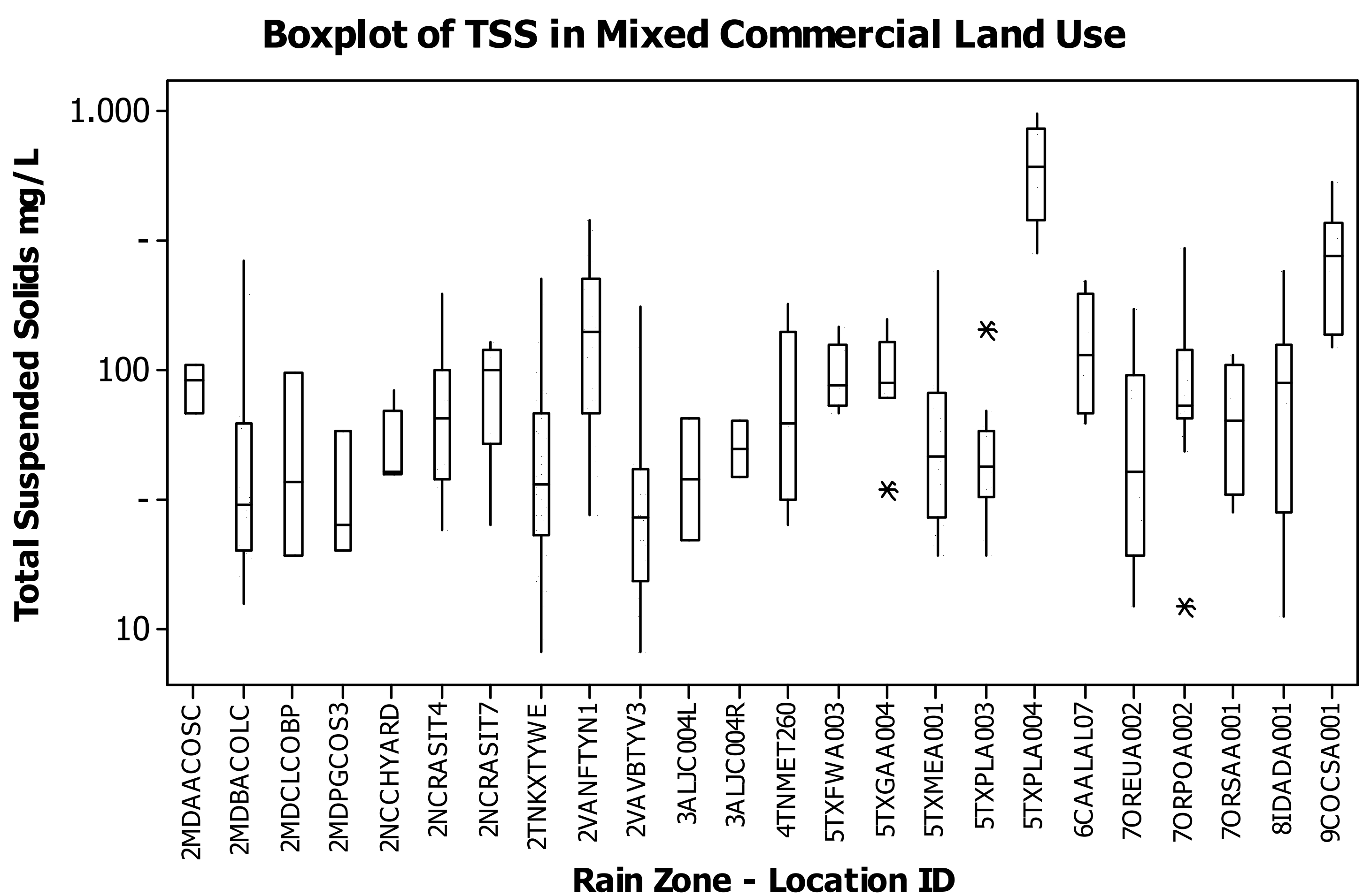


Figure D29. Box and whiskers plots for TSS concentrations at mixed commercial land use areas

The box plot indicates that there is a mixed commercial site located in Plano, Texas, and a site in Colorado with higher concentrations than the other sites in this category. Because the low number of sites sampled by geographical region, it is not possible to identify any trend by EPA Rain Zone. Table D17 lists those sites outside the control limits by EPA Rain Zone for all of the mixed commercial sites.

Table D17. Sites failing Xbar and S Chart Tests in Mixed Commercial Land Use Areas

EPA Rain Zone	Sites Failing Xbar Chart Test	Sites Failing S Chart Test
ALL	2TNKXTYWE(L) 2VANFTYN1(H) 2VAVBTYV3(L) 5TXPLA004(H) 9COCSA001(H)	None
1	None	None
2	2VANFTYN1(H) 2VAVBTYV3(L)	None
3	None	None
4	None	None
5	5TXPLA004(H)	None
6	None	None
7	None	None
9	None	None

The Xbar chart for all mixed commercial observations indicates that sites with high TSS concentrations occurred in EPA Rain Zones 5 and 9. In EPA Rain Zone 2, three sites were outside of the control limits, two below the lower control limit and one above the upper control limit. As in the commercial site analyses, EPA Rain Zone 9 seems to have higher TSS concentrations than the other EPA Rain Zones.

The analysis by EPA Rain Zone indicates that only EPA Rain Zones 2 and 5 have sites outside the control limits. In EPA Rain Zone 2, the site with high concentrations (VANFTYN1) is located at Armistead Avenue in Norfolk, Virginia. A total of 28 observations were collected at this site. The median TSS for this location was 117 mg/L. The site having unusually low median TSS concentration was at Haygood, Virginia Beach, Virginia (VAVBTYV3). A total of 33 storms were sampled at this site. The median TSS concentration at this location was 26 mg/L. This site is 79% commercial and 13% open space. The site having unusually high TSS concentrations in EPA Rain Zone 5 is located at Spring Creek, Plano, Texas (TXPLA004). There are 7 events from this site in the database. The median TSS concentration is 575 mg/L. No information was found to explain the elevated concentrations. Another site that appears to be outside the control limits compared to all the sites, but not in its group. It is located in Sixteenth Hole Valley, Colorado Springs, Colorado. The median concentration for this site was 251 mg/L. This site has two automobile dealerships and a gas station, along with evidence of erosion observed in the aerial photograph.
The ANOVA analysis indicates that there are significant differences among EPA Rain Zones (P-value = 0) in mixed commercial land uses. The Dunnett's comparison test, with a family error of 5%, indicates that TSS concentrations compared with EPA Rain Zone 2 (median TSS = 46 mg/L) are larger in EPA Rain Zones 5 (median TSS = 72 mg/L) and 9 (median TSS = 254 mg/L). The median TSS values in the remaining EPA Rain Zones are not statistically different than those observed in EPA Rain Zone 2.

Industrial and Mixed Industrial Locations

Box plots, Xbar, S charts, and ANOVA tests were used to examine the observations from sites located in industrial land use areas. Figure D30 shows the box plots by EPA Rain Zone and location. Sites located in EPA Rain Zones 6 and 9 seem to have higher concentrations than the remaining industrial sites. A site with two unusually low concentrations was located in Boston, Massachusetts.

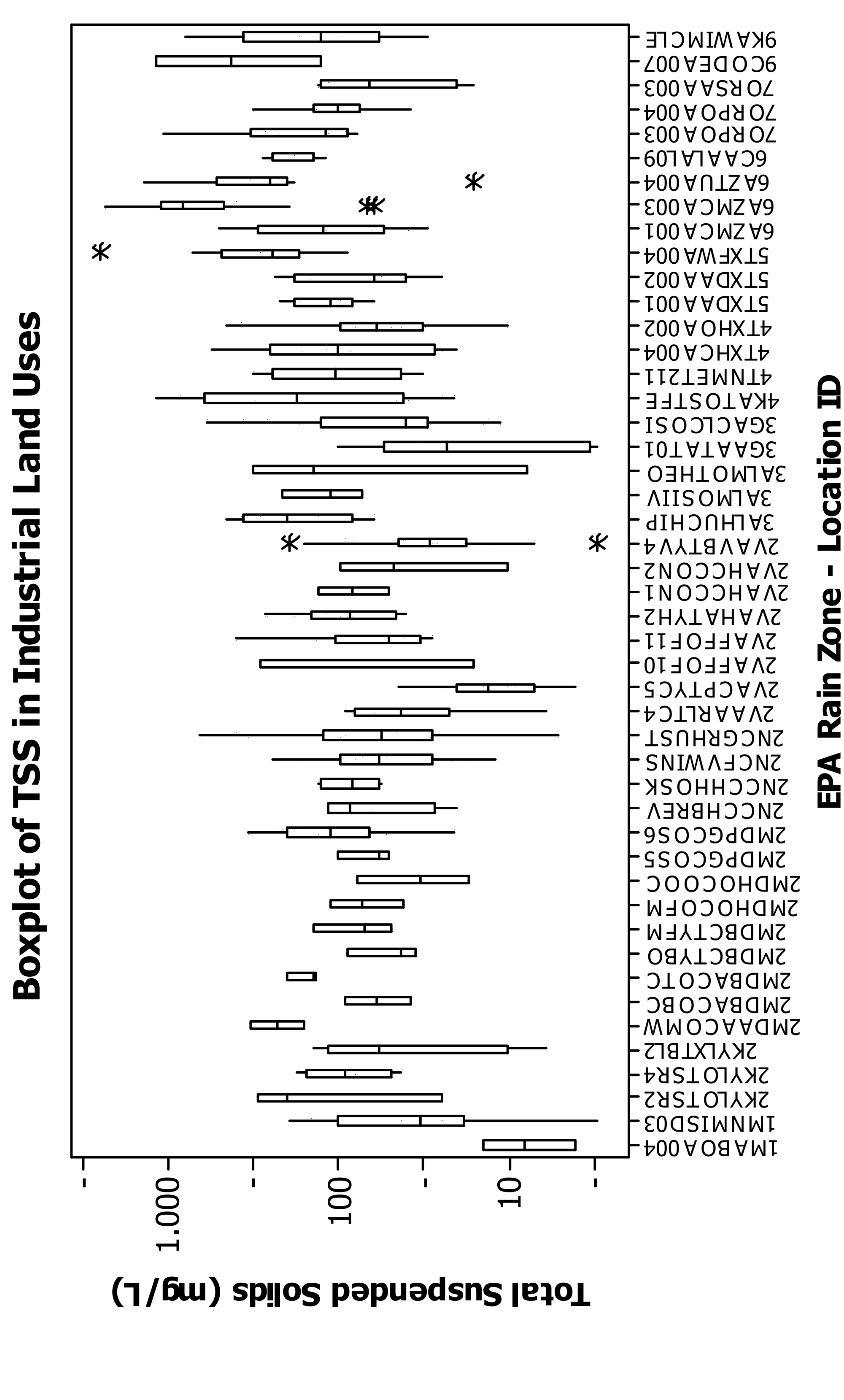


Figure D30. Box and whiskers plots for TSS concentrations at industrial land use areas

Table D18 shows those industrial sites that are outside the control limits of the pooled dataset and by each EPA Rain Zone.

Table D18. Sites failing Xbar and S Chart Tests in Industrial Land Use Areas

EPA Rain Zone	Sites Failing Xbar Chart Test	Sites Failing S Chart Test
ALL	1MABOA004(L) 2VACPTYC5(L) 2VAVBTYV4(L) 3GAATAT01(L) 5TXFWA004(H) 6AZMCA003(H) 6AZTUA004(H)	None
1	None	None
2	MDPGCOS6(H) VACPTYC5(L) VAVBTYV4(L)	None
3	None	None
4	None	None
5	TXFWA004(H)	None
6	AZMCA003(H)	None
7	None	None
9	None	None

As in the other land uses, sites with concentrations below the control limit were observed in EPA Rain Zones 1, 2 and 3. Sites with median concentrations larger than the upper control limit were located in EPA Rain Zones 5 and 6. Three sites were outside the control limits in EPA Rain Zone 2, one in EPA Rain Zone 5, and one in EPA Rain Zone 6. The two sites in EPA Rain Zone 2 with low concentrations were located in Virginia, and the site with high concentrations was located in Maryland. One of the sites located in Virginia was located in Cavalier Industrial Park in the city of Chesapeake (VACPTYC5). This 16 acres site is 92% industrial, with the remaining 8% open space. A total of 15 samples were collected from this site during the period 1993 to 1999. The median TSS concentration for this site is 13 mg/L. No additional information was observed in the aerial photos that might explain the low concentrations.

The second site was located in Viking Drive, Virginia Beach (VAVBTYV5). This 29-acre site was comprised of 55 percent impervious surfaces. There are 30 samples from this site in the database. The samples were collected between 1992 and 1999. The median TSS concentration is 29 mg/L.

The site with elevated concentrations in EPA Rain Zone 2 is located in Pennsy Drive in Riverdale, Prince George County, Maryland (MDPGCOS6). This 42.4-acre size site has a grass swale drainage system. There are 30 samples in the database from this location. The samples were collected between 1994 and 1997. The median TSS concentration is 98 mg/L. The site is located next to Glenridge Elementary School. The aerial photo shows construction activity in the northwest part of the watershed.

The site with high TSS concentrations in EPA Rain Zone 5 is located at Dry Branch, in Fort Worth, Texas (TXFWA004). A total of 21 samples were obtained at this site. The median TSS for this location is 288 mg/L. Several bare ground open space areas were observed in the aerial photograph. The site located in EPA Rain Zone 6 is at 27th Avenue at Salt River in Maricopa County, Arizona (AZMCA003). There are 27 samples from this location. The median TSS concentration is 660 mg/L. The scarce vegetation and the type of soils may be the reason of this elevated median value.

The ANOVA analysis indicates that there are significant differences among EPA Rain Zones (P-value = 0) for industrial land uses. The Dunnett's comparison test with a family error of 5%, indicates that TSS concentrations compared with EPA Rain Zone 2 (median TSS = 53 mg/L) are larger for EPA Rain Zones 4 (median TSS = 92 mg/L), 5 (median TSS = 147 mg/L), 6 (median TSS = 288 mg/L), 7 (median TSS = 120 mg/L), and 9 (median TSS

= 170 mg/L). The median TSS concentrations in EPA Rain Zones 1 and 3 are not statistically different from those observed in EPA Rain Zone 2.

The box plots in mixed industrial land uses are shown in Figure D31. Most of the box plots have the same median except for those located in EPA Rain Zone 9. The sites that fail the quality control charts are shown in Table D19. Three sites are outside the control limits for mixed industrial land uses. Two sites in Colorado and one site in North Carolina are out of control. This result is similar to those observed in the other land uses. When each EPA Rain Zone was analyzed individually, no sites were found to be out of control.

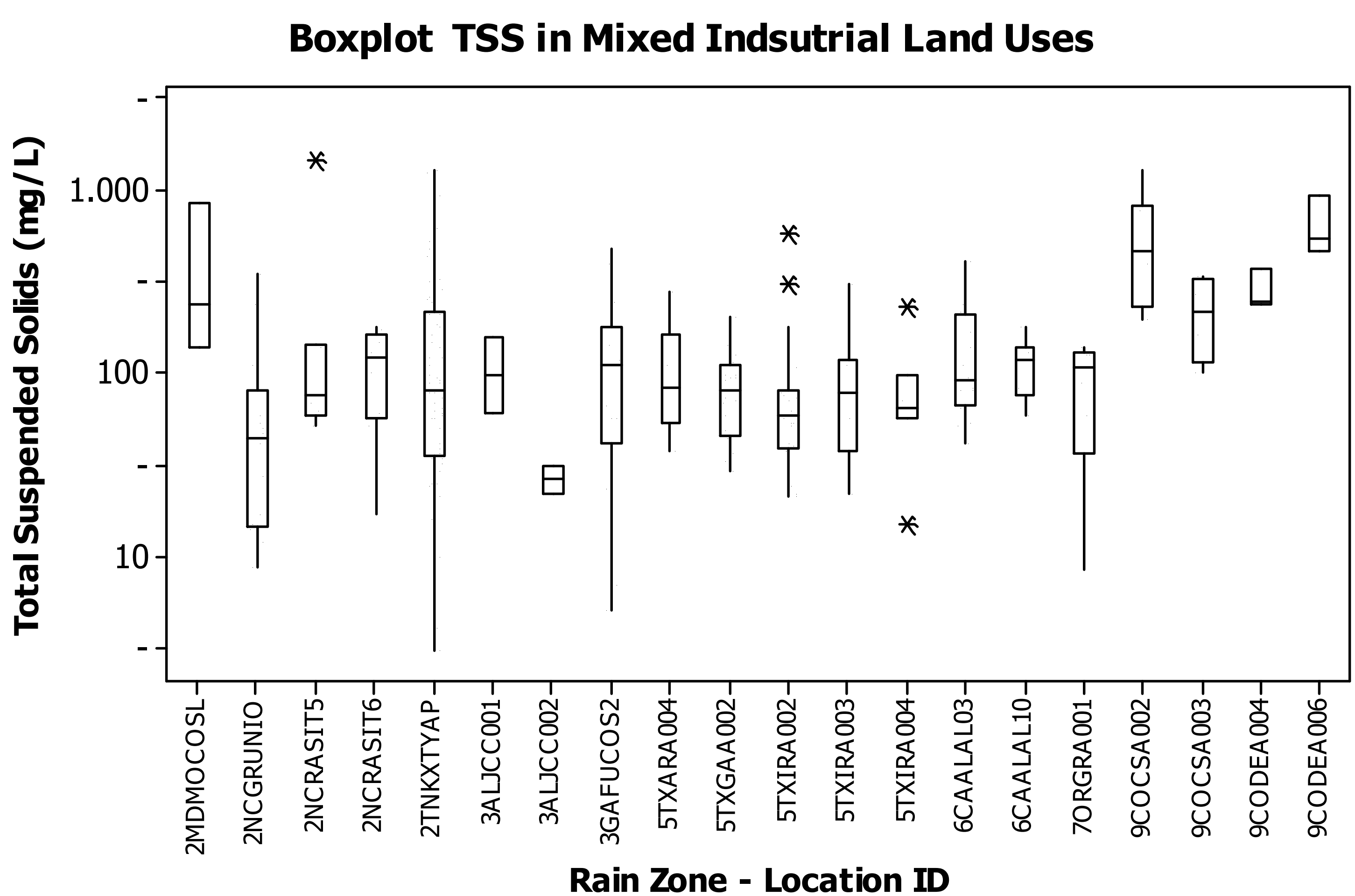


Figure D31. Box and whiskers plots for TSS concentrations at industrial land use areas

Table D19. Sites failing Xbar and S Chart Tests in Mixed Industrial Land Use Areas

EPA Rain Zone	Sites Failing Xbar Chart Test	Sites Failing S Chart Test
ALL	9COCSA002(H) 9CODEA006(H) 2NCGRUNIO(L)	None
2	None	None
3	None	None
5	None	None
6	None	None
7	None	None
9	None	None

The ANOVA analysis indicates that there are significant differences among EPA Rain Zones (P-value = 0) at mixed industrial land use sites. The Dunnett's comparison test with a family error of 5%, indicates that TSS concentrations compared with EPA Rain Zone 2 (median TSS = 82 mg/L) are larger only for EPA Rain Zone 9 (median TSS = 341 mg/L). The median TSS concentrations in EPA Rain Zones 3, 5, 6, and 7 are not statistically different from the median TSS concentrations found in EPA Rain Zone 2.

Appendix E: First Flush Tables

Description

The following table shows the summary statistic for each constituent included in the database.

Table E1. Results of Preliminary Statistical Analysis for Total Suspended Solids (TSS)

TSS (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	90	90	54.95	1.740	0.106	0.168	0.254	0.730	1.85
Commercial First Flush	90	90	101.86	2.008	0.200	-0.508	0.254	0.016	
Industrial Composite	83	83	66.07	1.820	0.186	-0.021	0.264	0.336	0.97
Industrial First Flush	83	83	63.97	1.806	0.374	-0.157	0.264	0.055	
Institutional Composite	18	18	16.48	1.217	0.110	-0.176	0.536	0.122	2.12
Institutional First Flush	18	18	34.99	1.544	0.145	-0.164	0.536	0.846	
Open Space Composite	32	32	21.98	1.342	0.424	-0.526	0.414	0.511	0.95
Open Space First Flush	32	32	20.89	1.320	0.563	-0.126	0.414	0.847	
Residential Composite	144	144	37.50	1.574	0.217	-0.033	0.202	0.282	1.84
Residential First Flush	144	144	69.02	1.839	0.302	-0.267	0.202	0.533	
All Land Uses Composite	372	372	44.36	1.647	0.226	-0.381	0.126	0.008	1.60
All Land Uses First Flush	372	372	70.96	1.851	0.335	0.457	0.126	0	

Table E2. Results of Preliminary Test Analysis for Total Suspended Solids (TSS)

TSS (mg/L	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 5.345 ; P = 0	0.014	N/A	Different (first flush)
Industrial	0.627	U = 0.483 ; P = 0.31	0.222	0.432	Same (no first flush)
Institutional	0.007	U = 3.095 ; P = 0	0.309	0.001	Different
Open Space	0.706	U = 0.39 ; P = 0.35	0.183	0.614	Same
Residential	N/A	U = 4.89 ; P = 0	0	N/A	Different
All Land Uses	N/A	U = 6.65 ; P = 0	0	N/A	Different

Table E3. Results of Preliminary Statistical Analysis for Turbidity (NTU)

Turbidity (NTU)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	11	11	19.68	1.294	0.062	-0.984	0.661	0.186	1.32
Commercial First Flush	11	11	26.00	1.415	0.078	0.523	0.661	0.564	
Residential Composite	12	12	23.44	1.370	0.163	0.213	0.637	0.721	1.24
Residential First Flush	12	12	28.97	1.462	0.148	1.407	0.637	0.168	
All Land Uses Composite	26	26	21.73	1.337	0.109	0.204	0.456	0.406	1.26
All Land Uses First Flush	26	26	27.48	1.439	0.105	1.197	0.456	0.108	

Table E4. Results of Preliminary Test Analysis for Turbidity (NTU)

Turbidity (NTU)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.224	U = 1.26 ; P > 0.1	0.652	0.219	Same (no first flush)
Residential	0.418	U = 0.853 ; P > 0.1	0.240	0.021	Same
All Land Uses	0.124	U = 0.673 ; P = 0.25	0.134	0	Same

Table E5. Results of Preliminary Statistical Analysis for pH

pH	Total Events	Selected Cases	Median	Var (Log)	Skew	SE Skew	Test Norm. p-value	Median Ratio
Commercial Composite	17	17	7.4	0.368	-0.299	0.550	0.527	1.03
Commercial First Flush	17	17	7.6	0.509	0.788	0.550	0.351	
Industrial Composite	16	16	6.755	0.194	0.482	0.564	0.179	1.00
Industrial First Flush	16	16	6.750	0.388	-0.854	0.564	0.307	
Residential Composite	26	26	7.213	0.195	-0.520	0.456	0.447	1.01
Residential First Flush	26	26	7.250	0.212	-0.283	0.456	0.408	
All Composite	63	63	7.2	0.302	0.102	0.302	0.562	1.01
All First Flush	63	63	7.3	0.437	0.036	0.302	0.110	

Table E6. Results of Preliminary Test Analysis for pH

pH	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.208	U = 1.28 ; P = 0.10	0.007	N/A	Same
Industrial	N/A	U = 0.428 ; P = 0.33	0.341	0.828	Same
Residential	0.308	U = 1.32 ; P = 0.09	0	N/A	Same
All Land Uses	0.219	U = 1.68 ; P = 0.05	0	N/A	Same

Table E7. Results of Preliminary Statistical Analysis for BOD_5

BOD_5 (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	83	83	15.21	1.182	0.125	0.263	0.264	0.513	1.77
Commercial First Flush	83	83	26.98	1.431	0.153	-0.241	0.264	0.390	
Industrial Composite	80	80	15.14	1.18	0.188	0.190	0.269	0.013	1.58
Industrial First Flush	80	80	23.99	1.38	0.180	-0.502	0.269	0.044	
Institutional Composite	18	18	7.48	0.874	0.151	-0.737	0.536	0.247	1.67
Institutional First Flush	18	18	12.47	1.096	0.173	-0.732	0.536	0.281	
Open Space Composite	28	28	3.79	0.579	0.148	0.523	0.441	0.242	1.07
Open Space First Flush	28	28	4.05	0.607	0.197	0.449	0.441	0.077	
Residential Composite	133	133	12.59	1.100	0.154	0.314	0.210	0.137	1.67
Residential First Flush	133	133	20.99	1.322	0.220	-0.150	0.210	0.010	
All Land uses Composite	344	344	12.53	1.098	0.184	0.073	0.131	0.003	1.67
All Land Uses First Flush	344	344	20.89	1.320	0.233	-0.385	0.131	0	

Table E8. Results of Preliminary Test Analysis for BOD_5

BOD_5	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0	U = 4.85 ; P = 0	0.013	N/A	Different (first flush)
Industrial	0.007	U = 2.76 ; P = 0	0.434	0.012	Different
Institutional	0.027	U = 2.46 ; P = 0.01	0.056	0.001	Different
Open Space	0.706	U = 0.39 ; P = 0.35	0.183	0.614	Same (no first flush)
Residential	N/A	U = 4.89; P = 0	0	N/A	Different
All Land Uses	N/A	U = 6.65 ; P = 0	0	N/A	Different

Table E9. Results of Preliminary Statistical Analysis for COD

COD (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite)	91	91	71.94	1.857	0.075	0.261	0.253	0.022	2.29
Commercial First Flush	91	91	164.82	2.217	0.119	-0.201	0.253	0.877	
Industrial Composite	84	84	75.34	1.877	0.100	0.167	0.263	0.014	1.43
Industrial First Flush	84	84	107.40	2.031	0.151	-0.141	0.263	0.804	
Institutional Composite	18	18	43.85	1.642	0.220	-0.456	0.536	0.567	2.73
Institutional First Flush	18	18	119.67	2.078	0.151	-0.969	0.536	0.105	
Open Space Composite	28	28	20.00	1.301	0.130	0.441	0.441	0.084	0.67
Open Space First Flush	28	28	13.43	1.128	0.211	0.731	0.441	0.013	
Residential Composite	140	140	67.92	1.832	0.095	0.271	0.205	0.008	1.63
Residential First Flush	140	140	110.41	2.043	0.138	-0.831	0.205	0.005	
All Land Uses Composite	363	363	65.92	1.819	0.123	-0.293	0.128	0	1.71
All Land Uses First Flush	363	363	112.98	2.053	0.194	-0.710	0.128	0	

Table E10. Results of Preliminary Test Analysis for COD

COD	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 4.83 ; P = 0	0.269	0	Different (first flush)
Industrial	N/A	U = 1.67 ; P = 0.05	0.691	0.01	Different
Institutional	0.01	U = 2.94 ; P = 0	0.677	0	Different
Open Space	N/A	U = 0.269 ; P = 0.39	0.004	N/A	Same (no first flush)
Residential	N/A	U = 6.715 ; P = 0	0	N/A	Different
All Land Uses	N/A	U =9.19 ; P = 0	0	N/A	Different

Table E11. Results of Preliminary Statistical Analysis for Total Dissolved Solids (TDS)

TDS (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	82	82	73.28	1.865	0.064	-0.338	0.266	0.263	1.83
Commercial First Flush	82	82	133.97	2.127	0.065	-0.219	0.266	0.115	
Industrial Composite	82	81	97.72	1.990	0.093	-0.482	0.267	0.341	1.32
Industrial First Flush	82	81	128.82	2.110	0.126	-0.513	0.267	0.109	
Institutional Composite	18	18	52.48	1.720	0.068	-0.034	0.536	0.360	2.66
Institutional First Flush	18	18	139.64	2.145	0.090	-0.303	0.536	0.158	
Open Space Composite	31	30	69.98	1.845	0.051	0.617	0.427	0.376	1.07
Open Space First Flush	31	30	74.99	1.875	0.104	-1.483	0.427	0.005	
Residential Composite	137	133	70.31	1.870	0.119	-0.245	0.210	0.041	1.52
Residential First Flush	137	133	107.15	2.030	0.125	0.500	0.210	0.167	
All Land Uses Composite	354	342	77.62	1.890	0.083	0.188	0.132	0.334	1.55
All Land Uses First Flush	354	342	120.23	2.080	0.104	0.225	0.132	0.126	

Table E12. Results of Preliminary Test Analysis for Total Dissolved Solids (TDS)

TDS (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0	U = 7.33 ; P = 0	0.160	0	Different (first flush)
Industrial	0.0245	U = 2.28 ; P = 0.01	0.070	0.003	Different
Institutional	0.0118	U = 2.945 ; P = 0	0.544	0	Different
Open Space	N/A	U = 0.161 ; P = 0.44	0	N/A	Same (no first flush)
Residential	N/A	U = 4.89 ; P = 0	0	N/A	Different
All Land Uses	0	U = 7.58 ; P = 0	0	N/A	Different

Table E13. Results of Preliminary Statistical Analysis for O&G

O&G (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	10	10	5.19	0.715	0.068	-0.976	0.687	0.016	1.54
Commercial First Flush	10	10	8.00	0.027	0.903	1.641	0,687	0.019	
Residential Composite	8	4	5.00	0.699	0.066	1.985	1.014	0.013	2.05
Residential First Flush	8	4	10.23	1.010	0.134	0.003	1.014	0.056	
All Land Uses Composite	18	14	5.00	0.699	0.073	-0.370	0.597	0.015	1.60
All Land Uses First Flush	18	14	8.00	0.903	0.051	0.890	0.597	0.011	

Table E14. Results of Preliminary Test Analysis for O&G

O&G (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 6.198 ; P < 0.01	0.222	0.004	Different
Residential	N/A	U = 1.069 ; P > 0.1	0.049	0.306	Same
All Land Uses	N/A	U = 4.072 ; P = 0	0.036	N/A	Different

Table E15. Results of Preliminary Statistical Analysis for Fecal Coliforms

Fecal Coliforms (mpn/100 mL)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	12	12	67764	4.831	1.099	-0.691	0.637	0.627	0.87
Commercial First Flush	12	12	58884	4.770	1.732	-0.388	0.637	0.228	
Residential Composite	10	9	41976	4.623	0.292	0.485	0.717	0.276	0.98
Residential First Flush	10	9	41020	4.643	0.685	0.247	0.717	0.799	
All Land Uses Composite	22	21	46238	4.665	0.745	-0.886	-0.515	0.511	1.21
All Land Uses First Flush	22	21	55976	4.748	1.269	0.501	0.501	0.391	

Table E16. Results of Preliminary Test Analysis for Fecal Coliforms

Fecal Coliforms (mpn/100 mL)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 0 ; P > 0.10	0.833	0.583	Same
Residential	N/A	U = 0.289 ; P > 0.1	0.016	0.973	Same
All Land Uses	N/A	U = 0.181 ; P = 0.43	0.086	0.665	Same

Table E17. Results of Preliminary Statistical Analysis for Fecal Streptococcus

Fecal Streptococcus (mpn/100 mL)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	12	11	37153	4.570	0.780	-0.255	0.661	0.948	1.05
Commercial First Flush	12	11	38904	4.590	1.094	0.009	0.661	0.722	
Residential Composite	11	8	77625	4.890	0.231	-0.223	0.752	0.426	1.30
Residential First Flush	11	8	101158	5.005	0.327	-0.659	0.752	0.319	
All Land Uses Composite	26	22	43651	4.640	0.536	-0.513	0.491	0.713	1.11
All Land Uses First Flush	26	22	48417	4.685	0.705	-0.188	0.491	0.802	

Table E18. Results of Preliminary Test Analysis for Fecal Streptococcus

Fecal Streptococcus (mpn/100mL)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 0.281 ; P > 0.10	0.027	N/A	Same (no first flush)
Residential	N/A	U = 0.344 ; P > 0.10	0.109	0.905	Same
All Land Uses	N/A	U = 0.309 ; P = 0.38	0.033	N/A	Same

Table E19. Results of Preliminary Statistical Analysis for Ammonia

Ammonia (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	70	52	0.76	-0.122	0.147	-0.245	0.330	0.237	2.11
Commercial First Flush	70	52	1.60	0.204	0.117	-0.718	0.330	0.027	
Industrial Composite	40	33	0.62	-0.208	0.166	-0.399	0.409	0.284	1.08
Industrial First Flush	40	33	0.67	-0.174	0.201	-0.535	0.409	0.046	
Institutional Composite	18	16	0.31	-0.509	0.058	-0.038	0.564	0.273	1.66
Institutional First Flush	18	16	0.51	-0.290	0.077	0.284	0.564	0.384	
Residential Composite	119	86	0.50	-0.301	0.370	0.779	0.260	0.001	1.36
Residential First Flush	119	86	0.68	-0.168	0.172	0.195	0.260	0.519	
All Land Uses Composite	269	190	0.52	-0.284	0.251	0.501	0.176	0.002	1.54
All Land Uses First Flush	269	190	0.80	-0.097	0.176	-0.197	0.176	0.713	

* Ammonia in Open Space was found in 22 events. Only 3 events had values above the detection limit

Table E20. Results of Preliminary Test Analysis for Ammonia

Ammonia (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 4.467 ; P = 0	0.028	N/A	Different
Industrial	N/A	U = 0.113 ; P = 0.46	0.262	0.985	Same
Institutional	0.0287	U = 2.484 ; P = 0.01	0.254	0	Different
Residential	N/A	U = 2.283 ; P = 0.01	0	N/A	Different
All Land Uses	N/A	U = 4.092 ; P = 0	0	N/A	Different

Table E21. Results of Preliminary Statistical Analysis for $NO_2 + NO_3$

$NO_2 + NO_3$ (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	84	82	0.75	-0.125	0.095	-0.092	0.266	0.188	1.73
Commercial First Flush	84	82	1.30	0.114	0.166	-0.790	0.266	0.007	
Industrial Composite	72	71	0.90	-0.046	0.073	-0.240	0.285	0.807	1.31
Industrial First Flush	72	71	1.18	0.072	0.116	-0.839	0.285	0.030	
Institutional	18	18	0.60	-0.222	0.122	-0.714	0.536	0.117	1.70
Institutional First Flush	18	18	1.02	0.009	0.151	0.268	0.536	0.381	
Open Space Composite	30	21	0.24	-0.620	0.290	0.468	0.501	0.141	0.96
Open Space First Flush	30	21	0.23	-0.638	0.356	0.823	0.501	0.030	
Residential Composite	121	118	0.60	-0.222	0.104	-0.196	0.223	0.504	1.66
Residential First Flush	121	118	1.00	-0.002	0.125	-0.292	0.223	0.102	
All Land Uses Composite	324	310	0.70	-0.155	0.124	-0.497	0.138	0	1.50
All Land Uses First Flush	324	310	1.05	0.021	0.162	-0.584	0.138	0	

Table E22. Results of Preliminary Test Analysis for $NO_2 + NO_3$

$NO_2 + NO_3$	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 3.286 ; P = 0	0	N/A	Different (first flush)
Industrial	N/A	U = 1.836 ; P = 0.03	0.941	0.034	Different
Institutional	0.043	U = 2.242 ; P = 0.01	0.026	N/A	Different
Open Space	N/A	U = 0.209 ; P = 0.42	0.023	N/A	Same (no first flush)
Residential	0	U = 4.769 ; P = 0	0.023	N/A	Different
All Land Uses	N/A	U = 5.834 ; P = 0	0	N/A	Different

Table E23. Results of Preliminary Statistical Analysis for Total Nitrogen

Total N (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	19	19	1.42	0.152	0.180	-0.133	0.524	0.215	1.35
Commercial First Flush	19	19	1.91	0.281	0.203	-0.617	0.524	0.337	
Industrial Composite	19	16	2.01	0.303	0.286	-0.306	0.564	0.431	1.79
Industrial First Flush	19	16	3.61	0.557	0.349	-0.452	0.564	0.029	
Open Space Composite	6	6	1.39	0.142	0.112	-0.150	0.845	0.330	1.53
Open Space First Flush	6	6	2.12	0.326	0.248	-0.100	0.845	0.221	
Residential Composite	31	30	1.67	0.222	0.325	1.22	0.427	0.009	0.88
Residential First Flush	31	30	1.47	0.166	0.447	-0.587	0.427	0.367	
All Land Uses Composite	77	73	1.60	0.204	0.253	0.769	0.281	0.136	1.22
All Land Uses First Flush	77	73	1.95	0.290	0.331	0.599	0.281	0.071	

Table E24. Results of Preliminary Test Analysis for Total Nitrogen

Total N (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.220	U = 1.234 ; P = 0.11	0.329	0.013	Same
Industrial	N/A	U = 0.460 ; P = 0.32	0.759	0.161	Same
Open Space	N/A	U = 0 ; P > 0.104	0.339	0.703	Same
Residential	N/A	U = 0.106 ; P = 0.46	0.002	N/A	Same
All Land Uses	N/A	U = 0.919 ; P = 0.18	0	N/A	Same

Table E25. Results of Preliminary Statistical Analysis for TKN

TKN (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	93	86	1.63	0.213	0.085	-0.275	0.260	0.003	1.71
Commercial First Flush	93	86	2.80	0.447	0.120	-0.117	0.260	0.714	
Industrial Composite	77	76	1.69	0.227	0.116	1.157	0.276	0	1.35
Industrial First Flush	77	76	2.27	0.356	0.130	0.536	0.276	0.232	
Open Space Composite	32	14	0.61	-0.215	0.142	0.585	0.597	0.109	1.28
Open Space First Flush	32	14	0.78	-0.107	0.269	0.948	0.597	0.139	
Residential Composite	131	123	1.40	0.146	0.110	1.752	0.218	0	1.65
Residential First Flush	131	123	2.31	0.364	0.115	0.309	0.218	0.076	
All Land Uses Composite	335	301	1.50	0.176	0.114	0.856	0.140	0	1.60
All Land Uses First Flush	335	301	2.40	0.380	0.139	0.088	0.140	0	

Table E26. Results of Preliminary Test Analysis for TKN

TKN (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 6.499 ; P = 0	0.126	0	Different (first flush)
Industrial	N/A	U = 1.698 ; P = 0.04	0.054	0.063	Different
Open Space	N/A	U = 0.374 ; P = 0.35	0.116	0.364	Same (no first flush)
Residential	N/A	U = 6.079 ; P = 0	0	N/A	Different
All Land Uses	N/A	U = 7.68 ; P = 0	0	N/A	Different

Table E27. Results of Preliminary Statistical Analysis for Total Phosphorus

Total P (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	89	77	0.34	-0.469	0.160	-0.454	0.274	0.129	1.44
Commercial First Flush	89	77	0.49	-0.310	0.205	0.033	0.274	0.035	
Industrial Composite	84	71	0.29	-0.538	0.130	0.495	0.285	0.003	1.42
Industrial First Flush	84	71	0.41	-0.387	0.257	-0.441	0.285	0.397	
Institutional Composite	17	17	0.17	-0.770	0.203	-0.736	0.550	0.374	1.24
Institutional First Flush	17	17	0.21	-0.678	0.066	-0.177	0.550	0.704	
Open Space Composite	32	20	0.09	-1.023	0.147	0.613	0.512	0.218	1.05
Open Space First Flush	32	20	0.10	-1.000	0.381	0.833	0.512	0.288	
Residential Composite	140	128	0.28	-0.553	0.252	1.232	0.214	0	1.46
Residential First Flush	140	128	0.41	-0.389	0.188	-0.335	0.214	0.042	
All Land Uses Composite	363	313	0.28	-0.553	0.209	0.605	0.138	0	1.45
All Land Uses First Flush	363	313	0.41	-0.391	0.238	-0.258	0.138	0.003	

Table E28. Results of Preliminary Test Analysis for Total Phosphorus

Total P (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 3.089 ; P = 0	0.594	0	Different (first flush)
Industrial	N/A	U = 0.864 ; P = 0.19	0.194	0.667	Same (no first flush)
Institutional	N/A	U = 0.774 ; P = 0.22	0.044	N/A	Same
Open Space	N/A	U = 0.142 ; P = 0.44	0.091	0.527	Same
Residential	N/A	U = 2.671 ; P = 0	0	N/A	Different
All Land Uses	N/A	U = 3.641 ; P = 0	0	N/A	Different

Table E29. Results of Preliminary Statistical Analysis for Dissolved Phosphorus

Dissolved P (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	91	69	0.16	-0.788	0.152	0.467	0.289	0	1.23
Commercial First Flush	91	69	0.20	-0.699	0.212	0.904	0.289	0.005	
Industrial Composite	77	50	0.14	-0.854	0.142	1.248	0.337	0.093	1.04
Industrial First Flush	77	50	0.14	-0.839	0.160	0.406	0.337	0.043	
Institutional Composite	18	14	0.13	-0.891	0.066	-0.114	0.597	0.563	1.05
Institutional First Flush	18	14	0.13	-0.870	0.095	-0.770	0.597	0.122	
Open Space Composite	32	14	0.05	-1.301	0.111	-0.073	0.597	0.601	0.69
Open Space First Flush	32	14	0.03	-1.460	0.087	1.061	0.597	0.017	
Residential Composite	130	105	0.17	-0.770	0.117	0.152	0.236	0.458	1.24
Residential First Flush	130	105	0.21	-0.678	0.170	0.121	0.236	0.044	
All Land Uses Composite	350	254	0.15	-0.824	0.143	0.353	0.153	0.051	1.07
All Land Uses First Flush	350	254	0.16	-0.796	0.200	0.401	0.153	0.001	

Table E30. Results of Preliminary Test Analysis for Dissolved Phosphorus

Dissolved P (mg/L	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 1.582 ; P = 0.06	0.046	N/A	Same
Industrial	N/A	U = 0.051 ; P = 0.48	0.063	0.881	Same
Institutional	0.549	U = 0.605 ; P = 0.27	0.015	N/A	Same
Open Space	N/A	U = 0.760 ; P = 0.22	0.018	N/A	Same
Residential	N/A	U = 1.702 ; P = 0.04	0.039	N/A	Different
All Land Uses	N/A	U = 1.657 ; P = 0.05	0	N/A	Same

Table E31. Results of Preliminary Statistical Analysis for Orthophosphate

Orthophosphate (mg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Industrial Composite	6	6	0.16	-0.797	0.287	-0.047	0.845	0.838	1.55
Industrial First Flush	6	6	0.25	-0.607	0.356	-0.106	0.845	0.720	
Residential Composite	14	14	0.19	-0.714	0.554	2.557	0.597	0.001	0.95
Residential First Flush	14	14	0.18	-0.737	0.214	0.708	0.597	0.362	
All Land Uses Composite	22	22	0.19	-0.714	0.423	2.270	0.491	0.004	1.30
All Land Uses First Flush	22	22	0.25	-0.600	0.222	0.260	0.491	0.503	

Table E32. Results of Preliminary Test Analysis for Orthophosphate

Orthophosphate (mg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Industrial	0.471	U = 0.772 ; P > 0.104	0.071	0.611	Same
Residential	N/A	U = 0.022 ; P = 0.49	0	N/A	Same
All Land Uses	N/A	U = 0.460 ; P = 0.32	0	N/A	Same

Table E33. Results of Preliminary Statistical Analysis for Total Cadmium

Total Cadmium (µg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	74	48	0.56	-0.253	0.246	-0.325	0.343	0	2.15
Commercial First Flush	74	48	1.20	0.079	0.261	0.080	0.343	0.089	
Industrial Composite	80	41	1	0	0.124	-0.015	0.369	0.008	1.00
Industrial First Flush	80	41	1	0	0.130	0.261	0.369	0.065	
Open Space Composite	30	15	0.23	-0.638	0.282	1.074	0.580	0.183	1.30
Open Space First Flush	30	15	0.30	-0.523	0.325	0.465	0.580	0.402	
Residential Composite	123	33	0.28	-0.553	0.359	0.693	0.409	0.002	2.00
Residential First Flush	123	33	0.56	-0.252	0.264	0.512	0.409	0.061	
All Land Uses Composite	325	139	0.60	-0.222	0.269	-0.065	0.206	0.071	1.62
All Land Uses First Flush	325	139	0.97	-0.013	0.249	0.041	0.206	0.241	

Table E34. Results of Preliminary Test Analysis for Total Cadmium

Total Cadmium (µg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.006	U = 2.797 ; P = 0	0.009	N/A	Different (first flush)
Industrial	0.922	U = 0.100 ; P = 0.46	0.118	0.529	Same (no first flush)
Open Space	0.442	U = 0.765 ; P = 0.22	0.292	0.191	Same
Residential	0.038	U = 2.131 ; P = 0.02	0.015	N/A	Different
All Land Uses	0.005	U = 2.839 ; P = 0	0	N/A	Different

Table E35. Results of Preliminary Statistical Analysis for Total Chromium

Total Chromium (µg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	47	22	6.81	0.833	0.086	-0.051	0.491	0.911	1.67
Commercial First Flush	47	22	11.40	1.057	0.134	-0.796	0.491	0.121	
Industrial Composite	54	25	8.79	0.944	0.111	0.338	0.464	0.456	1.36
Industrial First Flush	54	25	11.99	1.079	0.155	-0.307	0.464	0.784	
Open Space Composite	16	4	2.64	0.422	0.169	-0.556	1.014	0.492	1.70
Open Space First Flush	16	4	4.50	0.653	0.015	1.291	1.014	0.355	
Residential Composite	86	31	8.00	0.903	0.169	-0.077	0.421	0.612	1.24
Residential First Flush	86	31	9.91	0.996	0.137	0.326	0.421	0.904	
All Land Uses Composite	218	82	7.50	0.875	0.140	-0.104	0.266	0.591	1.47
All Land Uses First Flush	218	82	10.99	1.041	0.141	-0.056	0.266	0.803	

Table E36. Results of Preliminary Test Analysis for Total Chromium

Total Chromium (µg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.0513	U = 2.024 ; P = 0.02	0.283	0.036	Different
Industrial	0.3032	U = 1.023 ; P = 0.15	0.216	0.320	Same
Open Space	0.3032	U = 1.586 ; P = 0.10	0.160	0.199	Same
Residential	0.6023	U = 0.519 ; P = 0.30	0.007	N/A	Same
All Land Uses	0.0547	U = 1.939 ; P = 0.03	0.001	N/A	Different

Table E37. Results of Preliminary Statistical Analysis for Total Copper

Total Copper (µg/L)	Total Events	Selected Cases	*Median*	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	92	82	16.98	1.230	0.083	-0.038	0.266	0.117	1.62
Commercial First Flush	92	82	27.48	1.439	0.120	0.343	0.266	0.035	
Industrial Composite	84	76	25.00	1.398	0.079	0.184	0.276	0.344	1.24
Industrial First Flush	84	76	30.97	1.491	0.166	-0.014	0.276	0.007	
Institutional Composite	18	7	16.98	1.230	0.083	-0.228	0.794	0.167	0.94
Institutional First Flush	18	7	16.00	1.204	0.047	0.954	0.794	0.555	
Open Space Composite	30	22	5.14	0.711	0.103	0.085	0.491	0.252	0.78
Open Space First Flush	30	22	4.00	0.602	0.120	1.005	0.491	0.015	
Residential Composite	144	108	11.99	1.079	0.082	-0.677	0.233	0	1.33
Residential First Flush	144	108	16.00	1.204	0.087	0.023	0.233	0.256	
All Land Uses Composite	368	295	15.00	1.176	0.116	-0.268	0.142	0	1.33
All Land Uses First Flush	368	295	20.00	1.301	0.167	0.009	0.142	0	

Table E38. Results of Preliminary Test Analysis for Total Copper

Total Copper (µg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 5.160 ; P = 0	0.001	N/A	Different (first flush)
Industrial	N/A	U = 1.864 ; P = 0.03	0.329	0.012	Different
Institutional	0.5224	U = 0.665 ; P > 0.099	0.318	0.029	Same (no first flush)
Open Space	N/A	U = 0.846 ; P = 0.19	0.074	0.337	Same
Residential	N/A	U = 4.029 ; P = 0	0.292	0	Different
All Land Uses	N/A	U = 5.146 ; P = 0	0	N/A	Different

Table E39. Results of Preliminary Statistical Analysis for Total Lead

Total Lead (µg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	89	83	16.98	1.230	0.062	0.075	0.264	0.824	1.65
Commercial First Flush	89	83	27.99	1.447	0.123	0.070	0.264	0.476	
Industrial Composite	84	71	16.98	1.230	0.160	0.527	0.285	0.081	1.41
Industrial First Flush	84	71	23.99	1.380	0.240	0.319	0.285	0.608	
Institutional Composite	18	13	7.00	0.845	0.082	0.675	0.616	0.158	2.28
Institutional First Flush	18	13	15.96	1.203	0.051	0.128	0.616	0.228	
Open Space Composite	31	16	5.00	0.699	0.381	-0.303	0.564	0.199	0.90
Open Space First Flush	31	16	4.48	0.651	0.346	-0.466	0.564	0.563	
Residential Composite	140	93	8.79	0.944	0.231	0.084	0.250	0.884	1.48
Residential First Flush	140	93	13.00	1.114	0.204	0.130	0.250	0.105	
All Land Uses Composite	364	278	13.00	1.114	0.198	-0.365	0.146	0.006	1.50
All Land Uses First Flush	364	278	19.50	1.290	0.239	-0.307	0.146	0.401	

Table E40. Results of Preliminary Test Analysis for Total Lead

Total Lead (µg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0	U = 5.256 ; P = 0	0.794	0	Different
Industrial	0.083	U = 1.742 ; P = 0.04	0.167	0.016	Different
Institutional	0.004	U = 3.973 ; P = 0	0.680	0.000	Different
Open Space	0.771	U = 0.292 ; P = 0.39	0.008	0.578	Same
Residential	0.012	U = 2.59 ; P = 0	0.014	N/A	Different
All Land Uses	N/A	U = 4.77 ; P = 0	0	N/A	Different

Table E41. Results of Preliminary Statistical Analysis for Total Nickel

Total Nickel (µg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	47	23	5.00	0.699	0.094	0.660	0.481	0.254	2.40
Commercial First Flush	47	23	11.99	1.079	0.134	-0.606	0.481	0.523	
Industrial Composite	51	22	7.00	0.845	0.106	-0.293	0.491	0.229	1.00
Industrial First Flush	51	22	7.00	0.845	0.197	0.605	0.491	0.228	
Residential Composite	83	18	7.48	0.874	0.094	0.152	0.536	0.814	1.20
Residential First Flush	83	18	8.99	0.954	0.115	1.551	0.536	0.048	
All Land Uses Composite	213	64	6.00	0.778	0.104	0.146	0.299	0.161	1.50
All Land Uses First Flush	213	64	8.99	0.954	0.147	0.322	0.299	0.443	

Table E42. Results of Preliminary Test Analysis for Total Nickel

Total Nickel (µg/L	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	0.006	U = 3.005 ; P = 0	0.128	0.002	Different (first flush)
Industrial	0.715	U = 0.365 ; P = 0.36	0.203	0.484	Same (no first flush)
Residential	N/A	U = 1.143 ; P = 0.13	0.512	0.098	Same
All Land Uses	0.014	U = 2.539 ; P = 0.01	0.367	0.001	Different

Table E43. Results of Preliminary Statistical Analysis for Total Zinc

Total Zinc (µg/L)	Total Events	Selected Cases	Median	Median (Log)	Var (Log)	Skew (Log)	SE Skew (Log)	Test Norm. (Log) p-value	Median Ratio
Commercial Composite	90	90	149.97	2.176	0.089	-1.359	0.254	0	1.93
Commercial First Flush	90	90	289.07	2.461	0.139	-0.374	0.254	0.647	
Industrial Composite	83	83	225.94	2.354	0.184	0.828	0.264	0	1.54
Industrial First Flush	83	83	348.34	2.542	0.135	-0.181	0.264	0.930	
Institutional Composite	18	18	304.79	2.484	0.114	-0.227	0.536	0.878	2.48
Institutional First Flush	18	18	755.09	2.878	0.133	-0.696	0.536	0.055	
Open Space Composite	21	21	20.00	1.301	0.165	0.081	0.501	0.073	1.25
Open Space First Flush	21	21	25.00	1.398	0.075	-0.242	0.501	0.295	
Residential Composite	136	136	69.34	1.841	0.114	0.824	0.208	0.003	1.58
Residential First Flush	136	136	109.90	2.041	0.200	-0.232	0.208	0.014	
All Land Uses Composite	350	350	125.89	2.100	0.216	0.121	0.130	0.001	1.59
All Land Uses First Flush	350	350	199.99	2.301	0.268	0.437	0.130	0.020	

Table E44. Results of Preliminary Test Analysis for Total Zinc

Total Zinc(µg/L)	Mann Wittn. p-value	Fligner Policello	Normality for t-Test p-value	Paired t – Test	Result
Commercial	N/A	U = 6.156 ; P = 0	0	N/A	Different
Industrial	N/A	U = 2.087 ; P = 0.02	0.006	N/A	Different
Institutional	0.007	U = 3.1 ; P = 0	0.498	0	Different
Open Space	N/A	U = 0.023 ; P = 0.49	0.667	0.977	Same
Residential	N/A	U = 4.329 ; P = 0	0	N/A	Different
All Land Uses	N/A	U = 5.374 ; P = 0	0	N/A	Different

Appendix F: Detailed Statistical Test Results to Identify Significant Land Use and Geographical Interactions

pH

Summary
Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	0	123	*	*	*	*
3	0	6	*	*	*	*
4	0	16	*	*	*	*
5	41	1	7.7039	0.438	6.9	9.1
6	27	7	6.737	0.632	5.7	8.6
7	27	14	7.363	0.662	5.3	8.38

Industrial

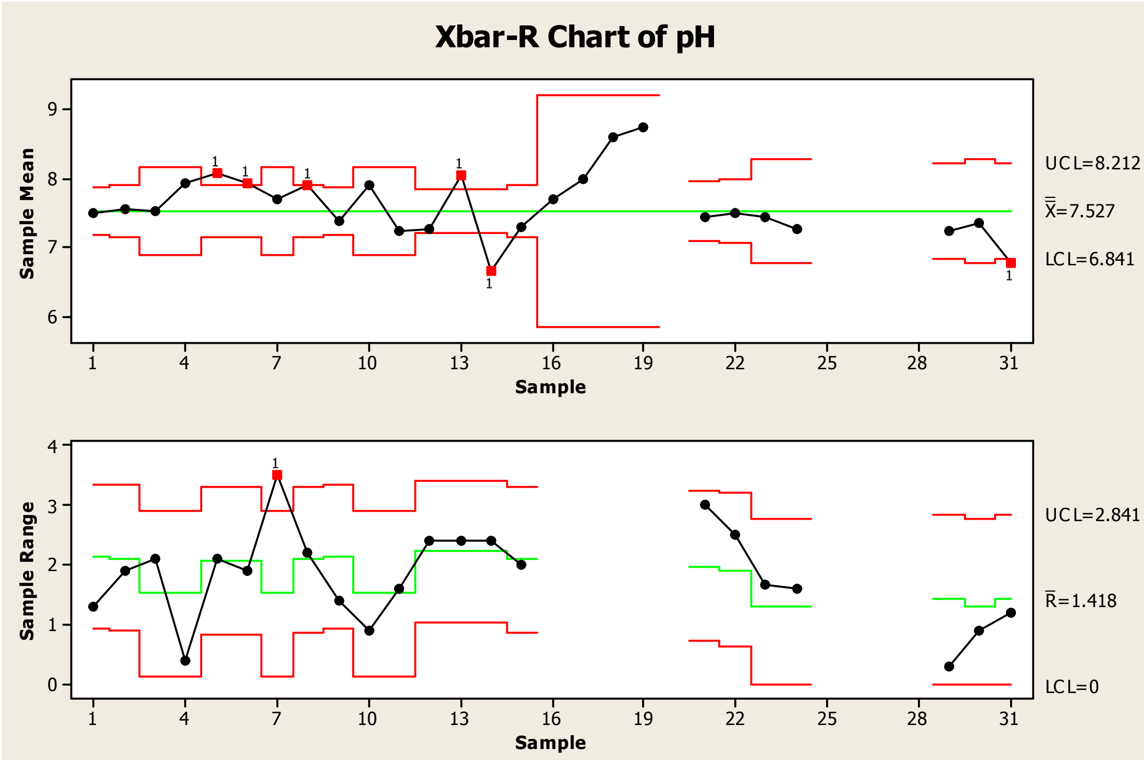
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	18	91	7.319	0.533	6.4	8.65
3	9	7	6.6111	0.2205	6.5	7
4	0	17	*	*	*	*
5	46	1	7.9822	0.504	6.6	9
6	53	17	7.6836	0.7219	5.9	9.3
7	5	28	7.34	0.358	6.8	7.7

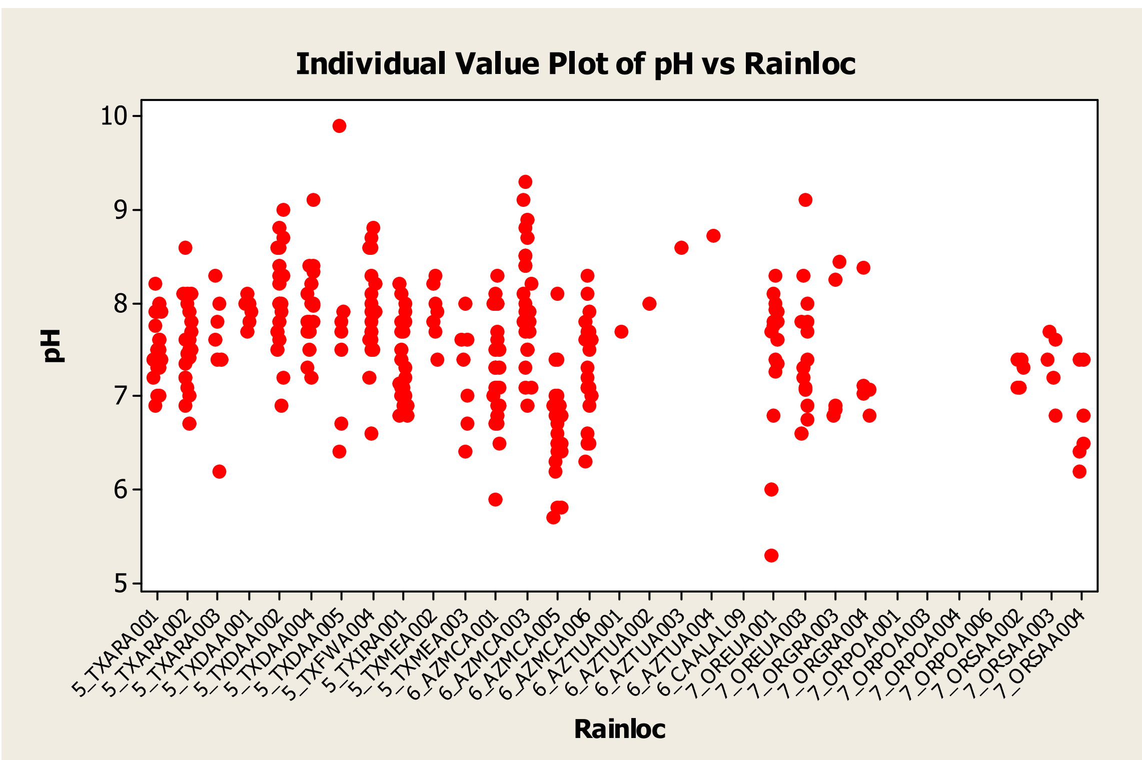
Residential

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	14	317	7.364	0.389	6.7	7.87
3	8	10	6.788	0.416	6.5	7.4
4	0	31	*	*	*	*
5	71	0	7.5179	0.5816	6.2	9.9
6	22	16	7.341	0.559	6.3	8.3
7	25	15	7.319	0.713	6.2	9.1

Highlighted cells indicated groups with enough observations to calculate mean and standard deviation.

