UNIVERSITY OF CALIFORNIA

Los Angeles

Characteristics of Pollutants in Highway Runoff: Regression, Representativeness, and First Flush

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Civil Engineering

by

Jiun-Shiu Ma

2002

The dissertation of Jiun-Shiu Ma is approved.

Irwin H. Suffet

Keith D. Stolzenbach

William W-G., Yeh

Michael K. Stenstrom, Committee Chair

University of California, Los Angeles

2002

To my parents

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Without the support and encouragement of many people, the completion of this dissertation would be extremely difficult.

I would like to thank my parents first. It was them who financially supported me in the beginning of this study. Without their sacrifice, I would have never thought of pursuing higher degree abroad. Especially, after my father retired, they deserved enjoying their retirement life more. Additionally, my grandmother, my aunt, and my uncle all helped me materially and morally in this period.

I am particularly grateful to my dissertation advisor, Professor Michael K Stenstrom. Under his supervision throughout my study, I learned so much from him on many aspects, academically or non-academically. His liberal style of supervision gave me the most freedom to develop my ability and skills. Additionally, I especially thank for his kindness and patience toward foreign students, regardless of the language and culture barriers in the beginning.

I would like to express my appreciation to the task committee, Professor William Yeh, Keith Stolzenbach, and Mel Suffet. Their suggestions and advice were extremely valuable on the investigation of topics in this dissertation. Additionally, I also owe a deep gratitude to the staff of the Civil and Environmental Engineering Department at UCLA, especially to Deeona Columbia. Her everlasting penchant rescued me from numerous administrative messes.

A special acknowledgement is given to *Caltrans* (California Department of Transportation). Three years of the highway stormwater monitoring project let me

collect the required data for this dissertation, and also financially supported me. The stormwater project under *IoE* (the Institute of Environment) also needs to be acknowledged. The opportunity they provided let me begin to investigate the stormwater field.

The whole stormwater group under Professor Michael K Stenstrom will be remembered for working together in every cold, rainy night while sampling. Although I screwed up a couple of times, they covered me really well and successfully had the jobs done. I would like to share a special thought with my crews: "People haven't sampled themselves cannot picture how tough this job is, so please don't blame persons occasionally screwed up sampling".

Finally, I would like to thank all emotional support and encouragement from my friends in Taiwan and the States in this period. I cherish every helpful moment you provided to me.

VITA

March 9, 1970	Born, Taipei, Taiwan
1992	B.S., Environmental Engineering National Cheng-Kung University, Tainan, Taiwan
1994-1995	Research Assistant Institute of the Environment National Taiwan University
1996	M.S. Civil and Environmental Engineering University of California, Los Angeles
1997-2002	Research Assistant Department of Civil and Environmental Engineering University of California, Los Angeles
1998	Teaching Assistant Department of Civil and Environmental Engineering University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

Jiun-Shiu Ma Lee-Hyung Kim, Haejin Ha, Ying-Xia Li, Lee-Hyung Kim, Haejin Ha, Sim-Lin Lau, Masoud Kayhanian, and Michael K. Stenstrom, "Implication of Oil and Grease Measurement in Stormwater Management Systems", Global Solutions for Urban Drainage, 9th International Conference on Urban Drainage, September, 2002, Portland, Oregon.

Jiun-Shiu Ma, Sabbir Khan, Ying-Xia Li, Lee-Hyung Kim, Haejin Ha, Sim-Lin Lau, Masoud Kayhanian, and Michael K. Stenstrom, "First Flush Phenomena for Highways: How it can be meaningfully defined", Global Solutions for Urban Drainage, 9th International Conference on Urban Drainage, September, 2002, Portland, Oregon.

Sim-Lin Lau, Jiun-Shiu Ma, Masoud Kayhanian and Michael K. Stenstrom, "First Flush of Organics in Highway Runoff", Global Solutions for Urban Drainage, 9th International Conference on Urban Drainage, September, 2002, Portland, Oregon.

Michael K. Stenstrom, Sim-Lin Lau, Hsueh-Hwa Lee, Jiun-Shiu Ma, Lee-Hyung Kim, Haejin Ha, Sabbir Khan, and Masoud Kayhanian, “First Flush Stormwater Runoff from Highways”, World Water & Environmental Resources Congress, May, 2001, Orlando, Florida.

Hsueh-Hwa Lee, Jiun-Shiu Ma, and Michael K. Stenstrom, “Calibration of a GIS-Based Empirical Urban Runoff Model on Ballona Creek Watershed”, American Geophysical Union, May, 1998, Boston.

Michael K. Stenstrom, Hsueh-Hwa Lee, and Jiun-Shiu Ma, “Mathematical Modeling and Computer Simulation”, Chapter 2 in *Dynamics and Control of Wastewater System*, Technomic Publishing Company, Inc., Pennsylvania, 1998.

ABSTRACT OF THE DISSERTATION

Characteristics of Pollutants in Highway Runoff:

Regression, Representativeness, and First Flush

by

Jiun-Shiu Ma

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2002

Professor Michael K. Stenstrom, Chair

Stormwater pollution control is the new environmental problem, even though it was recognized in the 1972 Clean Water Act Amendments. Impervious surfaces are the most problematic for stormwater management, and highways are especially important because of vehicular activities. To better understand the nature of highway runoff, three sites along major highways were monitored for 40 storm events over two seasons.

Water quality parameters were measured and chemical oxygen demand (COD), total suspended solids, oil and grease and two metals were selected for basic statistical analysis and correlation. The sampling distributions were skewed, and logarithmic transformations were useful to remove skewness. The logarithmically transformed concentrations were used as basic data forms for further analysis.

Regression was performed to relate the COD concentrations to the following predictors: the corresponding accumulative rainfall (CumR), the event's antecedent dry days (AtDry), and the previous event's precipitation (AtR). After evaluation a site-pooled COD regression model was developed regressing log COD with log CumR, log AtDry and log AtR and a constant. This model shows a strong agreement with the observations. The site effect is not important in the further analysis. This model structure can be applied to any parameter that correlates well with COD.

The flow-weighted average concentration, called the event mean concentration (EMC) is currently used for estimating stormwater mass loading, and total maximum daily loads are established using EMCs. EMCs can be estimated from a series of grab samples or by flow-weighted composite samplers. Simulations show that a composite sample is a better way to obtain the EMCs due to the large number of individual samples.

Simulations show that treating the entire first flush volume is a better strategy if mass removal efficiency is to be maximized. Computer simulations show that as many as forty samples may be necessary to construct a reliable mass-load curve to characterize the first flush. An improved, concentration-based first flush notation is proposed as an alternate way of defining first flush phenomenon. A hypothetical scenario shows how this new notation could benefit best management practice selection.

1. INTRODUCTION

1.1 Dissertation Organization

This dissertation is divided into seven chapters. The first chapter includes a statement of the problem and scope of the work. Chapter 2 describes a monitoring procedure that we used in a *Caltrans* (California Department of Transportation) highway runoff project. Chapter 3 provides an overview of experimental data from highway runoff monitoring. The subsequent chapters, from chapter 4 to 6, contain three independent and complete research works, which are the core of this dissertation. Basically, these three works use the monitoring results as the basis to investigate some popular topics in stormwater field. Chapter 4 contains a regression analysis, which describes the concentration variation for selected water quality constituents during highway runoff processes. Chapter 5 contains a reliability analysis of EMC (Event Mean Concentration) data, which provides guidelines for measuring EMCs and evaluating existing EMC data. Chapter 6 contains a study of first flush notations and criteria, which gives detailed descriptions and discussion to current first flush notations and, furthermore, presents a new and more feasible first flush notation. Therefore, although these three works are independent, the regression results obtained from Chapter 4 are used as the true process mechanism in the computer simulations in the EMC and the first flush studies. Finally, Chapter 7 includes an overall conclusion of this dissertation.