1. First Analysis. EPA Rain zones 5,6, and 7 Landuse: Residential Commercial Industrial



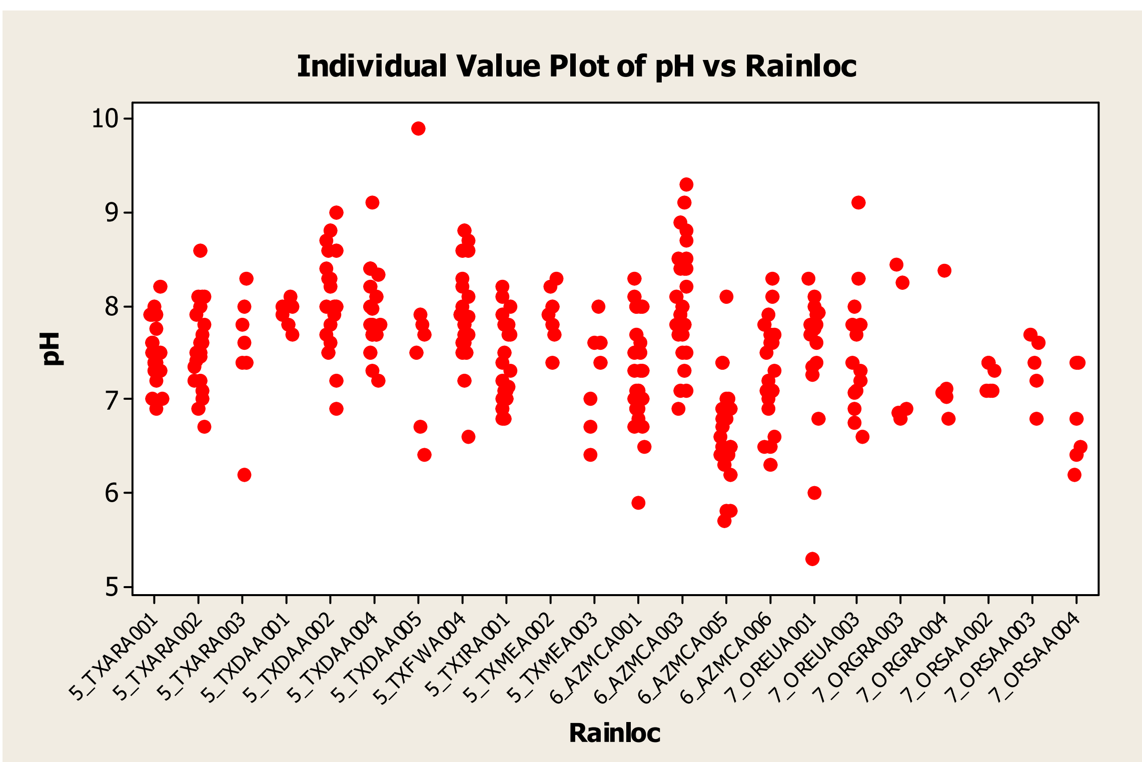


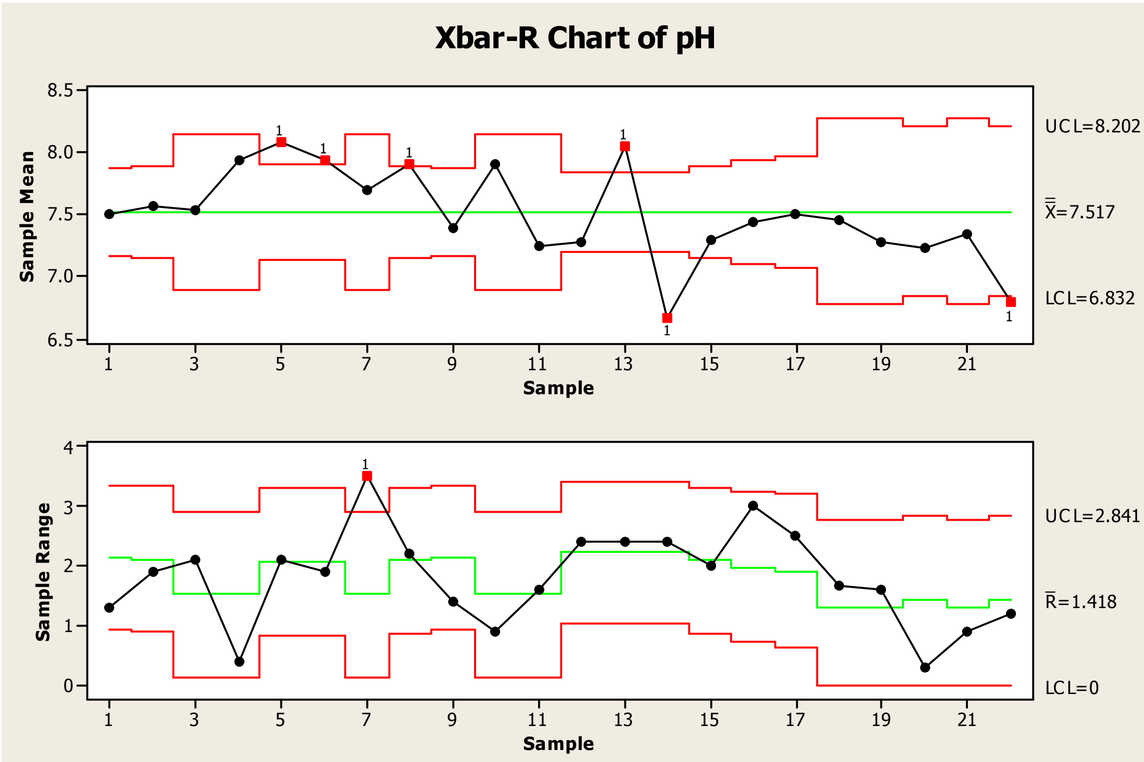
Sites with one observation or less will not be included in the analyses.

Variable	Rainloc	N	N*	Mean	Variance	Minimum	Maximum
pH	5_TXARA001	22	0	7.5073	0.1132	6.9000	8.2000
	5_TXARA002	21	0	7.563	0.220	6.700	8.600

5_TXARA003	7	0	7.529	0.449	6.200	8.300
5_TXDAA001	7	0	7.9286	0.0190	7.7000	8.1000
5_TXDAA002	19	0	8.079	0.306	6.900	9.000
5_TXDAA004	19	1	7.932	0.192	7.200	9.100
5_TXDAA005	7	0	7.700	1.270	6.400	9.900
5_TXFWA004	20	1	7.909	0.289	6.600	8.800
5_TXIRA001	22	0	7.3795	0.2017	6.8000	8.2000
5_TXMEA002	7	0	7.900	0.0933	7.400	8.300
5_TXMEA003	7	0	7.243	0.320	6.400	8.000
6_AZMCA001	26	1	7.277	0.331	5.900	8.300
6_AZMCA003	26	1	8.050	0.398	6.900	9.300
6_AZMCA005	26	0	6.665	0.271	5.700	8.100
6_AZMCA006	20	0	7.290	0.313	6.300	8.300
6_AZTUA001	1	9	7.7000	*	7.7000	7.7000
6_AZTUA002	1	7	8.0000	*	8.0000	8.0000
6_AZTUA003	1	7	8.6000	*	8.6000	8.6000
6_AZTUA004	1	6	8.7300	*	8.7300	8.7300
6_CAALAL09	0	9	*	*	*	*
7_OREUA001	16	0	7.437	0.633	5.300	8.300
7_OREUA003	14	1	7.502	0.452	6.600	9.100
7_ORGRA003	5	1	7.450	0.681	6.800	8.450
7_ORGRA004	5	1	7.280	0.393	6.800	8.380
7_ORPOA001	0	13	*	*	*	*
7_ORPOA003	0	14	*	*	*	*
7_ORPOA004	0	13	*	*	*	*
7_ORPOA006	0	13	*	*	*	*
7_ORSAA002	6	0	7.2333	0.0227	7.1000	7.4000
7_ORSAA003	5	1	7.340	0.128	6.800	7.700
7_ORSAA004	6	0	6.783	0.266	6.200	7.400

The following two plots will identify unusual sites among all the sites with pH observations. These sites will be included in the analysis, however they must be analyzed to identify potential conditions that produce these unusual observations.





The GLM will be used to identify if there is a significant difference in PH among land use and EPA rain zone. GLM were used instead of the ANOVA model because the number of observations is not the same in each combination land use – EPA rain zone.

RESULTS:

General Linear Model: pH versus Landuse, EPA_Rain_Zone

```
Factor          Type   Levels  Values
Landuse         fixed       3  CO, ID, RE
EPA_Rain_Zone   fixed       3  5, 6, 7


Analysis of Variance for pH, using Adjusted SS for Tests

Source                  DF    Seq SS    Adj SS  Adj MS      F      P
Landuse                  2   12.1467    4.6160  2.3080   6.64  0.001
EPA_Rain_Zone            2   14.0505   16.1575  8.0788  23.26  0.000
Landuse*EPA_Rain_Zone    4    7.9699    7.9699  1.9925   5.74  0.000
Error                  304  105.5977  105.5977  0.3474
Total                  312  139.7647
```

In this case the main factors and interaction term are considered significant. Now simultaneous tests will be used to evaluate if there is a significant difference among the land uses or the EPA rain zones. Bonferroni is the most conservative method (conservative means “true error rate is less than the stated one”).

```
Bonferroni Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of            Adjusted
Landuse    of Means  Difference  T-Value    P-Value
```

ID	0.4179	0.11470	3.644	0.0009
RE	0.1317	0.08873	1.485	0.4160

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.2862	0.1154	-2.480	0.0411

Tukey Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.4179	0.11470	3.644	0.0008
RE	0.1317	0.08873	1.485	0.2982

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.2862	0.1154	-2.480	0.0351

pH in industrial land use is significantly different than in residential and commercial land uses. pH in commercial and residential land uses is not significantly different.

Bonferroni Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	-0.5284	0.08050	-6.564	0.0000
7	-0.3941	0.11409	-3.454	0.0019

EPA_Rain_Zone = 6 subtracted from:

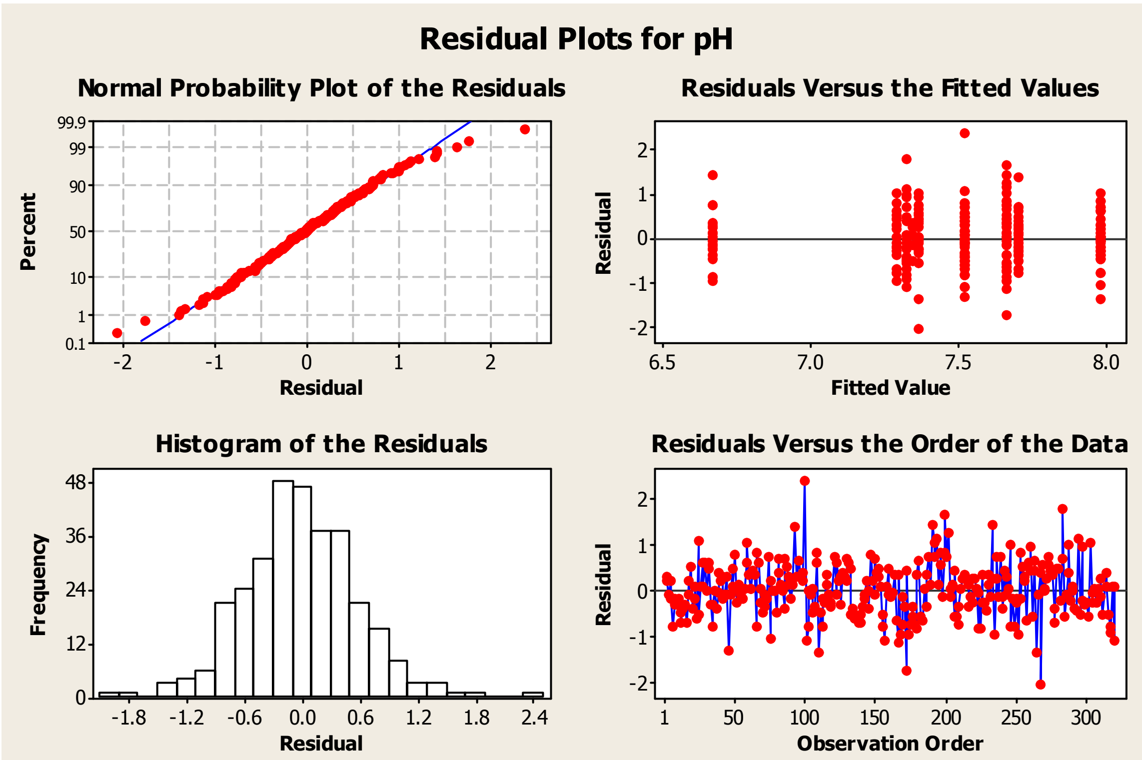
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	0.1343	0.1219	1.102	0.8137

Tukey Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	-0.5284	0.08050	-6.564	0.0000
7	-0.3941	0.11409	-3.454	0.0016

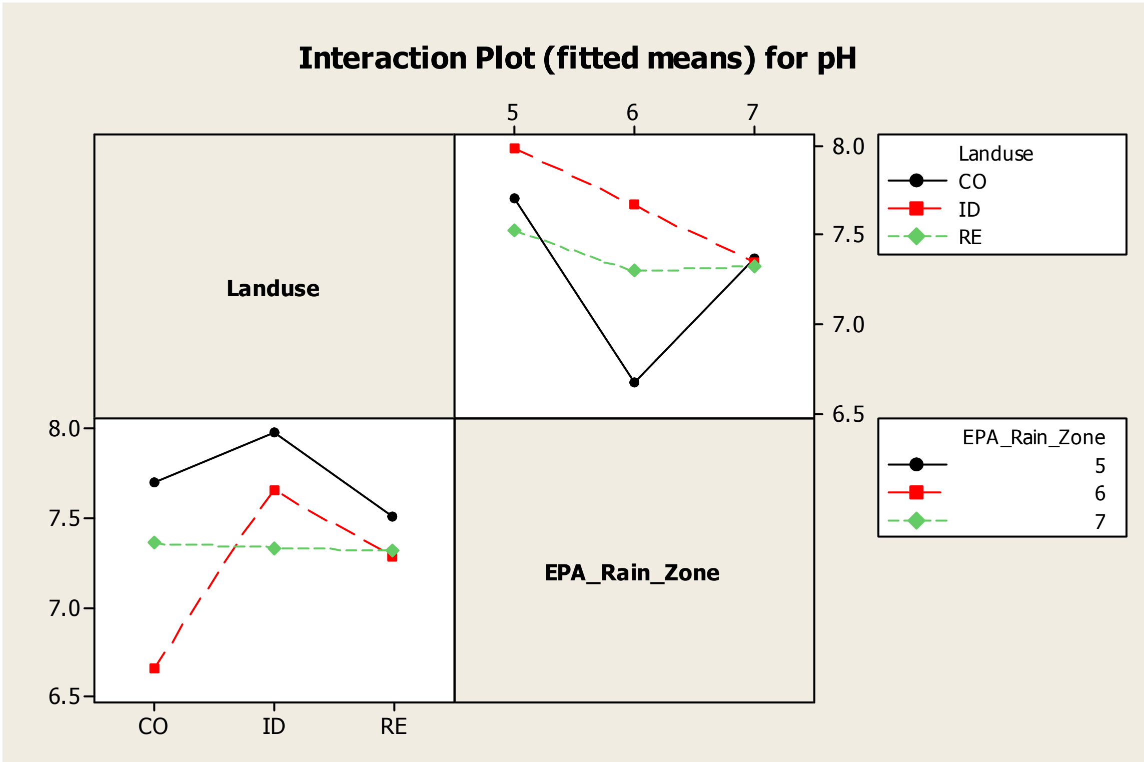
```
EPA_Rain_Zone = 6  subtracted from:

                Difference        SE of             Adjusted
EPA_Rain_Zone     of Means   Difference   T-Value    P-Value
7                   0.1343       0.1219     1.102     0.5127
```

pH in EPA rain zone 5 is significantly different than zones 6 and 7. pH in rain zones 6 and 7 is not significantly different.



The last analysis is to inspect the residual plots to confirm that they are normally distributed around zero and there is not a specific pattern. In this case the plots indicate normality and no trend against observation order.

The final plot is the interaction plot. In this plot it is possible to identify if there in a difference in the pattern of the levels when change from one level to another. In this case, for example, it was observed that pH in region 7 is not affected by land use. In the other two rain zones, when the land use changes from residential or commercial to industrial, there is an increase in the pH.

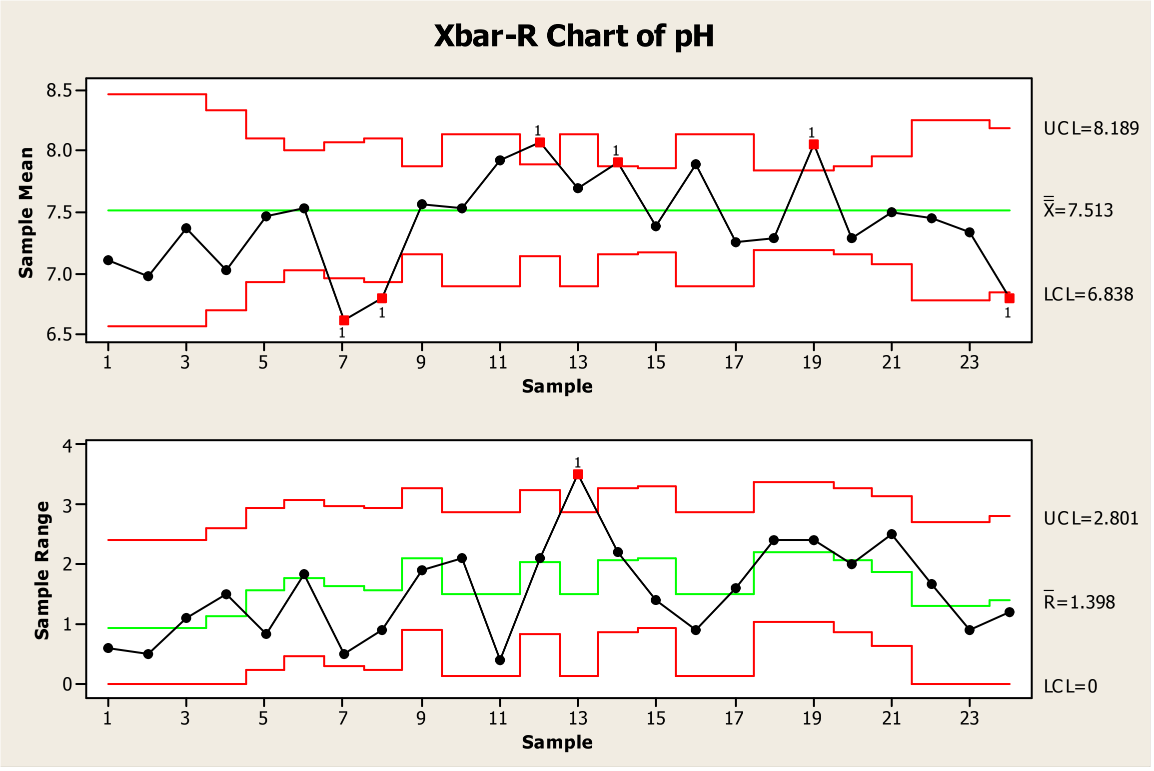


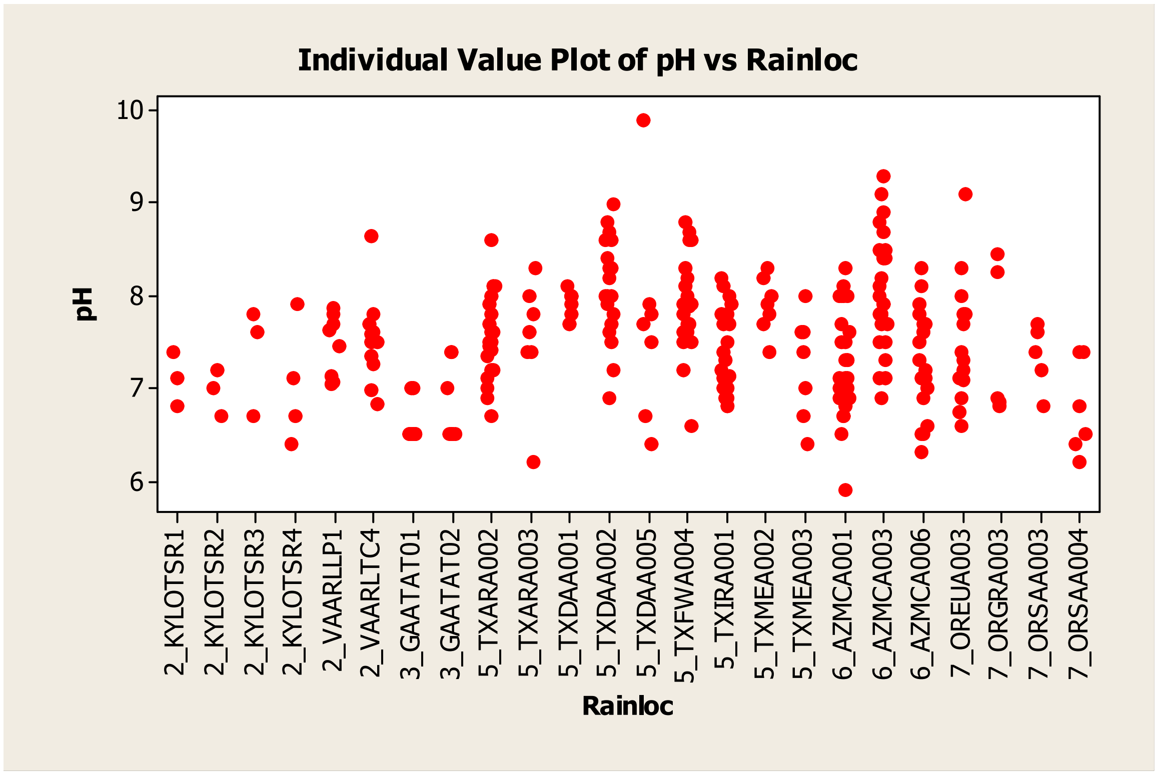
2. Second Analysis. EPA Rain zones 2, 3, 5, 6, and 7 Landuse: Residential and Industrial

Areas in Yellow are not included in the analysis

Variable	Rainloc	N	N*	Mean	Variance	Minimum	Maximum
pH	2_KYLOTSR1	3	0	7.100	0.0900	6.800	7.400
	2_KYLOTSR2	3	0	6.967	0.0633	6.700	7.200
	2_KYLOTSR3	3	0	7.367	0.343	6.700	7.800
	2_KYLOTSR4	4	0	7.025	0.422	6.400	7.900
	2_MDAACOMW	0	3	*	*	*	*
	2_MDAACOOD	0	3	*	*	*	*
	2_MDAACORK	0	3	*	*	*	*
	2_MDBACOBC	0	3	*	*	*	*
	2_MDBACOSC	0	26	*	*	*	*
	2_MDBACOTC	0	3	*	*	*	*
	2_MDBACOWC	0	3	*	*	*	*
	2_MDBCTYBO	0	3	*	*	*	*
	2_MDBCTYFM	0	3	*	*	*	*
	2_MDBCTYHO	0	3	*	*	*	*
	2_MDBCTYHR	0	3	*	*	*	*
	2_VAARLLP1	8	0	7.461	0.116	7.040	7.870
	2_VAARLTC4	11	2	7.523	0.226	6.830	8.650
	2_VACPTC1A	0	8	*	*	*	*
	2_VACPTSF2	0	3	*	*	*	*
	2_VACPTYC1	0	7	*	*	*	*
	2_VACPTYC3	0	14	*	*	*	*
	2_VACPTYC5	0	14	*	*	*	*
	2_VACPTYO1	0	3	*	*	*	*
	2_VAHATYH2	0	19	*	*	*	*
	2_VAHATYH3	0	17	*	*	*	*
	2_VAHATYH4	0	17	*	*	*	*
	2_VAHATYH5	0	17	*	*	*	*
	2_VAHCCON1	0	2	*	*	*	*

2_VAHCCON2	0	3	*	*	*	*
2_VAHCCOR1	0	3	*	*	*	*
2_VAHCCOR2	0	3	*	*	*	*
2_VANFTMS6	0	3	*	*	*	*
2_VANFTYN2	0	29	*	*	*	*
2_VANFTYN3	0	27	*	*	*	*
2_VANFTYN5	0	27	*	*	*	*
2_VANNTMF1	0	1	*	*	*	*
2_VANNTMF4	0	3	*	*	*	*
2_VANNTNN1	0	10	*	*	*	*
2_VANNTSF4	0	2	*	*	*	*
2_VANNTSF6	0	2	*	*	*	*
2_VAPMTYP2	0	17	*	*	*	*
2_VAPMTYP4	0	17	*	*	*	*
2_VAPMTYP5	0	17	*	*	*	*
2_VAVBTYI1	0	3	*	*	*	*
2_VAVBTYR1	0	5	*	*	*	*
2_VAVBTYV1	0	27	*	*	*	*
2_VAVBTYV4	0	30	*	*	*	*
3_ALMOCREO	0	3	*	*	*	*
3_ALMOSARA	0	3	*	*	*	*
3_ALMOSIIV	0	3	*	*	*	*
3_ALMOSIVI	0	3	*	*	*	*
3_ALMOTHEO	0	3	*	*	*	*
3_GAATAT01	9	1	6.6111	0.0486	6.5000	7.0000
3_GAATAT02	8	1	6.788	0.173	6.500	7.400
5_TXARA002	21	0	7.563	0.220	6.700	8.600
5_TXARA003	7	0	7.529	0.449	6.200	8.300
5_TXDAA001	7	0	7.9286	0.0190	7.7000	8.1000
5_TXDAA002	19	0	8.079	0.306	6.900	9.000
5_TXDAA005	7	0	7.700	1.270	6.400	9.900
5_TXFWA004	20	1	7.909	0.289	6.600	8.800
5_TXIRA001	22	0	7.3795	0.2017	6.8000	8.2000
5_TXMEA002	7	0	7.900	0.0933	7.400	8.300
5_TXMEA003	7	0	7.243	0.320	6.400	8.000
6_AZMCA001	26	1	7.277	0.331	5.900	8.300
6_AZMCA003	26	1	8.050	0.398	6.900	9.300
6_AZMCA006	20	0	7.290	0.313	6.300	8.300
6_AZTUA001	1	9	7.7000	*	7.7000	7.7000
6_AZTUA002	1	7	8.0000	*	8.0000	8.0000
6_AZTUA004	1	6	8.7300	*	8.7300	8.7300
6_CAALAL09	0	9	*	*	*	*
7_OREUA003	14	1	7.502	0.452	6.600	9.100
7_ORGRA003	5	1	7.450	0.681	6.800	8.450
7_ORPOA003	0	14	*	*	*	*
7_ORPOA004	0	13	*	*	*	*
7_ORPOA006	0	13	*	*	*	*
7_ORSAA003	5	1	7.340	0.128	6.800	7.700
7_ORSAA004	6	0	6.783	0.266	6.200	7.400





pH about 7.5 but some sites can have mean pH as low as 6.5 or high as 8.0

Results from general linear model.

General Linear Model: pH versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	2	ID, RE
EPA_Rain_Zone	fixed	5	2, 3, 5, 6, 7

Analysis of Variance for pH, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	1	4.2658	0.5749	0.5749	1.68	0.196
EPA_Rain_Zone	4	18.5915	19.0338	4.7584	13.93	0.000
Landuse*EPA_Rain_Zone	4	3.0489	3.0489	0.7622	2.23	0.066
Error	258	88.1302	88.1302	0.3416		
Total	267	114.0363				

S = 0.584457   R-Sq = 22.72%   R-Sq(adj) = 20.02%

No significant differences by land use or interactions. Significant differences by EPA rain zone.

Bonferroni Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of Landuse
Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1276	0.09837	-1.297	0.1957

Tukey Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of Landuse
Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1276	0.09837	-1.297	0.1945

No significant differences in pH between industrial and residential land uses

Bonferroni Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.6422	0.1761	-3.647	0.0032
5	0.4085	0.1179	3.465	0.0062
6	0.1352	0.1294	1.045	1.0000
7	-0.0119	0.1770	-0.067	1.0000

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	1.0507	0.1524	6.895	0.0000
6	0.7774	0.1615	4.814	0.0000
7	0.6303	0.2016	3.126	0.0198

EPA_Rain_Zone = 5 subtracted from:

	Difference	SE of		Adjusted

EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
6	-0.2733	0.09472	-2.885	0.0424
7	-0.4204	0.15348	-2.739	0.0658

EPA_Rain_Zone = 6 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
7	-0.1471	0.1625	-0.9054	1.000

pH in rain zones 3 and 5 is different than in region 2
No significant differences in pH between region 2 and 6 and 2 and 7
pH in rain zones 5, 6 and 7 is different than in region 3
No significant differences in pH between region 6 and 7 compared with region 5
No significant differences in pH between regions 6 and 7

Tukey Simultaneous Tests
Response Variable pH
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
3	-0.6422	0.1761	-3.647	0.0025
5	0.4085	0.1179	3.465	0.0048
6	0.1352	0.1294	1.045	0.8345
7	-0.0119	0.1770	-0.067	1.0000

EPA_Rain_Zone = 3 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
5	1.0507	0.1524	6.895	0.0000
6	0.7774	0.1615	4.814	0.0000
7	0.6303	0.2016	3.126	0.0153

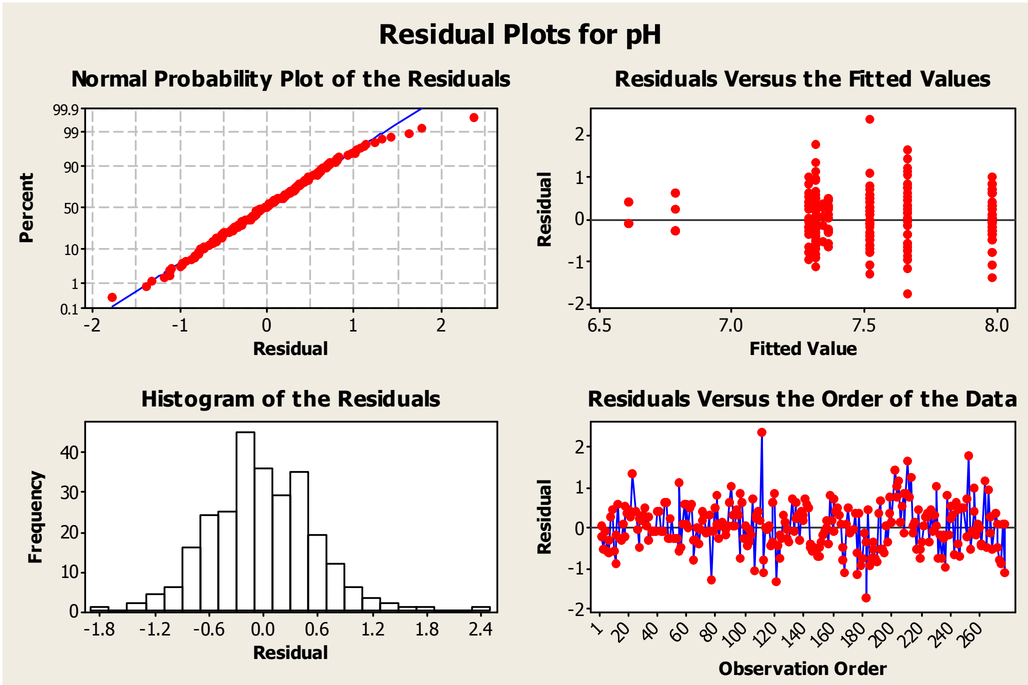
EPA_Rain_Zone = 5 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
6	-0.2733	0.09472	-2.885	0.0320
7	-0.4204	0.15348	-2.739	0.0484

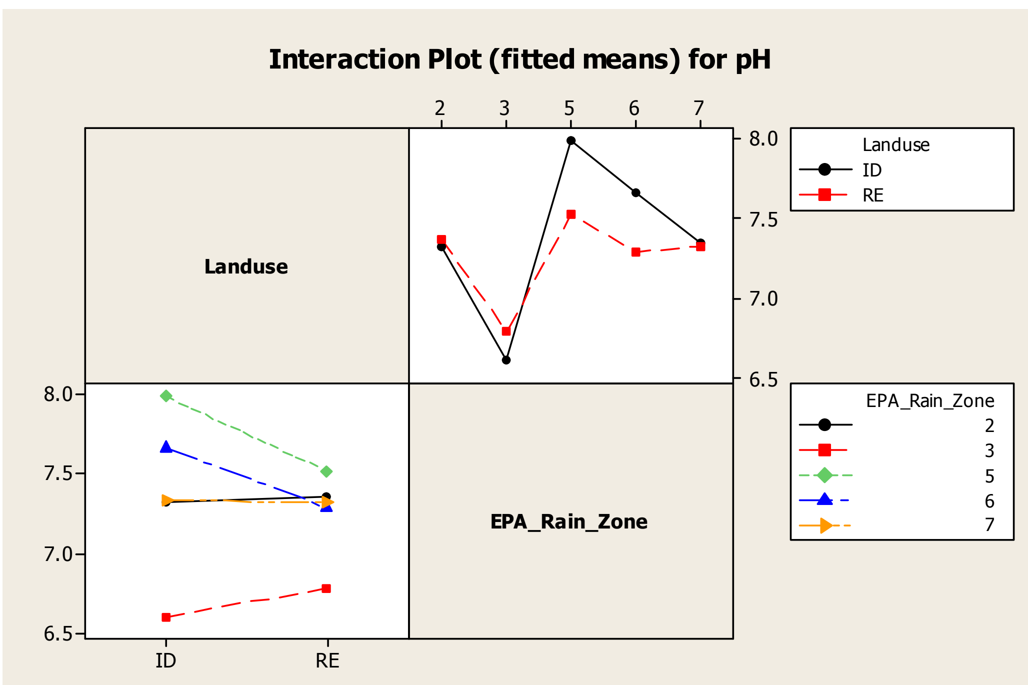
EPA_Rain_Zone = 6 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
7	-0.1471	0.1625	-0.9054	0.8949

pH in rain zones 3 and 5 is different than in region 2
No significant differences in pH between region 2 and 6 and 2 and 7
pH in rain zones 5, 6 and 7 is different than in region 3
pH in rain zones 6 and 7 is different than in region 5
No significant differences in pH between regions 6 and 7



Normality of the residuals, except for three or four unusual observations. No unusual patterns in the residuals



Temperature

Summary
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	0	123	*	*	*	*
3	0	6	*	*	*	*
4	0	16	*	*	*	*
5	41	1	16.229	4.762	6.5	24
6	26	8	20.87	7.13	11	30
7	21	20	12.724	4.54	5	21

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	0	109	*	*	*	*
3	0	16	*	*	*	*
4	0	17	*	*	*	*
5	47	0	19.47	5.837	7	30
6	52	18	20.213	6.395	11.5	31.5
7	5	28	9.9	2.7	7	14

Land use: Residential

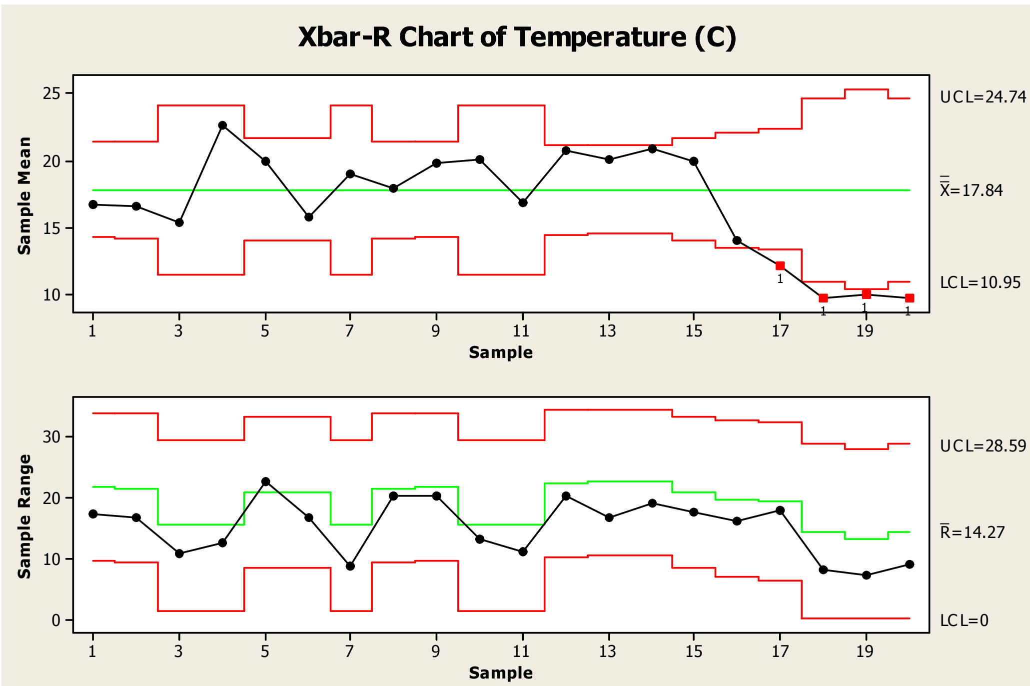
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	0	331	*	*	*	*
3	0	18	*	*	*	*
4	0	31	*	*	*	*
5	71	0	18.099	5.191	7.5	29
6	20	18	19.54	7.12	11	28.5
7	20	20	11.345	4.413	5	23

Analysis: Rain zone 5, 6 and 7. Land use RE, CO and ID

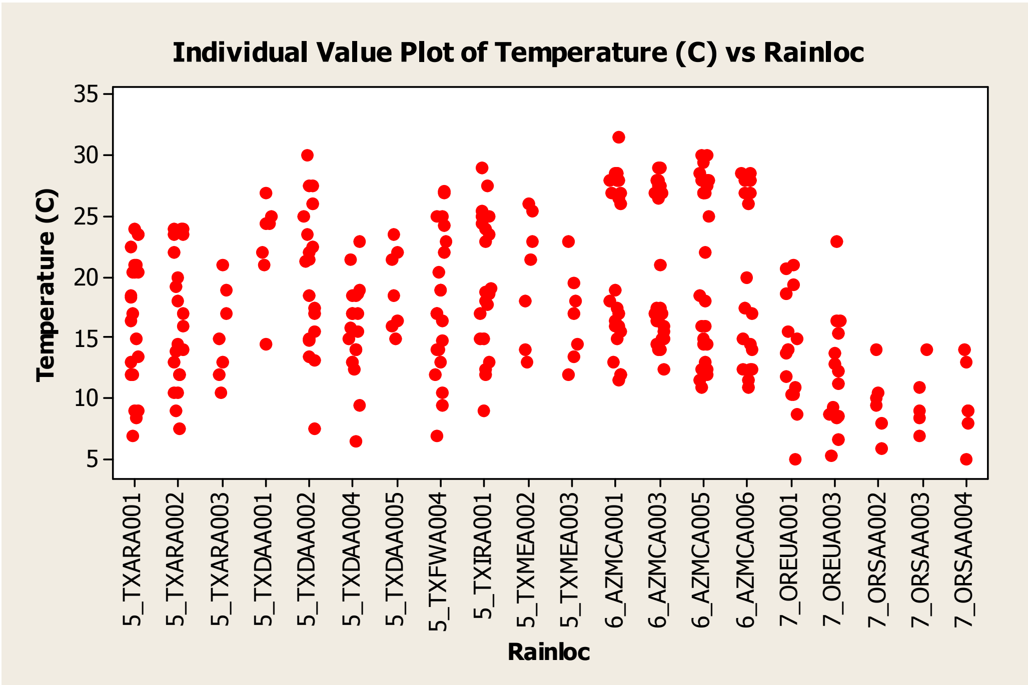
Descriptive Statistics: Temperature (C)

Variable	Rainloc	N	N*	Mean	Variance	Minimum	Maximum
Temperature (C)	5_TXARA001	22	0	16.65	29.37	7.00	24.00
	5_TXARA002	21	0	16.58	29.91	7.50	24.00
	5_TXARA003	7	0	15.36	14.73	10.50	21.00
	5_TXDAA001	7	0	22.64	16.81	14.50	27.00
	5_TXDAA002	19	0	19.97	35.33	7.50	30.00
	5_TXDAA004	19	1	15.737	15.654	6.500	23.000
	5_TXDAA005	7	0	19.00	11.17	15.00	23.50
	5_TXFWA004	21	0	17.96	35.39	7.00	27.10
	5_TXIRA001	22	0	19.90	31.47	9.00	29.00
	5_TXMEA002	7	0	20.14	27.73	13.00	26.00
	5_TXMEA003	7	0	16.79	14.40	12.00	23.00
	6_AZMCA001	25	2	20.74	46.94	11.50	31.50
	6_AZMCA003	26	1	20.04	35.10	12.50	29.00
	6_AZMCA005	26	0	20.87	50.79	11.00	30.00
	6_AZMCA006	19	1	19.95	49.94	11.00	28.50
	6_AZTUA001	1	9	11.800	*	11.800	11.800
	6_AZTUA002	0	8	*	*	*	*
	6_AZTUA003	0	8	*	*	*	*

6_AZTUA004	1	6	11.600	*	11.600	11.600
6_CAALAL09	0	9	*	*	*	*
7_OREUA001	15	1	13.95	21.28	5.00	21.00
7_OREUA003	14	1	12.06	22.35	5.30	23.00
7_ORGRA003	0	6	*	*	*	*
7_ORGRA004	0	6	*	*	*	*
7_ORPOA001	0	13	*	*	*	*
7_ORPOA003	0	14	*	*	*	*
7_ORPOA004	0	13	*	*	*	*
7_ORPOA006	0	13	*	*	*	*
7_ORSAA002	6	0	9.67	7.17	6.00	14.00
7_ORSAA003	5	1	9.90	7.30	7.00	14.00
7_ORSAA004	6	0	9.67	11.07	5.00	14.00



Samples in Rain zone 7 are colder than in the rain zones 5 and 6



General Linear Model: Temperature (C) versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	3	5, 6, 7

Analysis of Variance for Temperature (C), using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	418.05	0.87	0.44	0.01	0.986
EPA_Rain_Zone	2	1955.86	1815.42	907.71	28.34	0.000
Landuse*EPA_Rain_Zone	4	222.71	222.71	55.68	1.74	0.141
Error	292	9351.72	9351.72	32.03		
Total	300	11948.35				

Land use and interaction are not significant. EPA rain zone is significant.

Bonferroni Simultaneous Tests
Response Variable Temperature (C)
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.0220	1.1182	-0.0196	1.000
RE	-0.1425	0.8991	-0.1585	1.000

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1205	1.128	-0.1069	1.000

No differences among the three land uses

Tukey Simultaneous Tests

Response Variable Temperature (C)
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.0220	1.1182	-0.0196	0.9998
RE	-0.1425	0.8991	-0.1585	0.9862

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1205	1.128	-0.1069	0.9937

No differences among the three land uses

Bonferroni Simultaneous Tests
Response Variable Temperature (C)
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	2.466	0.7788	3.166	0.0051
7	-6.610	1.1277	-5.861	0.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-9.075	1.205	-7.529	0.0000

Temperature in rain zone 5 is different than in rain zones 6 and 7. Temperature in zone 6 is different than in rain zone 7.

Tukey Simultaneous Tests
Response Variable Temperature (C)
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5 subtracted from:

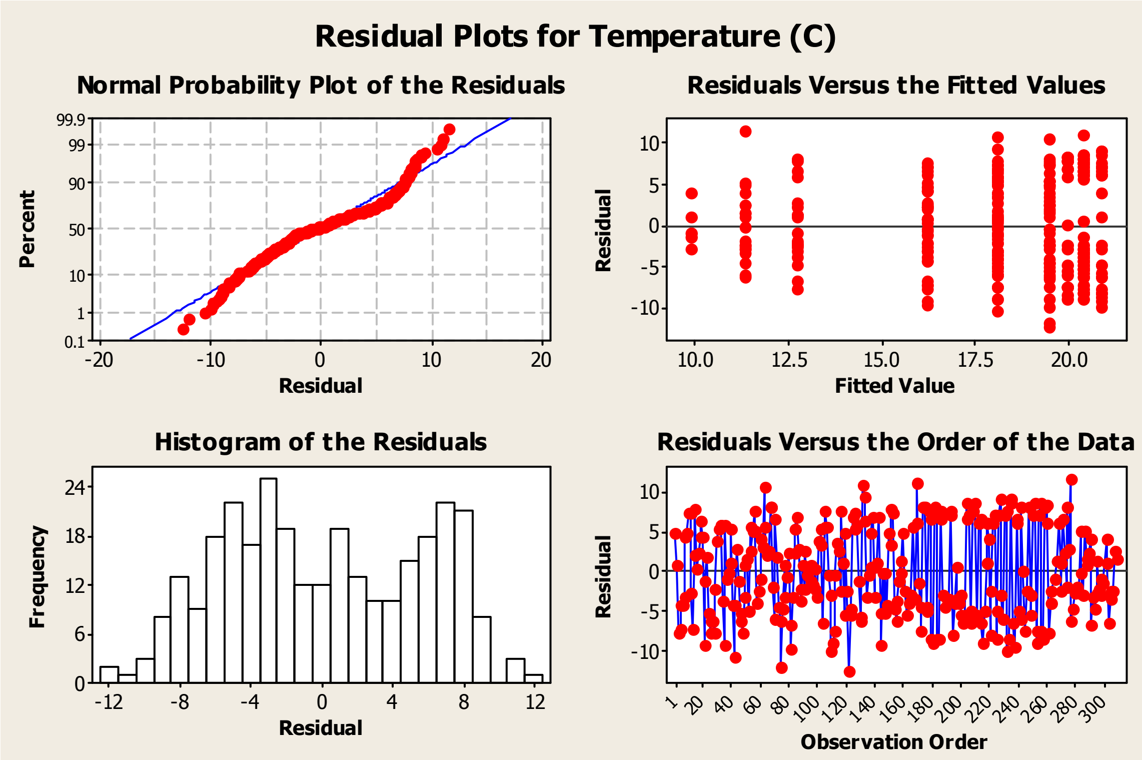
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	2.466	0.7788	3.166	0.0044
7	-6.610	1.1277	-5.861	0.0000

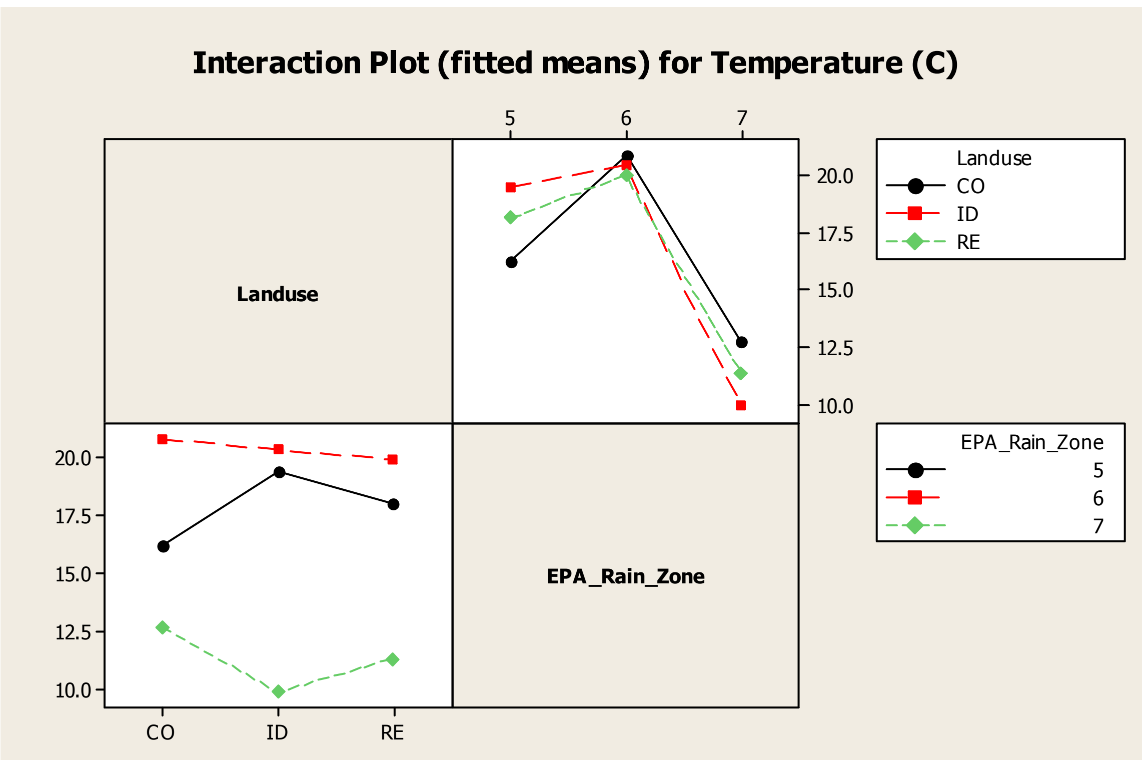
EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-9.075	1.205	-7.529	0.0000

Temperature in rain zone 5 is different than in rain zones 6 and 7. Temperature in zone 6 is different than in rain zone 7.

Bi modal probability plot. Normality assumption might not be valid. No specific pattern in the residuals. Season influence can be important.