1.2 Overview of the Problem

The investigation of stormwater pollution has a long history in the states. It can be traced back to the mid-1960s when government agencies had first identified stormwater discharge as a major pollution sources to the nation's waterways. In the 1972 Amendments to the Clean Water Act, the Nationwide Urban Runoff Program (NURP) was established to help investigate the urban runoff pollution. National Water Quality Inventory in a 1990 Report to Congress stated that 30% of identified cases of water quality impairment are attributable to storm water discharge or nonpoint source pollution (U.S. EPA 1990). Many case studies show that nonpoint sources that have surpassed point sources and are now the major cause of surface water quality degradation (Gilliland and Baxter-Potter 1987; Driscoll et al. 1990; National 1990, US EPA 1983). In addition, stormwater runoff may contain harmful toxic containments (Lau et at. 1993).

Even after much investigation, many uncertainties still exist in understanding stormwater pollution processes. These uncertainties reflect the lack of strong field verification. Corbitt (1989) pointed out that nonpoint sources processes such as stormwater runoff are inherently difficult to model due to their stochastic nature in both time and space domains. Like other complicated natural processes, the stochastic nature of stormwater pollution processes is possibly due to the complexity of the system. Thus, it is almost foreseeable that stormwater pollution processes cannot be fully predicted in a purely deterministic way. Therefore, from the engineering or management point of view, partial deterministic models are still useful. Minimizing its stochastic nature can increase the performance of invented controls to mitigate stormwater pollution.

Comprehensive stormwater management should address both quantity and quality of runoff. The quantity controls, such as flood control, have reached maturity due to continuing past effort. The quality controls are still in the early stage of development. Human activity is so far recognized as the most important factor that affects the quality of runoff, such as urbanization or agriculture. In fact, most human activities seriously impact runoff quality too due to changing perviousness of surfaces. However, the step following the qualitative recognition is to be able to reasonably quantify each effect from different human activities. The success of this task has to rely on large amounts of data collection and persuasive modeling on stormwater pollution processes.

1.3 Objectives

The goal of this dissertation is to gain more information about stormwater pollution processes and, furthermore, to deliver useful messages for management level decisions. Through a monitoring procedure on selected highway runoff systems, which comes with simple conveyance paths and purified catchment surfaces by single human activity, the complexity of the problem has been reduced. Thus, the study can avoid unnecessary interferences. The specific objectives are as follows,

1. Provide information on proper stormwater monitoring procedures.
2. Present highway monitoring results.
3. Model stormwater pollution processes.
4. Provide a method of evaluating and accurately measuring EMCs (Event Mean Concentrations).

5. Clarify the first flush phenomenon through improved notations and concepts.

Determining the EMCs and quantifying the first flush phenomenon has been the subject of previous research. The form of EMCs is a calculation factor from the NURP studies (USEPA, 1983) that was used to evaluate urban runoff pollution loading. By its name, an EMC represents the average concentration of a pollutant emitted from a site throughout the whole storm event. The subject of first flush is now very controversial among stormwater researchers and managers. A neutral and consensus qualitative understanding of the first flush phenomenon is that "the first part of runoff in a storm event is the most polluted". Understanding the first flush phenomenon is believed to assist in selecting BMPs (Best Management Practice).

2. EXPERIMENT PROCEDURE FOR HIGHWAY STORMWATER MONITORING PROGRAMS

2.1 Introduction

In this section, we will briefly review the experimental procedure used in our *Caltrans* highway monitoring project. This project has continued for four seasons since September 1999. The objective was originally to study first flush phenomenon from highway runoff, although trash and particle size investigation were added later. However, only the first flush part will be reviewed here. The experimental procedure includes the monitoring and analytical procedures. The monitoring method basically covers the field sampling work, such as equipment setup and operation, and sample collection. The whole stormwater team accomplished this part of the work. The analytical procedure basically covers the lab analysis of collected samples. The lab group accomplished this part. The writer participated only in the monitoring work.

2.2 Monitoring Procedures

2.2.1 Sites

Three monitoring sites were selected at the beginning of the project. All of the sites are adjacent to the outlets of highway storm drains, and thus, receive typical highway runoff. The considerations for site selection include personnel safety, proximity to UCLA, and access to the storm drain. *Caltrans*'s technical manual (*Caltrans*, 2000)

provides more detailed descriptions of general considerations about site selection. Figure 2.1 shows the locations of these sites.

There are several site characteristics that could affect flow rate and water quality in runoff, such as catchment area, average daily traffic (ADT), antecedent dry days, and other environmental conditions, such as wind. Table 2.1 shows some of these site characteristics.

2.2.2 Monitoring Equipment

The monitoring equipment includes flow meters, rain gauges, automated samplers, and power supplies. Figure 2.2 shows the photos of each site. A brief description of component is provided below.

2.2.2.1 Power Supply

Generally, commercially available stormwater equipment is capable of running either AC or DC power. When AC power is available, the preferred setup is to operate the equipment by AC power with battery backup. Among our three sites, only Site 3 has AC power. The other two sites have to rely on batteries. Two types of batteries were used for our monitoring. One is the gel cell battery, which is lighter and used for the flow meter. The other is a lead acid battery, which is much heavier and used for the automated sampler. Regular maintenance is required to recharge these batteries when the voltage is low. These batteries were drained out more quickly than reported in the manual during the rainy season.

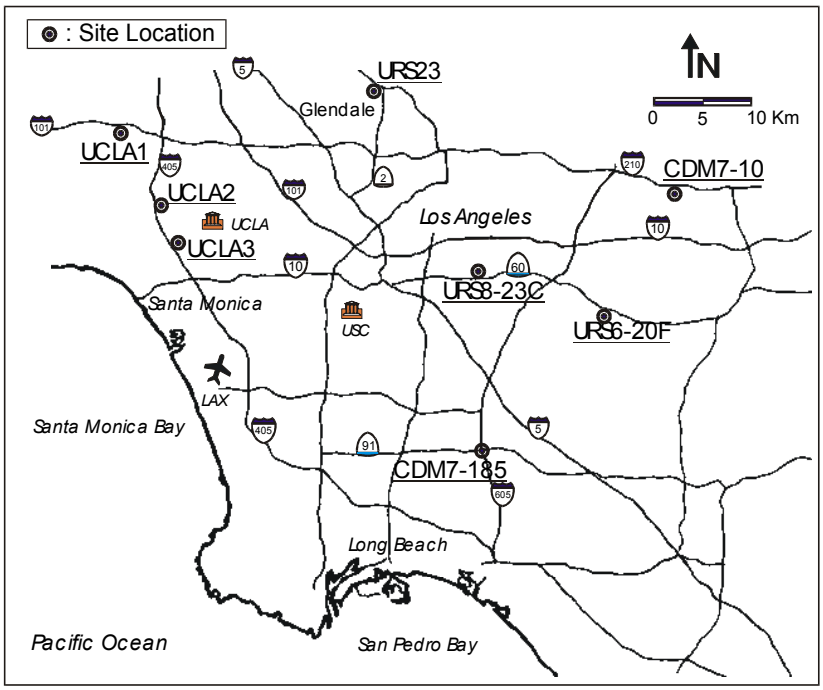


Figure 2.1 Locations of Monitoring Sites Including UCLA and Other Agencies

Table 2.1 Site Descriptions

Site Name	Site Location	ADT (cars/day)	Catchment Area (hect)	Approximate % Impervious
Site 1	101 Freeway	328,000	1.28	100
Site 2	405 Freeway	260,000	1.69	95
Site 3	405 Freeway	322,000	0.39	100

2.2.2.2 Flow Meter

There are two basic types of flow measuring devices typically used for flow weighted stormwater sampling (American Sigma, 2002): (1) depth sensors, which convert level measurements to flow rates based on the know pipe or channel geometry and Manning's flow equation; and (2) area velocity sensors, which measure both the depth and velocity of flow to produce more accurate flow measurements. Among our three monitoring sites, the depth-only sensor was originally installed at one site. All of sensors were changed to area velocity sensors after the first season (1999-2000). The model is *Ultrasonic 950 Area-Velocity Flow Meter* from American Sigma (Oakville, Ontario, Canada).