Residual Plots for Temperature (C)
Normal Probability Plot of the Residuals
Percent
Residual
Residuals Versus the Fitted Values
Residual
Fitted Value
Histogram of the Residuals
Frequency
Residual
Residuals Versus the Order of the Data
Residual
Observation Order


Interaction Plot (fitted means) for Temperature (C)
Landuse
EPA_Rain_Zone
Landuse
CO
ID
RE
EPA_Rain_Zone
5
6
7

Hardness

Summary statistics in LOG base 10 scale

Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	0	123	*	*	*	*
3	0	6	*	*	*	*
4	0	16	*	*	*	*
5	39	3	1.475	0.149	1.146	1.845
6	0	34	*	*	*	*
7	26	15	1.099	0.431	0.279	1.792

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	11	98	1.790	0.491	1.000	2.646
3	0	16	*	*	*	*
4	0	17	*	*	*	*
5	43	4	1.599	0.181	1.255	2.137
6	8	62	0.957	0.379	0.699	1.531
7	20	13	1.246	0.265	0.740	1.881

Land use: Residential

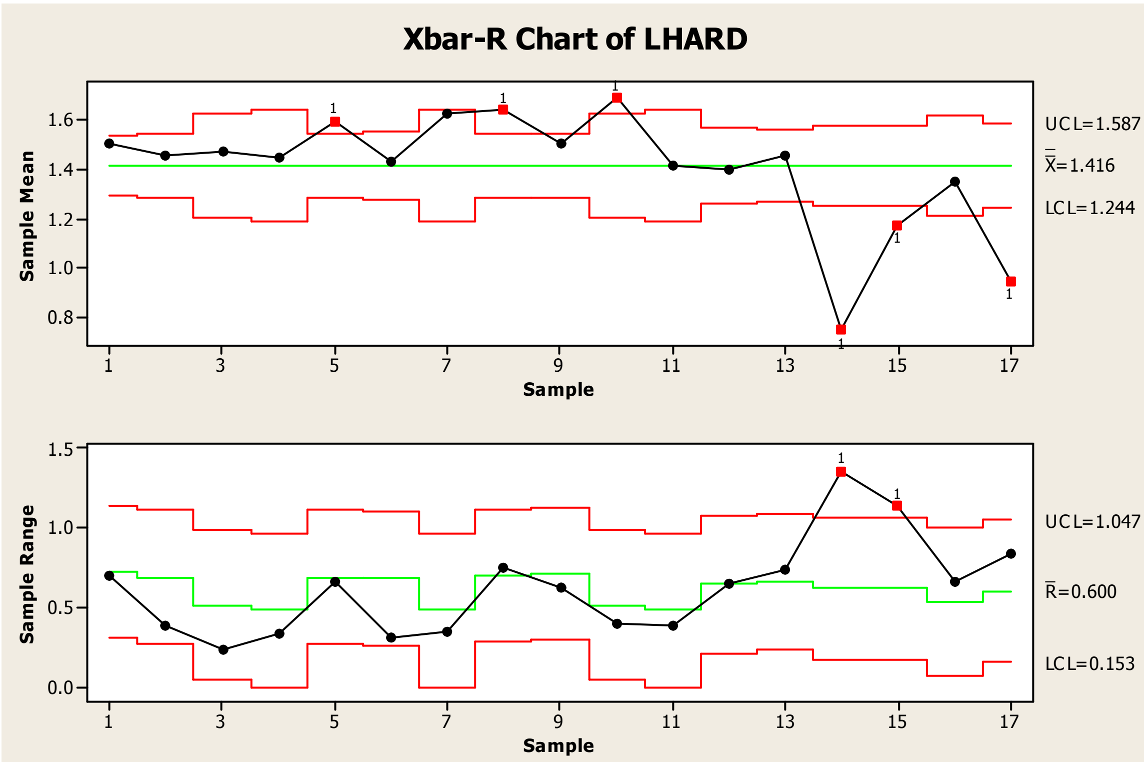
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	8	323	1.632	0.319	1.104	2.114
3	0	18	*	*	*	*
4	0	31	*	*	*	*
5	64	7	1.511	0.154	1.176	1.892
6	0	38	*	*	*	*
7	26	14	1.237	0.332	0.699	1.820

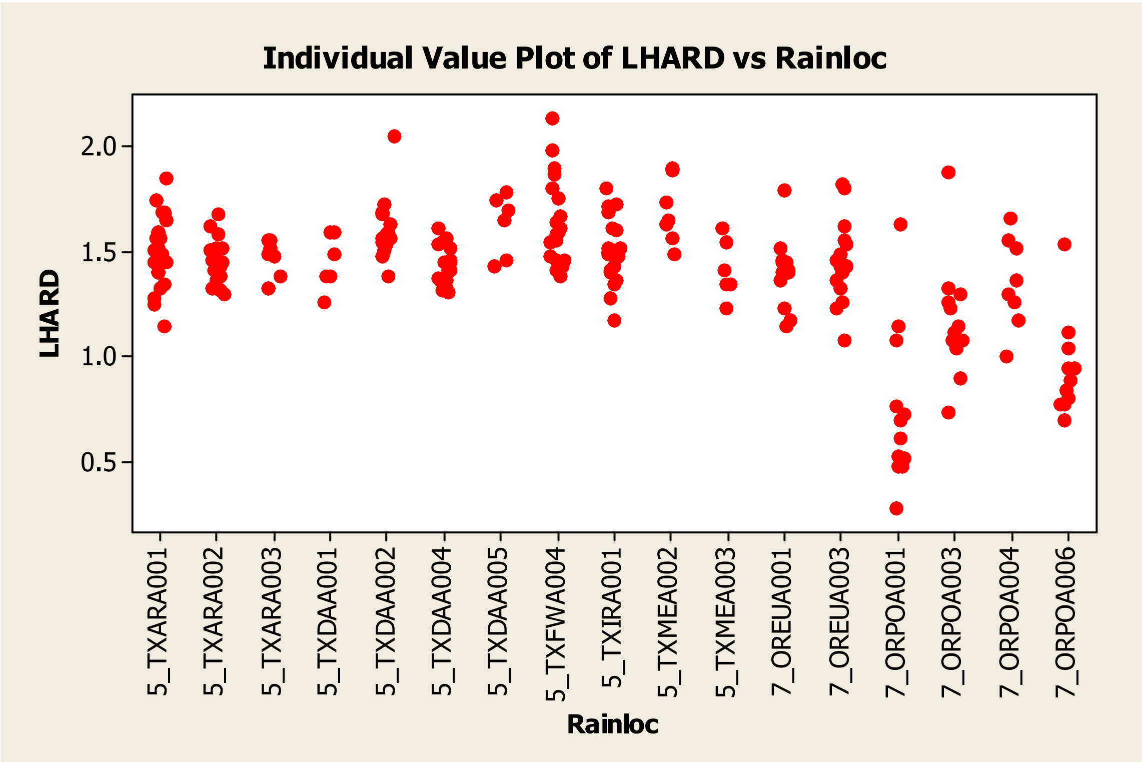
Analysis 1. Commercial, Residential, and industrial in rain zones 5 and 7

Descriptive Statistics: LHARD

Variable	Rainloc	N	N*	Mean	Variance	Minimum	Maximum
LHARD	5_TXARA001	22	0	1.5071	0.0309	1.1461	1.8445
	5_TXARA002	18	3	1.4559	0.0109	1.3010	1.6821
	5_TXARA003	7	0	1.4717	0.00794	1.3222	1.5563
	5_TXDAA001	6	1	1.4482	0.0178	1.2553	1.5911
	5_TXDAA002	18	1	1.5986	0.0187	1.3802	2.0453
	5_TXDAA004	17	3	1.4335	0.00894	1.3075	1.6128
	5_TXDAA005	6	1	1.6287	0.0217	1.4314	1.7782
	5_TXFWA004	19	2	1.6458	0.0440	1.3838	2.1367
	5_TXIRA001	20	2	1.5032	0.0248	1.1761	1.8007
	5_TXMEA002	7	0	1.6939	0.0233	1.4914	1.8921
	5_TXMEA003	6	1	1.4145	0.0201	1.2304	1.6128
	7_OREUA001	14	2	1.4016	0.0248	1.1461	1.7924
	7_OREUA003	15	0	1.4529	0.0400	1.0792	1.8195
	7_ORGRA003	0	6	*	*	*	*
	7_ORGRA004	0	6	*	*	*	*
	7_ORPOA001	12	1	0.746	0.140	0.279	1.633
	7_ORPOA003	12	2	1.1744	0.0772	0.7404	1.8808
	7_ORPOA004	8	5	1.3540	0.0476	1.0000	1.6628

7_ORPOA006	11	2	0.9423	0.0534	0.6990	1.5315
7_ORSAA002	0	6	*	*	*	*
7_ORSAA003	0	6	*	*	*	*
7_ORSAA004	0	6	*	*	*	*





General Linear Model: LHARD versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	2	5, 7

Analysis of Variance for LHARD, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.8799	0.5593	0.2796	4.77	0.009
EPA_Rain_Zone	1	5.1893	5.2468	5.2468	89.42	0.000
Landuse*EPA_Rain_Zone	2	0.0981	0.0981	0.0490	0.84	0.435
Error	212	12.4389	12.4389	0.0587		
Total	217	18.6061				

Main factors are significant. Interaction is not significant.

General Linear Model: LHARD versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	2	5, 7

Analysis of Variance for LHARD, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.8799	0.5800	0.2900	4.95	0.008
EPA_Rain_Zone	1	5.1893	5.1893	5.1893	88.58	0.000
Error	214	12.5369	12.5369	0.0586		
Total	217	18.6061				

Main factors are significant.

```
Bonferroni Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID          0.13487     0.04289    3.145    0.0057
RE          0.07044     0.03959    1.779    0.2298

Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE         -0.06443     0.03977   -1.620    0.3201
```

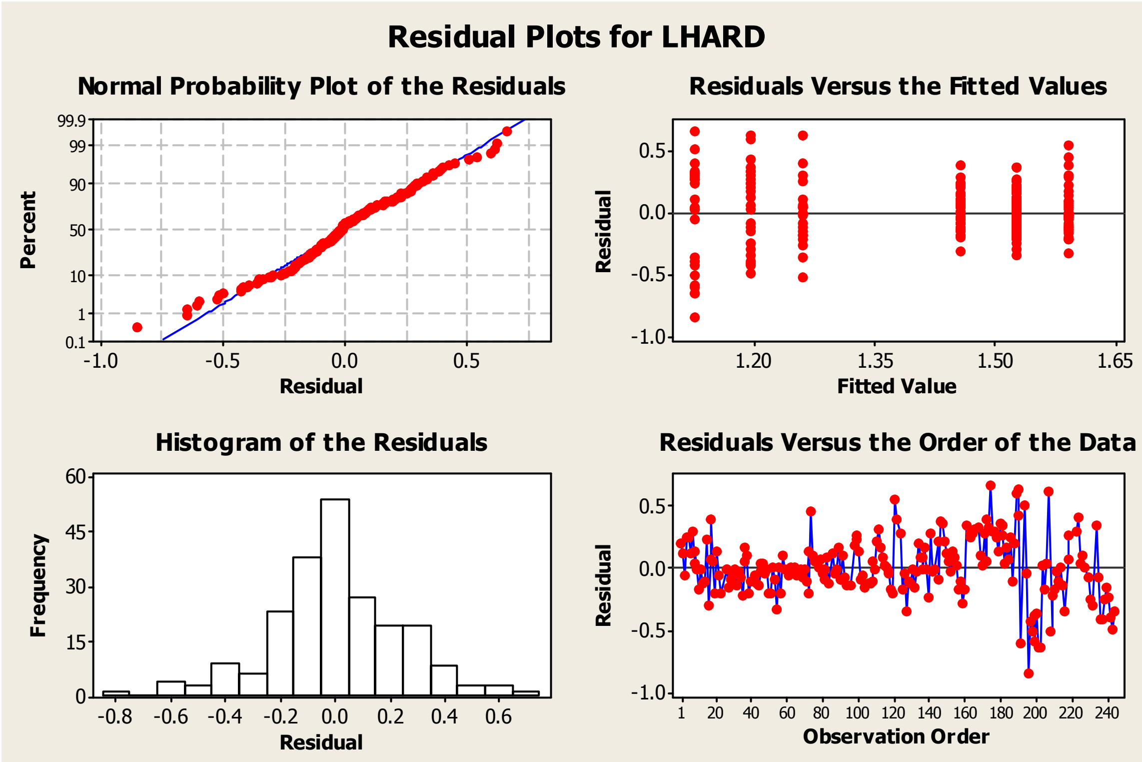
Hardness in rain commercial land use is different than in industrial land use. There are no differences in hardness between commercial and residential land use. There is no difference between residential and industrial land use for hardness concentrations (log)

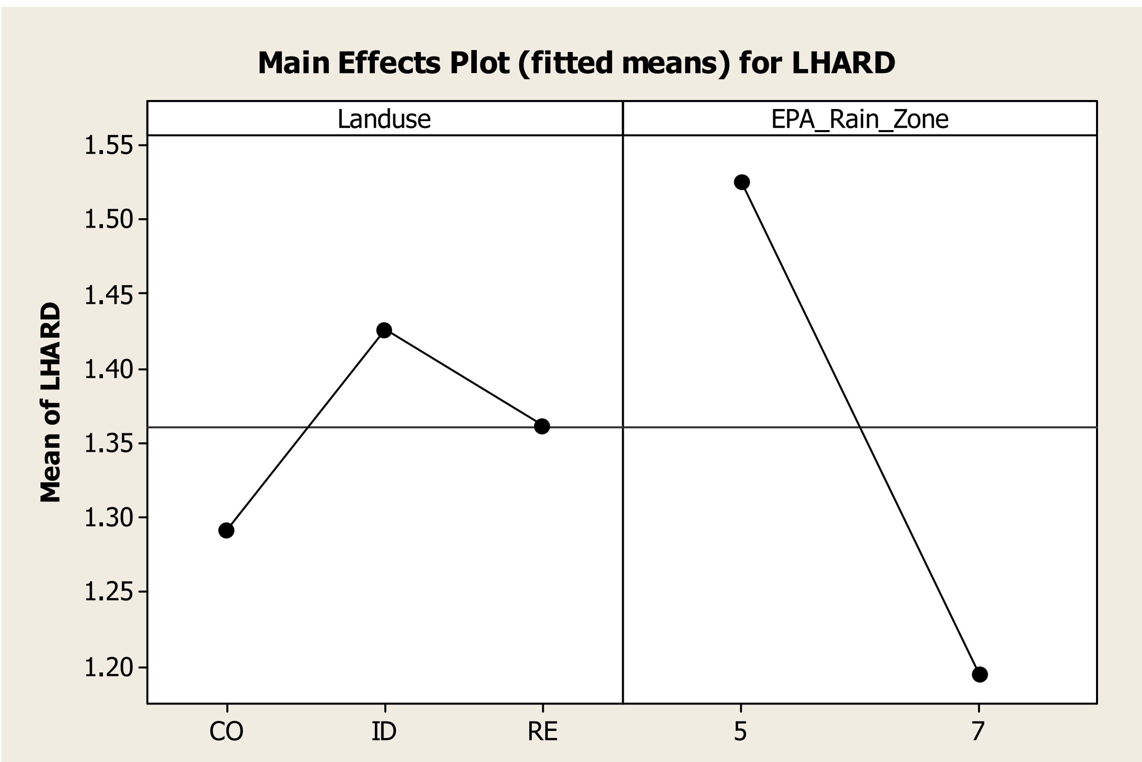
```
Tukey Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID          0.13487     0.04289    3.145    0.0054
RE          0.07044     0.03959    1.779    0.1790

Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE         -0.06443     0.03977   -1.620    0.2394
```

Same results using Tukey test.

```
Bonferroni Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
7                 -0.3297     0.03503   -9.412    0.0000

Tukey Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
7                 -0.3297     0.03503   -9.412    0.0000
```

Hardness concentrations in EPA rain zones 5 and 7 are significantly different.

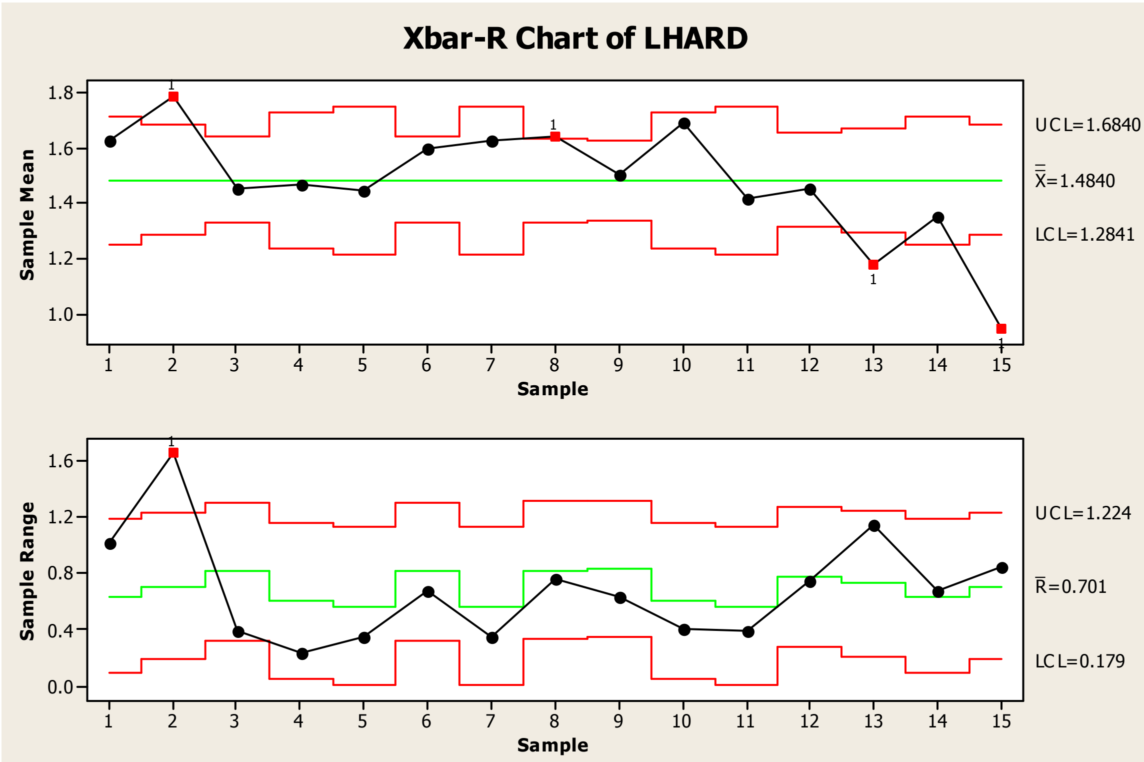


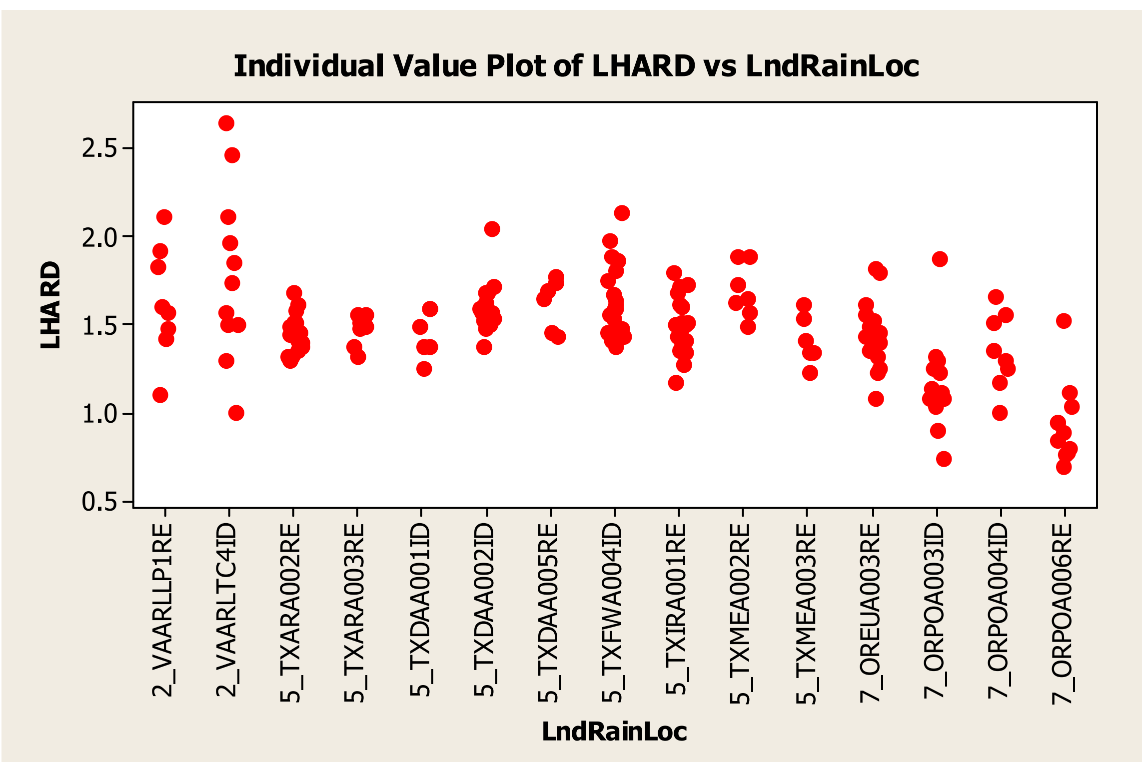


Analysis 2. Residential, and industrial in rain zones 2, 5 and 7

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LHARD	2_KYLOTSR1RE	0	3	*	*	*	*
	2_KYLOTSR2ID	0	3	*	*	*	*
	2_KYLOTSR3RE	0	3	*	*	*	*
	2_KYLOTSR4ID	0	4	*	*	*	*
	2_MDAACOMWID	0	3	*	*	*	*
	2_MDAACOODRE	0	3	*	*	*	*
	2_MDAACORKRE	0	3	*	*	*	*
	2_MDBACOBCID	0	3	*	*	*	*
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	0	3	*	*	*	*
	2_MDBACOWCRE	0	3	*	*	*	*
	2_MDBCTYBOID	0	3	*	*	*	*
	2_MDBCTYFMID	0	3	*	*	*	*
	2_MDBCTYHORE	0	3	*	*	*	*
	2_MDBCTYHRRE	0	3	*	*	*	*
	2_VAARLLP1RE	8	0	1.632	0.102	1.104	2.114
	2_VAARLTC4ID	11	2	1.790	0.241	1.000	2.646
	2_VACPTC1ARE	0	8	*	*	*	*
	2_VACPTSF2RE	0	3	*	*	*	*
	2_VACPTYC1RE	0	7	*	*	*	*
	2_VACPTYC3RE	0	14	*	*	*	*
	2_VACPTYC5ID	0	14	*	*	*	*
	2_VACPTYO1ID	0	3	*	*	*	*
	2_VAHATYH2ID	0	19	*	*	*	*
	2_VAHATYH3RE	0	17	*	*	*	*
	2_VAHATYH4RE	0	17	*	*	*	*
	2_VAHATYH5RE	0	17	*	*	*	*
	2_VAHCCON1ID	0	2	*	*	*	*
	2_VAHCCON2ID	0	3	*	*	*	*
	2_VAHCCOR1RE	0	3	*	*	*	*
	2_VAHCCOR2RE	0	3	*	*	*	*
	2_VANFTMS6RE	0	3	*	*	*	*
	2_VANFTYN2RE	0	29	*	*	*	*
	2_VANFTYN3RE	0	27	*	*	*	*
	2_VANFTYN5RE	0	27	*	*	*	*
	2_VANNTMF1RE	0	1	*	*	*	*
	2_VANNTMF4RE	0	3	*	*	*	*
	2_VANNTNN1RE	0	10	*	*	*	*
	2_VANNTSF4RE	0	2	*	*	*	*
	2_VANNTSF6RE	0	2	*	*	*	*
	2_VAPMTYP2RE	0	17	*	*	*	*
	2_VAPMTYP4RE	0	17	*	*	*	*
	2_VAPMTYP5RE	0	17	*	*	*	*
	2_VAVBTYI1ID	0	3	*	*	*	*
	2_VAVBTYR1RE	0	5	*	*	*	*
	2_VAVBTYV1RE	0	27	*	*	*	*
	2_VAVBTYV4ID	0	30	*	*	*	*
	5_TXARA002RE	18	3	1.4559	0.0109	1.3010	1.6821
	5_TXARA003RE	7	0	1.4717	0.00794	1.3222	1.5563
	5_TXDAA001ID	6	1	1.4482	0.0178	1.2553	1.5911
	5_TXDAA002ID	18	1	1.5986	0.0187	1.3802	2.0453
	5_TXDAA005RE	6	1	1.6287	0.0217	1.4314	1.7782
	5_TXFWA004ID	19	2	1.6458	0.0440	1.3838	2.1367
	5_TXIRA001RE	20	2	1.5032	0.0248	1.1761	1.8007
	5_TXMEA002RE	7	0	1.6939	0.0233	1.4914	1.8921
	5_TXMEA003RE	6	1	1.4145	0.0201	1.2304	1.6128
	7_OREUA003RE	15	0	1.4529	0.0400	1.0792	1.8195
	7_ORGRA003RE	0	6	*	*	*	*
	7_ORPOA003ID	12	2	1.1744	0.0772	0.7404	1.8808
	7_ORPOA004ID	8	5	1.3540	0.0476	1.0000	1.6628
	7_ORPOA006RE	11	2	0.9423	0.0534	0.6990	1.5315

```
          7_ORSAA003ID    0    6          *             *           *           *
          7_ORSAA004RE    0    6          *             *           *           *
```





General Linear Model: LHARD versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	2	ID, RE
EPA_Rain_Zone	fixed	3	2, 5, 7

Analysis of Variance for LHARD, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	1	0.2956	0.2296	0.2296	3.80	0.053
EPA_Rain_Zone	2	4.1525	4.1525	2.0762	34.33	0.000
Error	168	10.1597	10.1597	0.0605		
Total	171	14.6078				

Significant difference by rain zone but not by land use at 5%

Bonferroni Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of Landuse
Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.07424	0.03810	-1.948	0.0530

Tukey Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of Landuse
Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.07424	0.03810	-1.948	0.0530

No Significant difference by land use

Bonferroni Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.1643	0.06159	-2.667	0.0252
7	-0.4718	0.06729	-7.012	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.3075	0.04338	-7.090	0.0000

Tukey Simultaneous Tests
Response Variable LHARD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

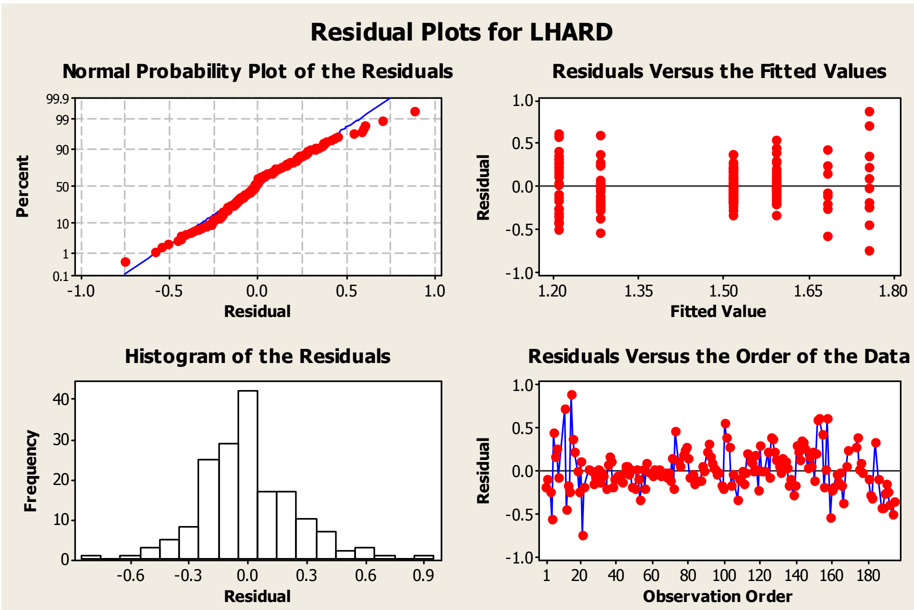
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.1643	0.06159	-2.667	0.0227
7	-0.4718	0.06729	-7.012	0.0000

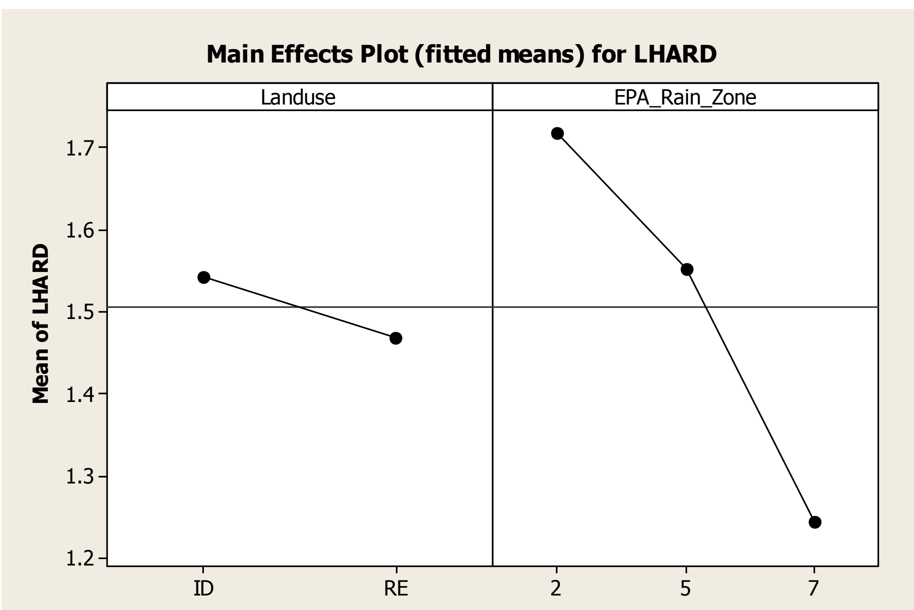
EPA_Rain_Zone = 5 subtracted from:

	Difference	SE of		Adjusted

```
EPA_Rain_Zone     of Means  Difference  T-Value   P-Value
7                  -0.3075     0.04338   -7.090    0.0000
```

Significant differences between rain zone 2 and rain zones 5 and 7. Significant differences between rain zone 5 and 7.





Oil and Grease

Summary statistics in LOG base 10 scale

Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	52	71	0.629	0.321	-0.097	1.556
3	0	6	*	*	*	*
4	14	2	0.683	0.416	-0.068	1.398
5	41	1	0.766	0.896	-0.962	2.555
6	29	5	0.620	0.487	-0.374	1.778
7	37	4	0.414	0.358	-0.476	1.255

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	41	68	0.531	0.417	-0.415	1.342
3	0	16	*	*	*	*
4	15	2	0.443	0.402	-0.291	1.146
5	46	1	0.191	1.074	-2.583	2.611
6	55	15	0.593	0.457	-0.571	1.380
7	29	4	0.569	0.314	-0.205	1.204

Land use: Residential

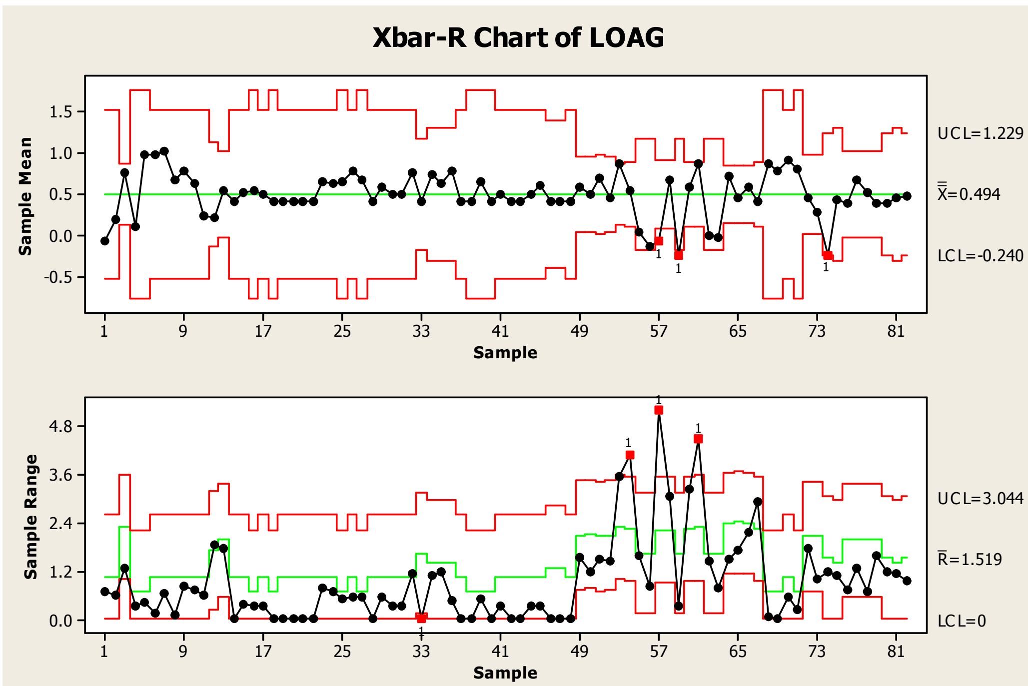
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	92	239	0.500	0.336	-0.742	1.505
3	0	18	*	*	*	*
4	31	0	0.531	0.401	-0.069	1.491
5	69	2	0.392	0.955	-1.470	3.474
6	24	14	0.474	0.573	-0.732	2.176
7	38	2	0.252	0.445	-0.804	1.491

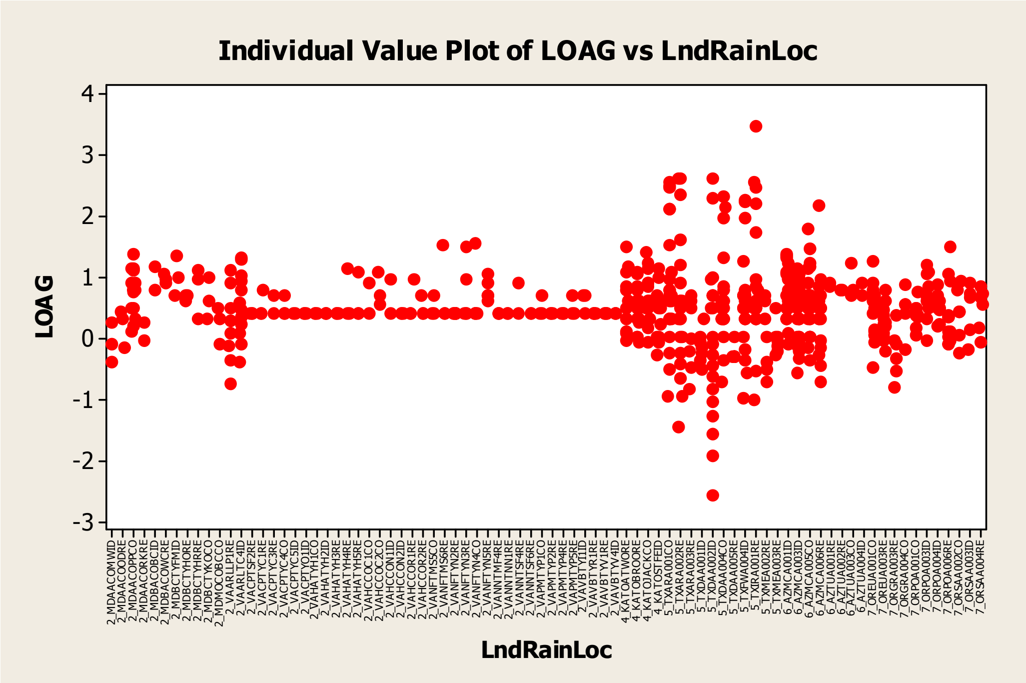
Descriptive Statistics: LOAG

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LOAG	2_KYLOTSR1RE	0	3	*	*	*	*
	2_KYLOTSR2ID	0	3	*	*	*	*
	2_KYLOTSR3RE	0	3	*	*	*	*
	2_KYLOTSR4ID	0	4	*	*	*	*
	2_MDAACOMWID	3	0	-0.0792	0.106	-0.396	0.255
	2_MDAACOODRE	3	0	0.187	0.0909	-0.155	0.415
	2_MDAACOPPCO	22	4	0.7647	0.0919	0.1139	1.3838
	2_MDAACORKRE	2	1	0.105	0.0453	-0.0458	0.255
	2_MDBACOBCID	2	1	0.977	0.0792	0.778	1.176
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	1	2	0.77815	*	0.77815	0.77815
	2_MDBACOWCRE	3	0	0.9662	0.00489	0.9031	1.0414
	2_MDBCTYBOID	1	2	1.1139	*	1.1139	1.1139
	2_MDBCTYFMID	3	0	1.014	0.104	0.699	1.342
	2_MDBCTYHORE	3	0	0.6667	0.00313	0.6021	0.6990
	2_MDBCTYHRRE	3	0	0.790	0.186	0.301	1.114
	2_MDBCTYKOCO	3	0	0.634	0.123	0.301	1.000
	2_MDMOCOBCCO	3	0	0.234	0.0882	-0.0969	0.477

2_VAARLLP1RE	8	0	0.201	0.392	-0.742	1.114
2_VAARLTC4ID	12	1	0.535	0.302	-0.415	1.322
2_VACPTC1ARE	0	8	*	*	*	*
2_VACPTSF2RE	3	0	0.39794	0.000000000	0.39794	0.39794
2_VACPTYC1RE	3	4	0.525	0.0482	0.398	0.778
2_VACPTYC3RE	2	12	0.548	0.0453	0.398	0.699
2_VACPTYC4CO	3	11	0.498	0.0302	0.398	0.699
2_VACPTYC5ID	2	12	0.39794	0.000000000	0.39794	0.39794
2_VACPTYO1ID	3	0	0.39794	0.000000000	0.39794	0.39794
2_VAHATYH1CO	3	15	0.39794	0.000000000	0.39794	0.39794
2_VAHATYH2ID	3	16	0.39794	0.000000000	0.39794	0.39794
2_VAHATYH3RE	3	14	0.39794	0.000000000	0.39794	0.39794
2_VAHATYH4RE	3	14	0.647	0.187	0.398	1.146
2_VAHATYH5RE	3	14	0.625	0.155	0.398	1.079
2_VAHCCOC1CO	2	0	0.651	0.128	0.398	0.903
2_VAHCCOC2CO	3	0	0.779	0.0724	0.559	1.079
2_VAHCCON1ID	2	0	0.676	0.155	0.398	0.954
2_VAHCCON2ID	3	0	0.39794	0.000000000	0.39794	0.39794
2_VAHCCOR1RE	3	0	0.583	0.103	0.398	0.954
2_VAHCCOR2RE	3	0	0.498	0.0302	0.398	0.699
2_VANFTMS5CO	3	0	0.498	0.0302	0.398	0.699
2_VANFTMS6RE	3	0	0.767	0.409	0.398	1.505
2_VANFTMS8CO	1	2	0.39794	*	0.39794	0.39794
2_VANFTMS9CO	1	2	0.39794	*	0.39794	0.39794
2_VANFTYN2RE	7	22	0.39794	0.000000000	0.39794	0.39794
2_VANFTYN3RE	5	22	0.725	0.235	0.398	1.477
2_VANFTYN4CO	5	22	0.630	0.268	0.398	1.556
2_VANFTYN5RE	5	22	0.7868	0.0330	0.5917	1.0414
2_VANNTMF1RE	1	0	0.39794	*	0.39794	0.39794
2_VANNTMF4RE	3	0	0.39794	0.000000000	0.39794	0.39794
2_VANNTNN1RE	2	8	0.39794	0.000000000	0.39794	0.39794
2_VANNTSF4RE	2	0	0.651	0.128	0.398	0.903
2_VANNTSF6RE	2	0	0.39794	0.000000000	0.39794	0.39794
2_VAPMTYP1CO	3	15	0.498	0.0302	0.398	0.699
2_VAPMTYP2RE	3	14	0.39794	0.000000000	0.39794	0.39794
2_VAPMTYP4RE	3	14	0.39794	0.000000000	0.39794	0.39794
2_VAPMTYP5RE	3	14	0.498	0.0302	0.398	0.699
2_VAVBTYI1ID	3	0	0.599	0.0302	0.398	0.699
2_VAVBTYR1RE	4	1	0.39794	0.000000000	0.39794	0.39794
2_VAVBTYV1RE	4	23	0.39794	0.000000000	0.39794	0.39794
2_VAVBTYV4ID	3	27	0.39794	0.000000000	0.39794	0.39794
4_KATOATWORE	15	0	0.572	0.211	-0.0429	1.491
4_KATOBROORE	16	0	0.4925	0.1225	-0.0689	1.0792
4_KATOJACKCO	14	2	0.683	0.173	-0.0679	1.398
4_KATOSTFEID	15	2	0.443	0.162	-0.291	1.146
5_TXARA001CO	22	0	0.858	0.973	-0.962	2.555
5_TXARA002RE	20	1	0.534	1.249	-1.470	2.622
5_TXARA003RE	7	0	0.0255	0.350	-0.855	0.699
5_TXDAA001ID	7	0	-0.151	0.0706	-0.519	0.301
5_TXDAA002ID	19	0	-0.0821	1.798	-2.583	2.611
5_TXDAA004CO	19	1	0.660	0.625	-0.719	2.322
5_TXDAA005RE	7	0	-0.2580	0.0129	-0.3010	0.000000000
5_TXFWA004ID	20	1	0.571	0.735	-0.981	2.265
5_TXIRA001RE	21	1	0.868	1.122	-1.012	3.474
5_TXMEA002RE	7	0	-0.00695	0.318	-0.732	0.699
5_TXMEA003RE	7	0	-0.0314	0.0623	-0.2963	0.4771
6_AZMCA001ID	26	1	0.7199	0.1590	-0.1160	1.3802
6_AZMCA003ID	27	0	0.4549	0.2412	-0.5713	1.1461
6_AZMCA005CO	26	0	0.5869	0.2482	-0.3743	1.7782
6_AZMCA006RE	20	0	0.403	0.365	-0.732	2.176
6_AZTUA001RE	2	8	0.8741	0.00168	0.8451	0.9031
6_AZTUA002RE	2	6	0.77815	0.000000000	0.77815	0.77815
6_AZTUA003CO	3	5	0.903	0.0822	0.699	1.230

6_AZTUA004ID	2	5	0.801	0.0208	0.699	0.903
6_CAALAL09ID	0	9	*	*	*	*
7_OREUA001CO	14	2	0.443	0.204	-0.476	1.255
7_OREUA003RE	14	1	0.2645	0.0742	-0.2218	0.7782
7_ORGRA003RE	6	0	-0.247	0.168	-0.804	0.362
7_ORGRA004CO	5	1	0.435	0.157	-0.203	0.869
7_ORPOA001CO	12	1	0.3800	0.0301	0.0562	0.7559
7_ORPOA003ID	12	2	0.670	0.120	-0.0486	1.204
7_ORPOA004ID	12	1	0.5210	0.0332	0.2041	0.8633
7_ORPOA006RE	12	1	0.374	0.263	-0.0969	1.491
7_ORSAA002CO	6	0	0.395	0.196	-0.253	0.934
7_ORSAA003ID	5	1	0.441	0.211	-0.205	0.903
7_ORSAA004RE	6	0	0.478	0.130	-0.0866	0.851





General Linear Model: LOAG versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	5	2, 4, 5, 6, 7

Analysis of Variance for LOAG, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	3.8655	3.3711	1.6856	4.60	0.010
EPA_Rain_Zone	4	3.1071	2.0783	0.5196	1.42	0.227
Landuse*EPA_Rain_Zone	8	6.2668	6.2668	0.7833	2.14	0.031
Error	593	217.4563	217.4563	0.3667		
Total	607	230.6956				

Land use and the interaction land use * EPA rain zone are significant.

Bonferroni Simultaneous Tests
Response Variable LOAG
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.1630	0.07104	-2.294	0.0664
RE	-0.1945	0.06667	-2.917	0.0110

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.03148	0.06549	-0.4807	1.000

```
Tukey Simultaneous Tests
Response Variable LOAG
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID          -0.1630     0.07104   -2.294    0.0566
RE          -0.1945     0.06667   -2.917    0.0099


Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE         -0.03148     0.06549  -0.4807    0.8804
```

Oil and grease in residential is different than in commercial land use. There is not enough evidence to prove a difference between commercial and industrial or between residential and industrial.

```
Bonferroni Simultaneous Tests
Response Variable LOAG
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
4                  0.0028     0.09617    0.029    1.0000
5                 -0.0998     0.06912   -1.443    1.0000
6                  0.0124     0.07843    0.158    1.0000
7                 -0.1381     0.07672   -1.800    0.7242

EPA_Rain_Zone = 4  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
5                 -0.1026     0.09701   -1.057     1.000
6                  0.0096     0.10385    0.092     1.000
7                 -0.1409     0.10256   -1.374     1.000

EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
6                 0.11214     0.07946   1.4113     1.000
7                -0.03832     0.07777  -0.4927     1.000

EPA_Rain_Zone = 6  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
7                 -0.1505     0.08615   -1.746    0.8125


Tukey Simultaneous Tests
Response Variable LOAG
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:
```

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.0028	0.09617	0.029	1.0000
5	-0.0998	0.06912	-1.443	0.5996
6	0.0124	0.07843	0.158	0.9999
7	-0.1381	0.07672	-1.800	0.3737

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.1026	0.09701	-1.057	0.8283
6	0.0096	0.10385	0.092	1.0000
7	-0.1409	0.10256	-1.374	0.6446

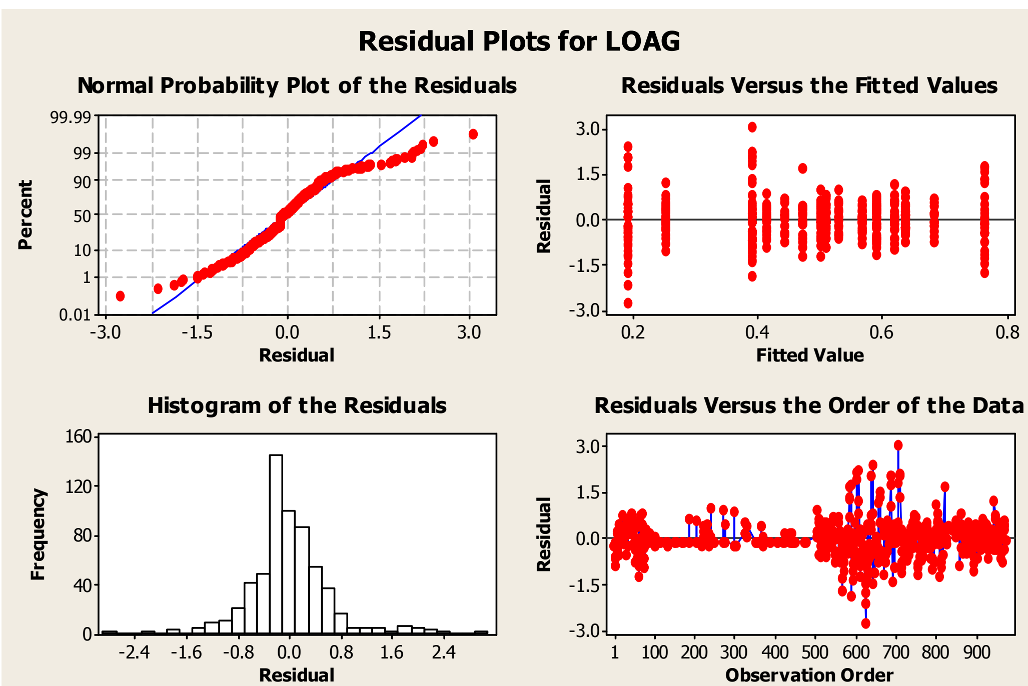
EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.11214	0.07946	1.4113	0.6203
7	-0.03832	0.07777	-0.4927	0.9881

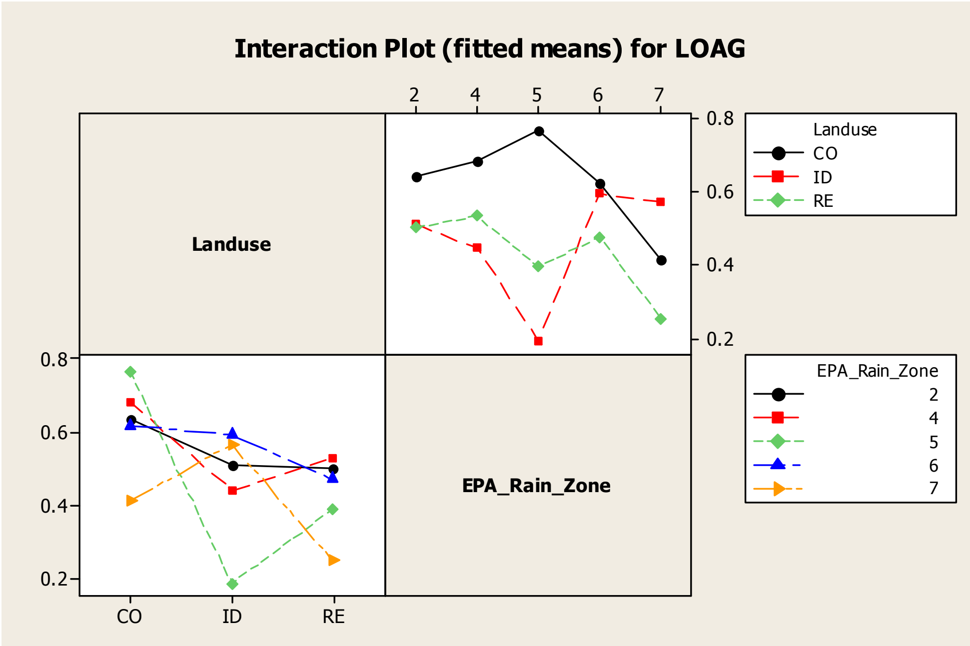
EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.1505	0.08615	-1.746	0.4054

No differences among EPA rain zones for oil and grease.



Residuals fail normality. Tails are larger compared with the normal distribution. A trend was observed in the residuals. Several observations were observed at the same value.