There is a known calibration issue with existing flow-weighted composite samplers that are associated with the flow sensor. The reference depth or datum can differ when the sensor is first activated, which produces a bias. The bias can be removed by correcting the datum, but this is usually only done after the storm event. The bias affects sample volume. The volume may be too large for the sample bottle, or too small to meet laboratory needs.

2.2.2.3 Automated Sampler

An automated sampler system is comprised of a peristaltic pump, a sample distribution, a housing that contains the composite bottles, and a sampler intake system. The peristaltic pump creates suction by compressing a flexible tube within a rotating roller, drawing the sample into the pump. As the roller turns, it pushes the sample out of the roller. As long as the suction side is not clogged, a peristaltic pump is a positive

displacement pump. According to the manufacturer's recommendation, the pump operates best when placed close to the source, which reduces the required suction head.

Proper installation of a sampler intake system assures the collection of representative samples. The intake strainer is used to mount the intake tubing to the bottom of a pipe or channel, and should be placed in the main flow. The vertical position of the intake strainer should avoid at very bottom or top of the flow in order to take a representative sample of heavy solids and floating material. The intake must also be located in a turbulent area to avoid stratified flows. Suitable protection for intake strainers can prevent rocks and debris from clogging or damaging the intake tubing and pump. Additionally, manufacturers recommend Teflon tubing due to its inert properties (American Sigma, 2002).

Automated samplers (*900MAX*, American Sigma) were installed at all sites after the first season. However, only a few composite samples were obtained in the first year. Based on our own learning experience, mastering automated samplers requires an extensive training period, including installation and programming.

2.2.2.4 Rain Gauge

An electronic "tipping bucket" rain gauge (from American Sigma) was used to measure precipitation at our sampling sites. This type of rain gauge is very accurate and electronically records rainfall in a function of time. It collects rainfall in a standard 8-inch cylinder in small increments, usually from 0.01 to 0.05 inch, and automatically tips, emptying the bucket. The rain gauge is connected to a datalogger (a flow meter), which

counts the number of tips. After events, all of rainfall data along with flow data can be retrieved from the datalogger.

2.3 Analytical Procedures

The selection of water quality parameters should be based on the potential pollution sources in the catchment area (Ma et al, 2002). For highway runoff, numerous constituents have been previously detected. In addition, the requirements listed in the monitoring plan for the Los Angeles County storm water/urban runoff NPDES permit are good references for parameter selection. The parameters in our study are divided into five groups, as follows:

conventional, such as TSS, COD, hardness;

nutrients, such as ammonia, nitrite, nitrate, phosphorous;

metals, such as copper, lead, zinc;

petroleum hydrocarbons, such as oil & grease, and

bacteria, such as total and fecal coliforms.

Polynuclear aromatic hydrocarbons (PAHs) were analyzed for a limited number of samples (Lau et al, 2002) but are not covered in this dissertation. Table 2.2 lists the selected water quality parameters and their corresponding laboratory methods, including method reference, laboratory reporting limits, maximum sample holding time and preservation, requirements.

Figure 2.2 Photos of Monitoring Sites

Table 2.2 List of Target Water Quality Parameters and Their Corresponding Analytical Methods

Parameters	Units	Reporting Limits	Analytical Method	Holding time and Preservation[3]
General				
Total suspended solids	mg/L	2	EPA[1] 160.2	7 days; refrigerated at 4°C
Turbidity	NTU	1	EPA 150.1	48 hours; refrigerated at 4°C
Conductivity	μmho/cm	1	EPA 180.1	28 days; refrigerated at 4°C
pH	pH	0.01	EPA 120.1	Analyze immediately
Hardness	mg/L as $CaCO_3$	2	EPA 130.2	6 months; acidify with HNO_3 to pH < 2
Chemical oxygen demand	mg/L	2	EPA 410.0	Analyze as soon as possible
Dissolved organic carbon	mg C/L	1	EPA 415.1	7 days; acidify to pH <2 with H_3PO_4
Nutrients				
Ammonia	mg N/L	0.01	EPA 350.3	Analyze as soon as possible
Nitrite	mg N/L	0.01	EPA 354.1	48 hours; refrigerated at 4°C
Nitrate	mg N/L	0.1	EPA 300.0	48 hours; refrigerated at 4°C
Total Kjeldahl Nitrogen	mg N/L	0.1	EPA 351.4	
Ortho-Phosphate	mg P/L	0.1	EPA 300.0	48 hours; refrigerated at 4°C
Phosphorus (Dissolved and Total)	mg P/L	0.03	EPA 200.7	48 hours; refrigerated at 4°C
Organics				
Oil and grease	mg/L	1	C18 SPE[2]	28 days; acidify to pH < 2 with HCl
Metals (Dissolved and particulate)			EPA 200.7	Filter immediately, acidity to pH < 2 with HNO_3
Cadmium, chromium, nickel, zinc	μg/L	1		
Copper		3		
Lead		5		
Microbiological				
Total coliform	MPN/100 ml	2	SM[3] B9221	24 hours
Fecal coliform	MPN/100 ml	2	SM C9221	24 hours

[1] EPA Methods and Guidance for Analysis of Water (USEPA, 1999)
[2] Lau and Stenstrom (1997)
[3] Standard Methods (1999)

3. HIGHWAY STORMWATER MONITORING RESULTS

This chapter provides an overview of experiment results from our highway runoff experiments. The results include field physical and lab analytical results for a period of two years (1999 to 2001). Over 40 storm events were monitored from three monitoring sites. All statistics analysis in this chapter uses the computer program, *Arc* (Cook and Weisberg, 1999)

3.1 Field Physical Results

3.1.1 Characteristics of Monitored Events

An event's important characteristics include date, precipitation, rain duration, runoff volume, runoff duration, antecedent dry days, and antecedent precipitation. An event's characteristics are usually a function of sampling site. For example, for the same event, different sites usually experience different precipitation due to their slightly different locations. Some of an event's characteristics are suspected to have physical meaning and affect runoff water quality. For example, antecedent dry days might reveal a pollutant's tendency to "build-up" during the dry days. Antecedent precipitation might be correlated to a pollutant's "leftover" amount from the previous event. Brief descriptions about monitored events' characteristics are covered here.

Figure 3.1 shows the histogram of antecedent dry days for the monitored storm events. The bins are skewed to the right, and most antecedent dry days were less than 10 days. Since we almost monitored every event in the rainy season, the skewed distribution seems to be natural.

Figure 3.2 shows the histogram of antecedent rainfall for the monitored storm events. The bins are skewed to the right, and most antecedent precipitation is less than 1 inch. The skewed distribution is also natural.

Figure 3.3 shows the histogram of event rainfalls for the monitored storm events. The bins are skewed to the right too. In fact, Figure 3 is a reflection of Figure 2. Most event rainfalls are less than 1 inch.

Figure 3.4 shows the histogram of rainfall duration for the monitored storm events. The bins are skewed to the right too, and most rain durations are less than 10 hours. The skewed distribution is natural.