Interaction Plot (fitted means) for LOAG
Landuse
EPA_Rain_Zone
2
4
5
6
7
0.8
0.6
0.4
0.2
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
4
5
6
7

Total Dissolved Solids

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N non detected	Mean	StDev	Minimum	Maximum
2	82	41	1.83	0.32	1.36	3.59
3	6	0	1.67	0.22	1.46	2.08
4	15	1	2.20	0.28	1.81	2.94
5	39	3	1.70	0.17	1.36	2.08
6	20	14	1.98	0.28	1.52	2.58
7	37	4	1.62	0.40	0.60	2.09

Land use: Industrial

EPA_Rain_Zone	N det/est	N non detected	Mean	StDev	Minimum	Maximum
2	86	23	1.81	0.50	0.65	4.05
3	16	0	1.97	0.30	1.15	2.35
4	16	1	2.00	0.28	1.48	2.72
5	43	4	1.88	0.20	1.43	2.38
6	56	14	2.07	0.24	1.20	2.57
7	24	9	1.79	0.32	0.85	2.19

Land use: Residential

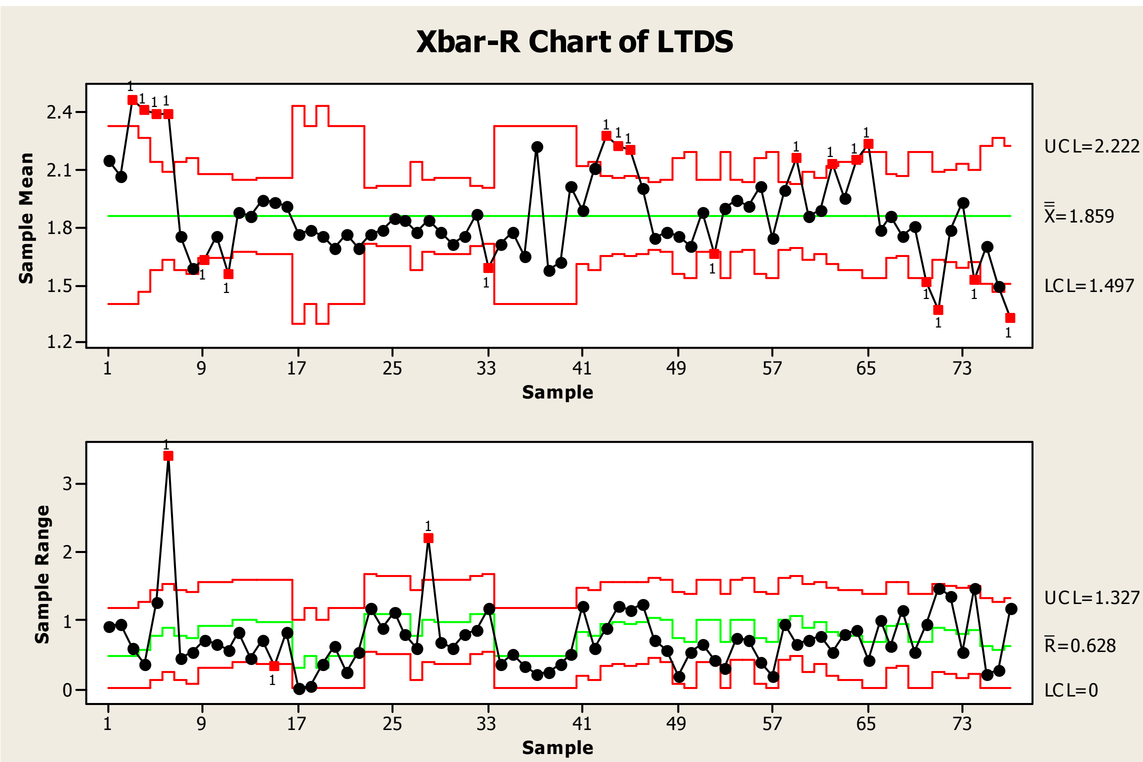
EPA_Rain_Zone	N det/est	N non detected	Mean	StDev	Minimum	Maximum
2	268	63	1.83	0.27	1.04	3.23
3	17	1	1.87	0.30	1.43	2.32
4	31	0	2.26	0.30	1.77	3.04
5	64	7	1.85	0.18	1.52	2.28
6	34	4	1.98	0.22	1.62	2.45
7	37	3	1.64	0.38	0.48	2.24

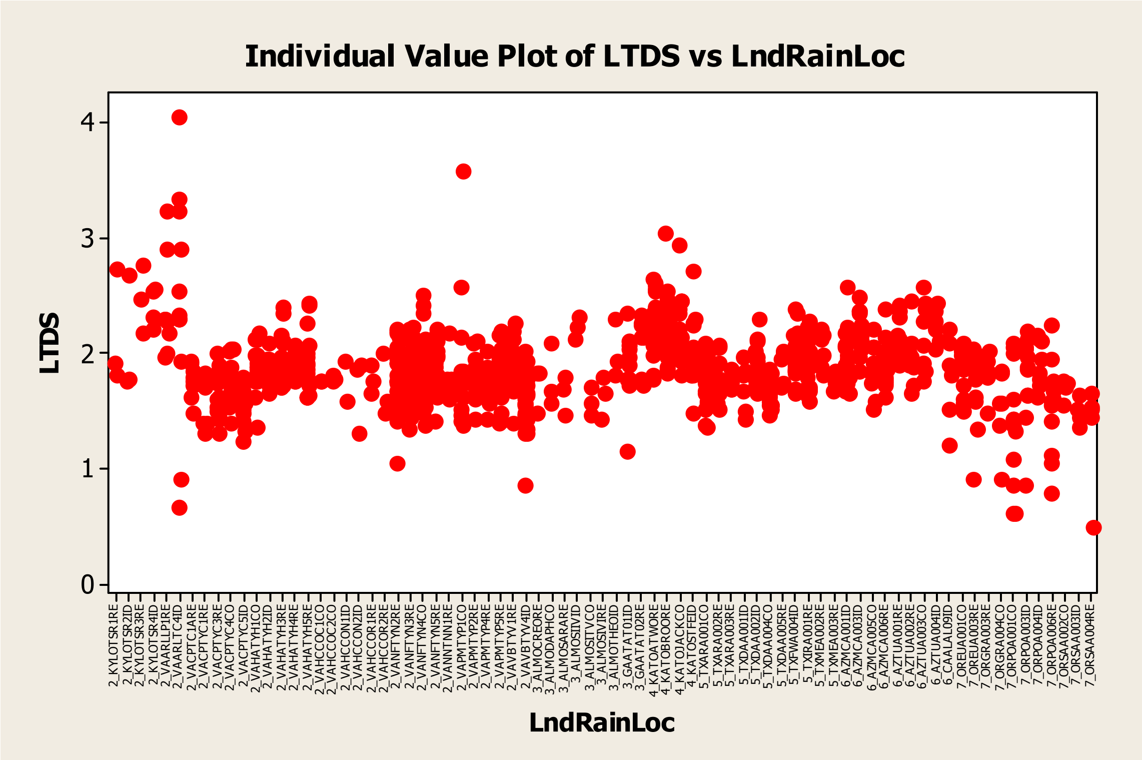
Descriptive Statistics: LTDS

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LTDS	2_KYLOTSR1RE	3	0	2.149	0.254	1.806	2.728
	2_KYLOTSR2ID	3	0	2.070	0.283	1.748	2.685
	2_KYLOTSR3RE	3	0	2.469	0.0871	2.173	2.763
	2_KYLOTSR4ID	4	0	2.4105	0.0289	2.2148	2.5635
	2_MDAACOMWID	1	2	1.5185	*	1.5185	1.5185
	2_MDAACOODRE	0	3	*	*	*	*
	2_MDAACOPPCO	0	26	*	*	*	*
	2_MDAACORKRE	0	3	*	*	*	*
	2_MDBACOBCID	0	3	*	*	*	*
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	0	3	*	*	*	*
	2_MDBACOWCRE	0	3	*	*	*	*
	2_MDBCTYBOID	0	3	*	*	*	*
	2_MDBCTYFMID	0	3	*	*	*	*
	2_MDBCTYHORE	0	3	*	*	*	*
	2_MDBCTYHRRE	0	3	*	*	*	*
	2_MDBCTYKOCO	0	3	*	*	*	*
	2_MDMOCOBCCO	0	3	*	*	*	*
	2_VAARLLP1RE	8	0	2.394	0.195	1.968	3.230
	2_VAARLTC4ID	12	1	2.397	0.920	0.653	4.049

2_VACPTC1ARE	8	0	1.7539	0.0230	1.4771	1.9243
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	1.5817	0.0433	1.3010	1.8195
2_VACPTYC3RE	14	0	1.6238	0.0404	1.3010	2.0043
2_VACPTYC4CO	14	0	1.7555	0.0379	1.3979	2.0374
2_VACPTYC5ID	14	0	1.5502	0.0264	1.2304	1.7853
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	1.8825	0.0395	1.3617	2.1703
2_VAHATYH2ID	18	1	1.8559	0.0104	1.6532	2.0864
2_VAHATYH3RE	17	0	1.9384	0.0498	1.6990	2.3962
2_VAHATYH4RE	17	0	1.9273	0.00830	1.7482	2.0682
2_VAHATYH5RE	17	0	1.9079	0.0691	1.6128	2.4440
2_VAHCCOC1CO	2	0	1.7634	0.000000000	1.7634	1.7634
2_VAHCCOC2CO	3	0	1.7826	0.000472	1.7634	1.8062
2_VAHCCON1ID	2	0	1.757	0.0629	1.580	1.934
2_VAHCCON2ID	3	0	1.687	0.112	1.301	1.903
2_VAHCCOR1RE	3	0	1.7663	0.0155	1.6435	1.8921
2_VAHCCOR2RE	3	0	1.689	0.0795	1.477	2.009
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	1.7616	0.0696	1.0414	2.2122
2_VANFTYN3RE	27	0	1.7805	0.0613	1.3424	2.2304
2_VANFTYN4CO	27	0	1.8441	0.0863	1.3802	2.4969
2_VANFTYN5RE	27	0	1.8353	0.0423	1.4150	2.2175
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	1.7778	0.0297	1.6021	2.1790
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	1.838	0.273	1.380	3.587
2_VAPMTYP2RE	17	0	1.7789	0.0284	1.4314	2.1106
2_VAPMTYP4RE	17	0	1.7073	0.0185	1.4314	2.0212
2_VAPMTYP5RE	17	0	1.7584	0.0541	1.3979	2.1875
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	1.8700	0.0339	1.3979	2.2553
2_VAVBTYV4ID	29	1	1.5823	0.0581	0.8451	2.0253
3_ALMOCREORE	3	0	1.712	0.0414	1.477	1.833
3_ALMODAPHCO	3	0	1.773	0.0729	1.568	2.079
3_ALMOSARARE	3	0	1.6477	0.0292	1.4624	1.7993
3_ALMOSIIVID	3	0	2.2222	0.0109	2.1139	2.3222
3_ALMOSITVCO	3	0	1.5765	0.0140	1.4624	1.6990
3_ALMOSIVIRE	3	0	1.620	0.0317	1.431	1.785
3_ALMOTHEOID	3	0	2.016	0.0646	1.813	2.301
3_GAATAT01ID	10	0	1.885	0.102	1.147	2.350
3_GAATAT02RE	8	1	2.1059	0.0522	1.7282	2.3242
4_KATOATWORE	15	0	2.2816	0.0828	1.7709	2.6474
4_KATOBROORE	16	0	2.2303	0.0962	1.8325	3.0398
4_KATOJACKCO	15	1	2.2009	0.0807	1.8062	2.9445
4_KATOSTFEID	16	1	2.0001	0.0786	1.4771	2.7193
5_TXARA001CO	21	1	1.7418	0.0397	1.3617	2.0792
5_TXARA002RE	18	3	1.7790	0.0263	1.5185	2.0719
5_TXARA003RE	7	0	1.7510	0.00375	1.6812	1.8633
5_TXDAA001ID	6	1	1.7012	0.0441	1.4314	1.9638
5_TXDAA002ID	18	1	1.8823	0.0268	1.6435	2.2967
5_TXDAA004CO	18	2	1.6562	0.0153	1.4624	1.8692
5_TXDAA005RE	6	1	1.9004	0.0201	1.7076	2.0128
5_TXFWA004ID	19	2	1.9418	0.0428	1.6532	2.3766
5_TXIRA001RE	20	2	1.9071	0.0409	1.5911	2.2810
5_TXMEA002RE	7	0	2.0181	0.0217	1.8129	2.2014
5_TXMEA003RE	6	1	1.7478	0.00446	1.6628	1.8388

6_AZMCA001ID	20	7	1.9987	0.0560	1.6435	2.5717
6_AZMCA003ID	24	3	2.1602	0.0191	1.8451	2.4857
6_AZMCA005CO	12	14	1.8548	0.0395	1.5185	2.2148
6_AZMCA006RE	16	4	1.8932	0.0345	1.6232	2.3909
6_AZTUA001RE	10	0	2.1314	0.0350	1.9031	2.4265
6_AZTUA002RE	8	0	1.9557	0.0591	1.6532	2.4472
6_AZTUA003CO	8	0	2.156	0.0933	1.748	2.583
6_AZTUA004ID	6	1	2.2375	0.0233	2.0414	2.4378
6_CAALAL09ID	6	3	1.785	0.139	1.204	2.204
7_OREUA001CO	14	2	1.8631	0.0375	1.4914	2.0934
7_OREUA003RE	15	0	1.7557	0.0890	0.9031	2.0414
7_ORGRA003RE	6	0	1.8070	0.0336	1.4771	2.0128
7_ORGRA004CO	6	0	1.515	0.121	0.903	1.845
7_ORPOA001CO	12	1	1.368	0.269	0.602	2.079
7_ORPOA003ID	11	3	1.788	0.151	0.845	2.188
7_ORPOA004ID	9	4	1.9331	0.0399	1.6232	2.1523
7_ORPOA006RE	11	2	1.519	0.175	0.775	2.243
7_ORSAA002CO	5	1	1.6984	0.00670	1.5563	1.7634
7_ORSAA003ID	4	2	1.4869	0.0130	1.3617	1.6335
7_ORSAA004RE	5	1	1.321	0.228	0.477	1.643





General Linear Model: LTDS versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	6	2, 3, 4, 5, 6, 7

Analysis of Variance for LTDS, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	1.13036	0.50905	0.25453	2.74	0.065
EPA_Rain_Zone	5	12.80378	10.39401	2.07880	22.37	0.000
Landuse*EPA_Rain_Zone	10	2.14609	2.14609	0.21461	2.31	0.011
Error	872	81.04647	81.04647	0.09294		
Total	889	97.12670				

Not significant by land use
Significant by rain zone
Significant by land use and rain zone interaction.

Bonferroni Simultaneous Tests
Response Variable LTDS
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.08739	0.03829	2.283	0.0681
RE	0.06801	0.03647	1.865	0.1875

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.01937	0.03159	-0.6134	1.000

Tukey Simultaneous Tests
Response Variable LTDS
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.08739	0.03829	2.283	0.0583
RE	0.06801	0.03647	1.865	0.1489

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.01937	0.03159	-0.6134	0.8128

No significant differences by land use

Bonferroni Simultaneous Tests
Response Variable LTDS
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	0.0143	0.05710	0.250	1.0000
4	0.3272	0.04419	7.403	0.0000
5	-0.0128	0.03086	-0.415	1.0000
6	0.1830	0.03592	5.094	0.0000
7	-0.1411	0.03570	-3.952	0.0013

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.3129	0.06813	4.593	0.0001
5	-0.0271	0.06034	-0.449	1.0000
6	0.1687	0.06308	2.675	0.1143
7	-0.1554	0.06295	-2.468	0.2066

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.3400	0.04830	-7.038	0.0000
6	-0.1442	0.05169	-2.789	0.0809
7	-0.4683	0.05153	-9.087	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value

```
6                  0.1958     0.04088    4.790    0.0000
7                 -0.1283     0.04068   -3.153    0.0250


EPA_Rain_Zone = 6  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
7                 -0.3241     0.04464   -7.259    0.0000
```

TDS in EPA rain zone 2 is significantly different than in rain zones 4, 6, and 7. No significant differences between EPA rain zone 2 and rain zones 3 and 5.

TDS in EPA rain zone 3 is significantly different than in rain zones 4, 6, and 7. No significant differences between EPA rain zone 3 and rain zone 5.

TDS in EPA rain zone 4 is significantly different than in rain zones 5, and 7. No significant differences between EPA rain zone 4 and rain zone 6.

TDS in EPA rain zone 5 is significantly different than in rain zones 6, and 7.

TDS in EPA rain zone 6 is significantly different than in rain zone 7.

```
Tukey Simultaneous Tests
Response Variable LTDS
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
3                  0.0143     0.05710    0.250    0.9999
4                  0.3272     0.04419    7.403    0.0000
5                 -0.0128     0.03086   -0.415    0.9984
6                  0.1830     0.03592    5.094    0.0000
7                 -0.1411     0.03570   -3.952    0.0011


EPA_Rain_Zone = 3  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
4                  0.3129     0.06813    4.593    0.0001
5                 -0.0271     0.06034   -0.449    0.9977
6                  0.1687     0.06308    2.675    0.0803
7                 -0.1554     0.06295   -2.468    0.1336


EPA_Rain_Zone = 4  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
5                 -0.3400     0.04830   -7.038    0.0000
6                 -0.1442     0.05169   -2.789    0.0591
7                 -0.4683     0.05153   -9.087    0.0000


EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
```

```
6                   0.1958     0.04088    4.790    0.0000
7                  -0.1283     0.04068   -3.153    0.0201


EPA_Rain_Zone = 6  subtracted from:

               Difference        SE of             Adjusted
EPA_Rain_Zone    of Means   Difference  T-Value     P-Value
7                 -0.3241      0.04464   -7.259      0.0000
```

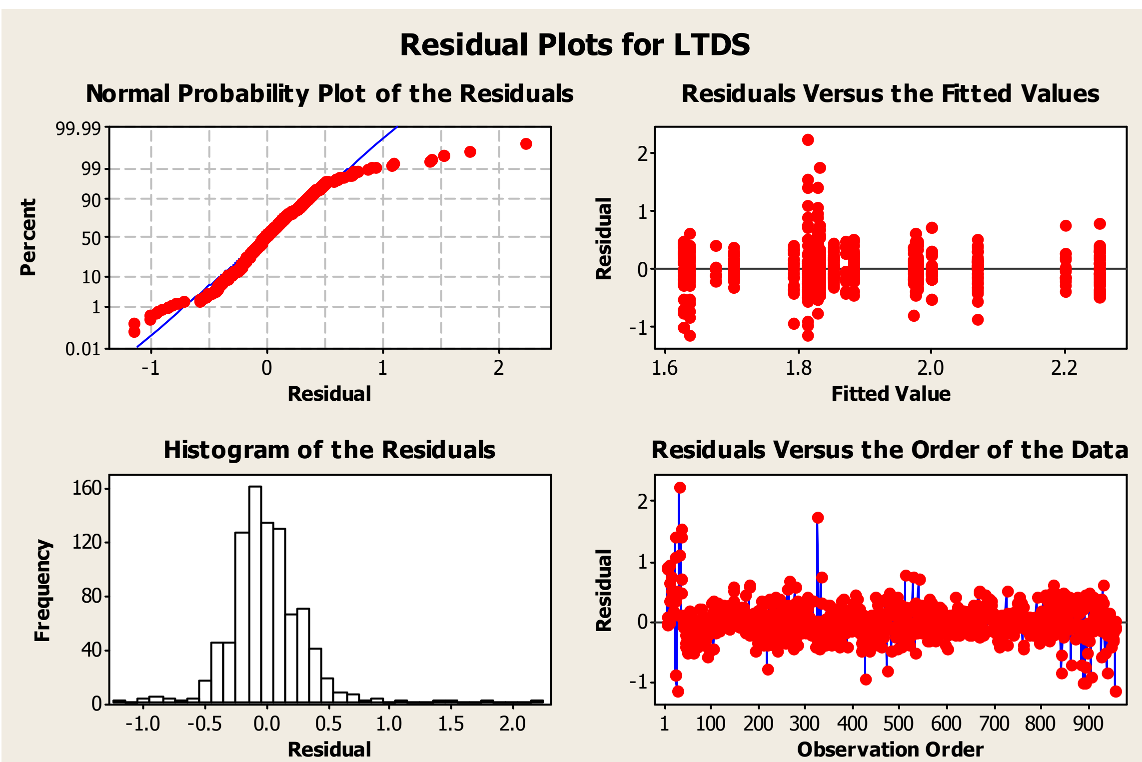
TDS in EPA rain zone 2 is significantly different than in rain zones 4, 6, and 7. No significant differences between EPA rain zone 2 and rain zones 3 and 5.

TDS in EPA rain zone 3 is significantly different than in rain zone 4. No significant differences between EPA rain zone 3 and rain zone 5, 6 and 7.

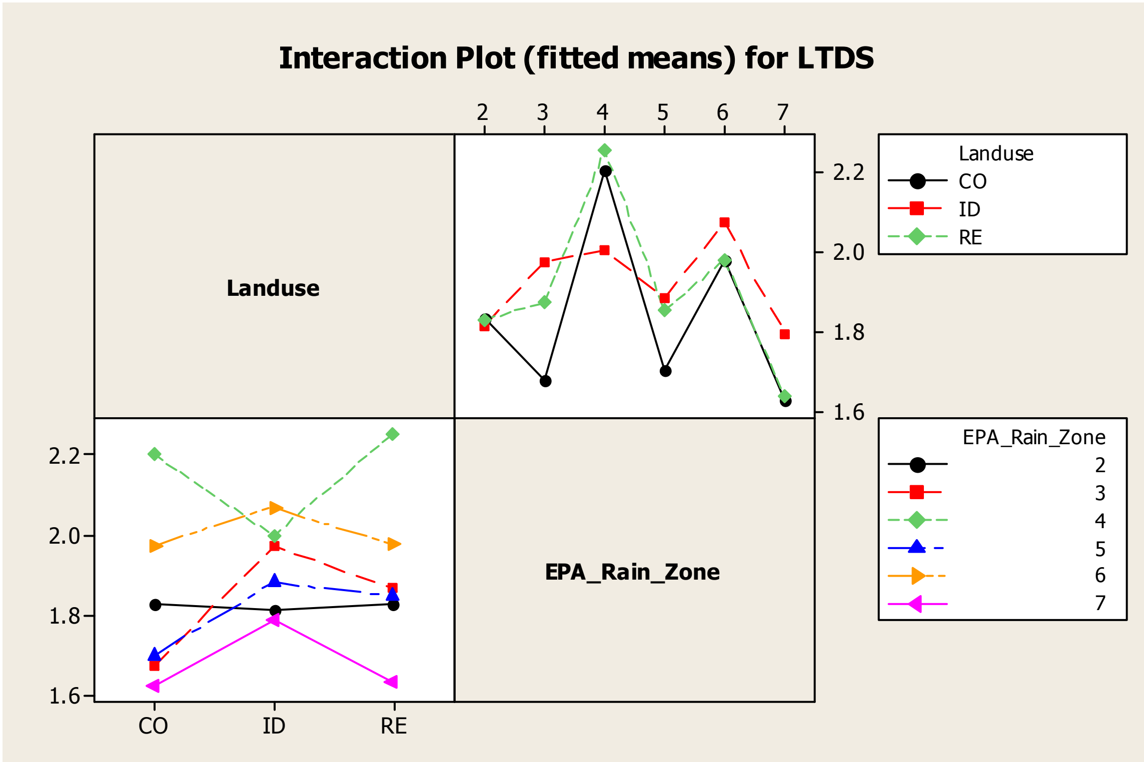
TDS in EPA rain zone 4 is significantly different than in rain zones 5, and 7. No significant differences between EPA rain zone 4 and rain zone 6.

TDS in EPA rain zone 5 is significantly different than in rain zones 6, and 7.

TDS in EPA rain zone 6 is significantly different than in rain zone 7.



Sites in Kentucky and Arlington Virginia are different than the remaining sites in the database.



Interaction Plot (fitted means) for LTDS
Landuse
EPA_Rain_Zone
2
3
4
5
6
7
2.2
2.0
1.8
1.6
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7

Total Suspended Solids

Summary statistics in LOG base 10 scale

Land use: Commercial

EPA_Rain_Zone	N det/est	N non detected	Mean	StDev	Minimum	Maximum
2	113	10	1.59	0.36	0.83	2.80
3	6	0	1.83	0.27	1.40	2.04
4	15	1	2.82	0.31	2.25	3.38
5	40	2	1.50	0.54	0.30	2.81
6	20	14	1.87	0.45	0.90	2.71
7	37	4	1.75	0.39	0.90	2.58

Land use: Industrial

EPA_Rain_Zone	N det/est	N non detected	Mean	StDev	Minimum	Maximum
2	101	8	1.6434	0.4385	0.4771	2.5185
3	16	0	1.351	0.809	-0.367	2.505
4	16	1	2.177	0.617	1.322	3.073
5	43	4	2.1667	0.4113	1.3802	3.3962
6	57	13	2.458	0.5143	1.2041	3.3664
7	24	9	2.0825	0.3837	1.2041	3.0334

Land use: Residential

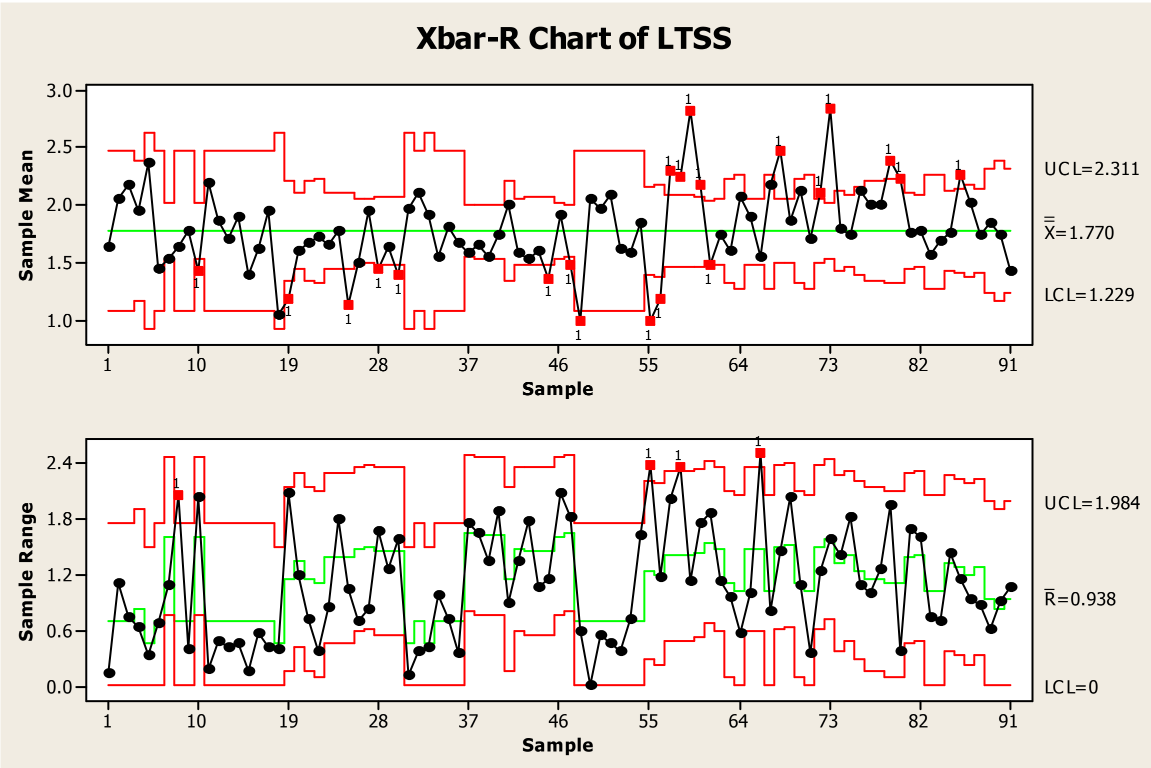
EPA_Rain_Zone	N det/est	N non sampled	Mean	StDev	Minimum	Maximum
2	309	22	1.6047	0.4328	0.4074	2.9154
3	18	0	1.346	0.454	0.623	2.255
4	30	1	2.2706	0.5227	1.0414	3.3913
5	64	7	1.8435	0.3878	0.6021	2.7839
6	33	5	1.9127	0.4257	0.4771	2.5441
7	37	3	1.681	0.3785	0.8451	2.8791

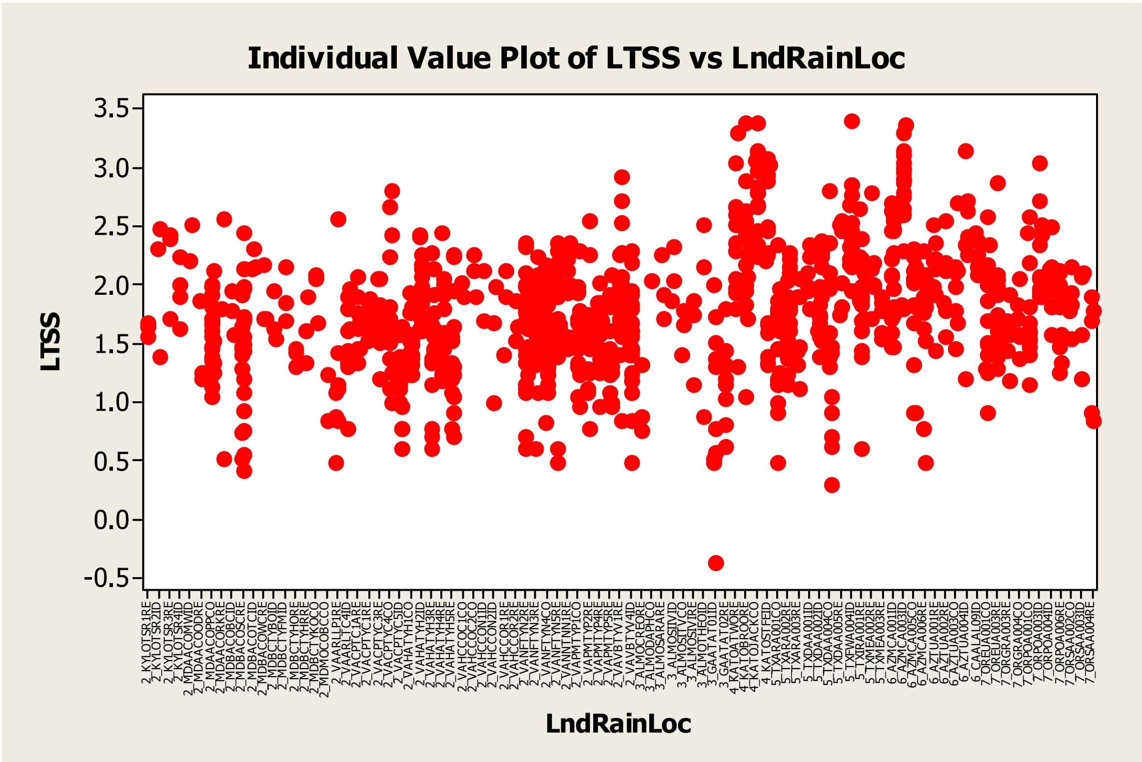
Descriptive Statistics: LTSS

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LTSS	2_KYLOTSR1RE	3	0	1.6270	0.00411	1.5563	1.6812
	2_KYLOTSR2ID	3	0	2.054	0.348	1.380	2.477
	2_KYLOTSR3RE	3	0	2.179	0.167	1.708	2.435
	2_KYLOTSR4ID	4	0	1.942	0.0652	1.623	2.241
	2_MDAACOMWID	2	1	2.361	0.0494	2.204	2.519
	2_MDAACOODRE	3	0	1.443	0.137	1.204	1.869
	2_MDAACOPPCO	26	0	1.5368	0.0803	1.0406	2.1276
	2_MDAACORKRE	3	0	1.625	1.077	0.513	2.568
	2_MDBACOBCID	3	0	1.764	0.0373	1.568	1.954
	2_MDBACOSCRE	26	0	1.427	0.280	0.407	2.443
	2_MDBACOTCID	3	0	2.1936	0.00870	2.1335	2.3010
	2_MDBACOWCRE	3	0	1.867	0.0719	1.708	2.176
	2_MDBCTYBOID	3	0	1.704	0.0450	1.544	1.944
	2_MDBCTYFMID	3	0	1.894	0.0537	1.690	2.146
	2_MDBCTYHORE	3	0	1.3877	0.00590	1.3010	1.4472
	2_MDBCTYHRRE	3	0	1.616	0.0787	1.342	1.903
	2_MDBCTYKOCO	3	0	1.937	0.0529	1.672	2.083
	2_MDMOCOBCCO	2	1	1.038	0.0742	0.845	1.230

2_VAARLLP1RE	8	0	1.189	0.380	0.477	2.556
2_VAARLTC4ID	13	0	1.5970	0.1206	0.7782	1.9638
2_VACPTC1ARE	8	0	1.6618	0.0595	1.3424	2.0645
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	1.7275	0.0159	1.5185	1.8808
2_VACPTYC3RE	14	0	1.6436	0.0620	1.2041	2.0569
2_VACPTYC4CO	14	0	1.766	0.316	1.000	2.799
2_VACPTYC5ID	14	0	1.1237	0.0902	0.6021	1.6435
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	1.4983	0.0383	1.2304	1.9243
2_VAHATYH2ID	19	0	1.9411	0.0732	1.6021	2.4265
2_VAHATYH3RE	17	0	1.437	0.212	0.603	2.274
2_VAHATYH4RE	17	0	1.6390	0.1167	1.1761	2.4378
2_VAHATYH5RE	17	0	1.383	0.206	0.699	2.265
2_VAHCCOC1CO	2	0	1.9552	0.00664	1.8976	2.0128
2_VAHCCOC2CO	3	0	2.094	0.0330	1.903	2.265
2_VAHCCON1ID	2	0	1.908	0.0875	1.699	2.117
2_VAHCCON2ID	3	0	1.551	0.252	1.000	1.982
2_VAHCCOR1RE	3	0	1.801	0.134	1.398	2.114
2_VAHCCOR2RE	3	0	1.675	0.0305	1.519	1.863
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	1.5865	0.1854	0.6021	2.3560
2_VANFTYN3RE	27	0	1.6417	0.1088	0.6021	2.2405
2_VANFTYN4CO	27	0	1.5526	0.1111	0.8264	2.1673
2_VANFTYN5RE	27	0	1.7382	0.2002	0.4771	2.3598
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	1.997	0.113	1.462	2.352
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	1.5758	0.1205	0.9542	2.2900
2_VAPMTYP2RE	17	0	1.535	0.178	0.778	2.548
2_VAPMTYP4RE	17	0	1.5970	0.0711	0.9542	2.0212
2_VAPMTYP5RE	17	0	1.3599	0.1114	0.9542	2.0934
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	1.9143	0.1752	0.8451	2.9154
2_VAVBTYV4ID	29	1	1.4765	0.1413	0.4771	2.2923
3_ALMOCREORE	3	0	0.984	0.0903	0.748	1.322
3_ALMODAPHCO	3	0	2.0414	0.000000000	2.0414	2.0414
3_ALMOSARARE	3	0	1.959	0.0765	1.708	2.255
3_ALMOSIIVID	3	0	2.076	0.0535	1.863	2.322
3_ALMOSITVCO	3	0	1.611	0.0368	1.398	1.771
3_ALMOSIVIRE	3	0	1.583	0.147	1.146	1.863
3_ALMOTHEOID	3	0	1.844	0.728	0.881	2.505
3_GAATAT01ID	10	0	0.986	0.513	-0.367	2.000
3_GAATAT02RE	9	0	1.183	0.119	0.623	1.794
4_KATOATWORE	15	0	2.293	0.269	1.301	3.299
4_KATOBROORE	15	1	2.248	0.296	1.041	3.391
4_KATOJACKCO	15	1	2.8203	0.0981	2.2455	3.3775
4_KATOSTFEID	16	1	2.177	0.381	1.322	3.073
5_TXARA001CO	22	0	1.4698	0.1748	0.4771	2.3385
5_TXARA002RE	18	3	1.7430	0.1094	1.1461	2.2742
5_TXARA003RE	7	0	1.598	0.123	1.114	2.057
5_TXDAA001ID	6	1	2.0691	0.0400	1.7924	2.3444
5_TXDAA002ID	18	1	1.8856	0.1023	1.3802	2.3729
5_TXDAA004CO	18	2	1.544	0.455	0.301	2.806
5_TXDAA005RE	6	1	2.179	0.134	1.740	2.547
5_TXFWA004ID	19	2	2.4638	0.1117	1.9445	3.3962
5_TXIRA001RE	20	2	1.8656	0.1851	0.6021	2.6415

5_TXMEA002RE	7	0	2.115	0.111	1.708	2.784
5_TXMEA003RE	6	1	1.7057	0.0222	1.5441	1.8921
6_AZMCA001ID	19	8	2.1062	0.1720	1.4771	2.6998
6_AZMCA003ID	24	3	2.8248	0.1505	1.7924	3.3664
6_AZMCA005CO	12	14	1.781	0.241	0.903	2.316
6_AZMCA006RE	15	5	1.730	0.240	0.477	2.294
6_AZTUA001RE	10	0	2.1178	0.0873	1.4314	2.5185
6_AZTUA002RE	8	0	1.998	0.0974	1.556	2.544
6_AZTUA003CO	8	0	1.997	0.141	1.462	2.708
6_AZTUA004ID	7	0	2.385	0.359	1.204	3.140
6_CAALAL09ID	7	2	2.2286	0.0190	2.0792	2.4472
7_OREUA001CO	14	2	1.758	0.225	0.903	2.580
7_OREUA003RE	15	0	1.774	0.166	1.279	2.879
7_ORGRA003RE	6	0	1.566	0.0616	1.176	1.914
7_ORGRA004CO	6	0	1.6843	0.0584	1.3617	2.0531
7_ORPOA001CO	12	1	1.748	0.188	1.146	2.580
7_ORPOA003ID	11	3	2.259	0.144	1.892	3.033
7_ORPOA004ID	9	4	2.0203	0.0659	1.5682	2.5011
7_ORPOA006RE	11	2	1.7331	0.0922	1.2553	2.1139
7_ORSAA002CO	5	1	1.8444	0.0486	1.5441	2.1492
7_ORSAA003ID	4	2	1.737	0.186	1.204	2.111
7_ORSAA004RE	5	1	1.427	0.261	0.845	1.903





General Linear Model: LTSS versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	6	2, 3, 4, 5, 6, 7

Analysis of Variance for LTSS, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	13.6706	3.9665	1.9833	10.20	0.000
EPA_Rain_Zone	5	54.4527	45.5820	9.1164	46.88	0.000
Landuse*EPA_Rain_Zone	10	18.7052	18.7052	1.8705	9.62	0.000
Error	961	186.8854	186.8854	0.1945		
Total	978	273.7139				

Both main factors and interactions are significant.

Bonferroni Simultaneous Tests
Response Variable LTSS
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.0857	0.05508	1.556	0.3600
RE	-0.1177	0.05245	-2.243	0.0753

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.2034	0.04545	-4.474	0.0000

Tukey Simultaneous Tests
Response Variable LTSS
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.0857	0.05508	1.556	0.2649
RE	-0.1177	0.05245	-2.243	0.0642

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.2034	0.04545	-4.474	0.0000

No difference in TSS concentrations between commercial and industrial land uses, and commercial and residential. TSS from industrial land uses is different than TSS from residential land use.

Bonferroni Simultaneous Tests
Response Variable LTSS
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1060	0.08141	-1.303	1.0000
4	0.8089	0.06314	12.812	0.0000
5	0.2241	0.04307	5.203	0.0000
6	0.4657	0.05089	9.151	0.0000
7	0.2256	0.05043	4.473	0.0001

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.9150	0.09830	9.307	0.0000
5	0.3302	0.08679	3.804	0.0023
6	0.5717	0.09092	6.288	0.0000
7	0.3316	0.09067	3.658	0.0040

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.5848	0.06994	-8.361	0.0000
6	-0.3432	0.07500	-4.576	0.0001
7	-0.5833	0.07470	-7.809	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.241542	0.05911	4.08597	0.0007
7	0.001458	0.05873	0.02483	1.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.2401	0.06467	-3.712	0.0033

TSS in EPA rain zone 2 is significantly different than in rain zones 4, 5, 6, and 7. No significant differences between EPA rain zone 2 and rain zone 3.

TSS in EPA rain zone 3 is significantly different than in rain zones 4, 5, 6, and 7

TSS in EPA rain zone 4 is significantly different than in rain zones 5, 6, and 7

TSS in EPA rain zone 5 is significantly different than in rain zone 6. No significant differences between EPA rain zone 5 and rain zone 7.

TSS in EPA rain zone 6 is significantly different than in rain zone 7.

Tukey Simultaneous Tests
Response Variable LTSS
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1060	0.08141	-1.303	0.7838
4	0.8089	0.06314	12.812	0.0000
5	0.2241	0.04307	5.203	0.0000
6	0.4657	0.05089	9.151	0.0000
7	0.2256	0.05043	4.473	0.0001

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.9150	0.09830	9.307	0.0000
5	0.3302	0.08679	3.804	0.0020
6	0.5717	0.09092	6.288	0.0000
7	0.3316	0.09067	3.658	0.0035

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.5848	0.06994	-8.361	0.0000
6	-0.3432	0.07500	-4.576	0.0001
7	-0.5833	0.07470	-7.809	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.241542	0.05911	4.08597	0.0006
7	0.001458	0.05873	0.02483	1.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.2401	0.06467	-3.712	0.0028

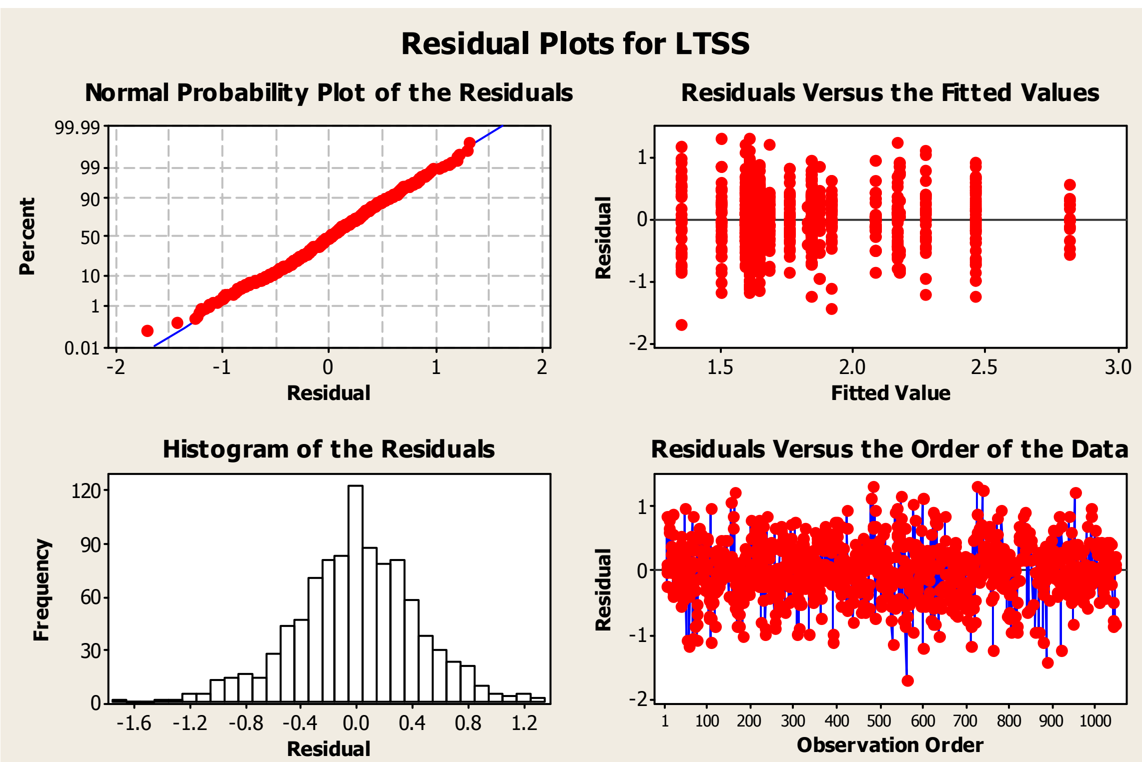
TSS in EPA rain zone 2 is significantly different than in rain zones 4, 5, 6, and 7. No significant differences between EPA rain zone 2 and rain zone 3.

TSS in EPA rain zone 3 is significantly different than in rain zones 4, 5, 6, and 7

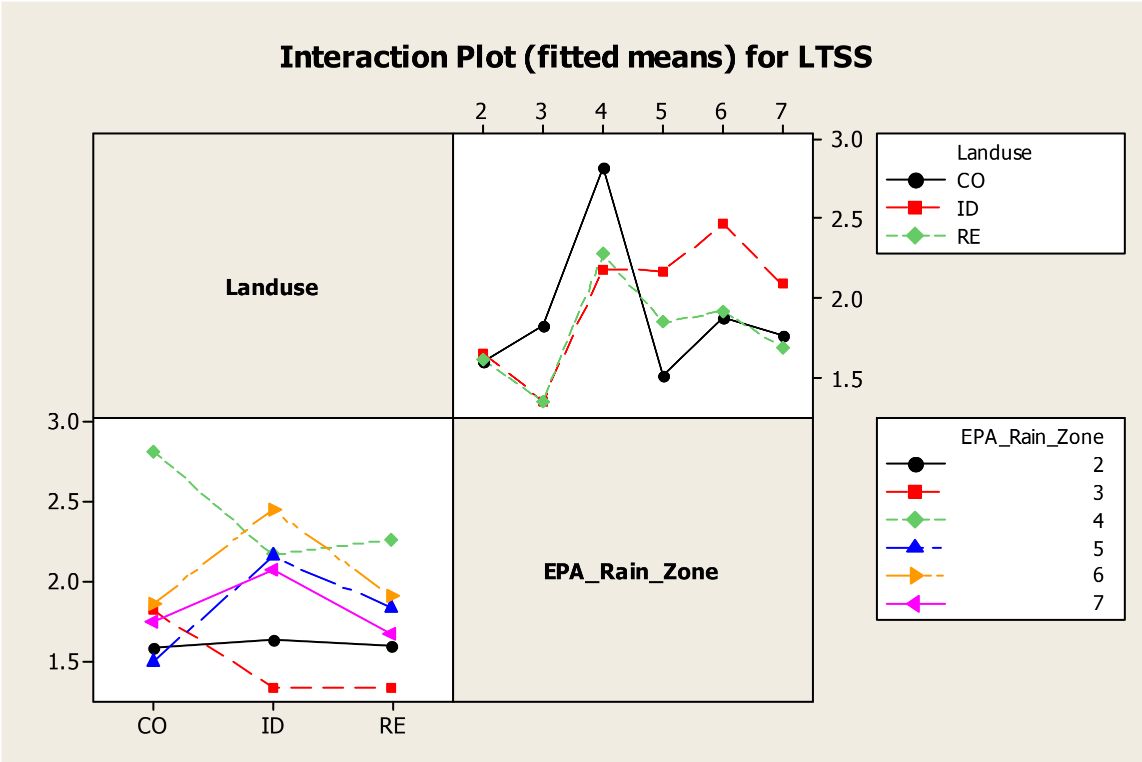
TSS in EPA rain zone 4 is significantly different than in rain zones 5, 6, and 7

TSS in EPA rain zone 5 is significantly different than in rain zone 6. No significant differences between EPA rain zone 5 and rain zone 7.

TSS in EPA rain zone 6 is significantly different than in rain zone 7.



The model satisfies normality of residuals. No specific trend was observed in the residuals.


Interaction Plot (fitted means) for LTSS
Landuse
EPA_Rain_Zone
2
3
4
5
6
7
3.0
2.5
2.0
1.5
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7

Biochemical Oxygen Demand, 5 day (BOD_5)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	114	9	1.10	0.29	0.47	1.98
3	6	0	1.30	0.42	0.86	1.92
4	15	1	1.25	0.31	0.60	1.76
5	40	2	0.80	0.20	0.40	1.23
6	13	21	1.55	0.35	0.95	2.00
7	37	4	0.85	0.37	-0.13	1.62

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	97	12	0.93	0.35	0.09	2.17
3	16	0	0.73	0.38	0.00	1.28
4	16	1	0.72	0.25	0.30	1.15
5	46	1	0.78	0.16	0.30	1.11
6	21	49	1.45	0.56	0.15	2.43
7	24	9	1.45	0.44	0.60	2.20

Land use: Residential

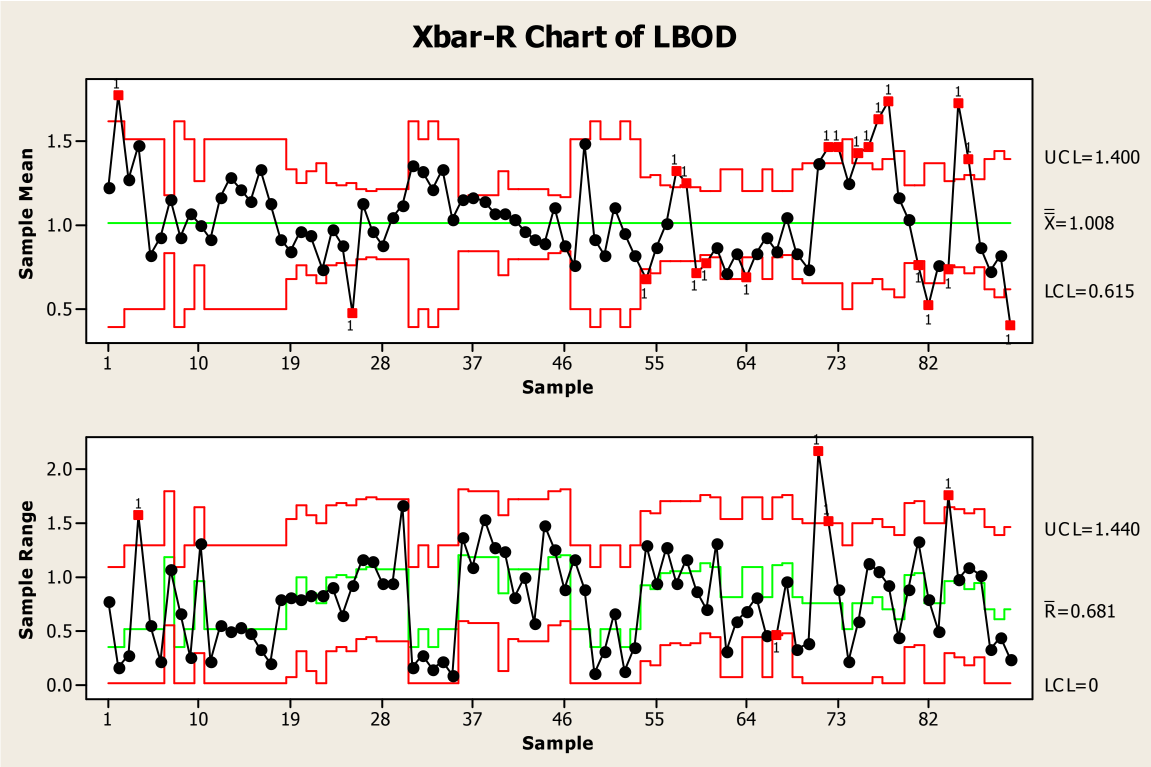
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	290	41	1.04	0.31	0.30	2.35
3	16	2	0.86	0.32	0.30	1.45
4	31	0	1.17	0.35	0.30	1.84
5	67	4	0.90	0.25	0.38	1.70
6	15	23	1.41	0.28	1.00	2.11
7	37	3	0.70	0.32	0.13	1.61

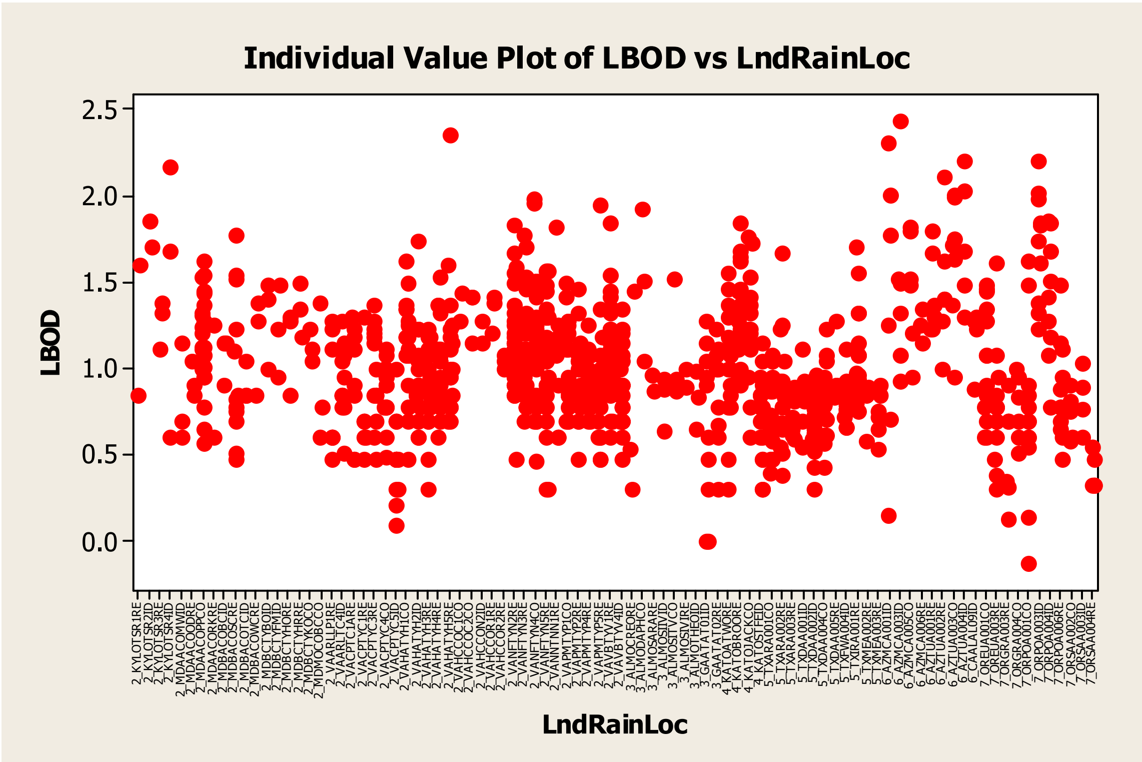
Descriptive Statistics: LBOD

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LBOD	2_KYLOTSR1RE	2	1	1.224	0.286	0.845	1.602
	2_KYLOTSR2ID	2	1	1.7794	0.0103	1.7076	1.8513
	2_KYLOTSR3RE	3	0	1.2721	0.0196	1.1139	1.3802
	2_KYLOTSR4ID	3	1	1.484	0.642	0.602	2.167
	2_MDAACOMWID	3	0	0.816	0.0842	0.602	1.146
	2_MDAACOODRE	3	0	0.9299	0.0102	0.8451	1.0414
	2_MDAACOPPCO	26	0	1.1532	0.0696	0.5658	1.6217
	2_MDAACORKRE	2	1	0.929	0.213	0.602	1.255
	2_MDBACOBCID	3	0	1.0651	0.0197	0.9031	1.1461
	2_MDBACOSCRE	12	14	1.002	0.182	0.477	1.771
	2_MDBACOTCID	3	0	0.9105	0.0128	0.8451	1.0414
	2_MDBACOWCRE	3	0	1.168	0.0808	0.845	1.380
	2_MDBCTYBOID	3	0	1.292	0.0654	1.000	1.477
	2_MDBCTYFMID	3	0	1.221	0.0684	0.954	1.477
	2_MDBCTYHORE	3	0	1.142	0.0661	0.845	1.301
	2_MDBCTYHRRE	3	0	1.3366	0.0249	1.1761	1.4914
	2_MDBCTYKOCO	3	0	1.1286	0.00910	1.0414	1.2304
	2_MDMOCOBCCO	3	0	0.920	0.167	0.602	1.380
	2_VAARLLP1RE	7	1	0.844	0.120	0.477	1.279
	2_VAARLTC4ID	13	0	0.9687	0.0439	0.5052	1.2788

2_VACPTC1ARE	8	0	0.940	0.106	0.477	1.301
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	6	1	0.730	0.0849	0.477	1.301
2_VACPTYC3RE	13	1	0.9777	0.0786	0.4771	1.3617
2_VACPTYC4CO	14	0	0.8804	0.0321	0.4808	1.1139
2_VACPTYC5ID	13	1	0.4737	0.0547	0.0931	1.0000
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	1.1351	0.0798	0.4771	1.6232
2_VAHATYH2ID	19	0	0.9623	0.0655	0.6021	1.7324
2_VAHATYH3RE	17	0	0.8799	0.0625	0.3010	1.2304
2_VAHATYH4RE	17	0	1.0430	0.0626	0.6021	1.5315
2_VAHATYH5RE	17	0	1.1146	0.1622	0.6990	2.3541
2_VAHCCOC1CO	2	0	1.3551	0.0116	1.2788	1.4314
2_VAHCCOC2CO	3	0	1.3254	0.0241	1.1461	1.4150
2_VAHCCON1ID	1	1	1.1139	*	1.1139	1.1139
2_VAHCCON2ID	2	1	1.2124	0.00879	1.1461	1.2788
2_VAHCCOR1RE	3	0	1.3331	0.0128	1.2041	1.4150
2_VAHCCOR2RE	3	0	1.0402	0.00157	1.0000	1.0792
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	1.1504	0.0818	0.4771	1.8325
2_VANFTYN3RE	27	0	1.1703	0.0699	0.6990	1.7782
2_VANFTYN4CO	27	0	1.1449	0.1274	0.4671	1.9823
2_VANFTYN5RE	27	0	1.0673	0.1126	0.3010	1.5682
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	1.069	0.154	0.602	1.820
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	1.0331	0.0657	0.6990	1.4914
2_VAPMTYP2RE	17	0	0.9654	0.0587	0.4771	1.4624
2_VAPMTYP4RE	17	0	0.9154	0.0213	0.6990	1.2553
2_VAPMTYP5RE	17	0	0.8928	0.1316	0.4771	1.9494
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	1.1097	0.0792	0.6021	1.8388
2_VAVBTYV4ID	29	1	0.8796	0.0448	0.4771	1.3424
3_ALMOCREORE	3	0	0.760	0.368	0.301	1.447
3_ALMODAPHCO	3	0	1.489	0.193	1.041	1.919
3_ALMOSARARE	2	1	0.9141	0.00403	0.8692	0.9590
3_ALMOSIIVID	3	0	0.8163	0.0258	0.6335	0.9345
3_ALMOSITVCO	3	0	1.105	0.129	0.863	1.519
3_ALMOSIVIRE	2	1	0.9460	0.00582	0.8921	1.0000
3_ALMOTHEOID	3	0	0.8242	0.0279	0.6532	0.9868
3_GAATAT01ID	10	0	0.670	0.216	0.000000000	1.279
3_GAATAT02RE	9	0	0.8641	0.0887	0.3010	1.2304
4_KATOATWORE	15	0	1.0089	0.1205	0.3010	1.5563
4_KATOBROORE	16	0	1.3221	0.0771	0.9031	1.8388
4_KATOJACKCO	15	1	1.2500	0.0989	0.6021	1.7559
4_KATOSTFEID	16	1	0.7156	0.0606	0.3010	1.1461
5_TXARA001CO	21	1	0.7764	0.0414	0.3979	1.0792
5_TXARA002RE	19	2	0.8653	0.0899	0.3802	1.6721
5_TXARA003RE	7	0	0.7098	0.0118	0.5911	0.8921
5_TXDAA001ID	7	0	0.8324	0.0319	0.5441	1.1139
5_TXDAA002ID	19	0	0.6886	0.0313	0.3010	0.9638
5_TXDAA004CO	19	1	0.8254	0.0364	0.4314	1.2304
5_TXDAA005RE	7	0	0.9312	0.0242	0.8325	1.2788
5_TXFWA004ID	20	1	0.8374	0.0109	0.6628	1.1139
5_TXIRA001RE	21	1	1.0430	0.0607	0.7559	1.6990
5_TXMEA002RE	7	0	0.8293	0.0123	0.5798	0.8921
5_TXMEA003RE	6	1	0.7328	0.0178	0.5315	0.9031

6_AZMCA001ID	6	21	1.365	0.675	0.148	2.301
6_AZMCA003ID	6	21	1.462	0.279	0.929	2.431
6_AZMCA005CO	6	20	1.462	0.114	0.954	1.820
6_AZMCA006RE	3	17	1.2479	0.00967	1.1461	1.3424
6_AZTUA001RE	6	4	1.4310	0.0583	1.2304	1.7993
6_AZTUA002RE	6	2	1.469	0.141	1.000	2.114
6_AZTUA003CO	7	1	1.629	0.136	0.954	2.000
6_AZTUA004ID	5	2	1.739	0.142	1.301	2.204
6_CAALAL09ID	4	5	1.1700	0.0397	0.8751	1.3010
7_OREUA001CO	14	2	1.0315	0.1038	0.6021	1.4771
7_OREUA003RE	15	0	0.7638	0.0964	0.3010	1.6128
7_ORGRA003RE	6	0	0.515	0.0889	0.126	0.908
7_ORGRA004CO	6	0	0.7651	0.0382	0.5051	0.9912
7_ORPOA001CO	12	1	0.732	0.232	-0.127	1.623
7_ORPOA003ID	11	3	1.7267	0.0956	1.2304	2.2041
7_ORPOA004ID	9	4	1.400	0.130	0.778	1.851
7_ORPOA006RE	11	2	0.8678	0.0835	0.4771	1.4771
7_ORSAA002CO	5	1	0.7268	0.0183	0.5798	0.8976
7_ORSAA003ID	4	2	0.8193	0.0324	0.6021	1.0253
7_ORSAA004RE	5	1	0.3976	0.0112	0.3222	0.5441





General Linear Model: LBOD versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	6	2, 3, 4, 5, 6, 7

Analysis of Variance for LBOD, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.6031	1.2713	0.6357	6.35	0.002
EPA_Rain_Zone	5	16.4957	15.1699	3.0340	30.30	0.000
Landuse*EPA_Rain_Zone	10	14.3460	14.3460	1.4346	14.33	0.000
Error	882	88.3176	88.3176	0.1001		
Total	899	119.7624				

Main Effects and Interaction are significant.

Bonferroni Simultaneous Tests
Response Variable LBOD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.1341	0.04144	-3.237	0.0038
RE	-0.1274	0.04009	-3.178	0.0046

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.006708	0.03551	0.1889	1.000

Tukey Simultaneous Tests
Response Variable LBOD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.1341	0.04144	-3.237	0.0035
RE	-0.1274	0.04009	-3.178	0.0042

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.006708	0.03551	0.1889	0.9805

BOD in commercial land uses is significantly different than in industrial or residential land uses. There is no difference in BOD concentrations between industrial and residential land uses.

Bonferroni Simultaneous Tests
Response Variable LBOD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.0594	0.05913	-1.004	1.0000
4	0.0245	0.04525	0.541	1.0000
5	-0.1973	0.03063	-6.442	0.0000
6	0.4479	0.04878	9.182	0.0000
7	-0.0183	0.03629	-0.503	1.0000

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.0838	0.07100	1.181	1.0000
5	-0.1379	0.06270	-2.200	0.4211
6	0.5073	0.07330	6.921	0.0000
7	0.0411	0.06565	0.626	1.0000

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.2218	0.04982	-4.451	0.0001
6	0.4234	0.06264	6.760	0.0000
7	-0.0427	0.05349	-0.799	1.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.6452	0.05304	12.164	0.0000
7	0.1790	0.04185	4.278	0.0003

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.4662	0.05650	-8.250	0.0000

BOD5 in EPA rain zone 2 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 2 and rain zones 3, 4, and 7.

BOD5 in EPA rain zone 3 is significantly different than in rain zone 6. No significant differences between EPA rain zone 3 and rain zones 4, 5 and 7.

BOD5 in EPA rain zone 4 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 4 and rain zone 7.

BOD5 in EPA rain zone 5 is significantly different than in rain zones 6, and 7.

BOD5 in EPA rain zone 6 is significantly different than in rain zone 7.

Tukey Simultaneous Tests
Response Variable LBOD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.0594	0.05913	-1.004	0.9168
4	0.0245	0.04525	0.541	0.9945
5	-0.1973	0.03063	-6.442	0.0000
6	0.4479	0.04878	9.182	0.0000
7	-0.0183	0.03629	-0.503	0.9961

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.0838	0.07100	1.181	0.8462
5	-0.1379	0.06270	-2.200	0.2377
6	0.5073	0.07330	6.921	0.0000
7	0.0411	0.06565	0.626	0.9891

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.2218	0.04982	-4.451	0.0001
6	0.4234	0.06264	6.760	0.0000
7	-0.0427	0.05349	-0.799	0.9677

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value

```
6                  0.6452     0.05304   12.164     0.0000
7                  0.1790     0.04185    4.278     0.0003


EPA_Rain_Zone = 6  subtracted from:

               Difference        SE of             Adjusted
EPA_Rain_Zone    of Means   Difference   T-Value    P-Value
7                 -0.4662      0.05650    -8.250     0.0000
```

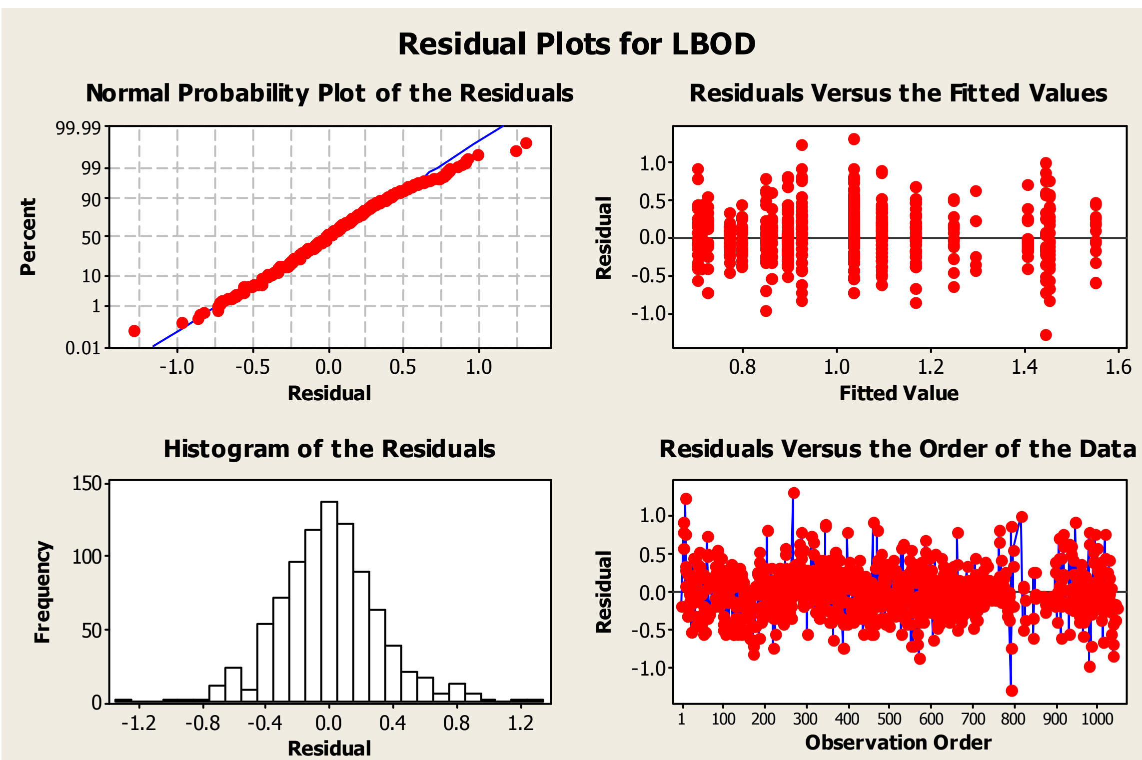
BOD5 in EPA rain zone 2 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 2 and rain zones 3, 4, and 7.

BOD5 in EPA rain zone 3 is significantly different than in rain zone 6. No significant differences between EPA rain zone 3 and rain zones 4, 5 and 7.

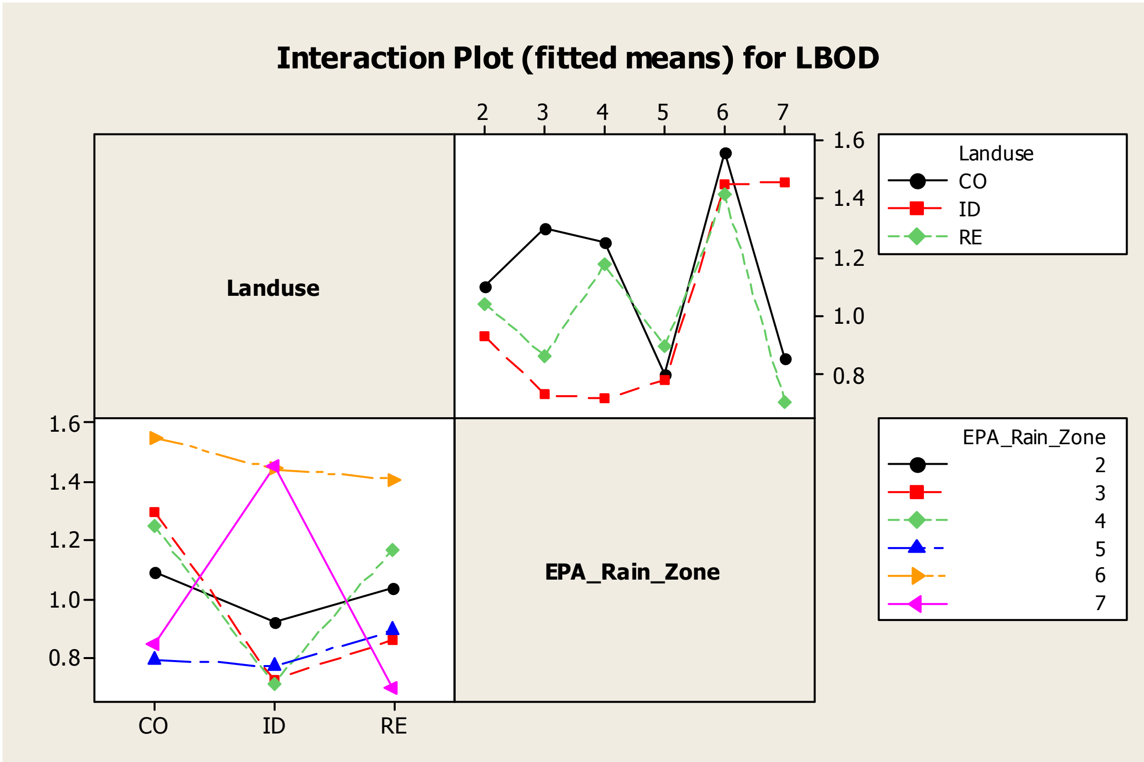
BOD5 in EPA rain zone 4 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 4 and rain zone 7.

BOD5 in EPA rain zone 5 is significantly different than in rain zones 6, and 7.

BOD5 in EPA rain zone 6 is significantly different than in rain zone 7.



The residuals are normally distributed and did not indicate any specific trend in any of the residuals plots.


Interaction Plot (fitted means) for LBOD
Landuse
EPA_Rain_Zone
2
3
4
5
6
7
1.6
1.4
1.2
1.0
0.8
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7

Chemical Oxygen Demand (COD)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	82	41	1.84	0.32	0.90	2.80
3	6	0	1.85	0.35	1.36	2.38
4	0	16 *	*	*	*	
5	41	1	1.65	0.22	1.15	2.18
6	33	1	2.29	0.24	1.89	2.76
7	37	4	1.65	0.30	0.90	2.52

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	85	24	1.72	0.32	0.30	2.53
3	16	0	1.52	0.36	0.60	2.06
4	0	17 *	*	*	*	
5	45	2	1.59	0.29	0.98	2.40
6	56	14	2.27	0.35	1.26	2.96
7	24	9	1.90	0.33	1.26	2.45

Land use: Residential

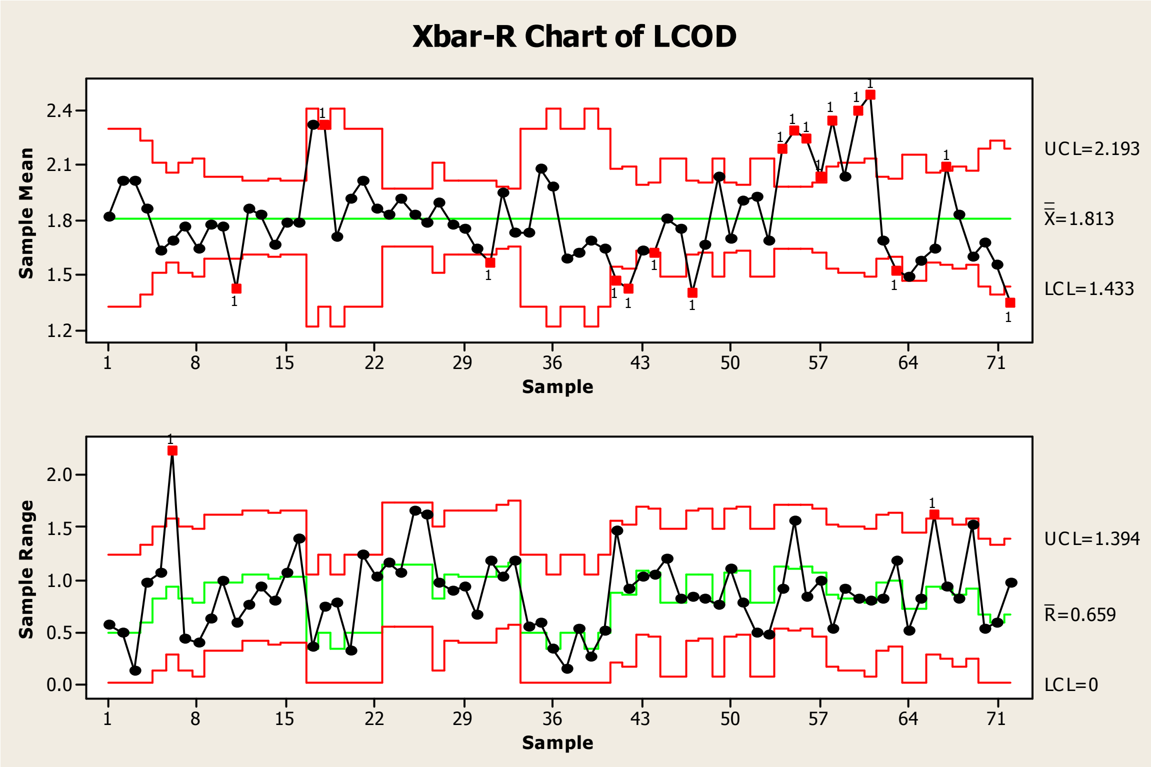
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	265	66	1.79	0.29	0.70	2.79
3	16	2	1.58	0.31	0.78	2.15
4	0	31 *	*	*	*	
5	69	2	1.81	0.29	1.00	2.68
6	37	1	2.12	0.28	1.51	2.57
7	37	3	1.52	0.35	0.95	2.48

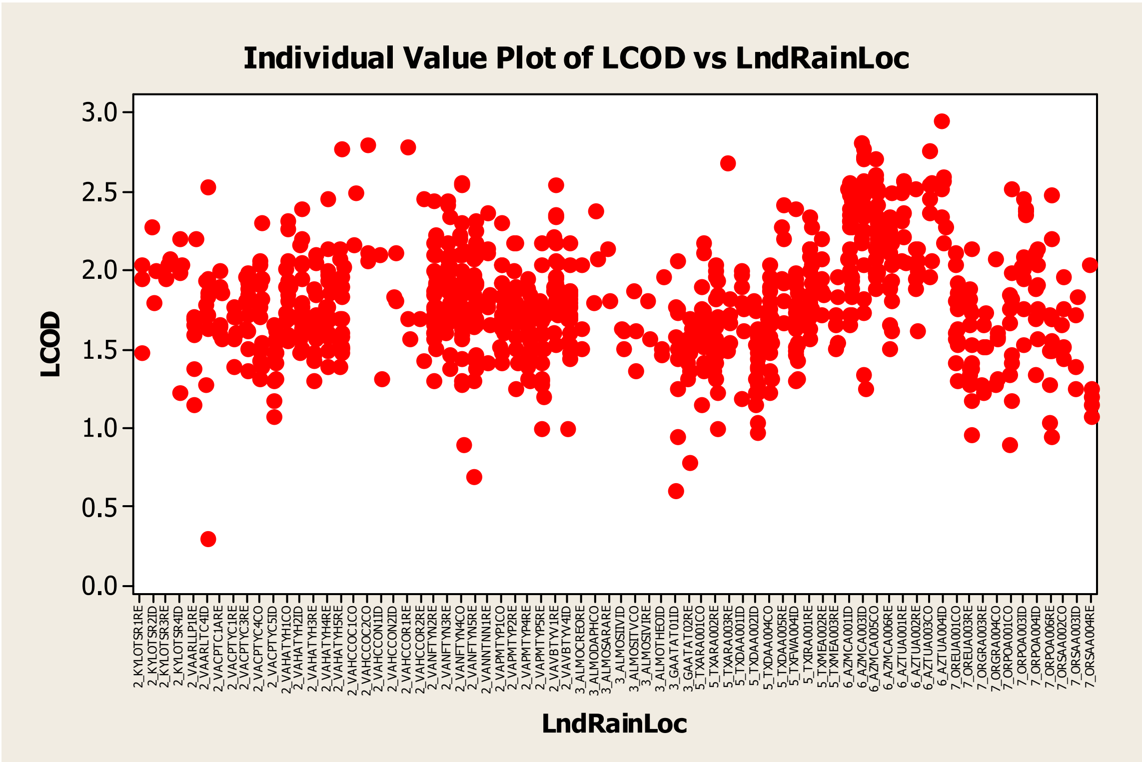
Descriptive Statistics: LCOD

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LCOD	2_KYLOTSR1RE	3	0	1.820	0.0899	1.477	2.033
	2_KYLOTSR2ID	3	0	2.021	0.0591	1.792	2.276
	2_KYLOTSR3RE	3	0	2.0211	0.00379	1.9542	2.0755
	2_KYLOTSR4ID	4	0	1.866	0.188	1.230	2.201
	2_MDAACOMWID	0	3	*	*	*	*
	2_MDAACOODRE	0	3	*	*	*	*
	2_MDAACOPPCO	0	26	*	*	*	*
	2_MDAACORKRE	0	3	*	*	*	*
	2_MDBACOBCID	0	3	*	*	*	*
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	0	3	*	*	*	*
	2_MDBACOWCRE	0	3	*	*	*	*
	2_MDBCTYBOID	0	3	*	*	*	*
	2_MDBCTYFMID	0	3	*	*	*	*
	2_MDBCTYHORE	0	3	*	*	*	*
	2_MDBCTYHRRE	0	3	*	*	*	*
	2_MDBCTYKOCO	0	3	*	*	*	*
	2_MDMOCOBCCO	0	3	*	*	*	*
	2_VAARLLP1RE	8	0	1.634	0.0917	1.146	2.204
	2_VAARLTC4ID	12	1	1.689	0.273	0.301	2.531

2_VACPTC1ARE	8	0	1.7616	0.0272	1.5682	1.9956
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	1.6384	0.0163	1.3979	1.7782
2_VACPTYC3RE	14	0	1.7743	0.0347	1.3617	1.9868
2_VACPTYC4CO	14	0	1.7656	0.0859	1.3222	2.3010
2_VACPTYC5ID	14	0	1.4269	0.0263	1.0792	1.6532
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	1.8627	0.0432	1.5563	2.3118
2_VAHATYH2ID	18	1	1.8318	0.0641	1.4771	2.3979
2_VAHATYH3RE	16	1	1.6684	0.0473	1.3010	2.0969
2_VAHATYH4RE	17	0	1.7869	0.0674	1.3979	2.4594
2_VAHATYH5RE	17	0	1.7828	0.1187	1.3979	2.7745
2_VAHCCOC1CO	2	0	2.331	0.0572	2.161	2.500
2_VAHCCOC2CO	3	0	2.331	0.168	2.068	2.803
2_VAHCCON1ID	2	0	1.711	0.303	1.322	2.100
2_VAHCCON2ID	3	0	1.9164	0.0284	1.8062	2.1106
2_VAHCCOR1RE	3	0	2.020	0.452	1.568	2.792
2_VAHCCOR2RE	3	0	1.861	0.281	1.431	2.453
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	28	1	1.8351	0.0593	1.3010	2.4472
2_VANFTYN3RE	27	0	1.9130	0.0667	1.3802	2.4393
2_VANFTYN4CO	27	0	1.8261	0.1420	0.9031	2.5563
2_VANFTYN5RE	27	0	1.7853	0.1228	0.6990	2.3222
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	1.892	0.0926	1.415	2.371
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	1.7730	0.0494	1.4150	2.3010
2_VAPMTYP2RE	17	0	1.7506	0.0579	1.2553	2.1761
2_VAPMTYP4RE	17	0	1.6401	0.0386	1.3010	1.9542
2_VAPMTYP5RE	17	0	1.5667	0.0907	1.0000	2.1761
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	25	2	1.9476	0.0561	1.5315	2.5416
2_VAVBTYV4ID	29	1	1.7279	0.0425	1.0000	2.1761
3_ALMOCREORE	3	0	1.727	0.0784	1.505	2.041
3_ALMODAPHCO	3	0	2.086	0.0844	1.799	2.380
3_ALMOSARARE	2	1	1.980	0.0555	1.813	2.146
3_ALMOSIIVID	3	0	1.5873	0.00509	1.5051	1.6335
3_ALMOSITVCO	3	0	1.620	0.0659	1.362	1.875
3_ALMOSIVIRE	2	1	1.691	0.0299	1.568	1.813
3_ALMOTHEOID	3	0	1.644	0.0773	1.462	1.964
3_GAATAT01ID	10	0	1.468	0.187	0.602	2.064
3_GAATAT02RE	9	0	1.4183	0.0723	0.7782	1.6902
5_TXARA001CO	22	0	1.6367	0.0515	1.1489	2.1761
5_TXARA002RE	20	1	1.6190	0.0758	1.0000	2.0414
5_TXARA003RE	7	0	1.807	0.161	1.491	2.681
5_TXDAA001ID	7	0	1.751	0.0788	1.187	2.000
5_TXDAA002ID	18	1	1.4042	0.0517	0.9802	1.8129
5_TXDAA004CO	19	1	1.6646	0.0464	1.2304	2.0414
5_TXDAA005RE	7	0	2.036	0.0725	1.672	2.415
5_TXFWA004ID	20	1	1.7008	0.0699	1.3075	2.3979
5_TXIRA001RE	21	1	1.9029	0.0494	1.5647	2.3424
5_TXMEA002RE	7	0	1.9342	0.0259	1.7243	2.2041
5_TXMEA003RE	7	0	1.6826	0.0298	1.5051	1.9685
6_AZMCA001ID	25	2	2.1968	0.0824	1.6532	2.5563
6_AZMCA003ID	24	3	2.2900	0.1624	1.2553	2.8062
6_AZMCA005CO	25	1	2.2521	0.0492	1.8865	2.7076
6_AZMCA006RE	20	0	2.0422	0.0677	1.5051	2.4914

6_AZTUA001RE	9	1	2.3530	0.0416	2.0531	2.5682
6_AZTUA002RE	8	0	2.0413	0.0656	1.6232	2.5185
6_AZTUA003CO	8	0	2.4071	0.0733	1.9590	2.7649
6_AZTUA004ID	7	0	2.4926	0.0667	2.1761	2.9571
6_CAALAL09ID	0	9	*	*	*	*
7_OREUA001CO	14	2	1.6907	0.0554	1.3010	2.1139
7_OREUA003RE	15	0	1.5191	0.0972	0.9638	2.1461
7_ORGRA003RE	6	0	1.4900	0.0406	1.2304	1.7404
7_ORGRA004CO	6	0	1.574	0.0822	1.279	2.083
7_ORPOA001CO	12	1	1.643	0.177	0.903	2.519
7_ORPOA003ID	11	3	2.0934	0.0832	1.5315	2.4533
7_ORPOA004ID	9	4	1.8285	0.0710	1.3424	2.1461
7_ORPOA006RE	11	2	1.601	0.202	0.954	2.477
7_ORSAA002CO	5	1	1.6702	0.0412	1.4472	1.9590
7_ORSAA003ID	4	2	1.554	0.0746	1.255	1.839
7_ORSAA004RE	5	1	1.344	0.153	1.079	2.033





General Linear Model: LCOD versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	5	2, 3, 5, 6, 7

Analysis of Variance for LCOD, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.5831	0.6005	0.3003	3.33	0.036
EPA_Rain_Zone	4	28.0369	25.5423	6.3856	70.89	0.000
Landuse*EPA_Rain_Zone	8	5.2090	5.2090	0.6511	7.23	0.000
Error	834	75.1295	75.1295	0.0901		
Total	848	108.9585				

Main factors and interactions are significant.

Bonferroni Simultaneous Tests
Response Variable LCOD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.05597	0.03874	-1.445	0.4468
RE	-0.09694	0.03775	-2.568	0.0312

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.04097	0.03237	-1.265	0.6182

Tukey Simultaneous Tests
Response Variable LCOD
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.05597	0.03874	-1.445	0.3180
RE	-0.09694	0.03775	-2.568	0.0276

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.04097	0.03237	-1.265	0.4148

COD in commercial land use is significantly different than COD in residential areas. There is not enough evidence that indicates a difference between commercial and industrial land uses or between residential and industrial areas.

Bonferroni Simultaneous Tests
Response Variable LCOD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1301	0.05654	-2.300	0.2169
5	-0.1008	0.02982	-3.380	0.0076
6	0.4446	0.03210	13.854	0.0000
7	-0.0914	0.03515	-2.600	0.0949

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.02927	0.05942	0.4926	1.0000
6	0.57470	0.06060	9.4841	0.0000
7	0.03866	0.06227	0.6209	1.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.545430	0.03693	14.7673	0.0000
7	0.009388	0.03962	0.2370	1.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value

7	-0.5360	0.04136	-12.96	0.0000

COD in EPA rain zone 2 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 2 and rain zones 3, and 7.

COD in EPA rain zone 3 is significantly different than in rain zone 6. No significant differences between EPA rain zone 3 and rain zones 5, and 7.

COD in EPA rain zone 5 is significantly different than in rain zone 6. No significant differences between EPA rain zone 5 and rain zone 7.

COD in EPA rain zone 6 is significantly different than in rain zone 7.

```
Tukey Simultaneous Tests
Response Variable LCOD
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:
```

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1301	0.05654	-2.300	0.1447
5	-0.1008	0.02982	-3.380	0.0065
6	0.4446	0.03210	13.854	0.0000
7	-0.0914	0.03515	-2.600	0.0703

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.02927	0.05942	0.4926	0.9881
6	0.57470	0.06060	9.4841	0.0000
7	0.03866	0.06227	0.6209	0.9718

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.545430	0.03693	14.7673	0.0000
7	0.009388	0.03962	0.2370	0.9993

EPA_Rain_Zone = 6 subtracted from:

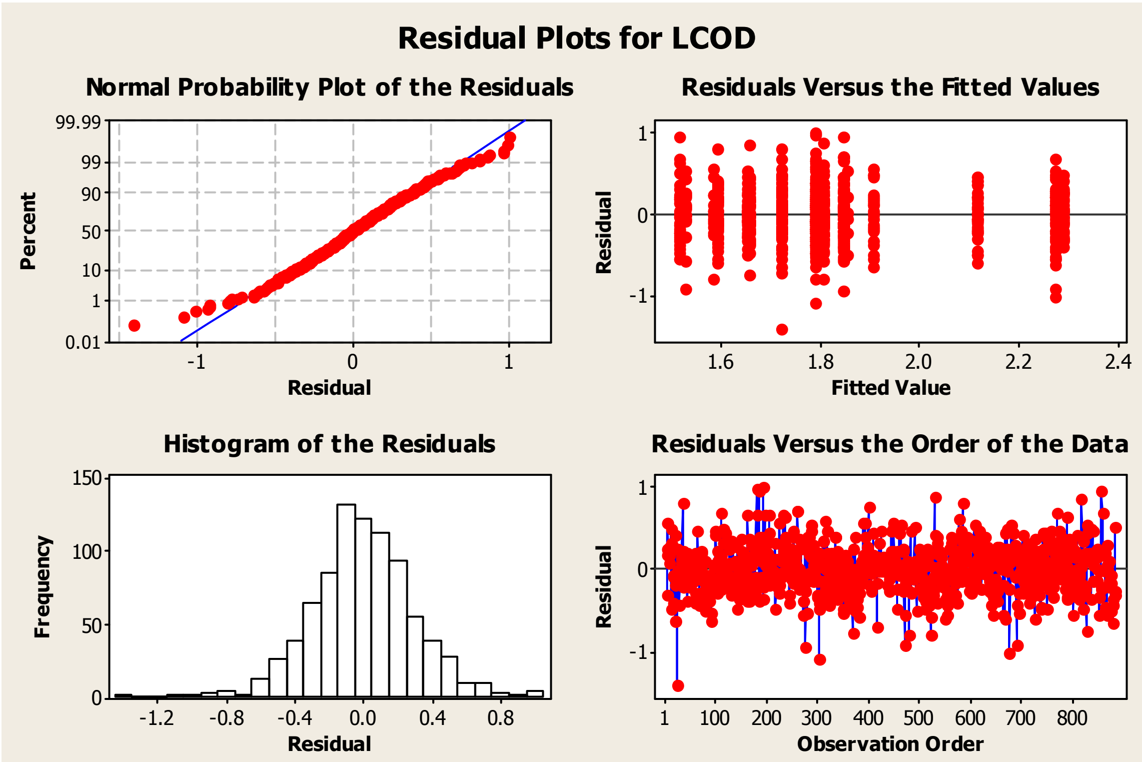
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.5360	0.04136	-12.96	0.0000

COD in EPA rain zone 2 is significantly different than in rain zones 5, and 6. No significant differences between EPA rain zone 2 and rain zones 3, and 7.

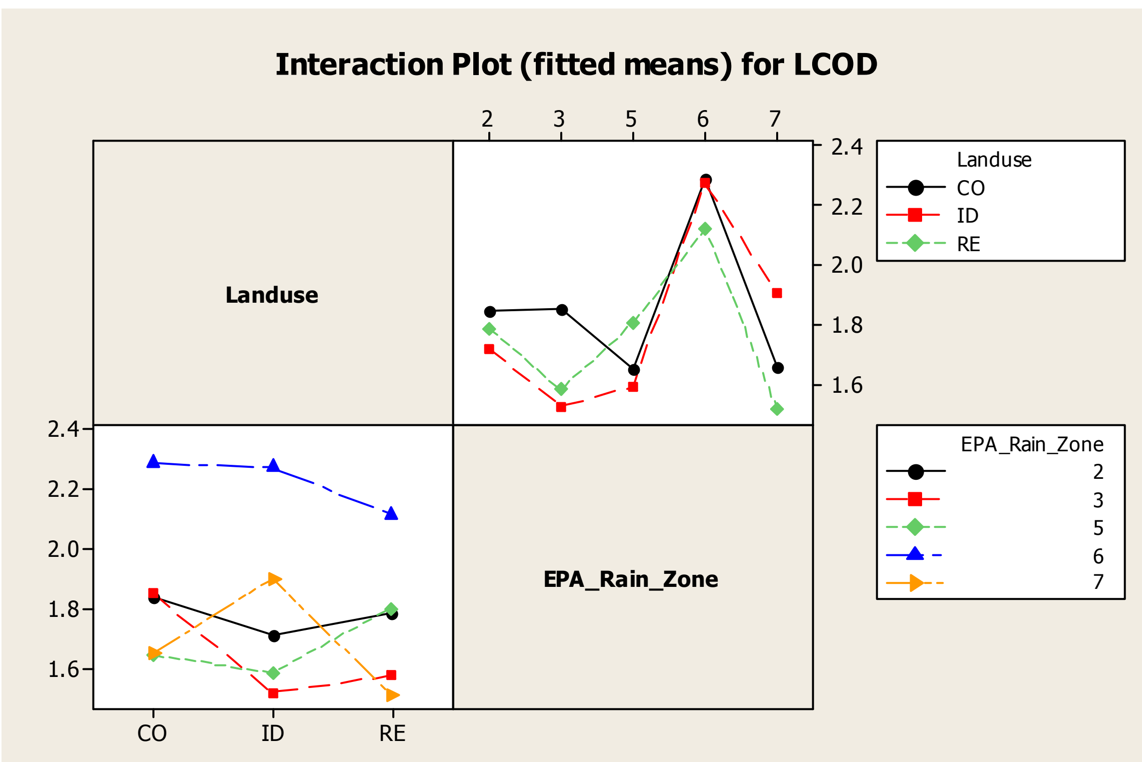
COD in EPA rain zone 3 is significantly different than in rain zone 6. No significant differences between EPA rain zone 3 and rain zones 5, and 7.

COD in EPA rain zone 5 is significantly different than in rain zone 6. No significant differences between EPA rain zone 5 and rain zone 7.

COD in EPA rain zone 6 is significantly different than in rain zone 7.



Residuals are normally distributed, no trend was observed in the residual plots.



Ammonia (NH_3)

Summary statistics in LOG base 10 scale

Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	82	41	-0.426	0.4129	-1.44	0.3997
3	0	6	*	*	*	*
4	0	16	*	*	*	*
5	0	42	*	*	*	*
6	23	11	0.2896	0.2415	-0.0362	0.8921
7	25	16	-0.812	0.618	-2.066	0.623

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	81	28	-0.616	0.3735	-1.675	0.2041
3	0	16	*	*	*	*
4	0	17	*	*	*	*
5	0	47	*	*	*	*
6	51	19	-0.132	0.4194	-1.5229	0.716
7	18	15	-0.591	0.469	-1.473	0.23

Land use: Residential

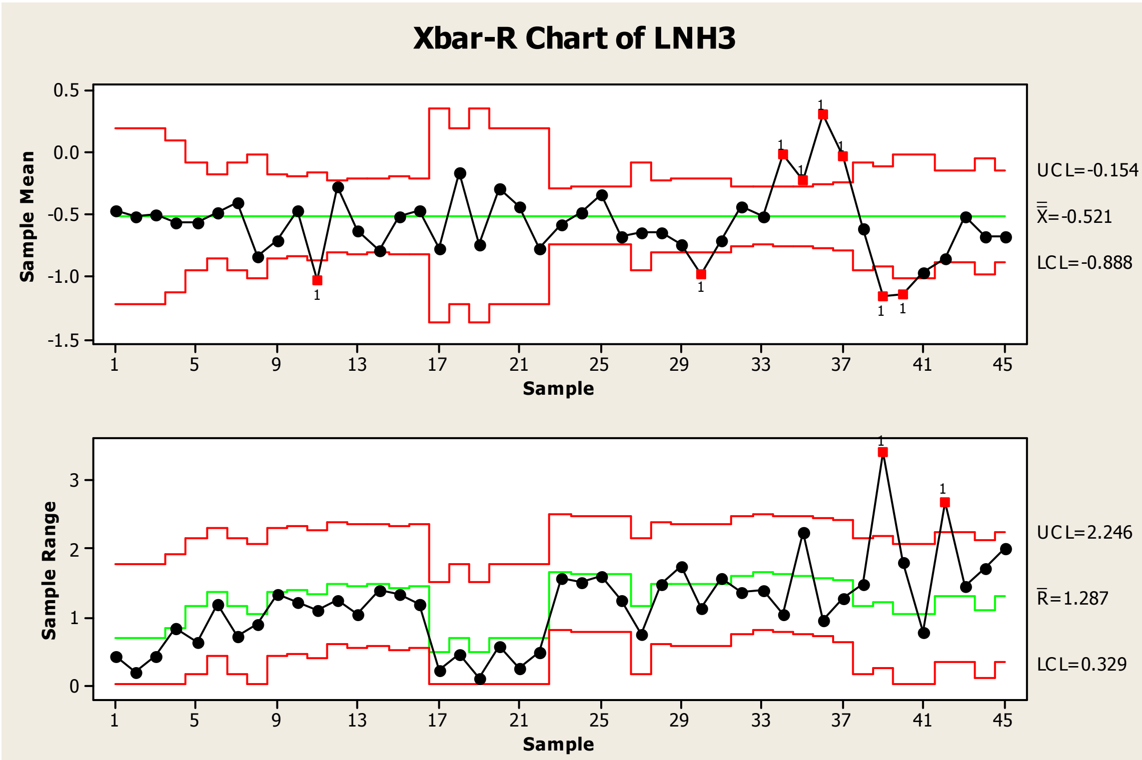
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	262	69	-0.629	0.3846	-1.6214	0.1732
3	0	18	*	*	*	*
4	0	31	*	*	*	*
5	0	71	*	*	*	*
6	20	18	-0.049	0.3025	-0.7212	0.5315
7	26	14	-0.962	0.771	-2.662	0.748

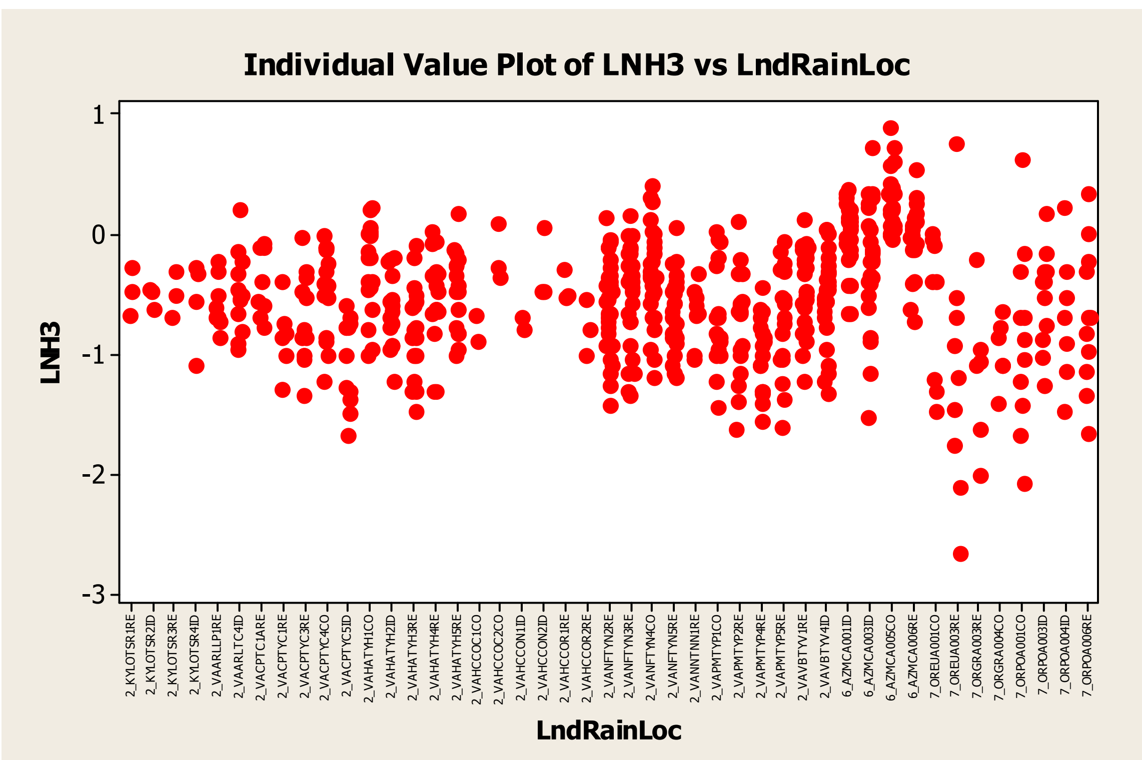
Descriptive Statistics: LNH3

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LNH3	2_KYLOTSR1RE	3	0	-0.476	0.0421	-0.678	-0.268
	2_KYLOTSR2ID	3	0	-0.5191	0.00777	-0.6198	-0.4559
	2_KYLOTSR3RE	3	0	-0.503	0.0396	-0.699	-0.301
	2_KYLOTSR4ID	4	0	-0.565	0.139	-1.097	-0.284
	2_MDAACOMWID	0	3	*	*	*	*
	2_MDAACOODRE	0	3	*	*	*	*
	2_MDAACOPPCO	0	26	*	*	*	*
	2_MDAACORKRE	0	3	*	*	*	*
	2_MDBACOBCID	0	3	*	*	*	*
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	0	3	*	*	*	*
	2_MDBACOWCRE	0	3	*	*	*	*
	2_MDBCTYBOID	0	3	*	*	*	*
	2_MDBCTYFMID	0	3	*	*	*	*
	2_MDBCTYHORE	0	3	*	*	*	*

2_MDBCTYHRRE	0	3	*	*	*	*
2_MDBCTYKOCO	0	3	*	*	*	*
2_MDMOCOBCCO	0	3	*	*	*	*
2_VAARLLP1RE	8	0	-0.5752	0.0480	-0.8539	-0.2218
2_VAARLTC4ID	13	0	-0.4880	0.1016	-0.9586	0.2041
2_VACPTC1ARE	8	0	-0.4120	0.0797	-0.7696	-0.0757
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	6	1	-0.852	0.0867	-1.291	-0.398
2_VACPTYC3RE	13	1	-0.724	0.132	-1.345	-0.0223
2_VACPTYC4CO	14	0	-0.480	0.142	-1.222	-0.0132
2_VACPTYC5ID	12	2	-1.038	0.134	-1.675	-0.585
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	-0.2818	0.1425	-1.0000	0.2304
2_VAHATYH2ID	16	3	-0.6322	0.0858	-1.2218	-0.2007
2_VAHATYH3RE	17	0	-0.7981	0.1527	-1.4811	-0.0969
2_VAHATYH4RE	15	2	-0.530	0.157	-1.301	0.0294
2_VAHATYH5RE	16	1	-0.4723	0.1126	-1.0000	0.1732
2_VAHCCOC1CO	2	0	-0.782	0.0217	-0.886	-0.678
2_VAHCCOC2CO	3	0	-0.176	0.0577	-0.357	0.0969
2_VAHCCON1ID	2	0	-0.7474	0.00470	-0.7959	-0.6990
2_VAHCCON2ID	3	0	-0.295	0.0970	-0.481	0.0645
2_VAHCCOR1RE	3	0	-0.4413	0.0167	-0.5229	-0.2924
2_VAHCCOR2RE	3	0	-0.778	0.0537	-1.000	-0.538
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	-0.5808	0.1551	-1.4320	0.1399
2_VANFTYN3RE	27	0	-0.4921	0.1784	-1.3355	0.1553
2_VANFTYN4CO	27	0	-0.3402	0.1611	-1.1902	0.3997
2_VANFTYN5RE	27	0	-0.6892	0.0991	-1.1902	0.0531
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	-0.6595	0.0632	-1.0458	-0.3188
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	-0.657	0.188	-1.440	0.0294
2_VAPMTYP2RE	17	0	-0.746	0.204	-1.621	0.107
2_VAPMTYP4RE	17	0	-0.9851	0.0992	-1.5647	-0.4437
2_VAPMTYP5RE	17	0	-0.721	0.202	-1.605	-0.0555
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	25	2	-0.4385	0.1138	-1.2218	0.1303
2_VAVBTYV4ID	28	2	-0.5267	0.1331	-1.3315	0.0414
6_AZMCA001ID	26	1	-0.0329	0.0850	-0.6576	0.3802
6_AZMCA003ID	25	2	-0.235	0.256	-1.523	0.716
6_AZMCA005CO	23	3	0.2896	0.0583	-0.0362	0.8921
6_AZMCA006RE	20	0	-0.0492	0.0915	-0.7212	0.5315
6_AZTUA001RE	0	10	*	*	*	*
6_AZTUA002RE	0	8	*	*	*	*
6_AZTUA003CO	0	8	*	*	*	*
6_AZTUA004ID	0	7	*	*	*	*
6_CAALAL09ID	0	9	*	*	*	*
7_OREUA001CO	8	8	-0.617	0.378	-1.472	0.000000000
7_OREUA003RE	9	6	-1.173	0.989	-2.662	0.748
7_ORGRA003RE	6	0	-1.157	0.378	-2.008	-0.208
7_ORGRA004CO	6	0	-0.974	0.0761	-1.403	-0.638
7_ORPOA001CO	11	2	-0.865	0.564	-2.066	0.623
7_ORPOA003ID	11	3	-0.531	0.175	-1.266	0.176
7_ORPOA004ID	7	6	-0.686	0.314	-1.473	0.230
7_ORPOA006RE	11	2	-0.682	0.355	-1.654	0.342
7_ORSAA002CO	0	6	*	*	*	*
7_ORSAA003ID	0	6	*	*	*	*

7_ORSAA004RE 0 6 * * * *





General Linear Model: LNH3 versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	3	2, 6, 7

Analysis of Variance for LNH3, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	6.6025	2.5582	1.2791	7.18	0.001
EPA_Rain_Zone	2	29.6281	29.4237	14.7118	82.60	0.000
Landuse*EPA_Rain_Zone	4	2.8313	2.8313	0.7078	3.97	0.003
Error	579	103.1278	103.1278	0.1781		
Total	587	142.1896				

Main factors and interactions are significant.

Bonferroni Simultaneous Tests
Response Variable LNH3
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.1302	0.06021	-2.163	0.0928
RE	-0.2306	0.06099	-3.782	0.0005

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1004	0.05965	-1.683	0.2786

Tukey Simultaneous Tests
Response Variable LNH3
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.1302	0.06021	-2.163	0.0777
RE	-0.2306	0.06099	-3.782	0.0005

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.1004	0.05965	-1.683	0.2116

There is a significant difference in ammonia concentrations between commercial and residential land uses. This difference was not observed between commercial and industrial areas nor between residential and industrial areas.

Bonferroni Simultaneous Tests
Response Variable LNH3
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.5929	0.05291	11.205	0.0000
7	-0.2317	0.05669	-4.088	0.0001

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.8246	0.06993	-11.79	0.0000

Ammonia in EPA rain zone 2 is significantly different than in rain zones 6, and 7.

Ammonia in EPA rain zone 6 is significantly different than in rain zone 7.

Tukey Simultaneous Tests
Response Variable LNH3
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

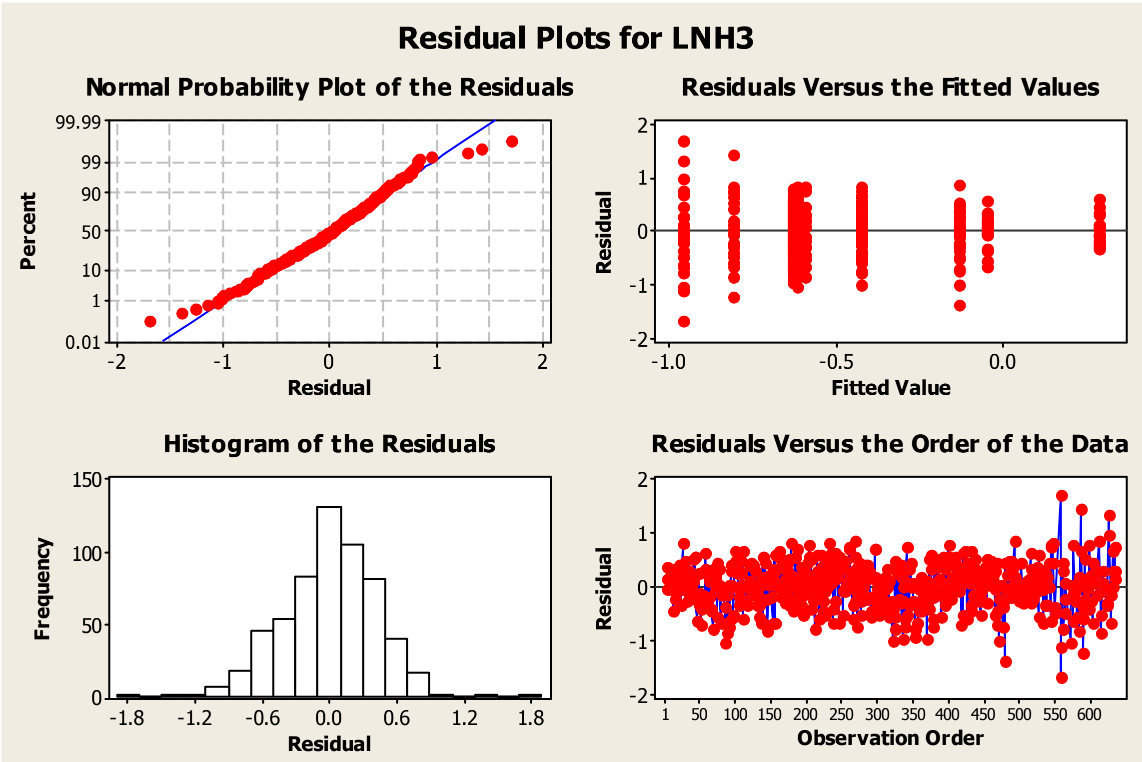
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.5929	0.05291	11.205	0.0000
7	-0.2317	0.05669	-4.088	0.0001

EPA_Rain_Zone = 6 subtracted from:

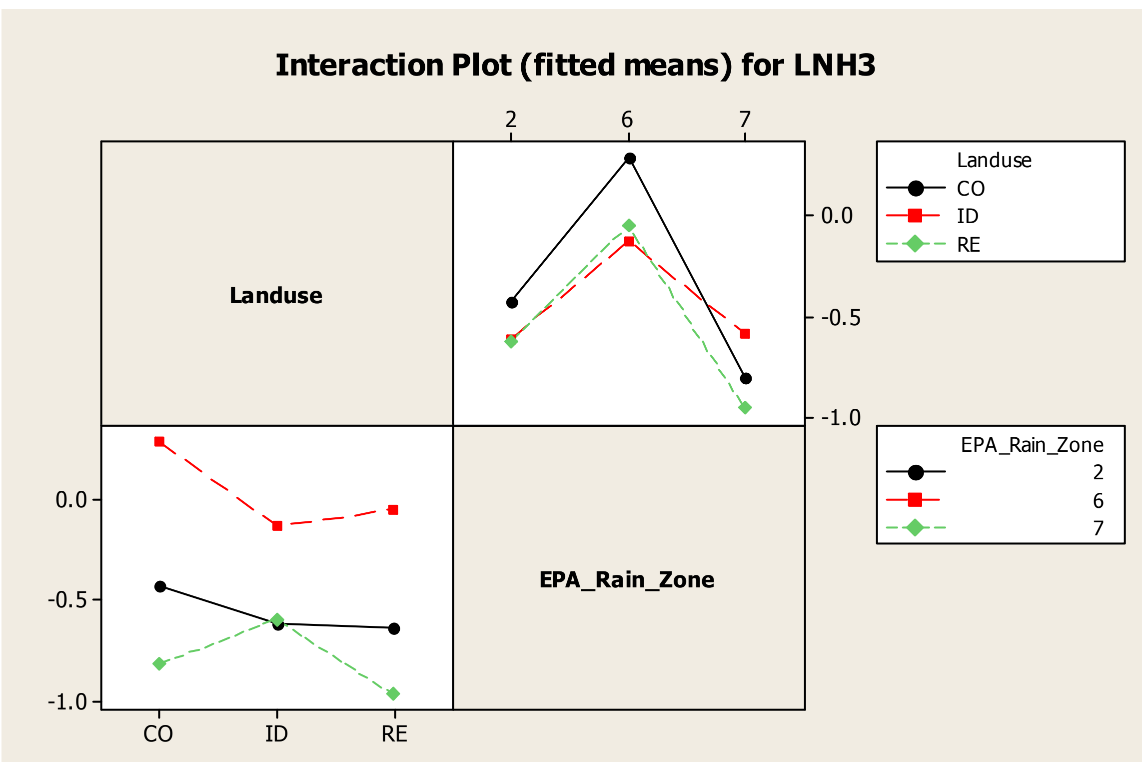
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.8246	0.06993	-11.79	0.0000

Ammonia in EPA rain zone 2 is significantly different than in rain zones 6, and 7.

Ammonia in EPA rain zone 6 is significantly different than in rain zone 7.



Residuals are normally distributed. There is no trend in any of the plots of residuals.