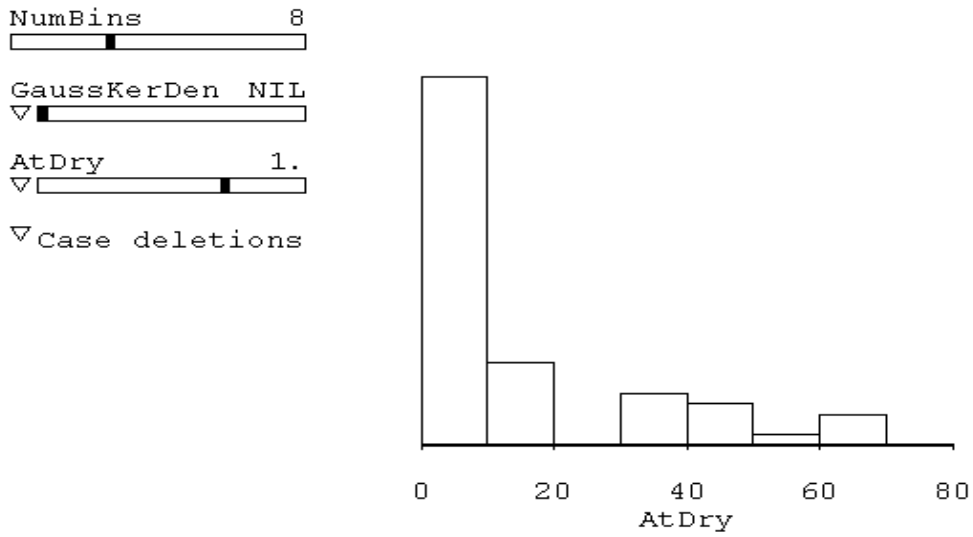


Figure 3.1 Histogram of Antecedent Dry Days of Monitored Events

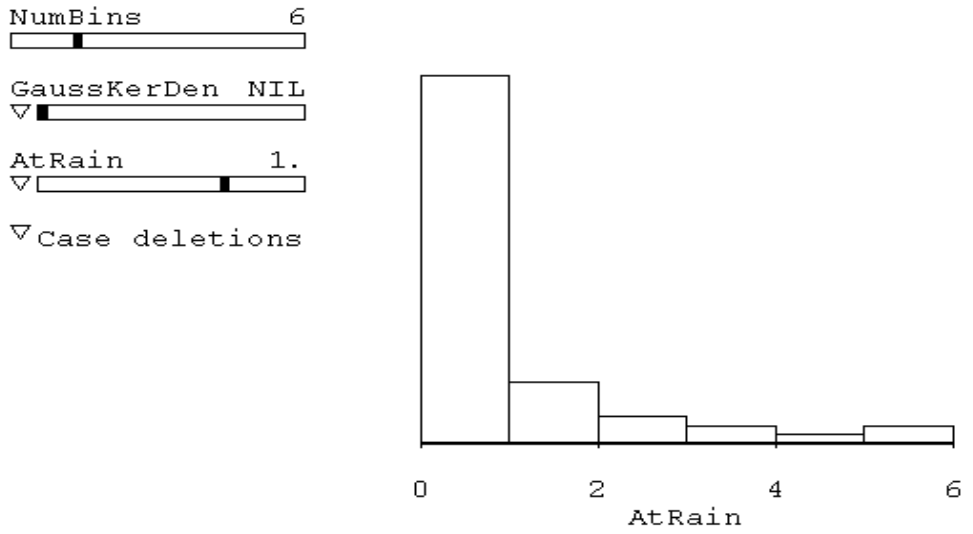


Figure 3.2 Histogram of Antecedent Rainfalls (inch) of Monitored Events

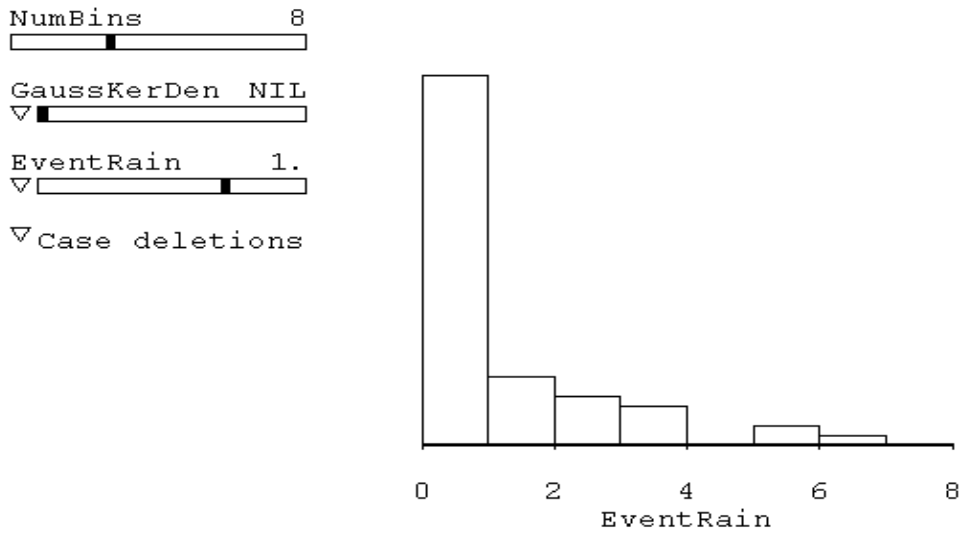


Figure 3.3 Histogram of Monitored Event Rainfalls (inch)

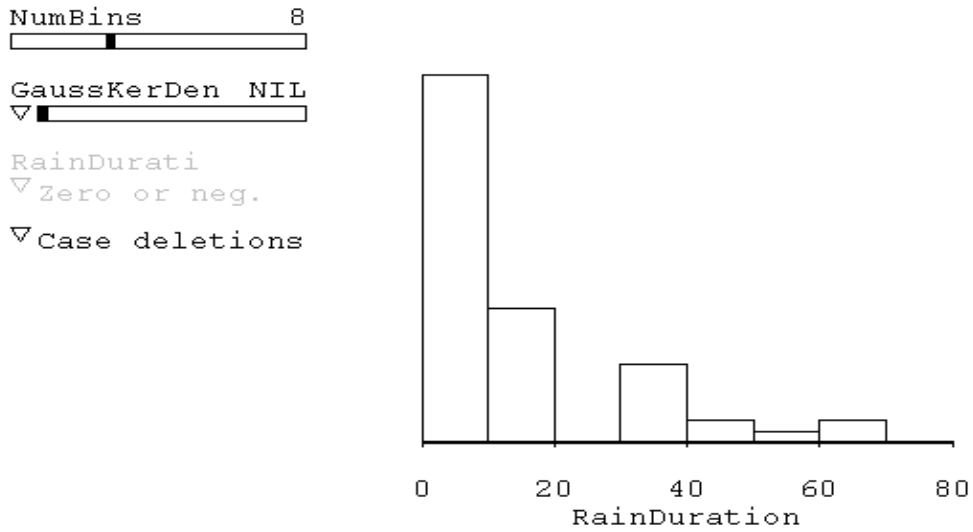


Figure 3.4 Histogram of Monitored Event Durations (hour)

3.1.2 Characteristics of Sampling Time

At each sampling time during an event, there are corresponding field conditions in flow and precipitation, such as flow rate, rainfall intensity, cumulative runoff volume, cumulative rainfall, and other environmental conditions. These corresponding field conditions provide important additional information to examine the lab analytical results. Among field conditions, the corresponding precipitation and flow conditions should be highly correlated to each other, based on principles of runoff hydrology. Thus, to avoid redundancy, the precipitation-related data are selected to represent corresponding field conditions.

Figure 3.5 shows the histogram of corresponding rainfall intensities (inch-rainfall/15-mins) at sampling time. The bins are skewed to the right, and more than half cases are less than 0.04 inch-rainfall per 15-min. There are two possible causes for this skewed distribution. One is that we did not often sample for high rainfall intensity conditions; the other is that the high rainfall intensity conditions are naturally rare. The second cause is more possible.

Figure 3.6 shows the histogram of corresponding cumulative rainfall at sampling time. The bins are skewed to the right, and most of cases are in the interval of 0 to 0.5 inch-rainfall. Both the nature (from the histogram of event rainfalls) and our sampling strategy (we intend to collect first flush samples) might cause this skewed distribution.