Nitrite plus Nitrate ($NO_2 + NO_3$)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	114	9	-0.297	0.424	-1.7096	0.8633
3	0	6	*	*	*	
4	0	16	*	*	*	
5	40	2	-0.313	0.2076	-0.8196	0.0531
6	30	4	-0.004	0.3379	-1.2218	0.5911
7	37	4	-0.552	0.3409	-1.0969	0.415

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	94	15	-0.308	0.3884	-1.5229	0.3729
3	9	7	-0.606	0.498	-1.347	0.22
4	0	17	*	*	*	
5	46	1	-0.212	0.2213	-1.1612	0.2279
6	58	12	0.2018	0.2552	-0.3372	0.6721
7	24	9	-0.849	0.4197	-1.661	-0.1549

Land use: Residential

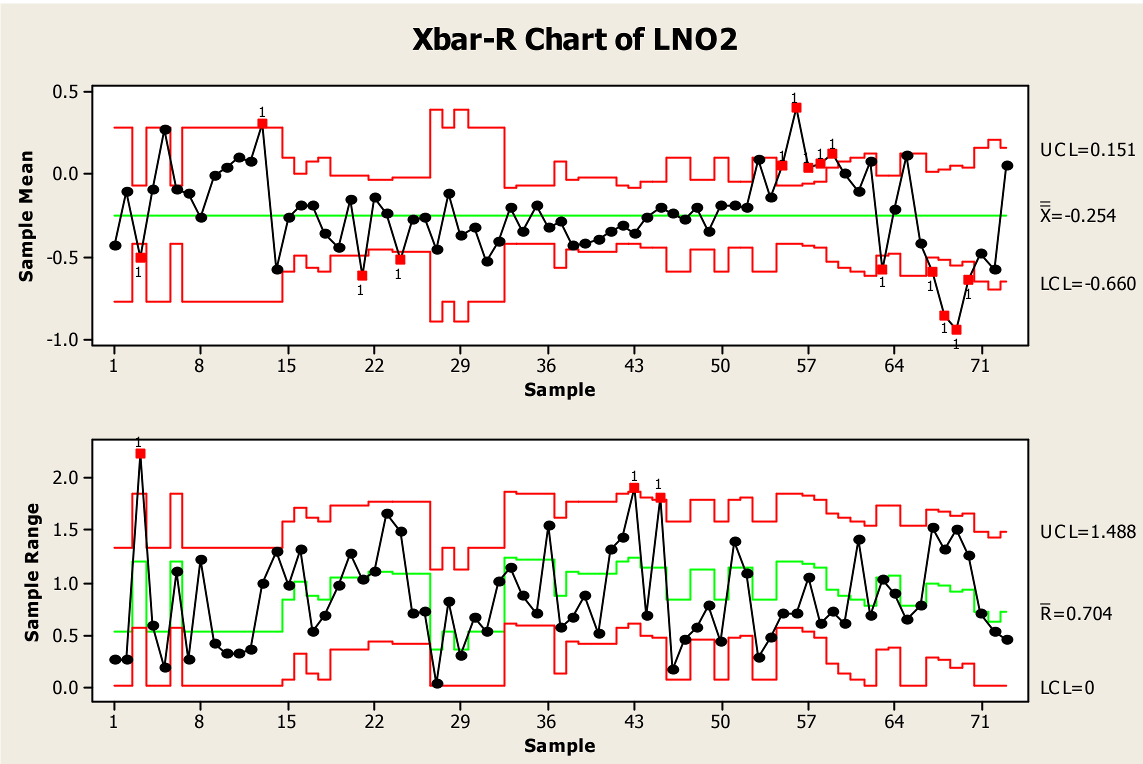
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	302	29	-0.307	0.3006	-1.3979	0.4624
3	7	11	-0.068	0.659	-1.102	0.5
4	0	31	*	*	*	
5	69	2	-0.177	0.26	-0.9586	0.8555
6	38	0	0.0566	0.1794	-0.301	0.4713
7	36	4	-0.248	0.4138	-1.1094	0.5441

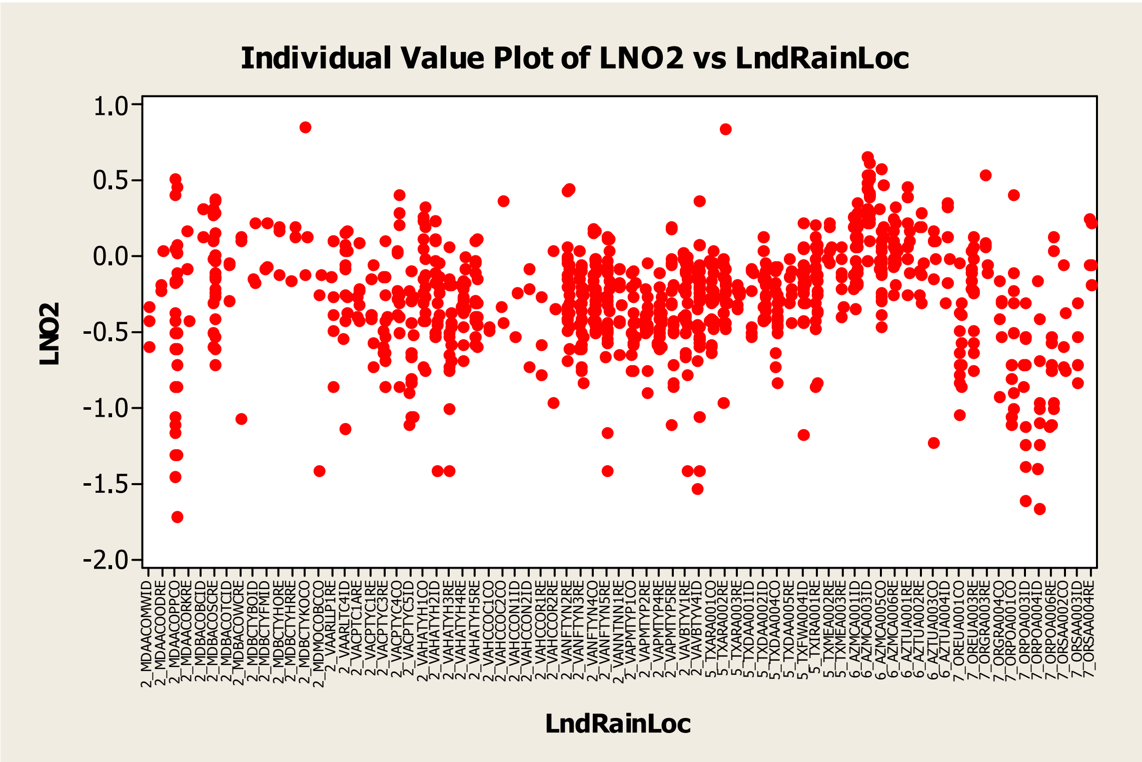
Descriptive Statistics: LNO2

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LNO2	2_KYLOTSR1RE	0	3	*	*	*	*
	2_KYLOTSR2ID	0	3	*	*	*	*
	2_KYLOTSR3RE	0	3	*	*	*	*
	2_KYLOTSR4ID	0	4	*	*	*	*
	2_MDAACOMWID	3	0	-0.4406	0.0173	-0.5850	-0.3279
	2_MDAACOODRE	3	0	-0.1157	0.0189	-0.2147	0.0414
	2_MDAACOPPCO	26	0	-0.516	0.373	-1.710	0.520
	2_MDAACORKRE	3	0	-0.103	0.0861	-0.409	0.176
	2_MDBACOBCID	3	0	0.2635	0.0103	0.1461	0.3222
	2_MDBACOSCRE	26	0	-0.1021	0.1018	-0.6990	0.3909
	2_MDBACOTCID	3	0	-0.1204	0.0201	-0.2840	-0.0315
	2_MDBACOWCRE	3	0	-0.269	0.478	-1.067	0.146
	2_MDBCTYBOID	3	0	-0.0220	0.0510	-0.167	0.238
	2_MDBCTYFMID	3	0	0.0348	0.0288	-0.0706	0.2304
	2_MDBCTYHORE	3	0	0.0908	0.0298	-0.1079	0.2041
	2_MDBCTYHRRE	3	0	0.0661	0.0356	-0.149	0.204

2_MDBCTYKOCO	3	0	0.299	0.254	-0.108	0.863
2_MDMOCOBCCO	3	0	-0.583	0.502	-1.398	-0.108
2_VAARLLP1RE	7	1	-0.267	0.119	-0.854	0.114
2_VAARLTC4ID	13	0	-0.195	0.131	-1.128	0.176
2_VACPTC1ARE	8	0	-0.1962	0.0402	-0.4089	0.1072
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	-0.3717	0.0533	-0.7212	-0.0458
2_VACPTYC3RE	14	0	-0.4563	0.0803	-0.8539	0.1139
2_VACPTYC4CO	14	0	-0.1653	0.1163	-0.8539	0.4166
2_VACPTYC5ID	14	0	-0.6254	0.1134	-1.0969	-0.0862
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	-0.1464	0.1036	-0.7447	0.3424
2_VAHATYH2ID	18	1	-0.2455	0.1313	-1.3979	0.2480
2_VAHATYH3RE	17	0	-0.5296	0.1191	-1.3979	0.0792
2_VAHATYH4RE	17	0	-0.2830	0.0298	-0.6778	0.0128
2_VAHATYH5RE	17	0	-0.2651	0.0506	-0.5850	0.1303
2_VAHCCOC1CO	2	0	-0.4687	0.000327	-0.4815	-0.4559
2_VAHCCOC2CO	3	0	-0.127	0.194	-0.432	0.378
2_VAHCCON1ID	2	0	-0.376	0.0431	-0.523	-0.229
2_VAHCCON2ID	3	0	-0.333	0.117	-0.721	-0.0757
2_VAHCCOR1RE	3	0	-0.533	0.0660	-0.770	-0.260
2_VAHCCOR2RE	3	0	-0.416	0.259	-0.959	0.0492
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	-0.2116	0.0684	-0.6778	0.4624
2_VANFTYN3RE	27	0	-0.3501	0.0563	-0.8239	0.0414
2_VANFTYN4CO	27	0	-0.1988	0.0427	-0.4949	0.1903
2_VANFTYN5RE	27	0	-0.3281	0.1299	-1.3979	0.1399
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	-0.3008	0.0339	-0.6383	-0.0757
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	-0.4410	0.0423	-0.7447	-0.0862
2_VAPMTYP2RE	17	0	-0.4302	0.0365	-0.8861	-0.0269
2_VAPMTYP4RE	17	0	-0.4084	0.0203	-0.6021	-0.1024
2_VAPMTYP5RE	17	0	-0.3531	0.1362	-1.0969	0.2068
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	-0.3223	0.1009	-1.3979	0.0334
2_VAVBTYV4ID	29	1	-0.3656	0.1690	-1.5229	0.3729
5_TXARA001CO	21	1	-0.2736	0.0403	-0.6198	0.0531
5_TXARA002RE	20	1	-0.2077	0.1057	-0.9586	0.8555
5_TXARA003RE	7	0	-0.2473	0.00324	-0.3372	-0.1739
5_TXDAA001ID	7	0	-0.2800	0.0417	-0.5229	-0.0757
5_TXDAA002ID	19	0	-0.2060	0.0255	-0.4101	0.1461
5_TXDAA004CO	19	1	-0.3557	0.0449	-0.8196	-0.0521
5_TXDAA005RE	7	0	-0.1957	0.0210	-0.4318	0.000000000
5_TXFWA004ID	20	1	-0.1942	0.0766	-1.1612	0.2279
5_TXIRA001RE	21	1	-0.2112	0.0834	-0.8539	0.2175
5_TXMEA002RE	7	0	0.0800	0.0102	-0.0362	0.2304
5_TXMEA003RE	7	0	-0.1509	0.0277	-0.3872	0.0792
6_AZMCA001ID	26	1	0.0415	0.0387	-0.3372	0.3617
6_AZMCA003ID	26	1	0.3936	0.0258	-0.0177	0.6721
6_AZMCA005CO	23	3	0.0309	0.0776	-0.4559	0.5911
6_AZMCA006RE	20	0	0.0536	0.0189	-0.2441	0.3424
6_AZTUA001RE	10	0	0.1134	0.0473	-0.2441	0.4713
6_AZTUA002RE	8	0	-0.00679	0.0487	-0.3010	0.3010
6_AZTUA003CO	7	1	-0.118	0.247	-1.222	0.185
6_AZTUA004ID	6	1	0.0653	0.0726	-0.301	0.365
6_CAALAL09ID	0	9	*	*	*	*

7_OREUA001CO	14	2	-0.5824	0.0691	-1.0402	-0.0269
7_OREUA003RE	15	0	-0.2205	0.0712	-0.7447	0.1461
7_ORGRA003RE	6	0	0.1092	0.0518	-0.0969	0.5441
7_ORGRA004CO	6	0	-0.427	0.0738	-0.921	-0.155
7_ORPOA001CO	12	1	-0.602	0.220	-1.097	0.415
7_ORPOA003ID	11	3	-0.860	0.171	-1.607	-0.301
7_ORPOA004ID	9	4	-0.951	0.227	-1.661	-0.155
7_ORPOA006RE	10	3	-0.650	0.197	-1.109	0.146
7_ORSAA002CO	5	1	-0.493	0.0850	-0.745	-0.0458
7_ORSAA003ID	4	2	-0.587	0.0515	-0.824	-0.301
7_ORSAA004RE	5	1	0.0437	0.0363	-0.1805	0.2553





General Linear Model: LNO3 versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	4	2, 5, 6, 7

Analysis of Variance for LNO3, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.8956	2.0365	1.0182	9.65	0.000
EPA_Rain_Zone	3	24.1930	22.5935	7.5312	71.35	0.000
Landuse*EPA_Rain_Zone	6	6.2276	6.2276	1.0379	9.83	0.000
Error	876	92.4598	92.4598	0.1055		
Total	887	123.7760				

Main effects and interactions are significant.

Bonferroni Simultaneous Tests
Response Variable LNO3
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.000404	0.03497	-0.01155	1.0000
RE	0.122472	0.03309	3.70067	0.0007

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.1229	0.03281	3.745	0.0006

Tukey Simultaneous Tests
Response Variable LNO3
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.000404	0.03497	-0.01155	0.9999
RE	0.122472	0.03309	3.70067	0.0006

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.1229	0.03281	3.745	0.0005

There is not a significant difference in nitrate concentrations between commercial and industrial land uses. Nitrates in commercial land use areas are significantly different than in residential land use areas. Nitrates in industrial land use areas are significantly different than in residential land use areas.

Bonferroni Simultaneous Tests
Response Variable LNO3
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.0701	0.03138	2.234	0.1545
6	0.3887	0.03418	11.372	0.0000
7	-0.2455	0.03739	-6.567	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.3186	0.04025	7.916	0.0000
7	-0.3156	0.04301	-7.339	0.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.6342	0.04509	-14.07	0.0000

Nitrates in EPA rain zone 2 are significantly different than in rain zones 6, and 7. Nitrates in EPA rain zone 2 are not significantly than in rain zone 5.

Nitrates in EPA rain zone 5 are significantly different than in rain zones 6, and 7.

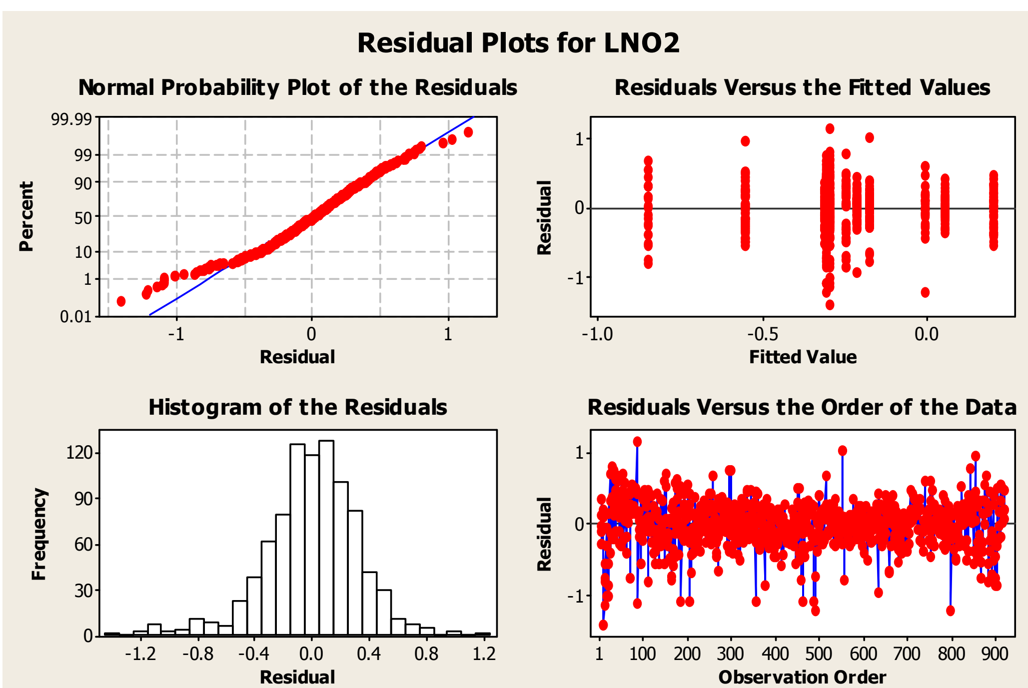
Nitrates in EPA rain zone 6 are significantly different than in rain zones 7.

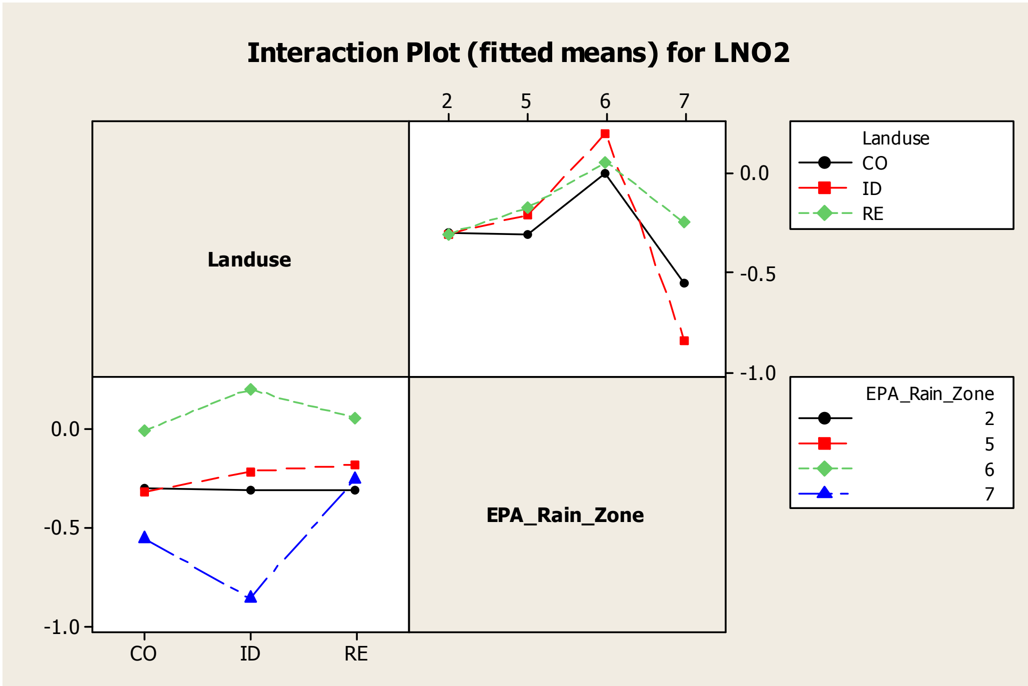
```
Tukey Simultaneous Tests
Response Variable LNO3
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value    P-Value
5                  0.0701     0.03138    2.234     0.1142
6                  0.3887     0.03418   11.372     0.0000
7                 -0.2455     0.03739   -6.567     0.0000


EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value    P-Value
6                  0.3186     0.04025    7.916     0.0000
7                 -0.3156     0.04301   -7.339     0.0000


EPA_Rain_Zone = 6  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value    P-Value
7                 -0.6342     0.04509   -14.07     0.0000
```

Nitrates in EPA rain zone 2 are significantly different than in rain zones 6, and 7. Nitrates in EPA rain zone 2 are not significantly than in rain zone 5.

Nitrates in EPA rain zone 5 are significantly different than in rain zones 6, and 7.

Nitrates in EPA rain zone 6 are significantly different than in rain zones 7.




Interaction Plot (fitted means) for LNO2
Landuse
EPA_Rain_Zone
2
5
6
7
0.0
-0.5
-1.0
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
5
6
7

Total Kjeldahl Nitrogen (TKN)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	114	9	0.174	0.331	-1.301	0.939
3	6	0	0.048	0.219	-0.143	0.477
4	0	16	*	*	*	*
5	40	2	-0.028	0.269	-0.586	0.602
6	34	0	0.567	0.264	0.017	1.079
7	37	4	-0.043	0.471	-1.696	0.919

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	94	15	0.001	0.307	-0.790	1.000
3	15	1	-0.194	0.415	-1.347	0.248
4	0	17	*	*	*	*
5	46	1	-0.074	0.294	-1.000	0.699
6	64	6	0.519	0.411	-0.602	1.204
7	24	9	0.192	0.487	-1.843	0.771

Land use: Residential

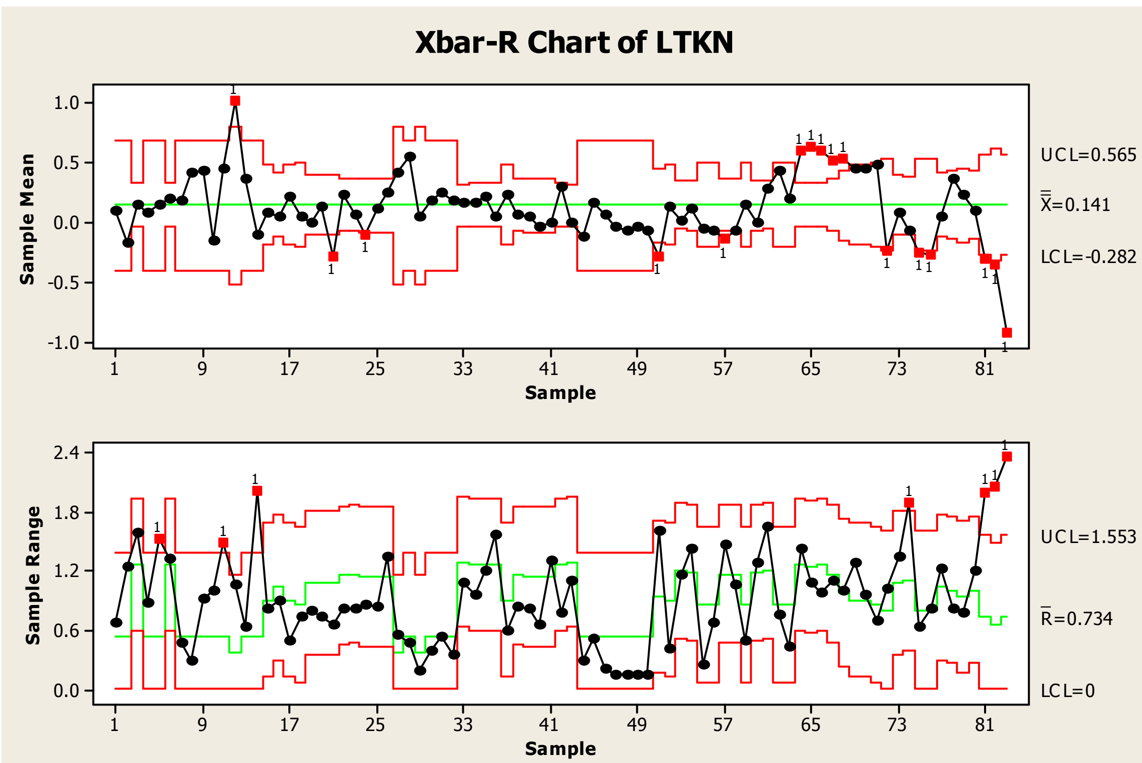
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	302	29	0.125	0.299	-1.000	1.556
3	17	1	0.048	0.161	-0.252	0.378
4	0	31	*	*	*	*
5	69	2	0.187	0.338	-0.691	1.000
6	37	1	0.504	0.300	-0.301	1.041
7	37	3	-0.167	0.579	-2.314	1.065

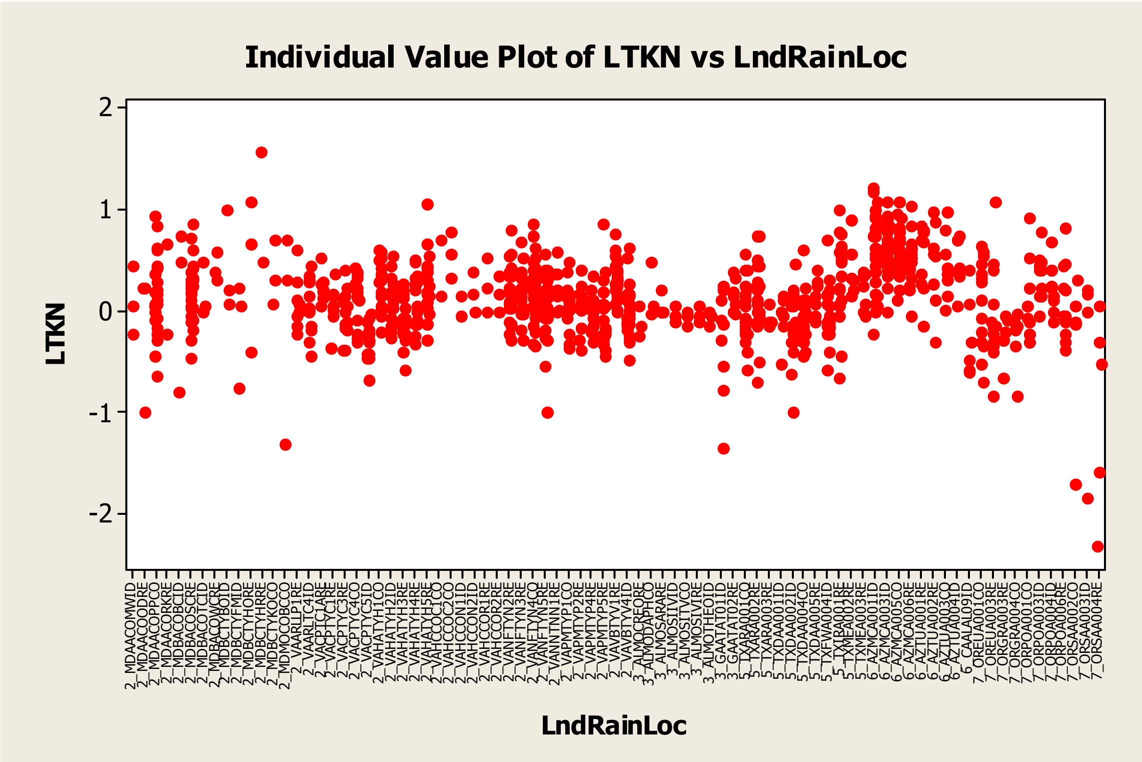
Descriptive Statistics: LTKN

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LTKN	2_KYLOTSR1RE	0	3	*	*	*	*
	2_KYLOTSR2ID	0	3	*	*	*	*
	2_KYLOTSR3RE	0	3	*	*	*	*
	2_KYLOTSR4ID	0	4	*	*	*	*
	2_MDAACOMWID	3	0	0.0889	0.114	-0.222	0.447
	2_MDAACOODRE	3	0	-0.180	0.505	-1.000	0.230
	2_MDAACOPPCO	26	0	0.1408	0.1412	-0.6353	0.9385
	2_MDAACORKRE	3	0	0.0698	0.255	-0.222	0.653
	2_MDBACOBCID	3	0	0.143	0.669	-0.789	0.740
	2_MDBACOSCRE	26	0	0.1970	0.0886	-0.4656	0.8513
	2_MDBACOTCID	3	0	0.173	0.0699	0.000000000	0.477
	2_MDBACOWCRE	3	0	0.4203	0.0206	0.3010	0.5798
	2_MDBCTYBOID	3	0	0.428	0.249	0.0792	1.000
	2_MDBCTYFMID	3	0	-0.162	0.276	-0.759	0.230
	2_MDBCTYHORE	3	0	0.441	0.587	-0.408	1.079
	2_MDBCTYHRRE	2	1	1.024	0.567	0.491	1.556
	2_MDBCTYKOCO	3	0	0.360	0.0986	0.0792	0.699
	2_MDMOCOBCCO	3	0	-0.100	1.121	-1.301	0.699
	2_VAARLLP1RE	8	0	0.0813	0.0678	-0.2147	0.6021
	2_VAARLTC4ID	12	1	0.0398	0.0767	-0.4451	0.4472

2_VACPTC1ARE	8	0	0.2131	0.0243	0.0414	0.5250
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	0.0479	0.0500	-0.3665	0.3579
2_VACPTYC3RE	14	0	-0.0127	0.0507	-0.3872	0.4014
2_VACPTYC4CO	14	0	0.1206	0.0613	-0.3098	0.4298
2_VACPTYC5ID	14	0	-0.2955	0.0314	-0.6778	-0.0362
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	0.2378	0.0554	-0.2007	0.6031
2_VAHATYH2ID	19	0	0.0597	0.0511	-0.2676	0.5353
2_VAHATYH3RE	17	0	-0.1125	0.0502	-0.5850	0.2625
2_VAHATYH4RE	17	0	0.1149	0.0647	-0.3188	0.5065
2_VAHATYH5RE	17	0	0.2432	0.1077	-0.2757	1.0550
2_VAHCCOC1CO	2	0	0.423	0.153	0.146	0.699
2_VAHCCOC2CO	3	0	0.555	0.0536	0.322	0.785
2_VAHCCON1ID	2	0	0.0502	0.0184	-0.0458	0.1461
2_VAHCCON2ID	3	0	0.185	0.0362	0.000000000	0.380
2_VAHCCOR1RE	3	0	0.254	0.0710	0.000000000	0.531
2_VAHCCOR2RE	3	0	0.1728	0.0293	0.000000000	0.3424
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	0.1693	0.0589	-0.2840	0.8000
2_VANFTYN3RE	27	0	0.1671	0.0461	-0.2757	0.6785
2_VANFTYN4CO	27	0	0.2207	0.0879	-0.3468	0.8500
2_VANFTYN5RE	27	0	0.0499	0.1049	-1.0000	0.5670
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	0.2322	0.0380	-0.0132	0.5752
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	0.0539	0.0448	-0.3565	0.4814
2_VAPMTYP2RE	17	0	0.0386	0.0486	-0.3872	0.4150
2_VAPMTYP4RE	17	0	-0.0385	0.0286	-0.3098	0.3385
2_VAPMTYP5RE	17	0	-0.00324	0.1173	-0.4437	0.8591
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	0.2966	0.0338	-0.00877	0.7574
2_VAVBTYV4ID	29	1	-0.00245	0.0548	-0.4815	0.6232
3_ALMOCREORE	3	0	-0.1197	0.0221	-0.2518	0.0414
3_ALMODAPHCO	3	0	0.165	0.0739	-0.0223	0.477
3_ALMOSARARE	3	0	0.0680	0.0139	0.000000000	0.2041
3_ALMOSIIVID	3	0	-0.0372	0.00530	-0.1024	0.0414
3_ALMOSITVCO	3	0	-0.0690	0.00462	-0.1427	-0.00877
3_ALMOSIVIRE	3	0	-0.0372	0.00530	-0.1024	0.0414
3_ALMOTHEOID	3	0	-0.0690	0.00462	-0.1427	-0.00877
3_GAATAT01ID	9	1	-0.288	0.273	-1.347	0.248
3_GAATAT02RE	8	1	0.1353	0.0234	-0.0278	0.3784
5_TXARA001CO	21	1	0.00765	0.0834	-0.5857	0.5635
5_TXARA002RE	20	1	0.1097	0.1304	-0.6905	0.7324
5_TXARA003RE	7	0	-0.0632	0.0101	-0.1549	0.0792
5_TXDAA001ID	7	0	-0.0662	0.0505	-0.5229	0.1461
5_TXDAA002ID	19	0	-0.1466	0.1066	-1.0000	0.4624
5_TXDAA004CO	19	1	-0.0675	0.0613	-0.4437	0.6021
5_TXDAA005RE	7	0	0.1449	0.0322	-0.0969	0.3979
5_TXFWA004ID	20	1	-0.00875	0.0778	-0.5807	0.6990
5_TXIRA001RE	21	1	0.2761	0.1655	-0.6517	1.0000
5_TXMEA002RE	7	0	0.4255	0.0675	0.1461	0.8921
5_TXMEA003RE	7	0	0.1962	0.0328	-0.0458	0.3802
6_AZMCA001ID	26	1	0.5937	0.1516	-0.2218	1.2041
6_AZMCA003ID	25	2	0.6372	0.0666	0.000000000	1.0792
6_AZMCA005CO	26	0	0.6018	0.0581	0.1139	1.0792
6_AZMCA006RE	19	1	0.5089	0.0538	-0.0458	1.0414

6_AZTUA001RE	10	0	0.532	0.113	-0.155	0.845
6_AZTUA002RE	8	0	0.457	0.176	-0.301	0.968
6_AZTUA003CO	8	0	0.452	0.102	0.0170	0.973
6_AZTUA004ID	7	0	0.4765	0.0649	0.0414	0.7340
6_CAALAL09ID	6	3	-0.246	0.169	-0.602	0.415
7_OREUA001CO	14	2	0.0821	0.194	-0.705	0.633
7_OREUA003RE	15	0	-0.0675	0.193	-0.827	1.064
7_ORGRA003RE	6	0	-0.2582	0.0490	-0.6655	-0.0400
7_ORGRA004CO	6	0	-0.276	0.0855	-0.833	-0.0223
7_ORPOA001CO	12	1	0.0416	0.140	-0.301	0.919
7_ORPOA003ID	11	3	0.3577	0.0498	-0.0458	0.7709
7_ORPOA004ID	9	4	0.2381	0.0558	-0.0969	0.6812
7_ORPOA006RE	11	2	0.0972	0.136	-0.377	0.813
7_ORSAA002CO	5	1	-0.315	0.624	-1.696	0.301
7_ORSAA003ID	4	2	-0.366	0.978	-1.843	0.204
7_ORSAA004RE	5	1	-0.936	0.961	-2.314	0.0414





General Linear Model: LTKN versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	5	2, 3, 5, 6, 7

Analysis of Variance for LTKN, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	0.1067	0.2805	0.1403	1.20	0.300
EPA_Rain_Zone	4	26.9717	24.3602	6.0901	52.29	0.000
Landuse*EPA_Rain_Zone	8	5.2809	5.2809	0.6601	5.67	0.000
Error	921	107.2734	107.2734	0.1165		
Total	935	139.6327				

Land use is not significant. EPA rain zone and interaction are significant.

Bonferroni Simultaneous Tests
Response Variable LTKN
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.05489	0.04388	-1.251	0.6338
RE	-0.00419	0.04249	-0.099	1.0000

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.05069	0.03657	1.386	0.4980

Tukey Simultaneous Tests
Response Variable LTKN
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	-0.05489	0.04388	-1.251	0.4232
RE	-0.00419	0.04249	-0.099	0.9946

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	0.05069	0.03657	1.386	0.3481

No significant differences among land uses.

Bonferroni Simultaneous Tests
Response Variable LTKN
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1325	0.06384	-2.076	0.3814
5	-0.0716	0.03296	-2.174	0.3000
6	0.4300	0.03502	12.276	0.0000
7	-0.1056	0.03915	-2.698	0.0711

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.06090	0.06763	0.9005	1.0000
6	0.56251	0.06866	8.1931	0.0000
7	0.02693	0.07085	0.3801	1.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.50161	0.04153	12.0770	0.0000
7	-0.03397	0.04507	-0.7537	1.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.5356	0.04660	-11.49	0.0000

TKN in rain zone 2 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 2 and zones 3, 5, and 7.

TKN in rain zone 3 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 3 and zones 5, and 7.

TKN in rain zone 5 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 5 and 7.

TKN in rain zone 6 is significantly different than TKN in rain zone 7.

Tukey Simultaneous Tests
Response Variable LTKN
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1325	0.06384	-2.076	0.2303
5	-0.0716	0.03296	-2.174	0.1897
6	0.4300	0.03502	12.276	0.0000
7	-0.1056	0.03915	-2.698	0.0543

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.06090	0.06763	0.9005	0.8968
6	0.56251	0.06866	8.1931	0.0000
7	0.02693	0.07085	0.3801	0.9956

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.50161	0.04153	12.0770	0.0000
7	-0.03397	0.04507	-0.7537	0.9436

EPA_Rain_Zone = 6 subtracted from:

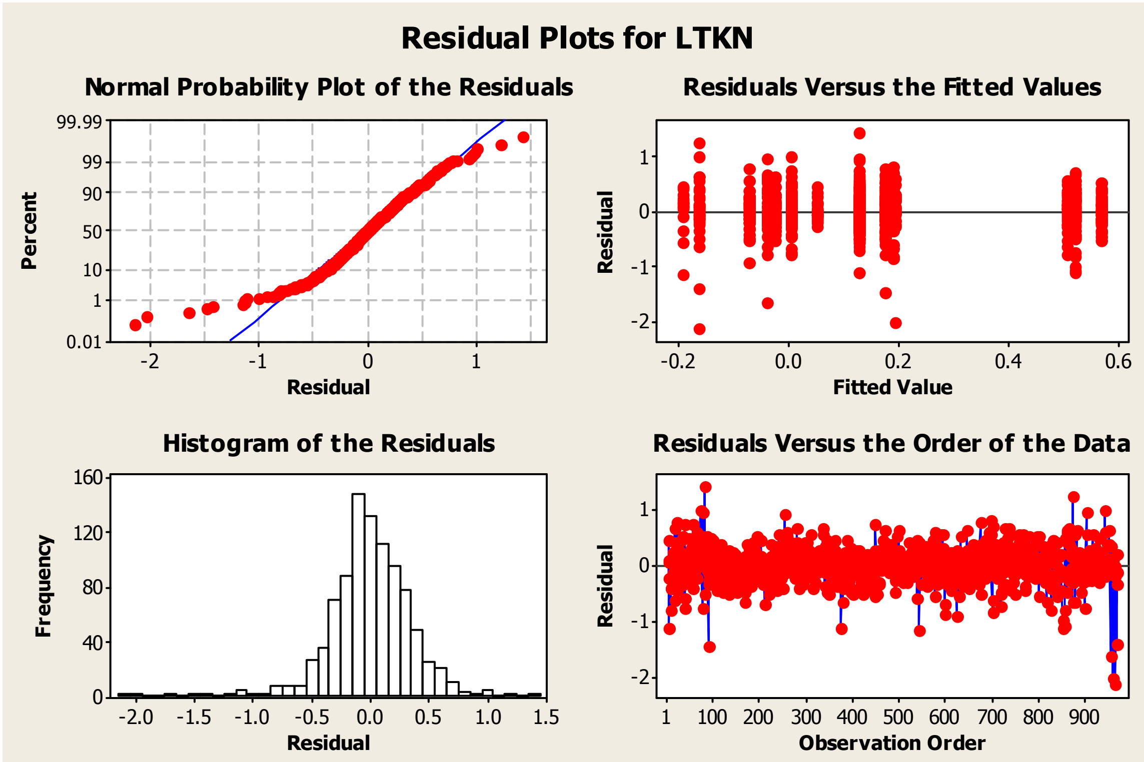
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.5356	0.04660	-11.49	0.0000

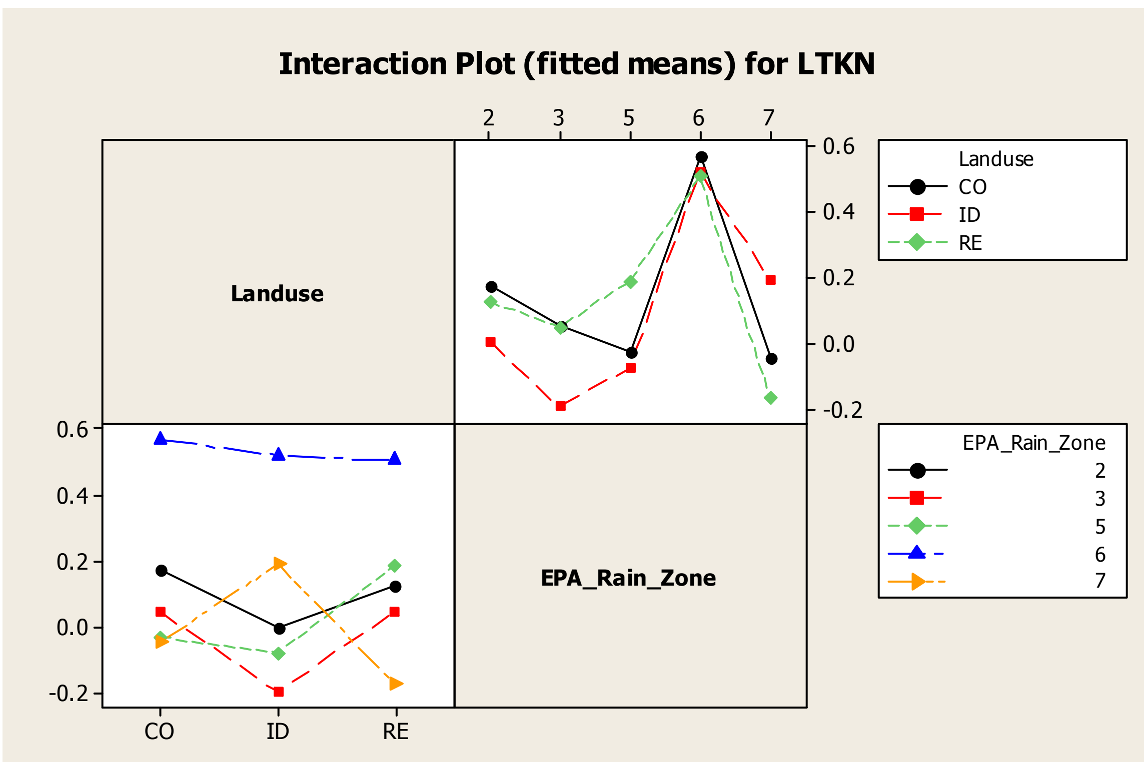
TKN in rain zone 2 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 2 and zones 3, 5, and 7.

TKN in rain zone 3 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 3 and zones 5, and 7.

TKN in rain zone 5 is significantly different than TKN in rain zone 6. No differences were observed between rain zone 5 and 7.

TKN in rain zone 6 is significantly different than TKN in rain zone 7.


Residual Plots for LTKN
Normal Probability Plot of the Residuals
Percent
99.99
99
90
50
10
1
0.01
-2
-1
0
1
Residual
Residuals Versus the Fitted Values
Residual
1
0
-1
-2
-0.2
0.0
0.2
0.4
0.6
Fitted Value
Histogram of the Residuals
Frequency
160
120
80
40
0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
Residual
Residuals Versus the Order of the Data
Residual
1
0
-1
-2
1
100
200
300
400
500
600
700
800
900
Observation Order


Interaction Plot (fitted means) for LTKN
2
3
5
6
7
0.6
0.4
0.2
0.0
-0.2
Landuse
CO
ID
RE
Landuse
EPA_Rain_Zone
2
3
5
6
7
EPA_Rain_Zone
CO
ID
RE

Total Phosphorus (P)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	114	9	-0.624	0.279	-1.222	0.243
3	6	0	-0.722	0.330	-1.187	-0.337
4	0	16*	*	*	*	
5	40	2	-0.869	0.411	-1.640	0.630
6	34	0	-0.344	0.243	-0.796	0.301
7	37	4	-0.591	0.362	-1.699	0.519

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	96	13	-0.706	0.387	-2.200	0.111
3	15	1	-0.793	0.403	-1.523	-0.046
4	0	17*	*	*	*	
5	45	2	-0.714	0.310	-1.337	0.107
6	59	11	-0.017	0.356	-0.854	0.898
7	24	9	-0.289	0.288	-1.194	0.146

Land use: Residential

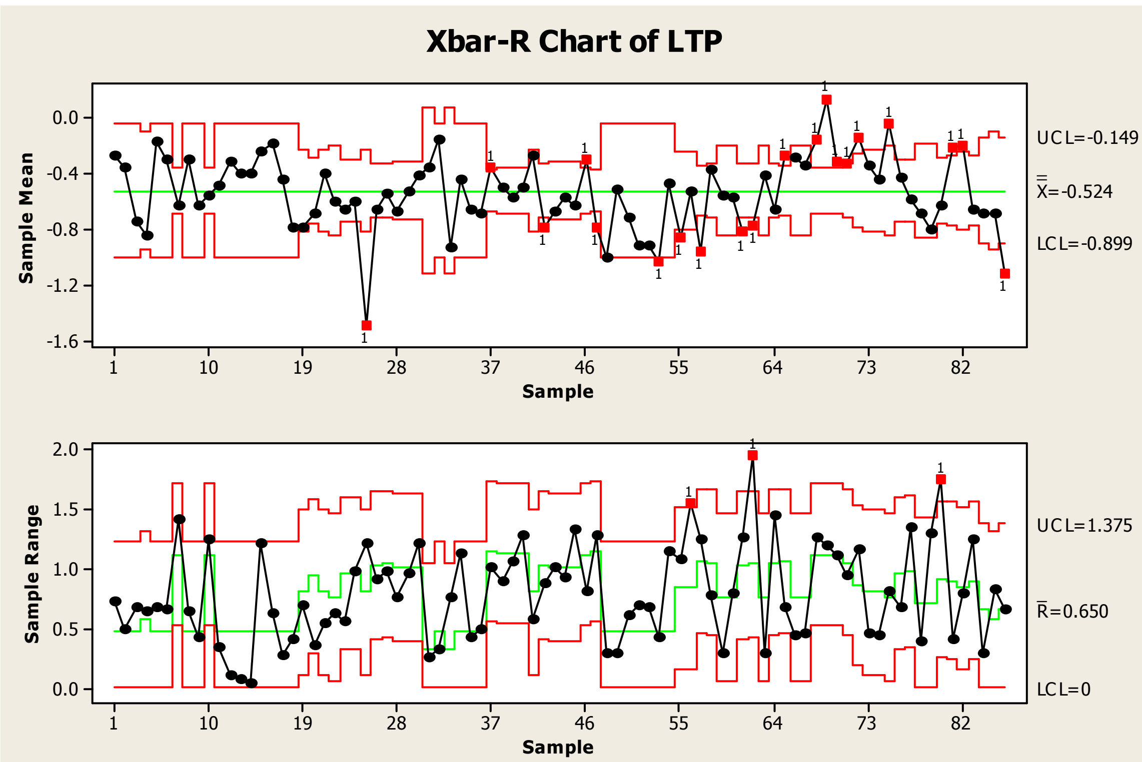
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	309	22	-0.503	0.281	-1.301	0.277
3	18	0	-0.724	0.409	-1.187	0.444
4	0	31*	*	*	*	
5	69	2	-0.354	0.190	-0.721	0.049
6	38	0	-0.288	0.263	-0.854	0.696
7	37	3	-0.700	0.337	-1.420	0.342

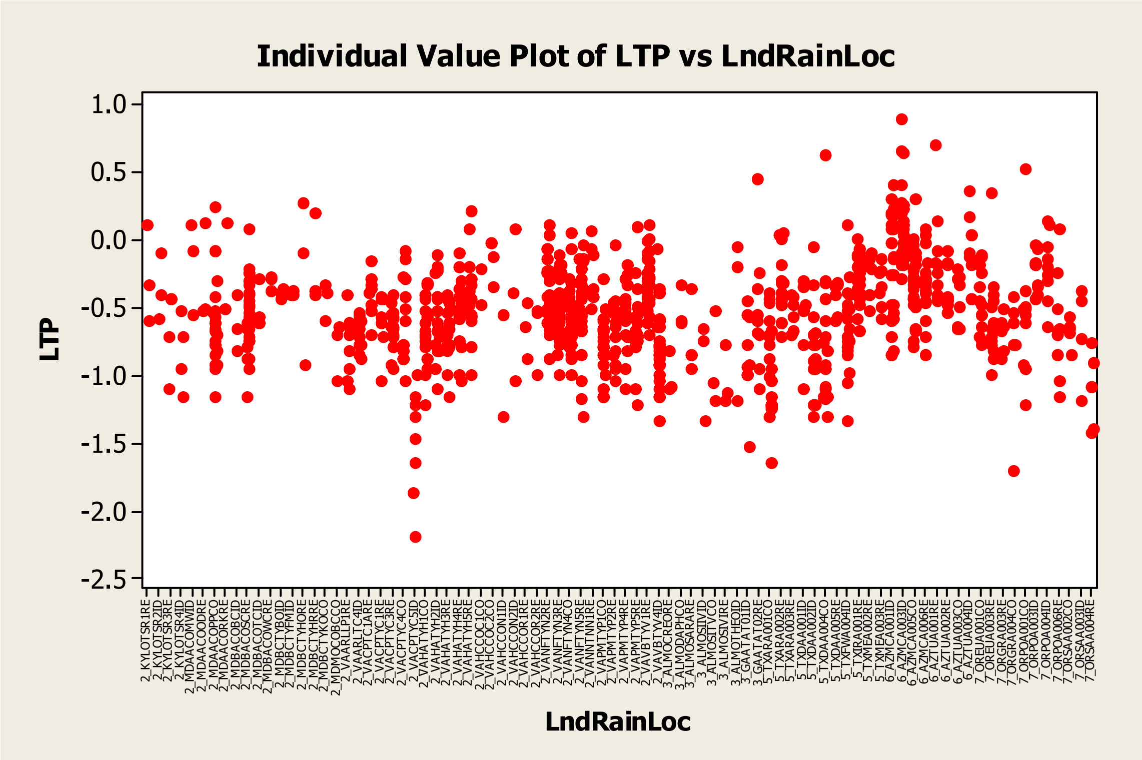
Descriptive Statistics: LTP

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LTP	2_KYLOTSR1RE	3	0	-0.275	0.131	-0.602	0.114
	2_KYLOTSR2ID	3	0	-0.362	0.0626	-0.585	-0.0915
	2_KYLOTSR3RE	3	0	-0.750	0.111	-1.097	-0.432
	2_KYLOTSR4ID	4	0	-0.839	0.0760	-1.155	-0.523
	2_MDAACOMWID	3	0	-0.176	0.116	-0.553	0.111
	2_MDAACOODRE	3	0	-0.301	0.138	-0.523	0.127
	2_MDAACOPPCO	26	0	-0.6324	0.0941	-1.1549	0.2430
	2_MDAACORKRE	3	0	-0.297	0.135	-0.509	0.127
	2_MDBACOBCID	3	0	-0.630	0.0436	-0.824	-0.409
	2_MDBACOSCRE	26	0	-0.5524	0.0746	-1.1549	0.0828
	2_MDBACOTCID	3	0	-0.494	0.0310	-0.620	-0.292
	2_MDBACOWCRE	3	0	-0.3122	0.00315	-0.3768	-0.2757
	2_MDBCTYBOID	3	0	-0.3988	0.00107	-0.4318	-0.3665
	2_MDBCTYFMID	3	0	-0.3982	0.000345	-0.4089	-0.3768
	2_MDBCTYHORE	3	0	-0.247	0.375	-0.921	0.276
	2_MDBCTYHRRE	3	0	-0.194	0.119	-0.409	0.204
	2_MDBCTYKOCO	3	0	-0.4426	0.0203	-0.6021	-0.3279
	2_MDMOCOBCCO	3	0	-0.794	0.0483	-1.046	-0.638
	2_VAARLLP1RE	8	0	-0.7864	0.0564	-1.0969	-0.4089
	2_VAARLTC4ID	13	0	-0.6859	0.0127	-0.8861	-0.5376

2_VACPTC1ARE	8	0	-0.4048	0.0266	-0.6990	-0.1612
2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	-0.6056	0.0506	-1.0458	-0.4202
2_VACPTYC3RE	14	0	-0.6537	0.0247	-0.9586	-0.4089
2_VACPTYC4CO	14	0	-0.5987	0.0982	-1.0458	-0.0809
2_VACPTYC5ID	8	6	-1.484	0.162	-2.200	-1.000
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	18	0	-0.6584	0.0618	-1.2218	-0.3188
2_VAHATYH2ID	19	0	-0.5477	0.0780	-1.0969	-0.1192
2_VAHATYH3RE	17	0	-0.6790	0.0416	-1.1549	-0.4089
2_VAHATYH4RE	17	0	-0.5309	0.0641	-1.0458	-0.0915
2_VAHATYH5RE	17	0	-0.4149	0.0828	-1.0000	0.2095
2_VAHCCOC1CO	2	0	-0.352	0.0337	-0.481	-0.222
2_VAHCCOC2CO	3	0	-0.1647	0.0275	-0.3468	-0.0223
2_VAHCCON1ID	2	0	-0.927	0.280	-1.301	-0.553
2_VAHCCON2ID	3	0	-0.450	0.321	-1.046	0.0828
2_VAHCCOR1RE	3	0	-0.664	0.0441	-0.886	-0.469
2_VAHCCOR2RE	3	0	-0.687	0.0736	-1.000	-0.523
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	-0.3548	0.0564	-0.8861	0.1106
2_VANFTYN3RE	27	0	-0.5054	0.0490	-1.0000	-0.1135
2_VANFTYN4CO	27	0	-0.5751	0.0563	-1.0000	0.0492
2_VANFTYN5RE	27	0	-0.4987	0.0971	-1.3010	-0.0315
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	8	2	-0.2777	0.0438	-0.5086	0.0607
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	-0.7819	0.0522	-1.1549	-0.2924
2_VAPMTYP2RE	17	0	-0.6677	0.0582	-1.0458	-0.0362
2_VAPMTYP4RE	17	0	-0.5667	0.0596	-1.0969	-0.1805
2_VAPMTYP5RE	17	0	-0.6325	0.0892	-1.2218	0.1004
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	-0.3085	0.0395	-0.6990	0.1038
2_VAVBTYV4ID	29	1	-0.7816	0.0742	-1.3379	-0.0655
3_ALMOCREORE	3	0	-1.0042	0.0244	-1.1024	-0.8239
3_ALMODAPHCO	3	0	-0.5197	0.0250	-0.6198	-0.3372
3_ALMOSARARE	3	0	-0.723	0.103	-0.959	-0.357
3_ALMOSIIVID	3	0	-0.914	0.138	-1.340	-0.658
3_ALMOSITVCO	3	0	-0.923	0.124	-1.187	-0.523
3_ALMOSIVIRE	3	0	-1.031	0.0519	-1.187	-0.770
3_ALMOTHEOID	3	0	-0.476	0.385	-1.187	-0.0458
3_GAATAT01ID	9	1	-0.859	0.105	-1.523	-0.455
3_GAATAT02RE	9	0	-0.528	0.202	-1.097	0.444
5_TXARA001CO	21	1	-0.9610	0.1163	-1.6401	-0.3979
5_TXARA002RE	20	1	-0.3678	0.0441	-0.7212	0.0492
5_TXARA003RE	7	0	-0.5650	0.0168	-0.6990	-0.4089
5_TXDAA001ID	7	0	-0.577	0.0747	-1.097	-0.310
5_TXDAA002ID	18	1	-0.8226	0.0931	-1.3010	-0.0458
5_TXDAA004CO	19	1	-0.767	0.216	-1.310	0.630
5_TXDAA005RE	7	0	-0.4158	0.0144	-0.6021	-0.3188
5_TXFWA004ID	20	1	-0.6650	0.0943	-1.3372	0.1072
5_TXIRA001RE	21	1	-0.2756	0.0308	-0.6778	-0.000435
5_TXMEA002RE	7	0	-0.2835	0.0265	-0.5229	-0.0915
5_TXMEA003RE	7	0	-0.3473	0.0245	-0.5850	-0.1367
6_AZMCA001ID	26	1	-0.1577	0.1547	-0.8539	0.3979
6_AZMCA003ID	26	1	0.1312	0.0790	-0.2924	0.8976
6_AZMCA005CO	26	0	-0.3129	0.0636	-0.7959	0.3010
6_AZMCA006RE	20	0	-0.3353	0.0517	-0.8539	0.0792

6_AZTUA001RE	10	0	-0.146	0.121	-0.456	0.695
6_AZTUA002RE	8	0	-0.3468	0.0299	-0.5376	-0.0809
6_AZTUA003CO	8	0	-0.4447	0.0355	-0.6576	-0.2218
6_AZTUA004ID	7	0	-0.0404	0.0668	-0.4437	0.3560
6_CAALAL09ID	0	9	*	*	*	*
7_OREUA001CO	14	2	-0.4341	0.0579	-0.7696	-0.1079
7_OREUA003RE	15	0	-0.5909	0.0987	-1.0000	0.3424
7_ORGRA003RE	6	0	-0.6907	0.0195	-0.8861	-0.5086
7_ORGRA004CO	6	0	-0.804	0.211	-1.699	-0.418
7_ORPOA001CO	12	1	-0.630	0.200	-1.222	0.519
7_ORPOA003ID	11	3	-0.2166	0.0208	-0.4437	-0.0362
7_ORPOA004ID	9	4	-0.1989	0.0645	-0.6383	0.1461
7_ORPOA006RE	11	2	-0.665	0.121	-1.155	0.0792
7_ORSAA002CO	5	1	-0.6827	0.0113	-0.8539	-0.5654
7_ORSAA003ID	4	2	-0.688	0.138	-1.194	-0.374
7_ORSAA004RE	5	1	-1.116	0.0848	-1.420	-0.762





General Linear Model: LTP versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	5	2, 3, 5, 6, 7

Analysis of Variance for LTP, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	2.9935	1.1228	0.5614	5.85	0.003
EPA_Rain_Zone	4	20.6147	17.8961	4.4740	46.65	0.000
Landuse*EPA_Rain_Zone	8	12.3467	12.3467	1.5433	16.09	0.000
Error	926	88.8109	88.8109	0.0959		
Total	940	124.7659				

Main factors and interaction are significant in the model

Bonferroni Simultaneous Tests
Response Variable LTP
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.1263	0.03989	3.166	0.0048
RE	0.1163	0.03836	3.031	0.0075

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.01003	0.03304	-0.3036	1.000

Tukey Simultaneous Tests
Response Variable LTP
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.1263	0.03989	3.166	0.0044
RE	0.1163	0.03836	3.031	0.0069

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.01003	0.03304	-0.3036	0.9505

Total phosphorus in commercial land use areas is significantly different than total phosphorus in residential or industrial land use areas. There was no difference in total phosphorus concentration between residential and industrial land uses.