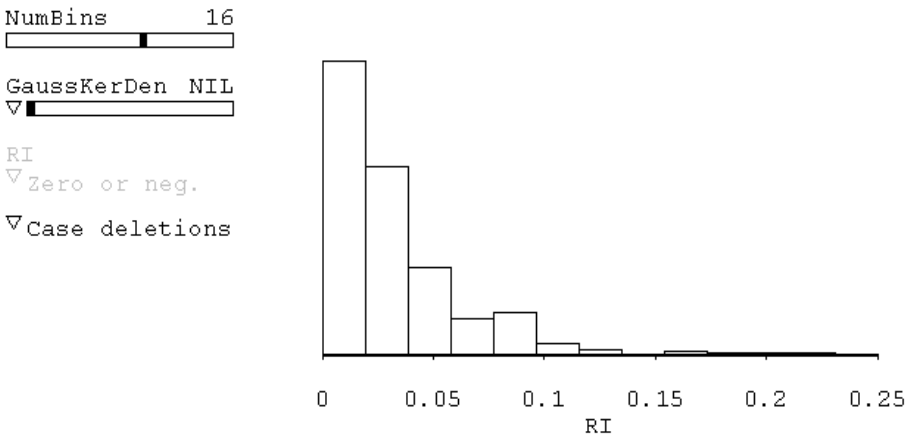


Figure 3.5 Histogram of Rainfall Intensity (inch/15-min) at Sampling Time

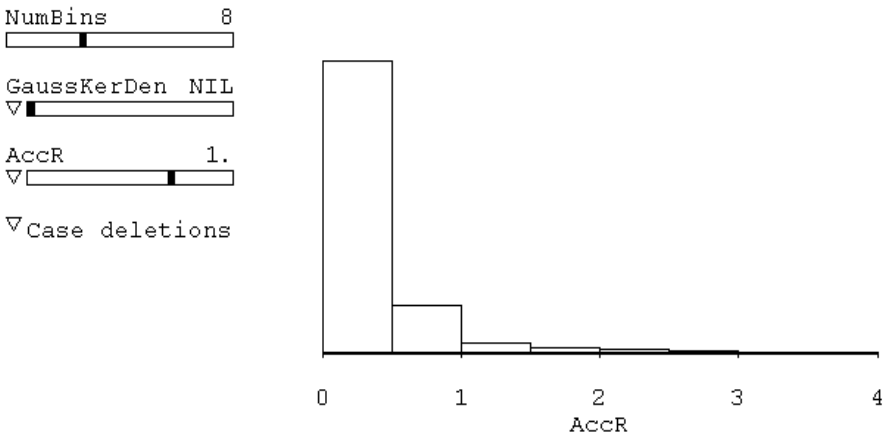


Figure 3.6 Histogram of Cumulative Rainfall at Sampling Time

3.2 Lab Analytical Results

3.2.1 Constituents in Absence

Table 3.1 lists constituents that are not often detected in highway runoff. These listed parameters show at least 20% that are under the reporting limits. The reporting limits are based on the machines' detection limits but with more confidence. Usually, we set the reporting limits at about three times higher than the detection limits. All of the reporting limits and detection limits can be shown in Table 2.2.

From Table 3.1, most listed metals shows significant absence in highway runoff, such as arsenic (As) in either dissolved phase (79%) or particular phase (92%), cadmium (Cd) in either dissolved phase (76%) or particular phases (92 %), lead (Pb) in dissolved phase (78%). Chromium (Cr) in dissolved phase shows a moderate absence (25%). Ortho-P is the only non-metal listed, showing really high absence (73%).

Table 3.1 Constituents Showing Significant Absence in Highway Runoff

Constituent	Phase	Total Cases	Undetected Cases	Undetected %
Cd	Dissolved	445	340	**76%**
Cd	Particulate	425	390	**92%**
Pb	Dissolved	445	349	**78%**
Cr	Dissolved	445	110	**25%**
As	Dissolved	228	179	**79%**
As	Particulate	228	209	**92%**
Ortho-P		449	327	**73%**

3.2.2 Summary Statistics and Histograms

It is convenient to describe the lab analytical results based on their parameter categories that are classified as conventional, nutrients, metals, and petroleum hydrocarbon. For some categories such as metals, we also describe constituents in two phases: dissolved and particulate phases. The summary statistics and histograms are described as a detailed manner for selected parameters.

Conventional/COD (Chemical Oxygen Demand):

Table 3.2 shows the pooled and the site-specific COD results. For pooled COD, there are total 441 cases reported from three sites. The average concentration is 187.6 mg/L. The median concentration is 80.6 mg/L. The fact that the median is much smaller than the mean implies an asymmetric distribution of COD results. The average difference between cases is 291.7 mg/L. The maximum COD is 1583.3 mg/L, and the minimum COD is 24.1 mg/L, a factor of more than 700. Figure 3.7a shows the histogram of pooled CODs. The density smoothing function, GKS (Gauss Kernel Density), with Parameter set as 0.8 was used to sketch the density curve for each site. Overall, the bins are skewed to the right, and there are several outliers found on the right with extreme high concentrations. Three sites' density curves are quite close except that Site 2's curve is a little lower on the mode. Since the bins are extremely skewed, a logarithmic transformation of COD responses is suggested to remove the skewness and display more information in COD results. Figure 3.7b shows the logarithmically transformed histogram.

Conventional/TSS (Total Suspended Solid):

Table 3.3 shows the pooled and the site-specific TSS results. For pooled TSS, there are 441 cases reported from three sites. The average concentration is 62.3 mg/L. The median concentration is 37.4 mg/L. The difference in the mean and median implies an asymmetric sampling distribution. The average difference between cases is 98.2 mg/L. The maximum TSS is 1534.7 mg/L, and the minimum TSS is 2.9 mg/L, a factor of more than 500. Figure 3.8a shows the histogram of pooled TSS, in which GKS with Parameter set as 0.8 is used to sketch the density curve for each site. Overall, the bins are skewed to the right, and there are several outliers on the right with extreme high concentrations. Site 2's density curve is obviously on the right of other two sites', representing a data set with higher TSS values. Since the bins are extremely skewed, a logarithmic transformation of TSS responses is suggested to remove the skewness and display more information in TSS results. Figure 3.8b shows the logarithmically transformed histogram.

Petroleum Hydrocarbon/Oil & Grease (O&G):

Table 3.4 shows the pooled and the site-specific O&G results. For pooled O&G, there are 437 cases reported from three sites. The average concentration is 12.6 mg/L. The median concentration is 6.2 mg/L. The difference in the mean and median implies an asymmetric sampling distribution. The average difference between cases is 16.7 mg/L. The maximum O&G is 108.0 mg/L, and the minimum O&G is under the reporting limit (1 mg/L). Figure 3.9a shows the histogram of pooled O&G, in which GKS with Parameter set as 0.8 is used to sketch the density curve for each site. Overall, the bins are

skewed to the right, and there are several outliers on the right with extreme high concentrations. The site density curves are quite close. Since the bins are extremely skewed, a logarithmic transformation of O&G responses is suggested to cure the skewness and display more information in O&G results. Figure 3.9b shows the logarithmically transformed histogram.

Nutrient/NH_3-N (Ammonia):

Table 3.5 shows the pooled and the site-specific NH_3-N results. For pooled NH_3-N, there are 441 cases reported from three sites. The average concentration is 3.04 mg/L. The median concentration is 1.07 mg/L. The difference in the mean and median implies an asymmetric sampling distribution. The average difference between cases is 5.93 mg/L. The maximum NH_3-N is 35.96 mg/L, and the minimum NH_3-N is 0.03 mg/L, a factor of more than 1000. Figure 3.10a shows the histogram of pooled NH_3-N, in which GKS with Parameter set as 0.8 is used to sketch the density curve for each site. Overall, the bins are skewed to the right, and there are several outliers on the right with extreme high concentrations. Thee sites' density curves are a little different. Since the bins are extremely skewed, a logarithmic transformation of NH_3-N responses is suggested to remove the skewness and display more information in NH_3-N results. Figure 3.10b shows the logarithmically transformed histogram.

Metal/Zn_dis (Zinc in Dissolved Phase):

Table 3.6 shows the pooled and the site-specific Zn_dis results. For pooled Zn_dis, there are 437 cases reported from three sites. The average concentration is 202.8 ug/L. The median concentration is 86 ug/L. The difference in the mean and median implies an asymmetric sampling distribution. The average difference between cases is 412.6 ug/L. The maximum Zn_dis is 6041 ug/L that is probably very unusual, and the minimum Zn_dis is 3.0 ug/L. Figure 3.11a shows the histogram of pooled Zn_dis, in which GKS with Parameter set as 0.8 is used to sketch the density curve for each site. Overall, there is only one obvious bin on the left. The sites' density curves are quite close. Since the original data are all gathered on the left, a logarithmic transformation of Zn_dis responses is suggested to spread them out. Figure 3.11b shows the logarithmically transformed histogram.

For other water quality parameters' summary statistics and histograms, Appendix A lists more results. Basically, almost every parameter except coliforms shows a skewed distribution, and thus, a logarithmic transformation is suggested to remove this skewness and display more information. Additionally, their site density curves show no significant difference, indicating that the site effect does not possibly exist.

Table 3.2 Summary Statistics of Analytical COD Results

COD (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	11.1	15.8	11.1	11.1
Maximum	2714.3	2714.3	2381.0	1800.0
Median	80.6	66.7	80.6	83.3
Mean	187.6	174.1	222.7	161.5
Standard Dev	295.9	332.4	322.6	225.9

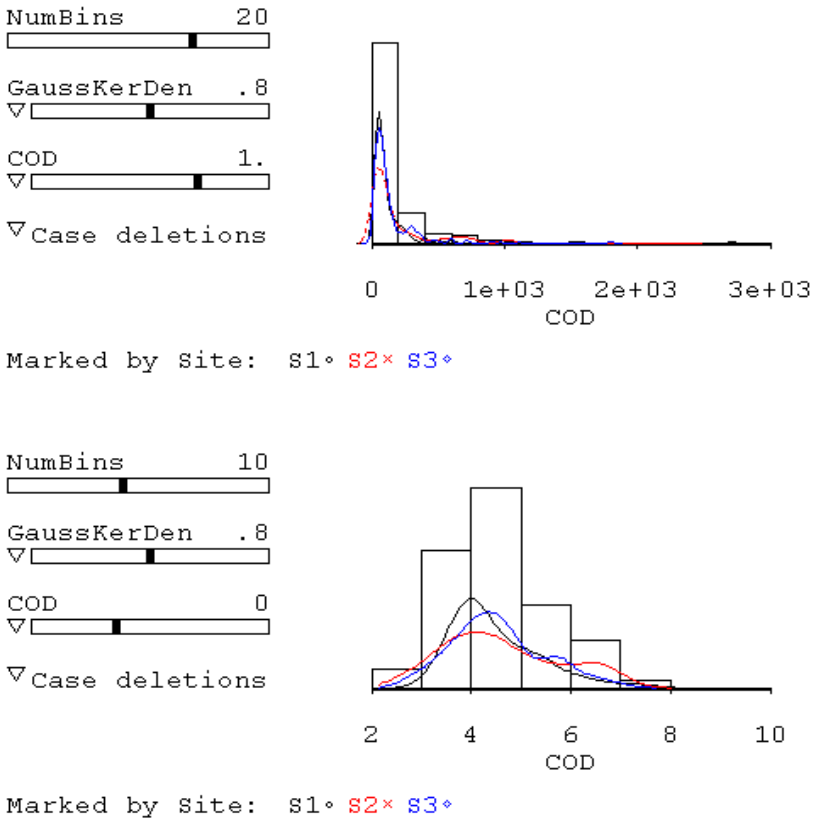
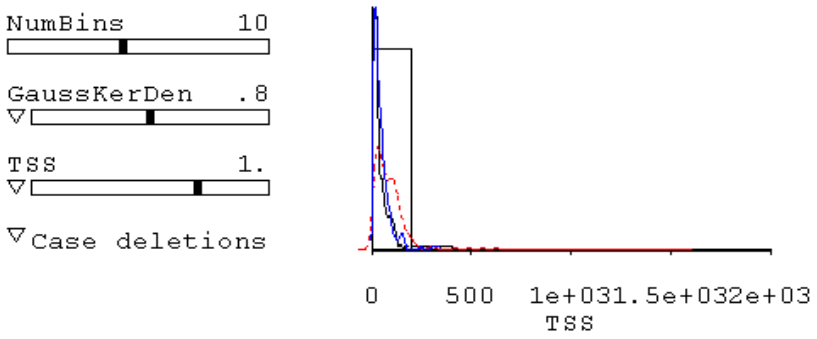


Figure 3.7a (top) & b (bottom) Histograms of Analytical COD Results plus Simulated Density Curves Using Normal and Log Scales

Table 3.3 Summary Statistics of Analytical TSS Results

TSS (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	2.9	2.9	4.5	6.3
Maximum	1534.7	174.7	1534.7	331.6
Median	37.4	20.6	68.3	31.8
Mean	62.3	33.4	97.9	47.7
Standard Dev	98.2	33.2	144.8	49.1



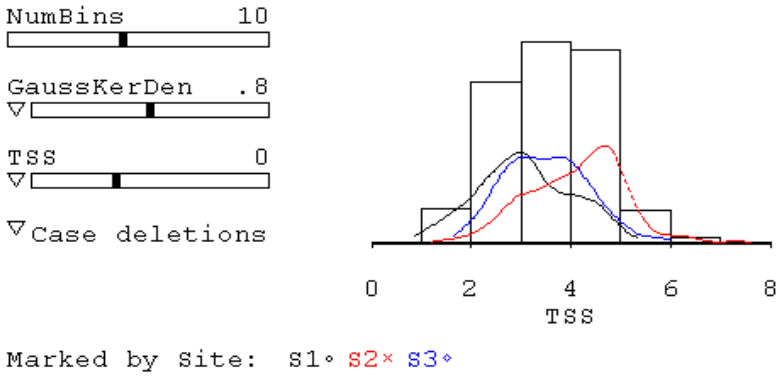
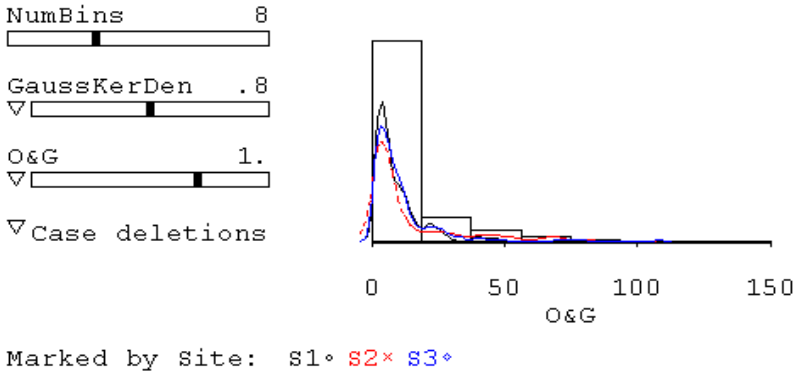


Figure 3.8a (top) & b (bottom) Histograms of Analytical TSS Results plus Simulated Density Curves Using Normal and Log Scales

Table 3.4 Summary Statistics of Analytical O&G Results

O&G (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	0.5*	1.3	0.5	0.5
Maximum	108.0	73.0	102.3	108.0
Median	6.2	5.6	5.9	7.3
Mean	12.6	9.8	15.5	11.7
Standard Dev	16.7	10.6	21.0	15.1