Bonferroni Simultaneous Tests
Response Variable LTP
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1349	0.05760	-2.342	0.1941
5	-0.0345	0.02994	-1.151	1.0000
6	0.3951	0.03183	12.413	0.0000
7	0.0846	0.03548	2.385	0.1729

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.1004	0.06112	1.643	1.0000
6	0.5300	0.06207	8.539	0.0000
7	0.2195	0.06402	3.429	0.0063

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.4296	0.03784	11.352	0.0000
7	0.1191	0.04096	2.907	0.0373

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value

7	-0.3105	0.04236	-7.330	0.0000

Total phosphorus in EPA rain zone 2 is different than total phosphorus in EPA rain zone 6. No difference was observed between EPA rain zone 2 and zones 3, 5, and 7

Total phosphorus in EPA rain zone 3 is different than total phosphorus in EPA rain zones 6 and 7. No difference was observed between EPA rain zone 3 and EPA rain zone 5

Total phosphorus in EPA rain zone 5 is different than total phosphorus in EPA rain zones 6 and 7.

Total phosphorus in EPA rain zone 6 is different than total phosphorus in EPA rain zone 7.

Tukey Simultaneous Tests
Response Variable LTP
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1349	0.05760	-2.342	0.1319
5	-0.0345	0.02994	-1.151	0.7792
6	0.3951	0.03183	12.413	0.0000
7	0.0846	0.03548	2.385	0.1194

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	0.1004	0.06112	1.643	0.4699
6	0.5300	0.06207	8.539	0.0000
7	0.2195	0.06402	3.429	0.0055

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.4296	0.03784	11.352	0.0000
7	0.1191	0.04096	2.907	0.0300

EPA_Rain_Zone = 6 subtracted from:

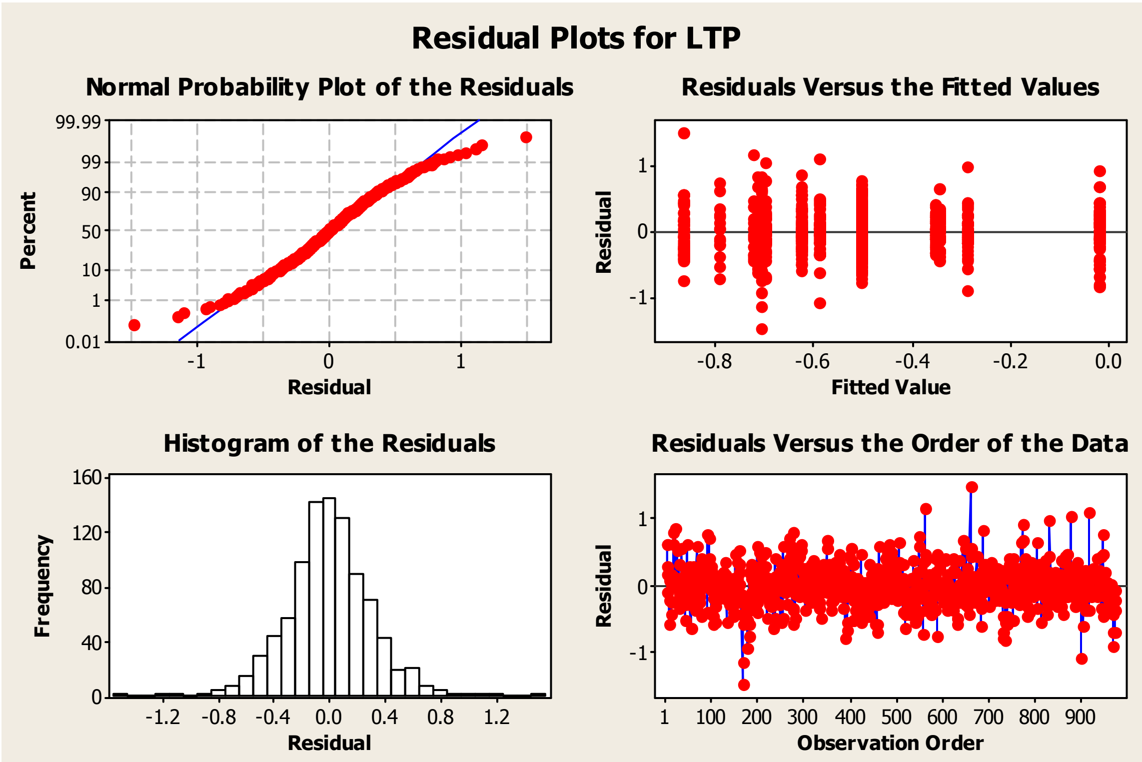
EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.3105	0.04236	-7.330	0.0000

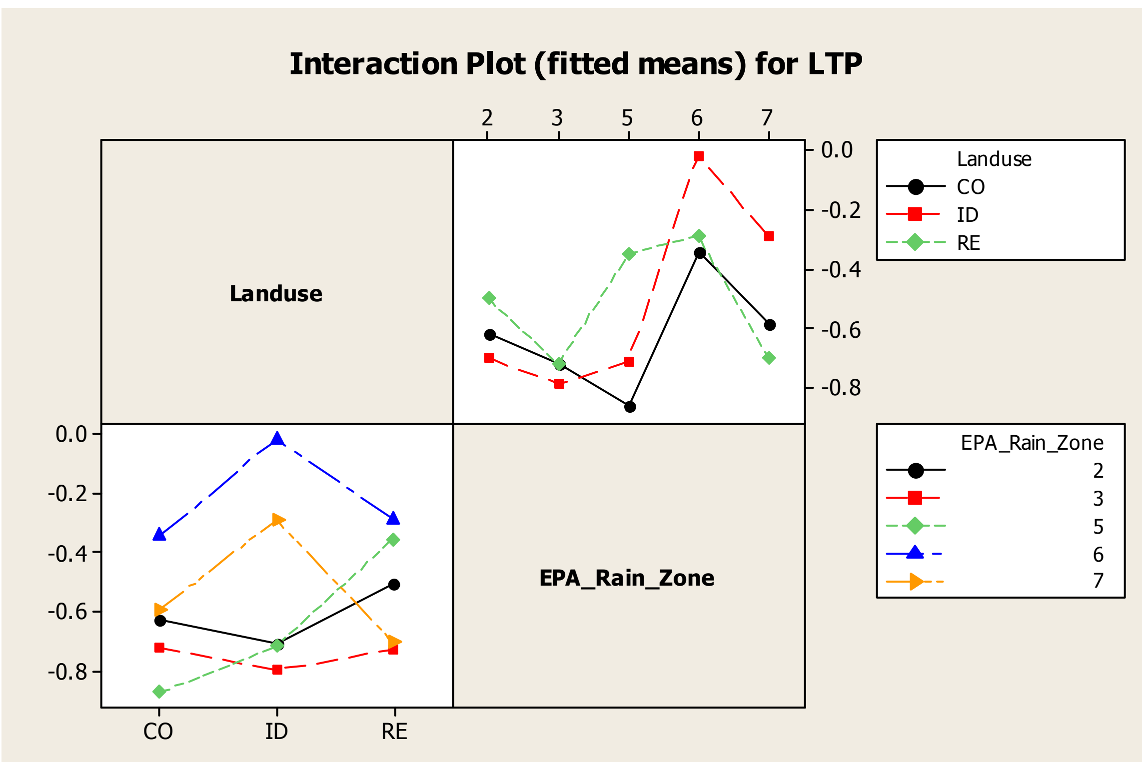
Total phosphorus in EPA rain zone 2 is different than total phosphorus in EPA rain zone 6. No difference was observed between EPA rain zone 2 and zones 3, 5, and 7

Total phosphorus in EPA rain zone 3 is different than total phosphorus in EPA rain zones 6 and 7. No difference was observed between EPA rain zone 3 and EPA rain zone 5

Total phosphorus in EPA rain zone 5 is different than total phosphorus in EPA rain zones 6 and 7.

Total phosphorus in EPA rain zone 6 is different than total phosphorus in EPA rain zone 7.


Residual Plots for LTP
Normal Probability Plot of the Residuals
Percent
Residual
Residuals Versus the Fitted Values
Residual
Fitted Value
Histogram of the Residuals
Frequency
Residual
Residuals Versus the Order of the Data
Residual
Observation Order


Interaction Plot (fitted means) for LTP
Landuse
EPA_Rain_Zone
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
5
6
7

Dissolved Phosphorus (dissolved – P)

Summary statistics in LOG base 10 scale

Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	73	50	-1.185	0.403	-2.168	-0.444
3	6	0	-1.168	0.438	-1.602	-0.602
4	0	16	*	*	*	*
5	40	2	-1.387	0.307	-2.043	-0.770
6	26	8	-0.500	0.354	-1.301	0.204
7	7	34	-1.546	0.317	-2.000	-1.051

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	67	42	-1.103	0.360	-2.000	-0.347
3	11	5	-1.460	0.213	-1.602	-0.991
4	0	17	*	*	*	*
5	46	1	-1.172	0.286	-1.745	-0.409
6	52	18	-0.654	0.310	-1.309	0.176
7	2	31	-1.150	0.415	-1.444	-0.857

Land use: Residential

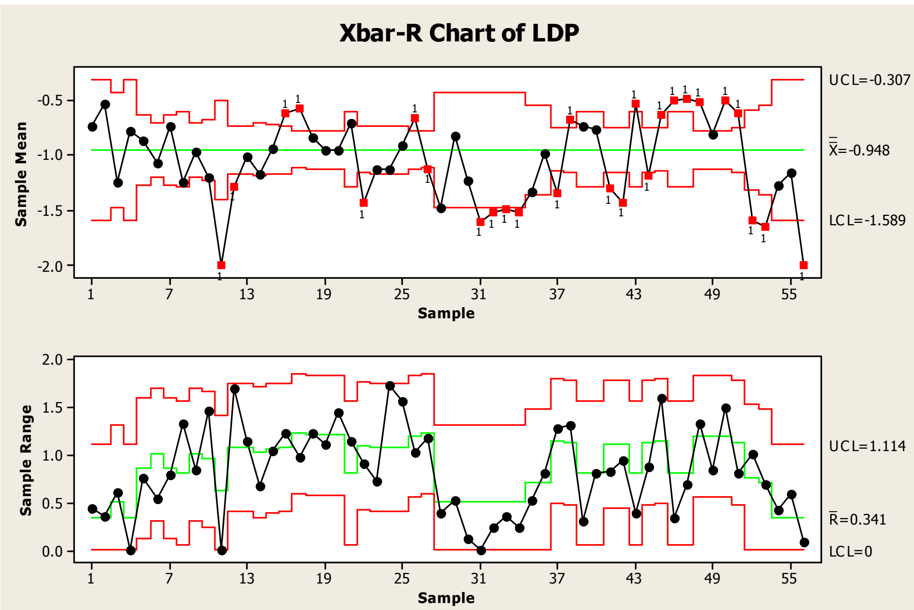
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	255	76	-0.869	0.370	-2.088	0.029
3	14	4	-1.247	0.320	-1.602	-0.585
4	0	31	*	*	*	*
5	69	2	-0.612	0.292	-1.770	-0.076
6	20	18	-0.609	0.213	-0.959	-0.155
7	8	32	-1.691	0.349	-2.060	-1.046

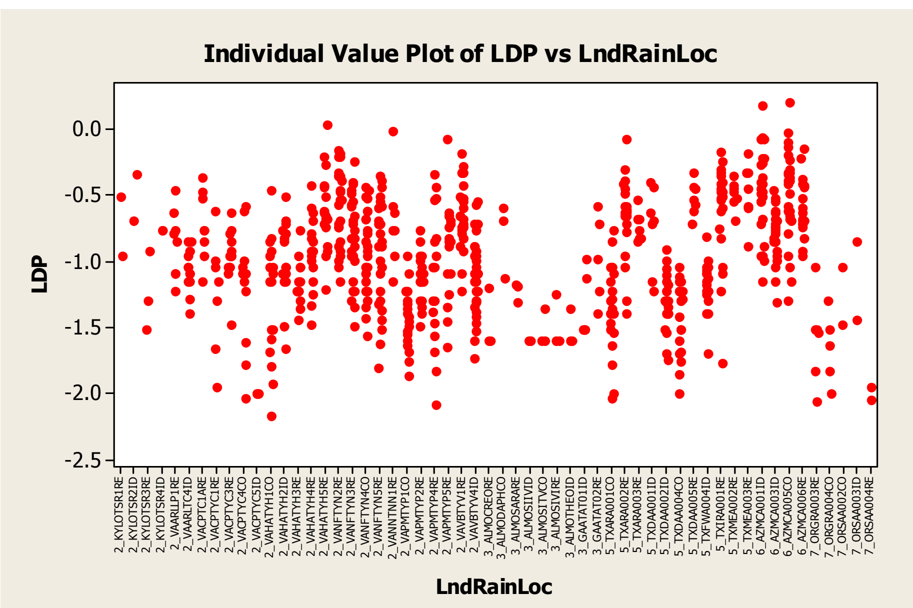
Descriptive Statistics: LDP

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LDP	2_KYLOTSR1RE	2	1	-0.734	0.101	-0.959	-0.509
	2_KYLOTSR2ID	2	1	-0.523	0.0620	-0.699	-0.347
	2_KYLOTSR3RE	3	0	-1.248	0.0927	-1.523	-0.921
	2_KYLOTSR4ID	2	2	-0.76955	0.000000000	-0.76955	-0.76955
	2_MDAACOMWID	0	3	*	*	*	*
	2_MDAACOODRE	0	3	*	*	*	*
	2_MDAACOPPCO	0	26	*	*	*	*
	2_MDAACORKRE	0	3	*	*	*	*
	2_MDBACOBCID	0	3	*	*	*	*
	2_MDBACOSCRE	0	26	*	*	*	*
	2_MDBACOTCID	0	3	*	*	*	*
	2_MDBACOWCRE	0	3	*	*	*	*
	2_MDBCTYBOID	0	3	*	*	*	*
	2_MDBCTYFMID	0	3	*	*	*	*
	2_MDBCTYHORE	0	3	*	*	*	*
	2_MDBCTYHRRE	0	3	*	*	*	*
	2_MDBCTYKOCO	0	3	*	*	*	*
	2_MDMOCOBCCO	0	3	*	*	*	*
	2_VAARLLP1RE	8	0	-0.8677	0.0654	-1.2218	-0.4685
	2_VAARLTC4ID	13	0	-1.0604	0.0266	-1.3979	-0.8539
	2_VACPTC1ARE	8	0	-0.7347	0.0697	-1.1549	-0.3665

2_VACPTSF2RE	0	3	*	*	*	*
2_VACPTYC1RE	7	0	-1.249	0.198	-1.958	-0.620
2_VACPTYC3RE	13	1	-0.9666	0.0406	-1.4776	-0.6383
2_VACPTYC4CO	11	3	-1.205	0.202	-2.044	-0.585
2_VACPTYC5ID	4	10	-2.0000	0.000000000	-2.0000	-2.0000
2_VACPTYO1ID	0	3	*	*	*	*
2_VAHATYH1CO	17	1	-1.279	0.208	-2.168	-0.469
2_VAHATYH2ID	17	2	-1.0041	0.0829	-1.6641	-0.5086
2_VAHATYH3RE	14	3	-1.1679	0.0293	-1.4466	-0.7696
2_VAHATYH4RE	16	1	-0.9326	0.0786	-1.4864	-0.4318
2_VAHATYH5RE	17	0	-0.6210	0.0901	-1.2100	0.0294
2_VAHCCOC1CO	0	2	*	*	*	*
2_VAHCCOC2CO	0	3	*	*	*	*
2_VAHCCON1ID	0	2	*	*	*	*
2_VAHCCON2ID	0	3	*	*	*	*
2_VAHCCOR1RE	0	3	*	*	*	*
2_VAHCCOR2RE	0	3	*	*	*	*
2_VANFTMS5CO	0	3	*	*	*	*
2_VANFTMS6RE	0	3	*	*	*	*
2_VANFTMS8CO	0	3	*	*	*	*
2_VANFTMS9CO	0	3	*	*	*	*
2_VANFTYN2RE	29	0	-0.5698	0.0803	-1.1549	-0.1675
2_VANFTYN3RE	27	0	-0.8331	0.1009	-1.4916	-0.2518
2_VANFTYN4CO	27	0	-0.9480	0.0990	-1.5633	-0.4437
2_VANFTYN5RE	27	0	-0.9455	0.1574	-1.8074	-0.3565
2_VANNTMF1RE	0	1	*	*	*	*
2_VANNTMF4RE	0	3	*	*	*	*
2_VANNTNN1RE	7	3	-0.698	0.129	-1.155	-0.0132
2_VANNTSF4RE	0	2	*	*	*	*
2_VANNTSF6RE	0	2	*	*	*	*
2_VAPMTYP1CO	18	0	-1.4390	0.0542	-1.8681	-0.9586
2_VAPMTYP2RE	17	0	-1.1321	0.0380	-1.4984	-0.7696
2_VAPMTYP4RE	17	0	-1.131	0.261	-2.088	-0.347
2_VAPMTYP5RE	17	0	-0.9125	0.1402	-1.6493	-0.0757
2_VAVBTYI1ID	0	3	*	*	*	*
2_VAVBTYR1RE	0	5	*	*	*	*
2_VAVBTYV1RE	26	1	-0.6572	0.0584	-1.2218	-0.1871
2_VAVBTYV4ID	29	1	-1.1185	0.0902	-1.7327	-0.5528
3_ALMOCREORE	3	0	-1.471	0.0519	-1.602	-1.208
3_ALMODAPHCO	3	0	-0.813	0.0811	-1.137	-0.602
3_ALMOSARARE	3	0	-1.2252	0.00516	-1.3080	-1.1805
3_ALMOSIIVID	3	0	-1.6021	0.000000000	-1.6021	-1.6021
3_ALMOSITVCO	3	0	-1.5236	0.0185	-1.6021	-1.3665
3_ALMOSIVIRE	3	0	-1.485	0.0409	-1.602	-1.252
3_ALMOTHEOID	3	0	-1.5236	0.0185	-1.6021	-1.3665
3_GAATAT01ID	5	5	-1.337	0.0670	-1.523	-0.991
3_GAATAT02RE	5	4	-0.982	0.115	-1.398	-0.585
5_TXARA001CO	21	1	-1.3512	0.1125	-2.0433	-0.7696
5_TXARA002RE	20	1	-0.6680	0.0992	-1.3979	-0.0757
5_TXARA003RE	7	0	-0.7378	0.0128	-0.8539	-0.5376
5_TXDAA001ID	7	0	-0.755	0.102	-1.222	-0.409
5_TXDAA002ID	19	0	-1.3062	0.0486	-1.7447	-0.9172
5_TXDAA004CO	19	1	-1.4267	0.0765	-2.0000	-1.0458
5_TXDAA005RE	7	0	-0.5263	0.0160	-0.7212	-0.3372
5_TXFWA004ID	20	1	-1.1893	0.0329	-1.6990	-0.8125
5_TXIRA001RE	21	1	-0.6270	0.1462	-1.7696	-0.1688
5_TXMEA002RE	7	0	-0.5001	0.0117	-0.6990	-0.3565
5_TXMEA003RE	7	0	-0.4830	0.0534	-0.8861	-0.1871
6_AZMCA001ID	26	1	-0.5093	0.1154	-1.1549	0.1761
6_AZMCA003ID	26	1	-0.7985	0.0364	-1.3086	-0.4685
6_AZMCA005CO	26	0	-0.5003	0.1250	-1.3010	0.2041
6_AZMCA006RE	20	0	-0.6088	0.0452	-0.9586	-0.1549
6_AZTUA001RE	0	10	*	*	*	*

6_AZTUA002RE	0	8	*	*	*	*
6_AZTUA003CO	0	8	*	*	*	*
6_AZTUA004ID	0	7	*	*	*	*
6_CAALAL09ID	0	9	*	*	*	*
7_OREUA001CO	0	16	*	*	*	*
7_OREUA003RE	0	15	*	*	*	*
7_ORGRA003RE	6	0	-1.587	0.118	-2.060	-1.046
7_ORGRA004CO	5	1	-1.658	0.0729	-2.000	-1.301
7_ORPOA001CO	0	13	*	*	*	*
7_ORPOA003ID	0	14	*	*	*	*
7_ORPOA004ID	0	13	*	*	*	*
7_ORPOA006RE	0	13	*	*	*	*
7_ORSAA002CO	2	4	-1.266	0.0928	-1.481	-1.051
7_ORSAA003ID	2	4	-1.150	0.172	-1.444	-0.857
7_ORSAA004RE	2	4	-2.0022	0.00380	-2.0458	-1.9586





General Linear Model: LDP versus Landuse, EPA_Rain_Zone

```
Factor         Type   Levels  Values
Landuse        fixed       3  CO, ID, RE
EPA_Rain_Zone  fixed       5  2, 3, 5, 6, 7


Analysis of Variance for LDP, using Adjusted SS for Tests

Source                  DF    Seq SS   Adj SS  Adj MS      F      P
Landuse                  2   10.5050   0.9326  0.4663   3.90  0.021
EPA_Rain_Zone            4   28.0949  20.2703  5.0676  42.40  0.000
Landuse*EPA_Rain_Zone    8    9.9010   9.9010  1.2376  10.36  0.000
Error                  681   81.3873  81.3873  0.1195
Total                  695  129.8882
```

Main factors and interaction are significant in the model

```
Bonferroni Simultaneous Tests
Response Variable LDP
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID          0.04954     0.07029   0.7048    1.0000
RE          0.15159     0.05581   2.7163    0.0203


Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE           0.1021     0.06602    1.546    0.3678



Tukey Simultaneous Tests
Response Variable LDP
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID          0.04954     0.07029   0.7048    0.7606
RE          0.15159     0.05581   2.7163    0.0181


Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE           0.1021     0.06602    1.546    0.2696
```

Dissolved phosphorus in commercial land use is significantly different than in residential land use areas. No significant difference was observed between commercial and industrial land uses. No significant difference was observed between industrial and residential land uses.

```
Bonferroni Simultaneous Tests
Response Variable LDP
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
3                 -0.2393     0.06929   -3.454    0.0059
5                 -0.0047     0.03529   -0.133    1.0000
6                  0.4646     0.04315   10.767    0.0000
7                 -0.4101     0.10310   -3.978    0.0008

EPA_Rain_Zone = 3  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
5                  0.2346     0.07199    3.259    0.0117
6                  0.7040     0.07615    9.244    0.0000
7                 -0.1708     0.12069   -1.415    1.0000


EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
6                  0.4693     0.04736    9.910    0.0000
7                 -0.4055     0.10493   -3.864    0.0012


EPA_Rain_Zone = 6  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
7                 -0.8748      0.1078   -8.113    0.0000
```

Dissolved phosphorus concentrations in EPA region 2 are significantly different than in regions 3, 6 and 7. No significant differences were observed between EPA regions 2 and 5.

Dissolved phosphorus concentrations in EPA region 3 are significantly different than in regions 5 and 6. No significant differences were observed between EPA regions 3 and 7.

Dissolved phosphorus concentrations in EPA region 5 are significantly different than in regions 6 and 7.

Dissolved phosphorus concentrations in EPA region 6 are significantly different than in region 7.

```
Tukey Simultaneous Tests
Response Variable LDP
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
3                 -0.2393     0.06929   -3.454    0.0050
5                 -0.0047     0.03529   -0.133    0.9999
6                  0.4646     0.04315   10.767    0.0000
7                 -0.4101     0.10310   -3.978    0.0007


EPA_Rain_Zone = 3  subtracted from:

               Difference       SE of           Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
```

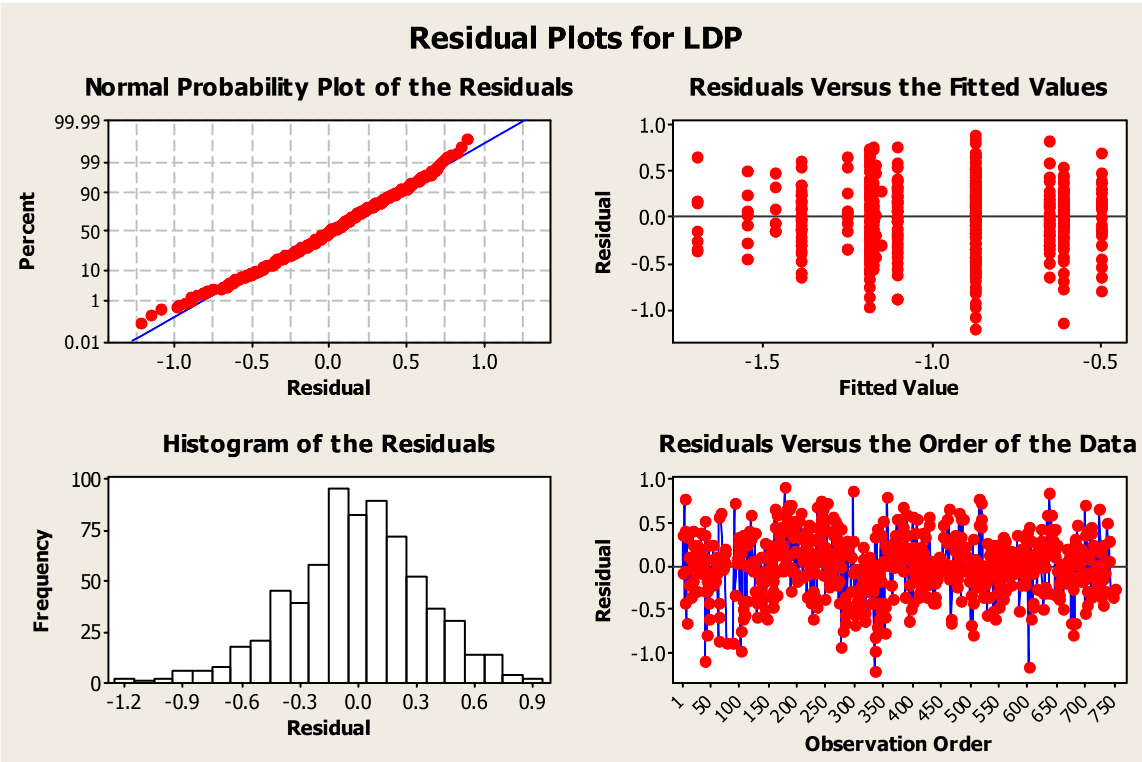
```
5                  0.2346     0.07199    3.259    0.0099
6                  0.7040     0.07615    9.244    0.0000
7                 -0.1708     0.12069   -1.415    0.6176


EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value    P-Value
6                  0.4693     0.04736    9.910     0.0000
7                 -0.4055     0.10493   -3.864     0.0011


EPA_Rain_Zone = 6  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value    P-Value
7                 -0.8748      0.1078   -8.113     0.0000
```

Dissolved phosphorus concentrations in EPA region 2 are significantly different than in regions 3, 6 and 7. No significant differences were observed between EPA regions 2 and 5.

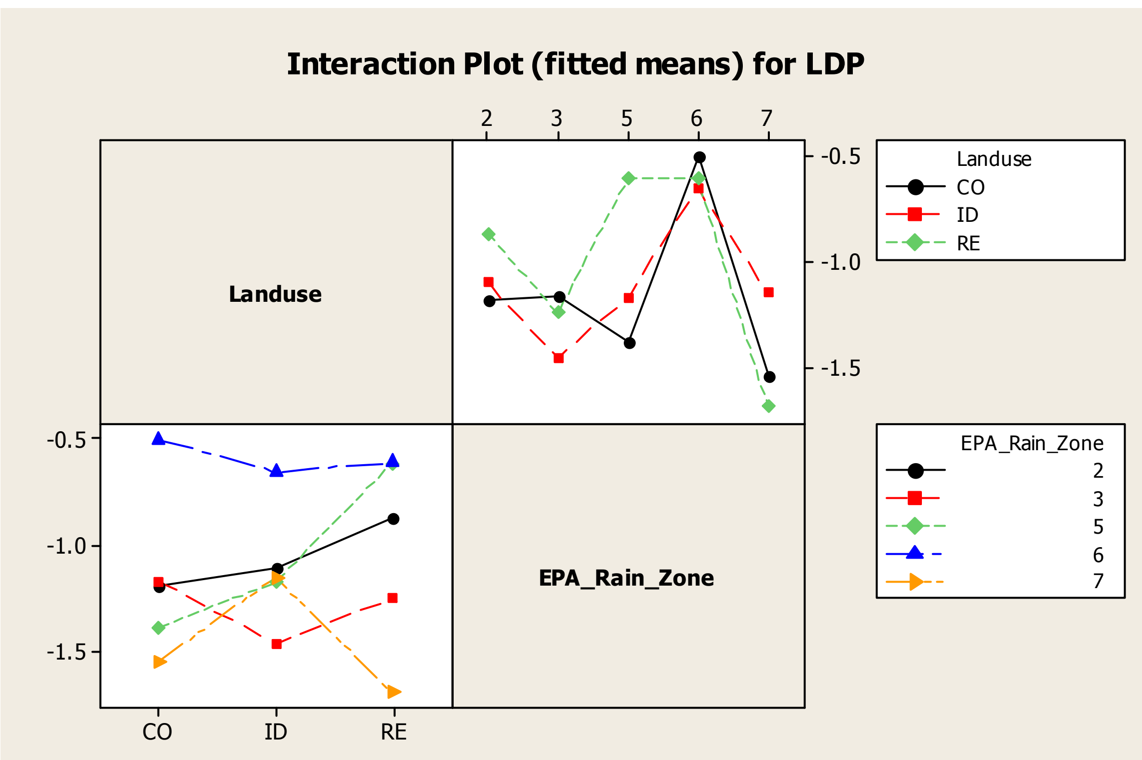
Dissolved phosphorus concentrations in EPA region 3 are significantly different than in regions 5 and 6. No significant differences were observed between EPA regions 3 and 7.

Dissolved phosphorus concentrations in EPA region 5 are significantly different than in regions 6 and 7.

Dissolved phosphorus concentrations in EPA region 6 are significantly different than in region 7.



The assumption of normality of residuals is valid. No specific trend was observed in the residuals.



Total Copper (µg/L) (Cu)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	66	57	1.203	0.254	0.230	1.688
3	6	0	0.944	0.212	0.699	1.230
4	15	1	2.010	0.454	0.699	2.584
5	38	4	0.946	0.359	0.301	1.914
6	34	0	1.108	0.374	0.176	1.799
7	37	4	1.265	0.349	0.477	2.114

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	54	55	1.152	0.359	0.342	1.940
3	15	1	1.099	0.476	0.301	1.771
4	15	2	2.064	0.597	1.000	3.134
5	44	3	1.263	0.234	0.602	1.940
6	66	4	1.778	0.456	0.301	2.532
7	25	8	1.562	0.309	1.041	2.079

Land use: Residential

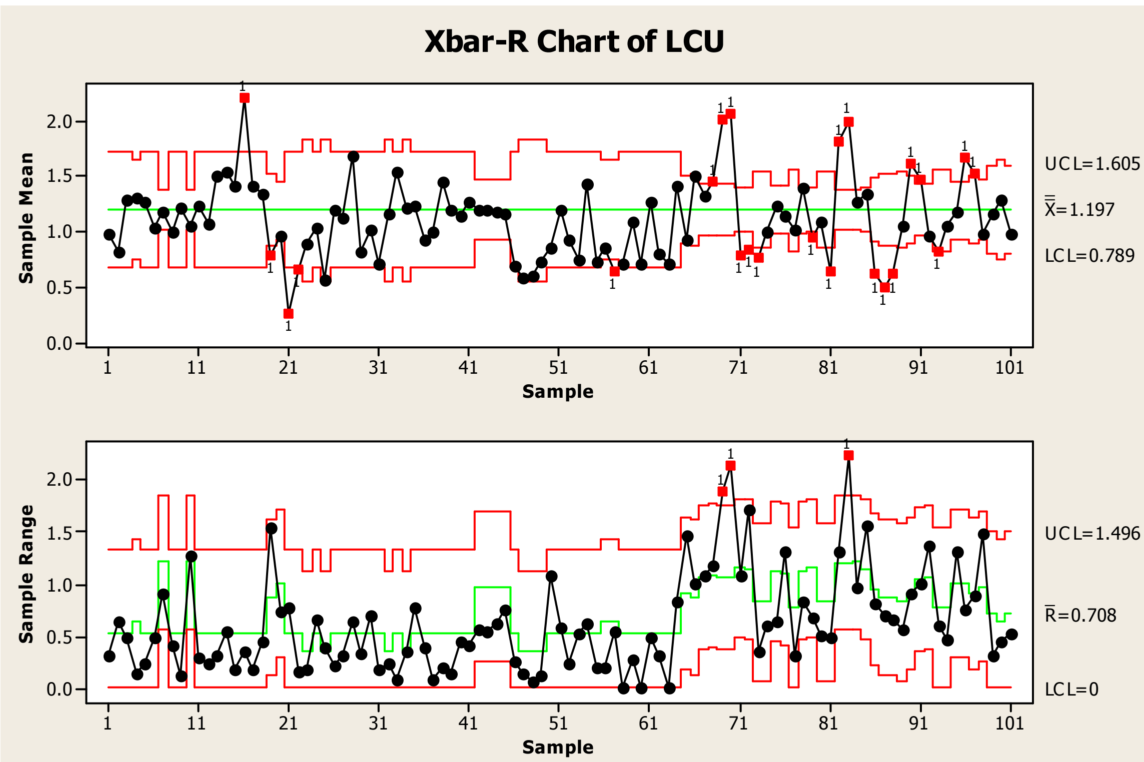
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	140	191	1.018	0.379	-0.149	2.380
3	17	1	1.079	0.463	0.699	2.009
4	31	0	1.386	0.345	0.845	2.013
5	68	3	0.883	0.272	0.095	1.799
6	38	0	0.972	0.496	0.146	2.255
7	37	3	0.948	0.312	0.211	1.909

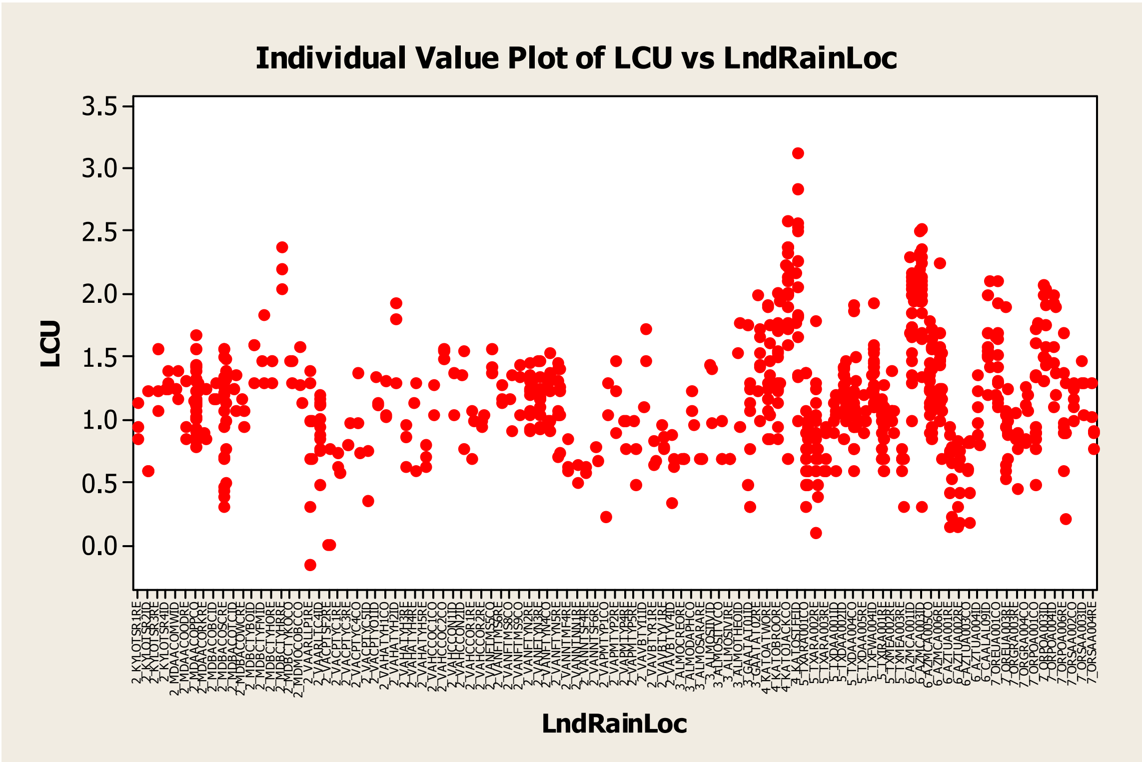
Descriptive Statistics: LCU

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LCU	2_KYLOTSR1RE	3	0	0.9813	0.0234	0.8437	1.1461
	2_KYLOTSR2ID	3	0	0.812	0.132	0.602	1.230
	2_KYLOTSR3RE	3	0	1.293	0.0627	1.079	1.568
	2_KYLOTSR4ID	4	0	1.3024	0.00452	1.2553	1.3979
	2_MDAACOMWID	3	0	1.2764	0.0126	1.1761	1.3979
	2_MDAACOODRE	3	0	1.041	0.0625	0.845	1.322
	2_MDAACOPPCO	26	0	1.1747	0.0491	0.7839	1.6876
	2_MDAACORKRE	3	0	1.001	0.0493	0.845	1.255
	2_MDBACOBCID	3	0	1.2177	0.00520	1.1761	1.3010
	2_MDBACOSCRE	26	0	1.0590	0.1588	0.3010	1.5641
	2_MDBACOTCID	3	0	1.2321	0.0204	1.0792	1.3617
	2_MDBACOWCRE	3	0	1.0698	0.0124	0.9542	1.1761
	2_MDBCTYBOID	3	0	1.502	0.0302	1.301	1.602
	2_MDBCTYFMID	3	0	1.541	0.0771	1.301	1.845
	2_MDBCTYHORE	3	0	1.4184	0.0103	1.3010	1.4771
	2_MDBCTYHRRE	3	0	2.2086	0.0287	2.0414	2.3802
	2_MDBCTYKOCO	3	0	1.4184	0.0103	1.3010	1.4771
	2_MDMOCOBCCO	3	0	1.335	0.0494	1.146	1.580
	2_VAARLLP1RE	8	0	0.779	0.265	-0.149	1.398
	2_VAARLTC4ID	12	1	0.9564	0.0466	0.4800	1.2041
	2_VACPTC1ARE	0	8	*	*	*	*

2_VACPTSF2RE	3	0	0.259	0.202	0.000000000	0.778
2_VACPTYC1RE	3	4	0.6512	0.00668	0.5798	0.7404
2_VACPTYC3RE	2	12	0.8885	0.0159	0.7993	0.9777
2_VACPTYC4CO	3	11	1.030	0.107	0.732	1.380
2_VACPTYC5ID	2	12	0.555	0.0747	0.362	0.748
2_VACPTYO1ID	3	0	1.2041	0.0145	1.1239	1.3424
2_VAHATYH1CO	3	15	1.1283	0.0283	1.0212	1.3222
2_VAHATYH2ID	3	16	1.684	0.114	1.301	1.940
2_VAHATYH3RE	3	14	0.8237	0.0296	0.6335	0.9685
2_VAHATYH4RE	3	14	1.016	0.135	0.602	1.301
2_VAHATYH5RE	3	14	0.7101	0.00776	0.6232	0.7993
2_VAHCCOC1CO	2	0	1.160	0.0282	1.041	1.279
2_VAHCCOC2CO	3	0	1.5386	0.00171	1.4914	1.5682
2_VAHCCON1ID	2	0	1.211	0.0574	1.041	1.380
2_VAHCCON2ID	3	0	1.232	0.164	0.778	1.556
2_VAHCCOR1RE	3	0	0.926	0.0402	0.699	1.079
2_VAHCCOR2RE	3	0	0.9985	0.00190	0.9542	1.0414
2_VANFTMS5CO	3	0	1.4599	0.00945	1.3802	1.5682
2_VANFTMS6RE	3	0	1.2003	0.00484	1.1461	1.2788
2_VANFTMS8CO	3	0	1.151	0.0506	0.914	1.362
2_VANFTMS9CO	3	0	1.277	0.0444	1.041	1.447
2_VANFTYN2RE	11	18	1.1917	0.0327	0.9085	1.4624
2_VANFTYN3RE	11	16	1.1920	0.0347	0.9365	1.4771
2_VANFTYN4CO	11	16	1.1798	0.0459	0.9085	1.5315
2_VANFTYN5RE	11	16	1.1569	0.0679	0.7099	1.4624
2_VANNTMF1RE	1	0	1.1461	*	1.1461	1.1461
2_VANNTMF4RE	3	0	0.6942	0.0201	0.6021	0.8573
2_VANNTNN1RE	2	8	0.5743	0.00956	0.5051	0.6435
2_VANNTSF4RE	2	0	0.6066	0.00144	0.5798	0.6335
2_VANNTSF6RE	2	0	0.7287	0.00641	0.6721	0.7853
2_VAPMTYP1CO	3	15	0.858	0.312	0.230	1.301
2_VAPMTYP2RE	3	14	1.204	0.0829	0.903	1.477
2_VAPMTYP4RE	3	14	0.9261	0.0164	0.7782	1.0000
2_VAPMTYP5RE	3	14	0.752	0.0689	0.477	1.000
2_VAVBTYI1ID	3	0	1.438	0.0942	1.114	1.724
2_VAVBTYR1RE	3	2	0.7181	0.0111	0.6435	0.8388
2_VAVBTYV1RE	4	23	0.8506	0.00764	0.7782	0.9638
2_VAVBTYV4ID	4	26	0.639	0.0500	0.342	0.881
3_ALMOCREORE	3	0	0.69897	0.000000000	0.69897	0.69897
3_ALMODAPHCO	3	0	1.0893	0.0186	0.9584	1.2304
3_ALMOSARARE	3	0	0.69897	0.000000000	0.69897	0.69897
3_ALMOSIIVID	3	0	1.279	0.0699	0.974	1.447
3_ALMOSITVCO	3	0	0.799	0.0302	0.699	1.000
3_ALMOSIVIRE	3	0	0.69897	0.000000000	0.69897	0.69897
3_ALMOTHEOID	3	0	1.419	0.184	0.942	1.771
3_GAATAT01ID	9	1	0.933	0.252	0.301	1.756
3_GAATAT02RE	8	1	1.507	0.0961	1.000	2.009
4_KATOATWORE	15	0	1.3209	0.1236	0.8451	1.9243
4_KATOBROORE	16	0	1.4468	0.1146	0.8451	2.0128
4_KATOJACKCO	15	1	2.010	0.206	0.699	2.584
4_KATOSTFEID	15	2	2.064	0.356	1.000	3.134
5_TXARA001CO	21	1	0.7802	0.0748	0.3010	1.3802
5_TXARA002RE	20	1	0.8408	0.1362	0.0951	1.7993
5_TXARA003RE	7	0	0.7522	0.0261	0.6021	0.9542
5_TXDAA001ID	7	0	0.9966	0.0398	0.6021	1.2041
5_TXDAA002ID	18	1	1.2273	0.0299	0.8451	1.4771
5_TXDAA004CO	17	3	1.1499	0.1237	0.6021	1.9138
5_TXDAA005RE	6	1	1.0149	0.0131	0.9031	1.2041
5_TXFWA004ID	19	2	1.3942	0.0419	1.1139	1.9395
5_TXIRA001RE	21	1	0.9402	0.0416	0.6021	1.2788
5_TXMEA002RE	7	0	1.0883	0.0231	0.9031	1.3979
5_TXMEA003RE	7	0	0.6396	0.0249	0.3010	0.7782
6_AZMCA001ID	25	2	1.8095	0.1090	1.0000	2.3010

6_AZMCA003ID	25	2	2.0065	0.1989	0.3010	2.5315
6_AZMCA005CO	26	0	1.2616	0.0659	0.8451	1.7993
6_AZMCA006RE	20	0	1.3454	0.1049	0.6990	2.2553
6_AZTUA001RE	10	0	0.6083	0.0739	0.1461	0.9545
6_AZTUA002RE	8	0	0.4918	0.0716	0.1461	0.8325
6_AZTUA003CO	8	0	0.6083	0.0512	0.1761	0.8325
6_AZTUA004ID	7	0	1.0509	0.0388	0.7993	1.3617
6_CAALAL09ID	9	0	1.6227	0.0779	1.2041	2.1072
7_OREUA001CO	14	2	1.4691	0.0940	1.1139	2.1139
7_OREUA003RE	15	0	0.9659	0.1075	0.5374	1.9085
7_ORGRA003RE	6	0	0.8230	0.0420	0.4472	1.0531
7_ORGRA004CO	6	0	1.0563	0.0353	0.8129	1.2695
7_ORPOA001CO	12	1	1.171	0.175	0.477	1.785
7_ORPOA003ID	12	2	1.6786	0.0705	1.3222	2.0792
7_ORPOA004ID	9	4	1.532	0.117	1.114	2.000
7_ORPOA006RE	11	2	0.975	0.153	0.211	1.690
7_ORSAA002CO	5	1	1.1697	0.0140	1.0019	1.3010
7_ORSAA003ID	4	2	1.2801	0.0322	1.0414	1.4771
7_ORSAA004RE	5	1	0.9829	0.0390	0.7782	1.3010





General Linear Model: LCU versus Landuse, EPA_Rain_Zone

```
Factor         Type   Levels  Values
Landuse        fixed       3  CO, ID, RE
EPA_Rain_Zone  fixed       6  2, 3, 4, 5, 6, 7


Analysis of Variance for LCU, using Adjusted SS for Tests

Source                  DF    Seq SS   Adj SS   Adj MS      F      P
Landuse                  2   27.1721  17.5268   8.7634  64.52  0.000
EPA_Rain_Zone            5   26.3863  28.2226   5.6445  41.56  0.000
Landuse*EPA_Rain_Zone   10   13.7000  13.7000   1.3700  10.09  0.000
Error                  727   98.7367  98.7367   0.1358
Total                  744  165.9951
```

Main effects and interaction are significant.

```
Bonferroni Simultaneous Tests
Response Variable LCU
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of           Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID           0.2403     0.04600    5.224    0.0000
RE          -0.1986     0.04333   -4.583    0.0000


Landuse = ID  subtracted from:

         Difference       SE of           Adjusted
```

Landuse	of Means	Difference	T-Value	P-Value
RE	-0.4389	0.03865	-11.35	0.0000

Tukey Simultaneous Tests
Response Variable LCU
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

	Difference	SE of		Adjusted
Landuse	of Means	Difference	T-Value	P-Value
ID	0.2403	0.04600	5.224	0.0000
RE	-0.1986	0.04333	-4.583	0.0000

Landuse = ID subtracted from:

	Difference	SE of		Adjusted
Landuse	of Means	Difference	T-Value	P-Value
RE	-0.4389	0.03865	-11.35	0.0000

Significant differences by land use. Copper concentrations are different among the three land uses.

Bonferroni Simultaneous Tests
Response Variable LCU
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
3	-0.08326	0.07089	-1.175	1.0000
4	0.69562	0.05582	12.463	0.0000
5	-0.09387	0.03973	-2.362	0.2763
6	0.16181	0.04106	3.940	0.0013
7	0.13418	0.04512	2.974	0.0456

EPA_Rain_Zone = 3 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
4	0.77889	0.08311	9.3715	0.0000
5	-0.01060	0.07329	-0.1446	1.0000
6	0.24508	0.07402	3.3111	0.0146
7	0.21745	0.07634	2.8484	0.0678

EPA_Rain_Zone = 4 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
5	-0.7895	0.05883	-13.42	0.0000
6	-0.5338	0.05974	-8.94	0.0000
7	-0.5614	0.06260	-8.97	0.0000

EPA_Rain_Zone = 5 subtracted from:

	Difference	SE of		Adjusted
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
6	0.2557	0.04507	5.672	0.0000
7	0.2280	0.04880	4.673	0.0001

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.02763	0.04989	-0.5538	1.000

Copper in EPA rain zone 2 is significantly different than in EPA rain zones 4, 6, and 7. No differences were observed between rain zone 2 and rain zones 3 and 5.

Copper in EPA rain zone 3 is significantly different than in EPA rain zones 4, 6, and 7. No differences were observed between rain zone 3 and rain zone 5.

Copper in EPA rain zone 4 is significantly different than in EPA rain zones 5, 6, and 7.

Copper in EPA rain zone 5 is significantly different than in EPA rain zones 6, and 7. No differences were observed between rain zones 6 and 7.

Tukey Simultaneous Tests
Response Variable LCU
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.08326	0.07089	-1.175	0.8491
4	0.69562	0.05582	12.463	0.0000
5	-0.09387	0.03973	-2.362	0.1696
6	0.16181	0.04106	3.940	0.0011
7	0.13418	0.04512	2.974	0.0349

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.77889	0.08311	9.3715	0.0000
5	-0.01060	0.07329	-0.1446	1.0000
6	0.24508	0.07402	3.3111	0.0120
7	0.21745	0.07634	2.8484	0.0502

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.7895	0.05883	-13.42	0.0000
6	-0.5338	0.05974	-8.94	0.0000
7	-0.5614	0.06260	-8.97	0.0000

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
6	0.2557	0.04507	5.672	0.0000
7	0.2280	0.04880	4.673	0.0000

EPA_Rain_Zone = 6 subtracted from:

Difference	SE of	Adjusted

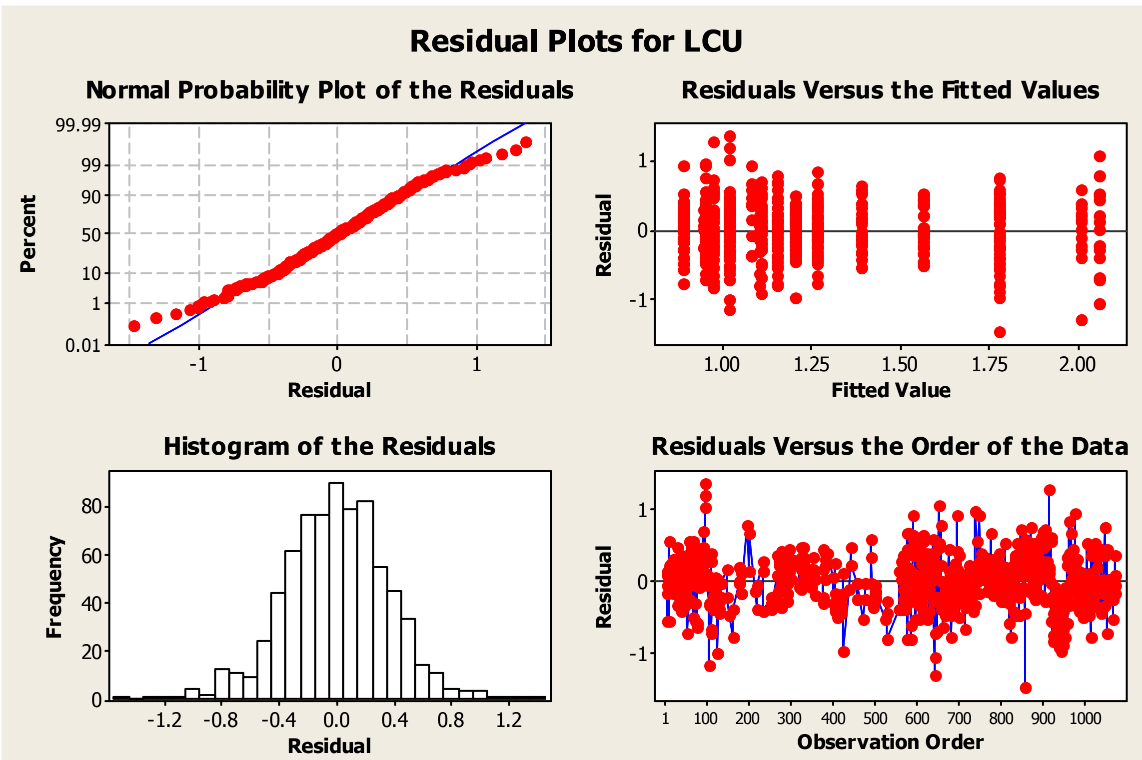
EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
7	-0.02763	0.04989	-0.5538	0.9938

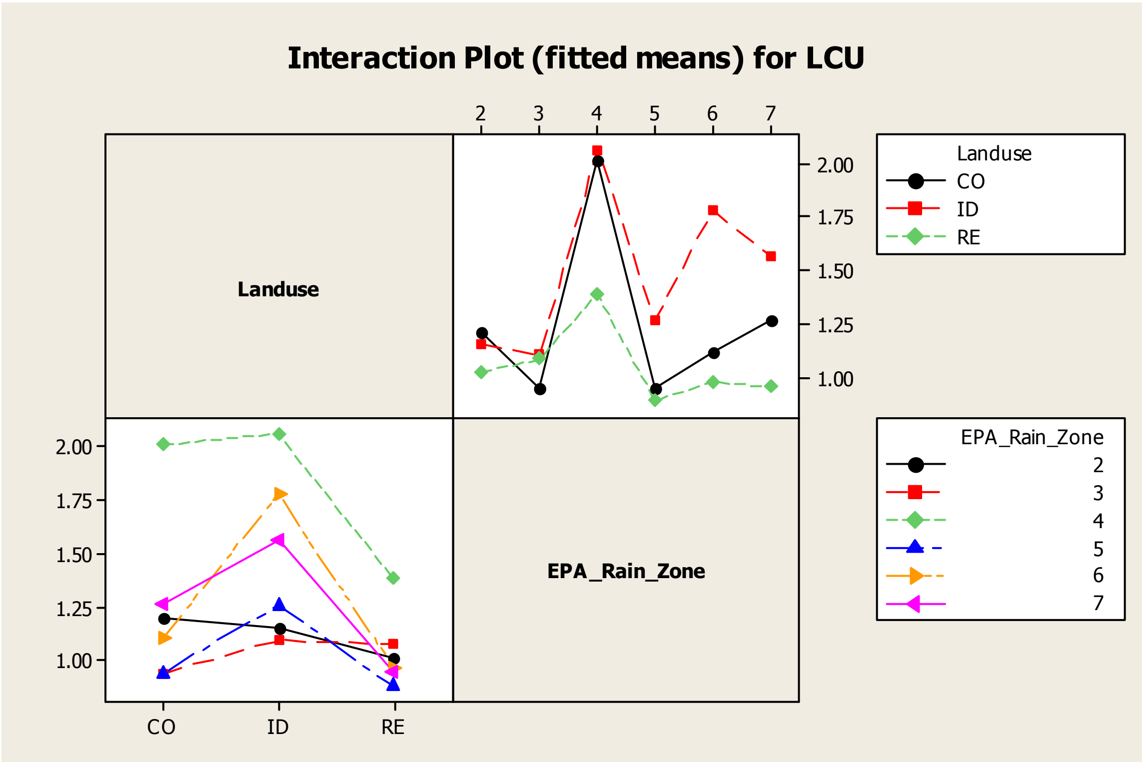
Copper in EPA rain zone 2 is significantly different than in EPA rain zones 4, 6, and 7. No differences were observed between rain zone 2 and rain zones 3 and 5.