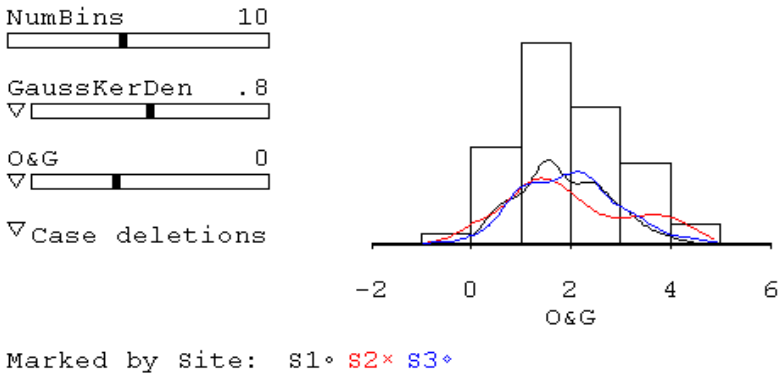
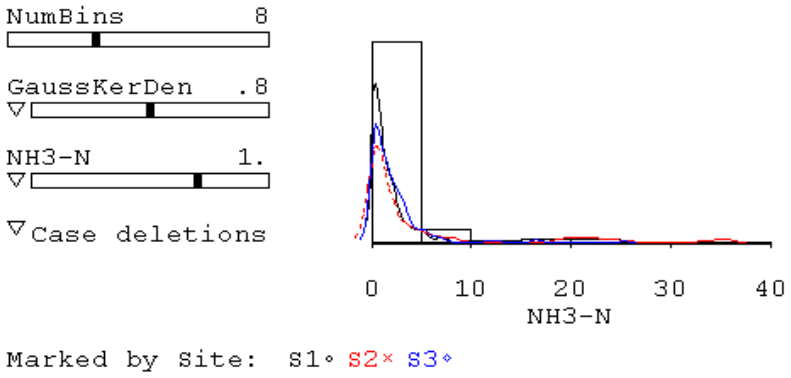


Figure 3.9a (top) &b (bottom) Histograms of Analytical O&G Results plus Simulated Density Curves Using Normal and Log Scales

Table 3.5 Summary Statistics of Analytical NH3-N Results

NH3-N (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	0.03	0.04	0.06	0.03
Maximum	35.96	13.69	35.96	25.41
Median	1.07	0.94	1.04	1.31
Mean	3.04	1.69	4.94	2.13
Standard Dev	5.93	2.27	8.72	3.21



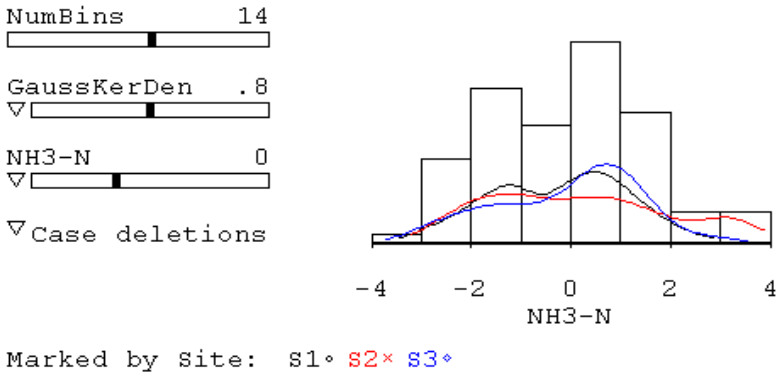
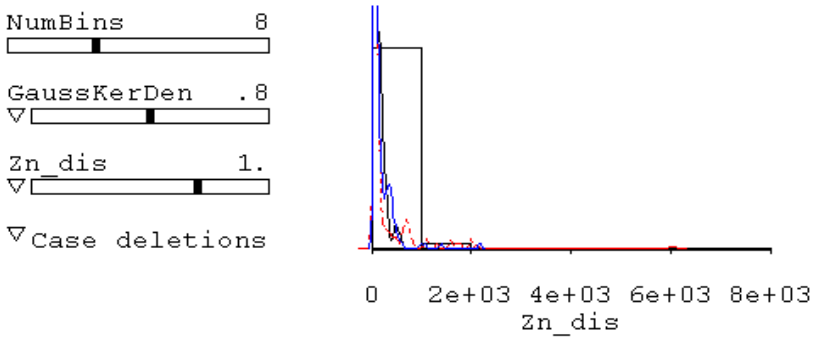


Figure 3.10a (top) & b (bottom) Histograms of Analytical NO3-N Results plus Simulated Density Curves Using Normal and Log Scales

Table 3.6 Summary Statistics of Analytical Zn_dis Results

Zn_dis (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	3.0	6.0	3.0	13.2
Maximum	6041.2	590.9	6041.2	2180.7
Median	86.0	105.5	69.0	102.2
Mean	202.8	142.4	268.9	180.5
Standard Dev	412.6	121.5	616.5	252.5



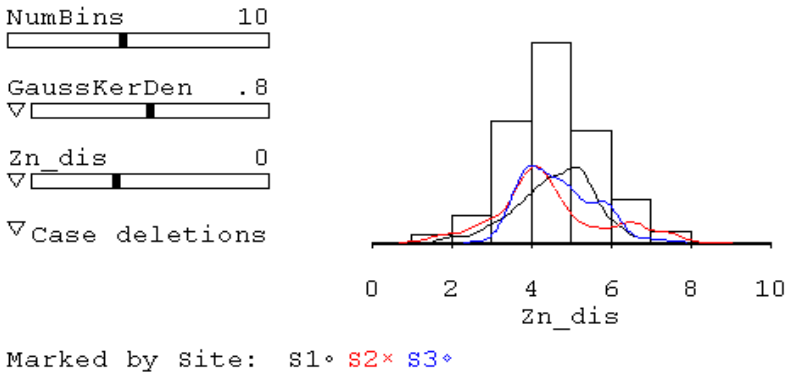


Figure 3.11a (top) & b (bottom) Histograms of Analytical Zn_dis Results plus Simulated Density Curves Using Normal and Log Scales

3.2.3 Correlations between Constituents

Correlations among the water quality parameters in the dissolved and particulate groups were investigated. Figure 3.12 shows the correlation of the dissolved phase parameters. The correlations are for logarithmically transformed data as suggested earlier. Table 3.7 shows the correlation coefficients, which range from 0.70 to 0.94.

Figure 3.13 shows the correlations among the particulate phase parameters. Table 3.8 shows the correlation coefficients. The coefficients range from 0.75 to 0.94 between Cu_prt, Ni_prt, Zn_prt, and Cr_prt. TSS correlates to the metal concentrations, but not as well as the metal's correlation among each other. The correlation coefficients range from 0.63 to 0.72. The correlation between TSS and VSS is 0.93.