Copper in EPA rain zone 3 is significantly different than in EPA rain zones 4, 6, and 7. No differences were observed between rain zone 3 and rain zone 5.

Copper in EPA rain zone 4 is significantly different than in EPA rain zones 5, 6, and 7.

Copper in EPA rain zone 5 is significantly different than in EPA rain zones 6, and 7. No differences were observed between rain zones 6 and 7.




Interaction Plot (fitted means) for LCU
Landuse
EPA_Rain_Zone
2
3
4
5
6
7
2.00
1.75
1.50
1.25
1.00
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7

Total Lead (µg/L) (Pb)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	66	57	1.011	0.601	-0.670	2.140
3	6	0	1.016	0.185	0.699	1.204
4	15	1	1.779	0.598	0.000	2.340
5	39	3	1.245	0.424	0.196	2.477
6	34	0	1.079	0.311	0.477	1.887
7	37	4	1.388	0.434	0.477	2.462

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	48	61	0.939	0.647	-0.254	2.114
3	10	6	0.989	0.538	0.000	1.602
4	15	2	1.917	0.705	0.301	3.079
5	45	2	1.370	0.397	0.347	2.431
6	69	1	1.892	0.529	0.301	2.792
7	25	8	1.518	0.389	0.602	2.230

Land use: Residential

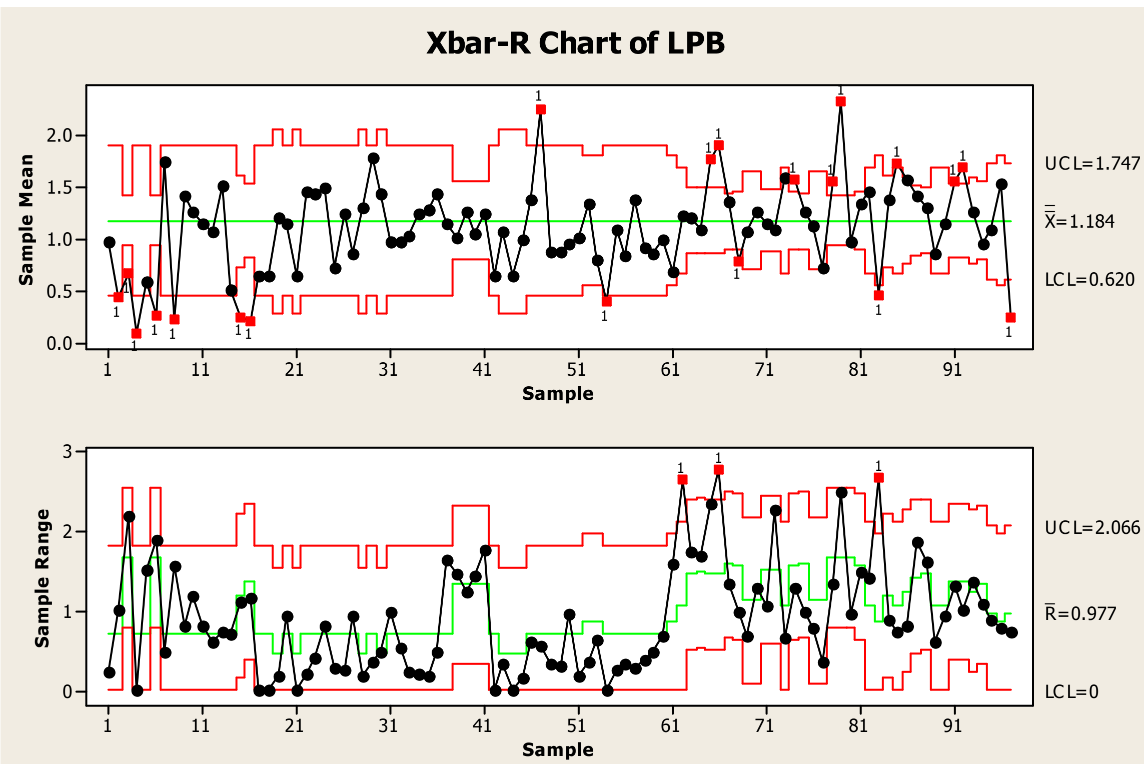
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	135	196	0.836	0.596	-0.341	2.566
3	15	3	0.915	0.639	0.000	2.653
4	31	0	1.155	0.476	0.000	2.340
5	68	3	1.073	0.357	0.224	1.949
6	30	8	1.254	0.591	-0.874	2.279
7	37	3	0.998	0.497	-0.222	2.322

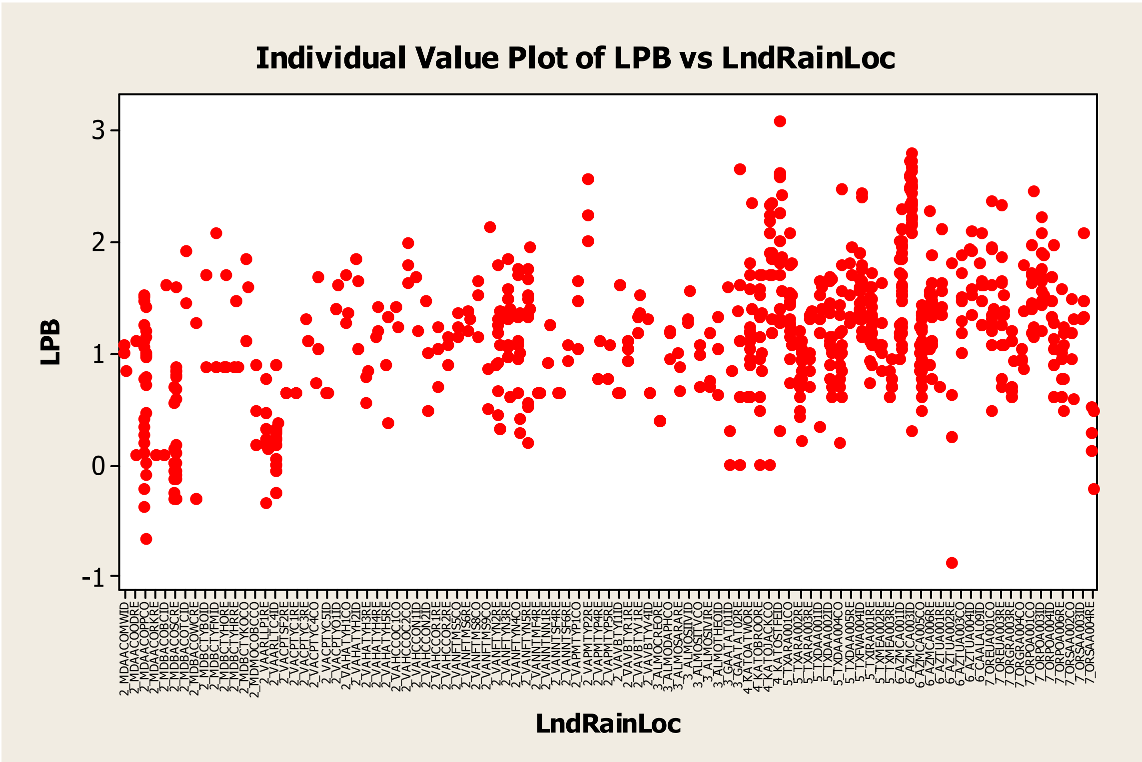
Descriptive Statistics: LPB

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LPB	2_KYLOTSR1RE	1	2	1.7782	*	1.7782	1.7782
	2_KYLOTSR2ID	0	3	*	*	*	*
	2_KYLOTSR3RE	0	3	*	*	*	*
	2_KYLOTSR4ID	1	3	2.1139	*	2.1139	2.1139
	2_MDAACOMWID	3	0	0.9748	0.0142	0.8451	1.0792
	2_MDAACOODRE	3	0	0.436	0.345	0.0969	1.114
	2_MDAACOPPCO	26	0	0.670	0.391	-0.670	1.531
	2_MDAACORKRE	3	0	0.096910	0.000000000	0.096910	0.096910
	2_MDBACOBCID	3	0	0.602	0.766	0.0969	1.613
	2_MDBACOSCRE	26	0	0.2642	0.2540	-0.3010	1.5911
	2_MDBACOTCID	3	0	1.762	0.0743	1.447	1.924
	2_MDBACOWCRE	3	0	0.226	0.832	-0.301	1.279
	2_MDBCTYBOID	3	0	1.424	0.226	0.875	1.699
	2_MDBCTYFMID	3	0	1.276	0.483	0.875	2.079
	2_MDBCTYHORE	3	0	1.150	0.226	0.875	1.699

2_MDBCTYHRRE	3	0	1.076	0.121	0.875	1.477
2_MDBCTYKOCO	3	0	1.522	0.137	1.117	1.845
2_MDMOCOBCCO	3	0	0.520	0.133	0.178	0.903
2_VAARLLP1RE	8	0	0.250	0.0997	-0.341	0.778
2_VAARLTC4ID	12	1	0.2128	0.0813	-0.2539	0.9031
2_VACPTC1ARE	0	8	*	*	*	*
2_VACPTSF2RE	3	0	0.65321	0.000000000	0.65321	0.65321
2_VACPTYC1RE	3	4	0.65321	0.000000000	0.65321	0.65321
2_VACPTYC3RE	2	12	1.2075	0.0175	1.1139	1.3010
2_VACPTYC4CO	3	11	1.154	0.231	0.740	1.681
2_VACPTYC5ID	2	12	0.65321	0.000000000	0.65321	0.65321
2_VACPTYO1ID	3	0	1.4696	0.0154	1.3979	1.6128
2_VAHATYH1CO	3	15	1.446	0.0495	1.279	1.699
2_VAHATYH2ID	3	16	1.513	0.176	1.041	1.845
2_VAHATYH3RE	3	14	0.7313	0.0237	0.5563	0.8451
2_VAHATYH4RE	3	14	1.2540	0.0204	1.1430	1.4150
2_VAHATYH5RE	3	14	0.862	0.227	0.371	1.322
2_VAHCCOC1CO	2	0	1.3227	0.0170	1.2304	1.4150
2_VAHCCOC2CO	3	0	1.801	0.0331	1.623	1.987
2_VAHCCON1ID	2	0	1.443	0.114	1.204	1.681
2_VAHCCON2ID	3	0	0.985	0.250	0.477	1.477
2_VAHCCOR1RE	3	0	0.990	0.0726	0.699	1.230
2_VAHCCOR2RE	3	0	1.0428	0.0158	0.9031	1.1461
2_VANFTMS5CO	3	0	1.2461	0.0118	1.1461	1.3617
2_VANFTMS6RE	3	0	1.2951	0.00778	1.2041	1.3802
2_VANFTMS8CO	3	0	1.440	0.0681	1.146	1.643
2_VANFTMS9CO	3	0	1.167	0.740	0.505	2.140
2_VANFTYN2RE	11	18	1.022	0.185	0.321	1.792
2_VANFTYN3RE	11	16	1.275	0.111	0.606	1.839
2_VANFTYN4CO	11	16	1.058	0.223	0.296	1.748
2_VANFTYN5RE	11	16	1.250	0.322	0.193	1.959
2_VANNTMF1RE	1	0	1.3802	*	1.3802	1.3802
2_VANNTMF4RE	3	0	0.65321	0.000000000	0.65321	0.65321
2_VANNTNN1RE	2	8	1.088	0.0562	0.920	1.255
2_VANNTSF4RE	2	0	0.65321	0.000000000	0.65321	0.65321
2_VANNTSF6RE	2	0	1.0051	0.0110	0.9310	1.0792
2_VAPMTYP1CO	3	15	1.382	0.0954	1.041	1.643
2_VAPMTYP2RE	3	14	2.271	0.0775	2.013	2.566
2_VAPMTYP4RE	3	14	0.890	0.0376	0.778	1.114
2_VAPMTYP5RE	3	14	0.878	0.0302	0.778	1.079
2_VAVBTYI1ID	3	0	0.973	0.307	0.653	1.613
2_VAVBTYR1RE	3	2	1.0307	0.00793	0.9368	1.1139
2_VAVBTYV1RE	4	23	1.3479	0.0214	1.1761	1.5315
2_VAVBTYV4ID	4	26	0.815	0.105	0.653	1.301
3_ALMOCREORE	3	0	0.39794	0.000000000	0.39794	0.39794
3_ALMODAPHCO	3	0	1.1099	0.0195	0.9494	1.2041
3_ALMOSARARE	3	0	0.8480	0.0279	0.6691	1.0000
3_ALMOSIIVID	3	0	1.3827	0.0259	1.2788	1.5682
3_ALMOSITVCO	3	0	0.922	0.0393	0.699	1.079
3_ALMOSIVIRE	3	0	0.877	0.0679	0.699	1.176
3_ALMOTHEOID	3	0	0.998	0.121	0.631	1.322
3_GAATAT01ID	4	6	0.687	0.494	0.000000000	1.602
3_GAATAT02RE	6	3	1.227	0.823	0.000000000	2.653
4_KATOATWORE	15	0	1.218	0.241	0.602	2.340
4_KATOBROORE	16	0	1.096	0.221	0.000000000	1.699
4_KATOJACKCO	15	1	1.779	0.357	0.000000000	2.340
4_KATOSTFEID	15	2	1.917	0.497	0.301	3.079
5_TXARA001CO	21	1	1.3634	0.0945	0.7298	2.0792
5_TXARA002RE	20	1	0.7964	0.0607	0.2243	1.2041
5_TXARA003RE	7	0	1.075	0.0717	0.699	1.380
5_TXDAA001ID	7	0	1.282	0.197	0.347	1.643
5_TXDAA002ID	18	1	1.1651	0.1032	0.6021	1.6812
5_TXDAA004CO	18	2	1.107	0.253	0.196	2.477

5_TXDAA005RE	6	1	1.599	0.0738	1.279	1.949
5_TXFWA004ID	20	1	1.5859	0.1196	1.1461	2.4314
5_TXIRA001RE	21	1	1.2757	0.0611	0.7381	1.7243
5_TXMEA002RE	7	0	1.1367	0.0644	0.8451	1.6335
5_TXMEA003RE	7	0	0.7399	0.0166	0.6021	0.9542
6_AZMCA001ID	27	0	1.5732	0.1349	0.9542	2.3010
6_AZMCA003ID	27	0	2.3438	0.2064	0.3010	2.7924
6_AZMCA005CO	26	0	0.9853	0.0598	0.4771	1.4314
6_AZMCA006RE	20	0	1.3528	0.1251	0.7782	2.2788
6_AZTUA001RE	6	4	1.460	0.212	0.699	2.114
6_AZTUA002RE	4	4	0.452	1.219	-0.874	1.806
6_AZTUA003CO	8	0	1.384	0.105	1.000	1.886
6_AZTUA004ID	6	1	1.735	0.0846	1.342	2.090
6_CAALAL09ID	9	0	1.5948	0.0671	1.2553	2.0792
7_OREUA001CO	14	2	1.439	0.230	0.477	2.362
7_OREUA003RE	15	0	1.320	0.195	0.699	2.322
7_ORGRA003RE	6	0	0.872	0.0631	0.602	1.204
7_ORGRA004CO	6	0	1.158	0.130	0.863	1.792
7_ORPOA001CO	12	1	1.563	0.140	1.146	2.462
7_ORPOA003ID	12	2	1.6963	0.0841	1.2041	2.2304
7_ORPOA004ID	9	4	1.268	0.170	0.602	1.968
7_ORPOA006RE	11	2	0.9724	0.0952	0.4771	1.5798
7_ORSAA002CO	5	1	1.103	0.120	0.591	1.491
7_ORSAA003ID	4	2	1.545	0.133	1.301	2.079
7_ORSAA004RE	5	1	0.237	0.0914	-0.222	0.522





General Linear Model: LPB versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	6	2, 3, 4, 5, 6, 7

Analysis of Variance for LPB, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	31.6472	12.9341	6.4671	24.05	0.000
EPA_Rain_Zone	5	33.0307	33.4042	6.6808	24.85	0.000
Landuse*EPA_Rain_Zone	10	16.0919	16.0919	1.6092	5.99	0.000
Error	704	189.2705	189.2705	0.2689		
Total	721	270.0403				

Main factors and interaction are significant.

Bonferroni Simultaneous Tests
Response Variable LPB
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.1802	0.06667	2.703	0.0211
RE	-0.2163	0.06184	-3.498	0.0015

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.3965	0.05772	-6.869	0.0000

Tukey Simultaneous Tests
Response Variable LPB
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.1802	0.06667	2.703	0.0188
RE	-0.2163	0.06184	-3.498	0.0014

Landuse = ID subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
RE	-0.3965	0.05772	-6.869	0.0000

Copper concentrations are different among commercial, industrial and residential land uses.

Bonferroni Simultaneous Tests
Response Variable LPB
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	0.05664	0.10616	0.5336	1.0000
4	0.70022	0.07912	8.8505	0.0000
5	0.31270	0.05641	5.5435	0.0000
6	0.49147	0.06017	8.1684	0.0000
7	0.38434	0.06421	5.9861	0.0000

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.6436	0.1221	5.272	0.0000
5	0.2561	0.1087	2.355	0.2823
6	0.4348	0.1107	3.926	0.0014
7	0.3277	0.1130	2.900	0.0577

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.3875	0.08256	-4.694	0.0000
6	-0.2087	0.08517	-2.451	0.2173
7	-0.3159	0.08807	-3.587	0.0054

EPA_Rain_Zone = 5 subtracted from:

Difference	SE of	Adjusted

EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
6	0.17877	0.06462	2.766	0.0873
7	0.07164	0.06840	1.047	1.0000

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.1071	0.07153	-1.498	1.000

Copper concentration in EPA rain zone 2 is different than in zones 4, 5, 6, and 7. No differences were observed between rain zones 2 and 3.

Copper concentration in EPA rain zone 3 is different than in zones 4, 6, and 7. No differences were observed between rain zones 3 and 5.

Copper concentration in EPA rain zone 4 is different than in zones 5 and 7. No differences were observed between rain zones 4 and 6.

Copper concentration in EPA rain zone 5 is different than in zone 6. No differences were observed between rain zones 5 and 7. No differences were observed between zones 6 and 7.

Tukey Simultaneous Tests
Response Variable LPB
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	0.05664	0.10616	0.5336	0.9948
4	0.70022	0.07912	8.8505	0.0000
5	0.31270	0.05641	5.5435	0.0000
6	0.49147	0.06017	8.1684	0.0000
7	0.38434	0.06421	5.9861	0.0000

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.6436	0.1221	5.272	0.0000
5	0.2561	0.1087	2.355	0.1725
6	0.4348	0.1107	3.926	0.0012
7	0.3277	0.1130	2.900	0.0433

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.3875	0.08256	-4.694	0.0000
6	-0.2087	0.08517	-2.451	0.1390
7	-0.3159	0.08807	-3.587	0.0045

EPA_Rain_Zone = 5 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value

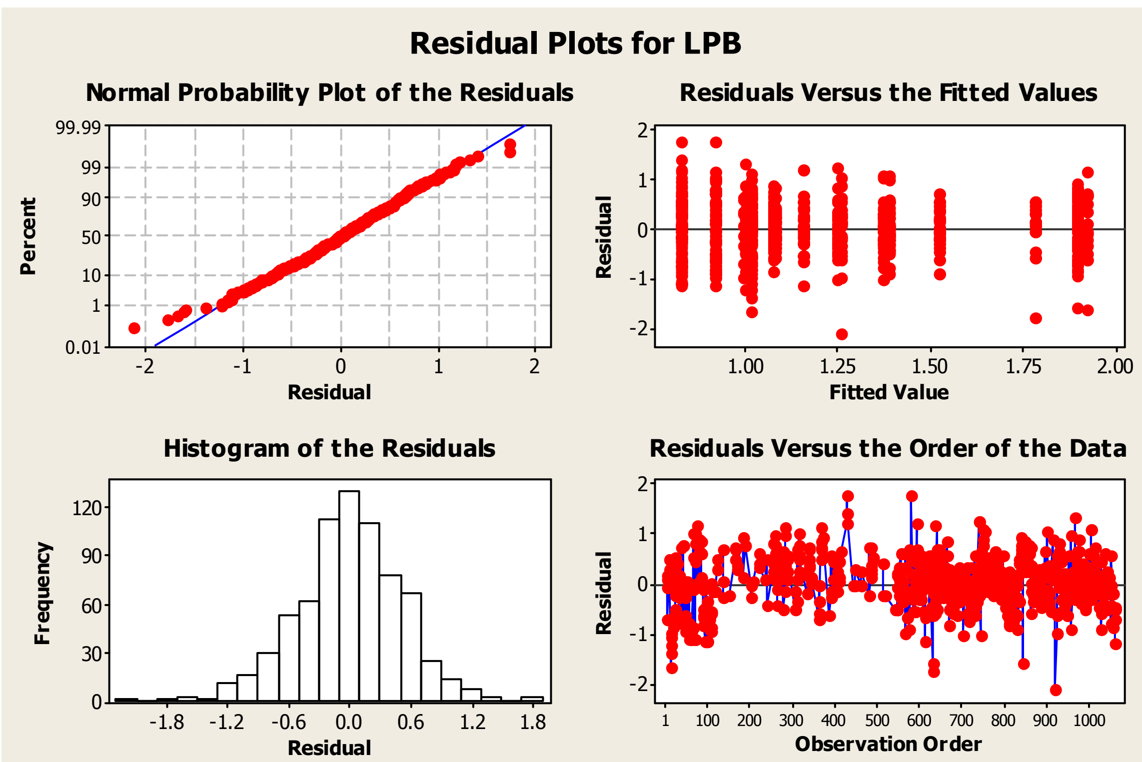
```
6                  0.17877     0.06462    2.766    0.0630
7                  0.07164     0.06840    1.047    0.9018


EPA_Rain_Zone = 6  subtracted from:

               Difference        SE of             Adjusted
EPA_Rain_Zone    of Means   Difference  T-Value     P-Value
7                 -0.1071      0.07153   -1.498      0.6659
```

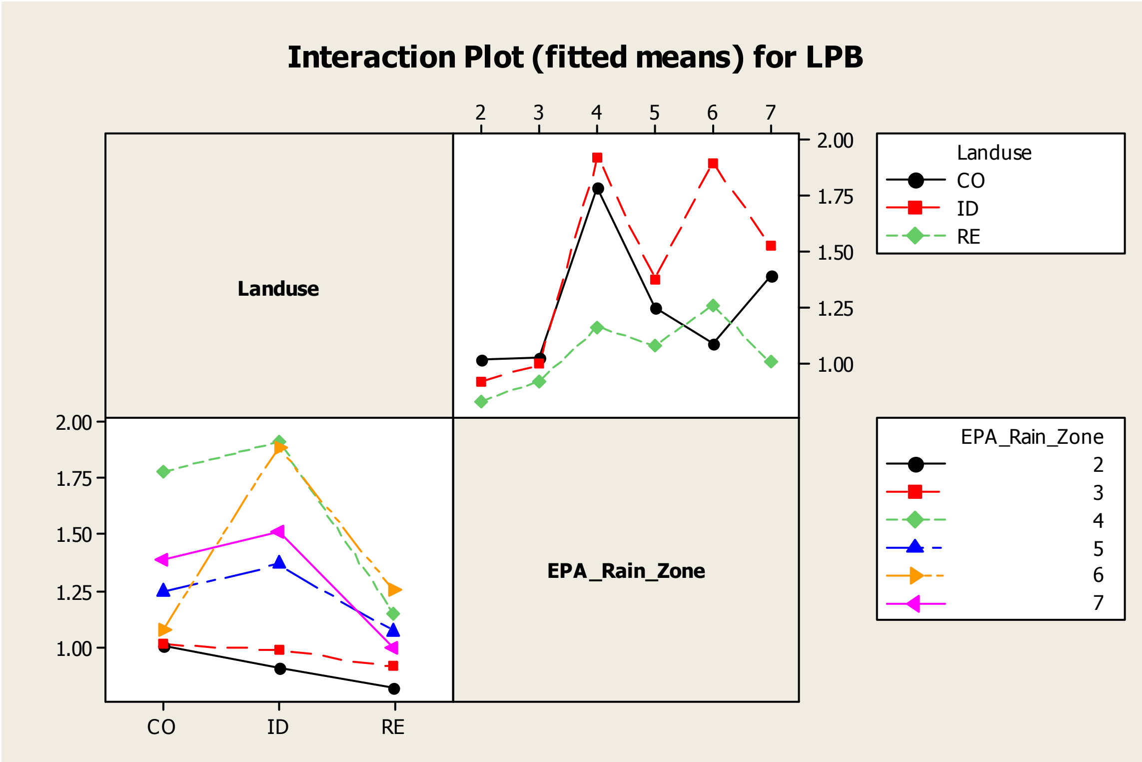
Copper concentration in EPA rain zone 2 is different than in zones 4, 5, 6, and 7. No differences were observed between rain zones 2 and 3.

Copper concentration in EPA rain zone 3 is different than in zones 4, 6, and 7. No differences were observed between rain zones 3 and 5.

Copper concentration in EPA rain zone 4 is different than in zones 5 and 7. No differences were observed between rain zones 4 and 6.

Copper concentration in EPA rain zone 5 is different than in zone 6. No differences were observed between rain zones 5 and 7. No difference was observed between zones 6 and 7




Interaction Plot (fitted means) for LPB
Landuse
EPA_Rain_Zone
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7
2.00
1.75
1.50
1.25
1.00
CO
ID
RE

Total Zinc (µg/L) (Zn)

Summary statistics in LOG base 10 scale
Land use: Commercial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	66	57	2.212	0.262	1.716	2.950
3	6	0	2.050	0.179	1.851	2.342
4	15	1	2.552	0.367	1.663	2.969
5	39	3	1.878	0.287	1.447	2.748
6	34	0	2.305	0.245	1.806	2.820
7	37	4	2.060	0.325	1.591	2.964

Land use: Industrial

EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	54	55	2.177	0.389	1.000	2.740
3	15	1	2.037	0.334	1.398	2.623
4	15	2	2.571	0.354	2.079	3.201
5	45	2	2.161	0.399	1.447	3.146
6	70	0	2.613	0.301	1.850	3.146
7	25	8	2.478	0.477	1.672	3.909

Land use: Residential

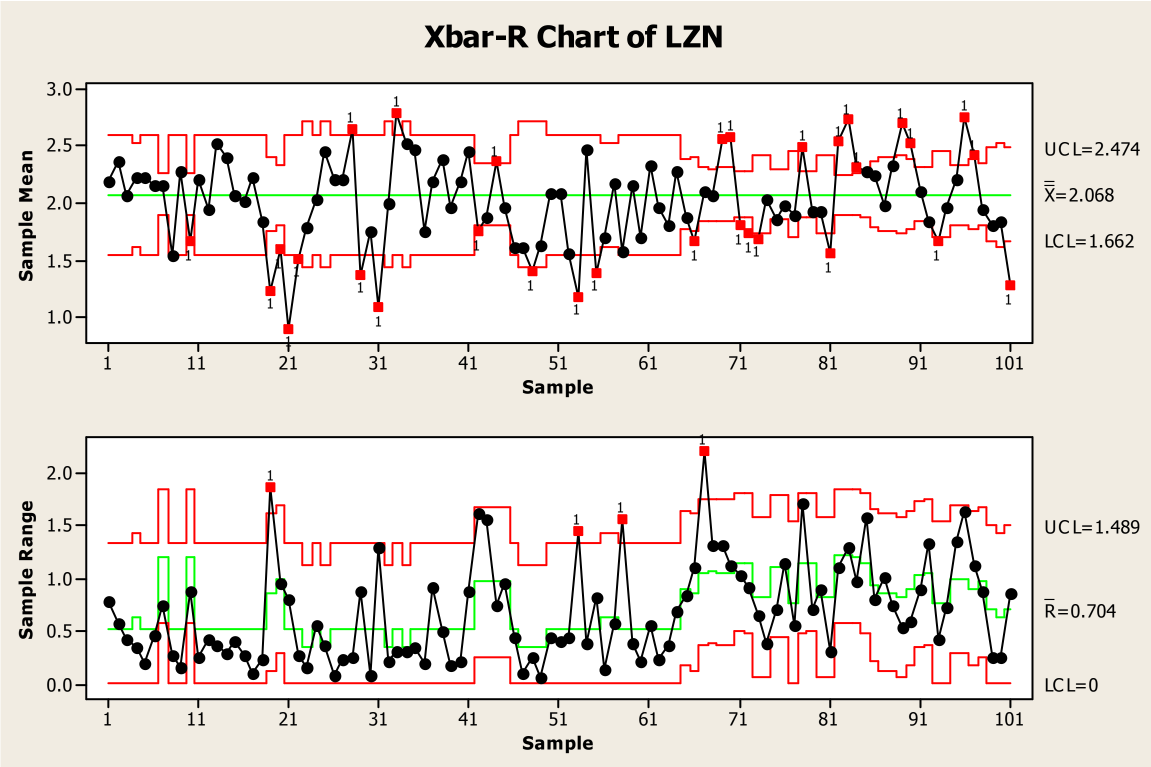
EPA_Rain_Zone	N det/est	N ND/NZ	Mean	StDev	Minimum	Maximum
2	140	191	1.702	0.426	0.047	2.726
3	17	1	1.673	0.411	0.912	2.462
4	31	0	2.073	0.441	1.000	3.199
5	68	3	1.796	0.241	1.204	2.362
6	38	0	2.195	0.342	1.544	3.176
7	37	3	1.760	0.352	0.840	2.813

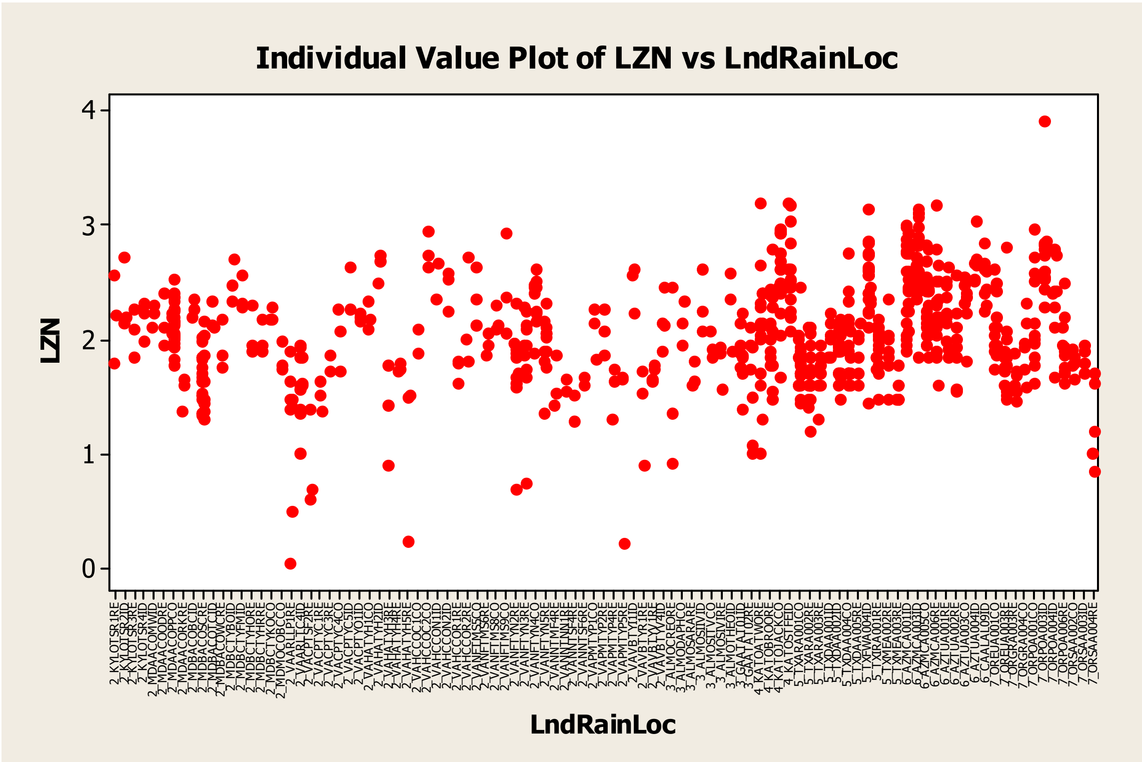
Descriptive Statistics: LZN

Variable	LndRainLoc	N	N*	Mean	Variance	Minimum	Maximum
LZN	2_KYLOTSR1RE	3	0	2.188	0.149	1.792	2.563
	2_KYLOTSR2ID	3	0	2.357	0.101	2.152	2.723
	2_KYLOTSR3RE	3	0	2.069	0.0444	1.851	2.272
	2_KYLOTSR4ID	4	0	2.2135	0.0235	1.9912	2.3243
	2_MDAACOMWID	3	0	2.2168	0.00901	2.1139	2.3010
	2_MDAACOODRE	3	0	2.155	0.0546	1.944	2.407
	2_MDAACOPPCO	26	0	2.1423	0.0330	1.7747	2.5221
	2_MDAACORKRE	3	0	1.5452	0.0211	1.3802	1.6532
	2_MDBACOBCID	3	0	2.2719	0.00602	2.1931	2.3483
	2_MDBACOSCRE	26	0	1.6532	0.0584	1.2967	2.1608
	2_MDBACOTCID	3	0	2.1928	0.0175	2.1004	2.3444
	2_MDBACOWCRE	3	0	1.931	0.0470	1.756	2.173
	2_MDBCTYBOID	3	0	2.506	0.0324	2.342	2.699
	2_MDBCTYFMID	3	0	2.3858	0.0223	2.2788	2.5563
	2_MDBCTYHORE	3	0	2.053	0.0469	1.903	2.301
	2_MDBCTYHRRE	3	0	2.0111	0.0211	1.9031	2.1761
	2_MDBCTYKOCO	3	0	2.2103	0.00351	2.1761	2.2788
	2_MDMOCOBCCO	3	0	1.8362	0.0162	1.7482	1.9823
	2_VAARLLP1RE	8	0	1.231	0.392	0.0465	1.903
	2_VAARLTC4ID	12	1	1.5903	0.0798	1.0000	1.9542

2_VACPTC1ARE	0	8	*	*	*	*
2_VACPTSF2RE	3	0	0.900	0.189	0.602	1.398
2_VACPTYC1RE	3	4	1.5096	0.0173	1.3802	1.6435
2_VACPTYC3RE	2	12	1.7897	0.0109	1.7160	1.8633
2_VACPTYC4CO	3	11	2.018	0.0775	1.716	2.265
2_VACPTYC5ID	2	12	2.449	0.0660	2.267	2.630
2_VACPTYO1ID	3	0	2.1956	0.00161	2.1584	2.2380
2_VAHATYH1CO	3	15	2.2005	0.0138	2.0969	2.3284
2_VAHATYH2ID	3	16	2.6408	0.0176	2.4900	2.7404
2_VAHATYH3RE	3	14	1.368	0.191	0.903	1.771
2_VAHATYH4RE	3	14	1.7492	0.00195	1.7160	1.7993
2_VAHATYH5RE	3	14	1.080	0.542	0.230	1.519
2_VAHCCOC1CO	2	0	1.988	0.0207	1.886	2.090
2_VAHCCOC2CO	3	0	2.7788	0.0251	2.6385	2.9504
2_VAHCCON1ID	2	0	2.507	0.0470	2.354	2.661
2_VAHCCON2ID	3	0	2.454	0.0339	2.243	2.584
2_VAHCCOR1RE	3	0	1.7427	0.0110	1.6232	1.8195
2_VAHCCOR2RE	3	0	2.182	0.230	1.820	2.726
2_VANFTMS5CO	3	0	2.371	0.0631	2.130	2.631
2_VANFTMS6RE	3	0	1.9543	0.00817	1.8692	2.0492
2_VANFTMS8CO	3	0	2.1748	0.0124	2.0899	2.3010
2_VANFTMS9CO	3	0	2.450	0.195	2.053	2.925
2_VANFTYN2RE	11	18	1.742	0.162	0.699	2.314
2_VANFTYN3RE	11	16	1.876	0.179	0.740	2.292
2_VANFTYN4CO	11	16	2.3598	0.0440	1.8808	2.6232
2_VANFTYN5RE	11	16	1.9504	0.0638	1.3617	2.3181
2_VANNTMF1RE	1	0	1.8325	*	1.8325	1.8325
2_VANNTMF4RE	3	0	1.609	0.0511	1.431	1.863
2_VANNTNN1RE	2	8	1.6048	0.00470	1.5563	1.6532
2_VANNTSF4RE	2	0	1.399	0.0287	1.279	1.519
2_VANNTSF6RE	2	0	1.6324	0.00184	1.6021	1.6628
2_VAPMTYP1CO	3	15	2.083	0.0503	1.833	2.265
2_VAPMTYP2RE	3	14	2.071	0.0392	1.869	2.265
2_VAPMTYP4RE	3	14	1.558	0.0525	1.301	1.740
2_VAPMTYP5RE	3	14	1.179	0.688	0.221	1.663
2_VAVBTYI1ID	3	0	2.466	0.0430	2.228	2.609
2_VAVBTYR1RE	3	2	1.384	0.182	0.903	1.716
2_VAVBTYV1RE	4	23	1.6995	0.00442	1.6335	1.7709
2_VAVBTYV4ID	4	26	2.159	0.0548	1.898	2.467
3_ALMOCREORE	3	0	1.579	0.636	0.912	2.462
3_ALMODAPHCO	3	0	2.149	0.0368	1.959	2.342
3_ALMOSARARE	3	0	1.6884	0.0133	1.6021	1.8195
3_ALMOSIIVID	3	0	2.319	0.0771	2.079	2.623
3_ALMOSITVCO	3	0	1.9498	0.0137	1.8513	2.0792
3_ALMOSIVIRE	3	0	1.795	0.0391	1.568	1.934
3_ALMOTHEOID	3	0	2.278	0.124	1.892	2.580
3_GAATAT01ID	9	1	1.8625	0.0591	1.3979	2.2304
3_GAATAT02RE	8	1	1.657	0.179	1.000	2.104
4_KATOATWORE	15	0	2.087	0.289	1.000	3.199
4_KATOBROORE	16	0	2.0592	0.1192	1.4771	2.7853
4_KATOJACKCO	15	1	2.5519	0.1345	1.6628	2.9685
4_KATOSTFEID	15	2	2.5707	0.1254	2.0792	3.2014
5_TXARA001CO	21	1	1.7920	0.0514	1.4472	2.4624
5_TXARA002RE	20	1	1.7226	0.0637	1.2041	2.1139
5_TXARA003RE	7	0	1.6831	0.0447	1.3010	1.9542
5_TXDAA001ID	7	0	2.0214	0.0204	1.8451	2.2304
5_TXDAA002ID	18	1	1.8553	0.0370	1.4771	2.1732
5_TXDAA004CO	18	2	1.9784	0.1037	1.6021	2.7482
5_TXDAA005RE	6	1	1.8889	0.0444	1.6021	2.1461
5_TXFWA004ID	20	1	2.4855	0.1225	1.4472	3.1461
5_TXIRA001RE	21	1	1.9164	0.0303	1.4771	2.1761
5_TXMEA002RE	7	0	1.915	0.0730	1.477	2.362
5_TXMEA003RE	7	0	1.5558	0.0131	1.4771	1.7782

6_AZMCA001ID	27	0	2.5259	0.1205	1.9031	3.0000
6_AZMCA003ID	27	0	2.7165	0.0785	1.8501	3.1461
6_AZMCA005CO	26	0	2.3006	0.0643	1.8451	2.8195
6_AZMCA006RE	20	0	2.2649	0.1261	1.6021	3.1761
6_AZTUA001RE	10	0	2.2332	0.0702	1.8451	2.6335
6_AZTUA002RE	8	0	1.974	0.113	1.544	2.556
6_AZTUA003CO	8	0	2.3194	0.0519	1.8062	2.5441
6_AZTUA004ID	7	0	2.6825	0.0302	2.5185	3.0414
6_CAALAL09ID	9	0	2.5108	0.0354	2.2553	2.8451
7_OREUA001CO	14	2	2.0880	0.0693	1.7324	2.6232
7_OREUA003RE	15	0	1.8384	0.1033	1.4771	2.8129
7_ORGRA003RE	6	0	1.6522	0.0216	1.4624	1.8865
7_ORGRA004CO	6	0	1.949	0.0690	1.591	2.312
7_ORPOA001CO	12	1	2.192	0.170	1.613	2.964
7_ORPOA003ID	12	2	2.747	0.165	2.288	3.908
7_ORPOA004ID	9	4	2.404	0.130	1.672	2.799
7_ORPOA006RE	11	2	1.9330	0.0652	1.6128	2.4914
7_ORSAA002CO	5	1	1.7955	0.00895	1.6532	1.9085
7_ORSAA003ID	4	2	1.8372	0.0131	1.6990	1.9542
7_ORSAA004RE	5	1	1.273	0.143	0.840	1.699





General Linear Model: LZN versus Landuse, EPA_Rain_Zone

Factor	Type	Levels	Values
Landuse	fixed	3	CO, ID, RE
EPA_Rain_Zone	fixed	6	2, 3, 4, 5, 6, 7

Analysis of Variance for LZN, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Landuse	2	41.3490	21.4483	10.7241	85.53	0.000
EPA_Rain_Zone	5	21.5544	19.7393	3.9479	31.49	0.000
Landuse*EPA_Rain_Zone	10	5.6755	5.6755	0.5676	4.53	0.000
Error	733	91.9066	91.9066	0.1254		
Total	750	160.4855				

Main factors and interaction are significant.

Bonferroni Simultaneous Tests
Response Variable LZN
All Pairwise Comparisons among Levels of Landuse
Landuse = CO subtracted from:

Landuse	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
ID	0.1634	0.04412	3.703	0.0007
RE	-0.3098	0.04160	-7.446	0.0000

Landuse = ID subtracted from:

	Difference	SE of		Adjusted

```
Landuse    of Means  Difference  T-Value   P-Value
RE          -0.4732     0.03708   -12.76    0.0000


Tukey Simultaneous Tests
Response Variable LZN
All Pairwise Comparisons among Levels of Landuse
Landuse = CO  subtracted from:

         Difference       SE of            Adjusted
Landuse    of Means  Difference  T-Value   P-Value
ID           0.1634     0.04412    3.703    0.0006
RE          -0.3098     0.04160   -7.446    0.0000


Landuse = ID  subtracted from:

         Difference       SE of            Adjusted
Landuse    of Means  Difference  T-Value   P-Value
RE          -0.4732     0.03708   -12.76    0.0000
```

Zinc concentration is different in commercial compared with residential and industrial. Zinc concentrations in residential are different than in industrial.

```
Bonferroni Simultaneous Tests
Response Variable LZN
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
3                 -0.1102     0.06811   -1.618    1.0000
4                  0.3684     0.05363    6.869    0.0000
5                 -0.0850     0.03796   -2.240    0.3813
6                  0.3412     0.03930    8.680    0.0000
7                  0.0691     0.04335    1.594    1.0000


EPA_Rain_Zone = 3  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
4                 0.47863     0.07986   5.9935    0.0000
5                 0.02521     0.07030   0.3586    1.0000
6                 0.45139     0.07103   6.3547    0.0000
7                 0.17934     0.07335   2.4449    0.2209


EPA_Rain_Zone = 4  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
5                 -0.4534     0.05638   -8.042    0.0000
6                 -0.0272     0.05729   -0.475    1.0000
7                 -0.2993     0.06014   -4.976    0.0000


EPA_Rain_Zone = 5  subtracted from:

               Difference       SE of            Adjusted
EPA_Rain_Zone    of Means  Difference  T-Value   P-Value
6                  0.4262     0.04298    9.916    0.0000
```

7	0.1541	0.04671	3.300	0.0152

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.2720	0.04781	-5.690	0.0000

Zinc concentration in EPA rain zone 2 is different than in rain zones 4 and 6. There were no differences between rain zone 2 and rain zones 3, 5 and 7.

Zinc concentration in EPA rain zone 3 is different than in rain zones 4 and 6. There was no differences between rain zone 3 and rain zones 5 and 7.

Zinc concentration in EPA rain zone 4 is different than in rain zones 5 and 7. There was no differences between rain zone 4 and rain zone 6.

Zinc concentration in EPA rain zone 5 is different than in rain zones 6 and 7.

Zinc concentration in EPA rain zone 6 is different than in rain zone 7.

Tukey Simultaneous Tests
Response Variable LZN
All Pairwise Comparisons among Levels of EPA_Rain_Zone
EPA_Rain_Zone = 2 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
3	-0.1102	0.06811	-1.618	0.5864
4	0.3684	0.05363	6.869	0.0000
5	-0.0850	0.03796	-2.240	0.2196
6	0.3412	0.03930	8.680	0.0000
7	0.0691	0.04335	1.594	0.6024

EPA_Rain_Zone = 3 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
4	0.47863	0.07986	5.9935	0.0000
5	0.02521	0.07030	0.3586	0.9992
6	0.45139	0.07103	6.3547	0.0000
7	0.17934	0.07335	2.4449	0.1410

EPA_Rain_Zone = 4 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
5	-0.4534	0.05638	-8.042	0.0000
6	-0.0272	0.05729	-0.475	0.9970
7	-0.2993	0.06014	-4.976	0.0000

EPA_Rain_Zone = 5 subtracted from:

Difference	SE of	Adjusted

EPA_Rain_Zone	of Means	Difference	T-Value	P-Value
6	0.4262	0.04298	9.916	0.0000
7	0.1541	0.04671	3.300	0.0124

EPA_Rain_Zone = 6 subtracted from:

EPA_Rain_Zone	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
7	-0.2720	0.04781	-5.690	0.0000

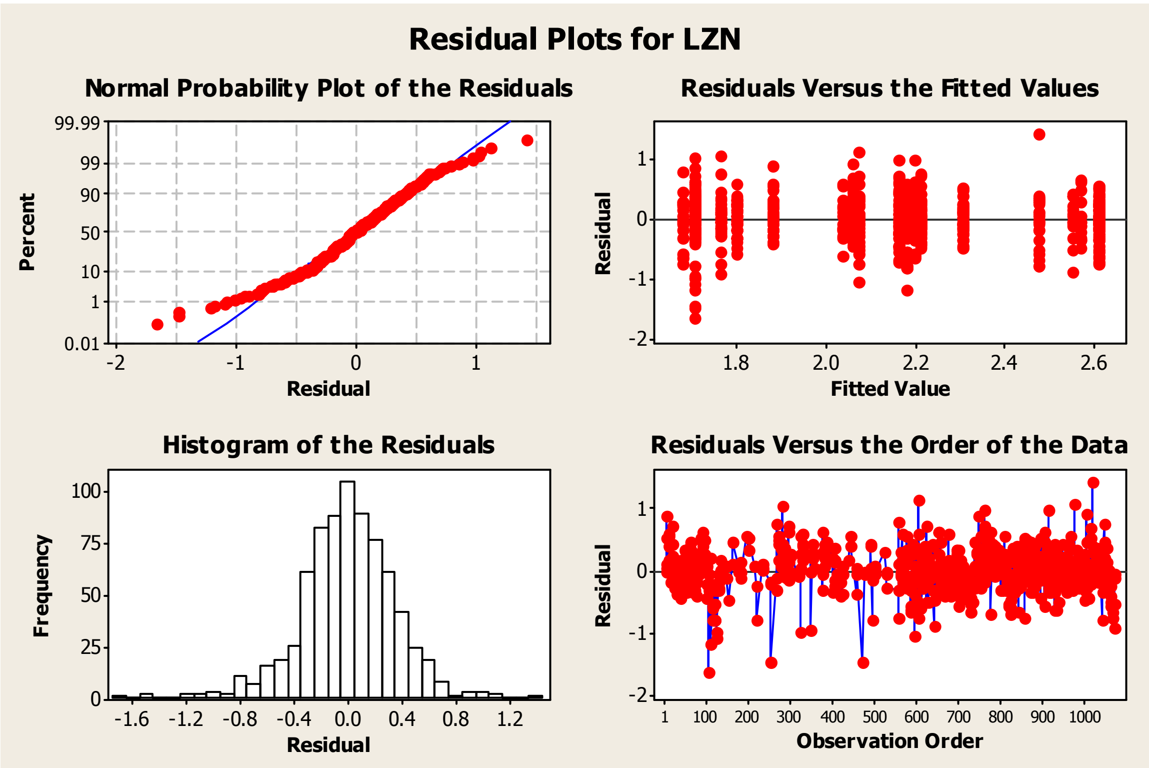
Zinc concentration in EPA rain zone 2 is different than in rain zones 4 and 6. There was no differences between rain zone 2 and rain zones 3, 5 and 7.

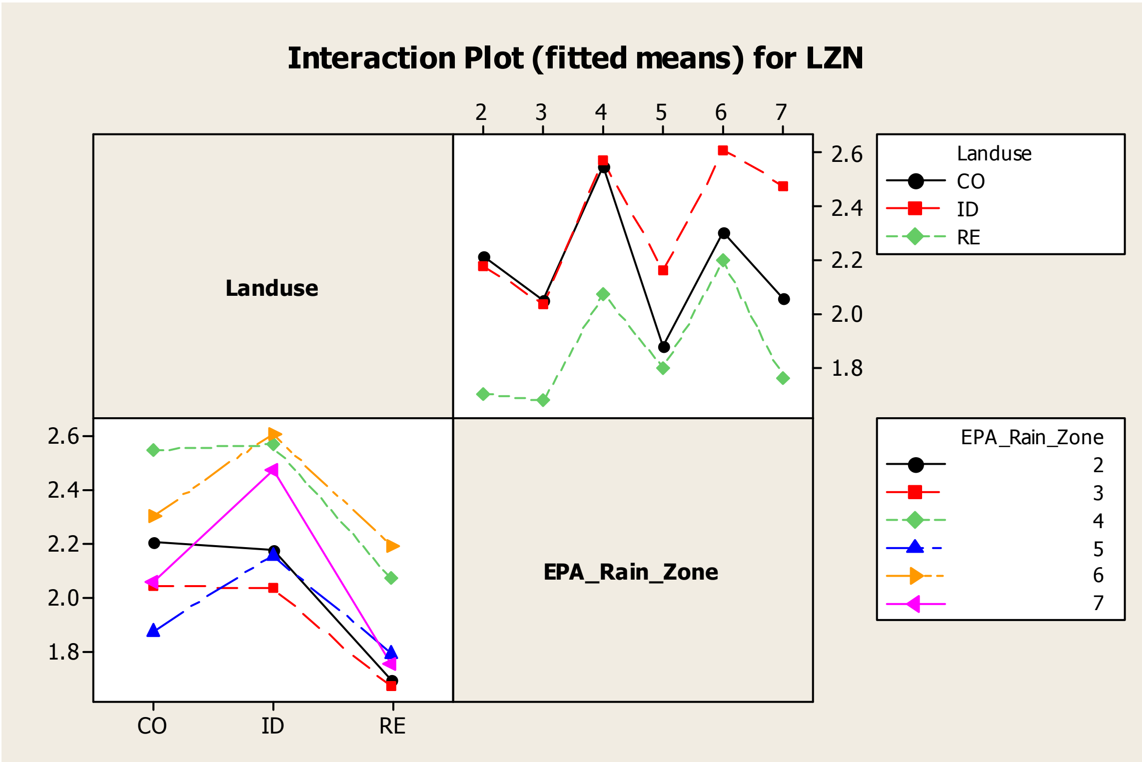
Zinc concentration in EPA rain zone 3 is different than in rain zones 4 and 6. There was no differences between rain zone 3 and rain zones 5 and 7.

Zinc concentration in EPA rain zone 4 is different than in rain zones 5 and 7. There was no differences between rain zone 4 and rain zone 6.

Zinc concentration in EPA rain zone 5 is different than in rain zones 6 and 7.

Zinc concentration in EPA rain zone 6 is different than in rain zone 7




Interaction Plot (fitted means) for LZN
Landuse
EPA_Rain_Zone
2
3
4
5
6
7
2.6
2.4
2.2
2.0
1.8
CO
ID
RE
Landuse
CO
ID
RE
EPA_Rain_Zone
2
3
4
5
6
7