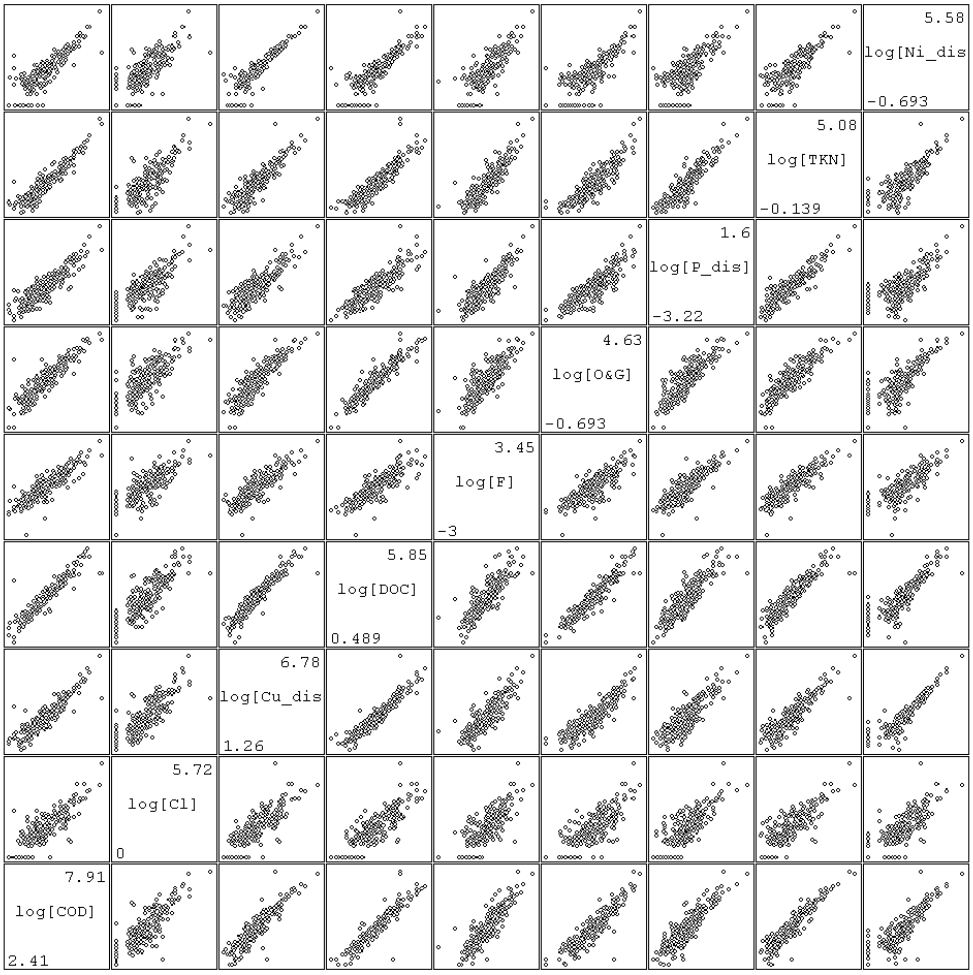


Figure 3.12 Scatterplot Matrix of Correlated Constituents in Dissolved Phase

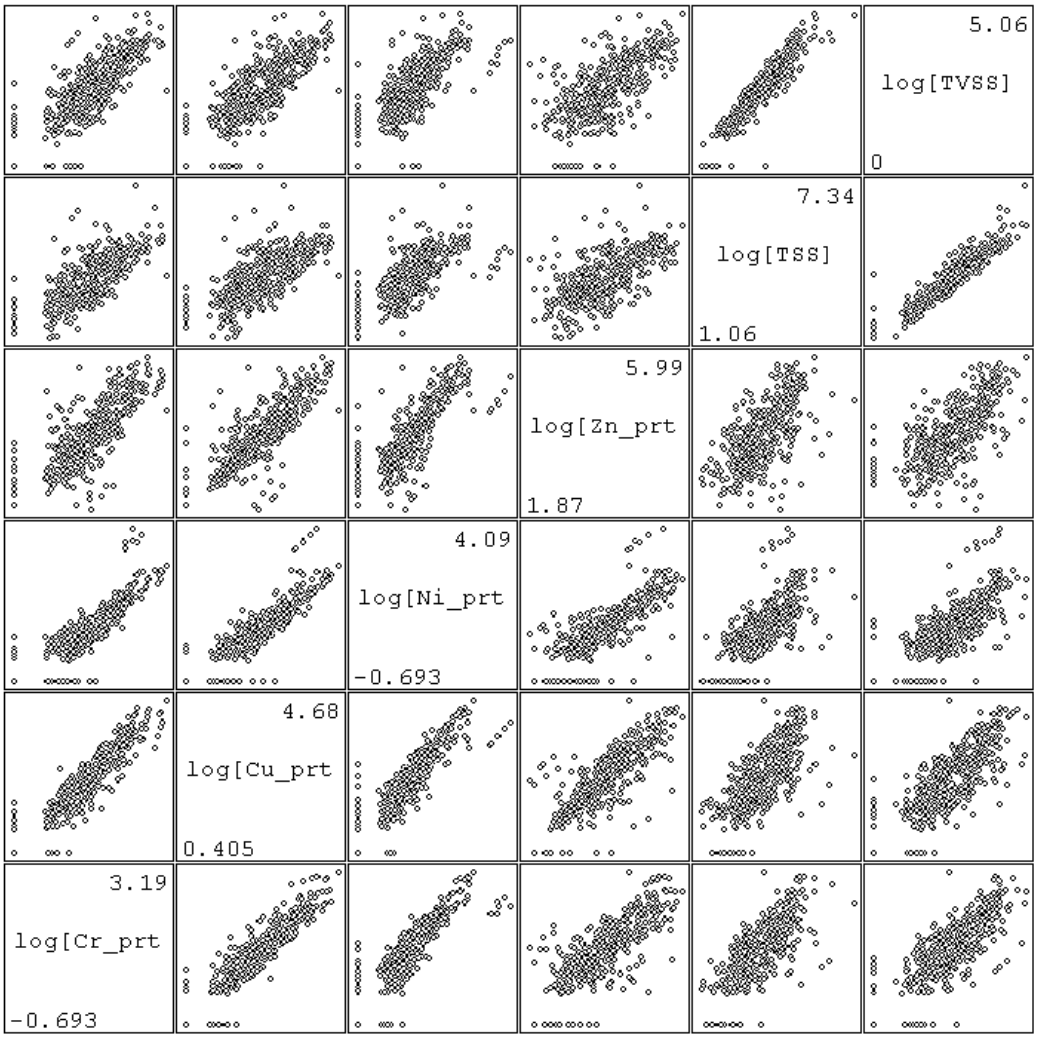


Figure 3.13 Scatterplot Matrix of Correlated Constituents in Particulate Phase

Table 3.7 Sample Correlation Coefficients in Dissolved Group

log[COD]	0.77	0.91	0.94	0.82	0.91	0.85	0.83	0.93
0.77	log[Cl-]	0.78	0.78	0.69	0.76	0.68	0.71	0.81
0.91	0.78	log[Cu_dis]	0.94	0.83	0.91	0.84	0.89	0.89
0.94	0.78	0.94	log[DOC]	0.81	0.94	0.85	0.84	0.91
0.82	0.69	0.83	0.81	log[F-]	0.80	0.77	0.80	0.88
0.91	0.76	0.91	0.94	0.80	log[O&G]	0.85	0.80	0.88
0.85	0.68	0.84	0.85	0.77	0.85	log[P_dis]	0.70	0.90
0.83	0.71	0.89	0.84	0.80	0.80	0.70	log[Ni_dis]	0.81
0.93	0.81	0.89	0.91	0.88	0.88	0.90	0.81	log[TKN]

Table 3.8 Sample Correlation Coefficients in Particulate Group

log[Cr_prt]	0.90	0.85	0.75	0.72	0.73
0.90	log[Cu_prt]	0.85	0.79	0.69	0.73
0.85	0.85	log[Ni_prt]	0.76	0.71	0.74
0.75	0.79	0.76	log[Zn_prt]	0.63	0.64
0.72	0.69	0.71	0.63	log[TSS]	0.93
0.73	0.73	0.74	0.64	0.93	log[TVSS]

4. REGRESSION ANALYSIS ON SELECTED WATER QUALITY PARAMETERS IN HIGHWAY RUNOFF

4.1 Regression Description

4.1.1 Introduction

Some people think that regression is the workhorse in statistics due to its popularity and functionality. However, it is also quite common for people misusing or overusing it. In order to avoid this problem, a brief overview is provided.

Basically, regression is a way to study relations among observations. These relations involve in responses (also called dependent variables), symbolized as $\boldsymbol{y}$, and several other explained variables (also called independent variables, or predictors), symbolized as $\mathbf{x}$. Usually for the single response case, $\boldsymbol{y}$ is in a scalar from (as y), and $\mathbf{x}$ is in a vector form. Regression models are approximations, which are obtained from these studying relations, to the mechanisms generating the observations (y and $\mathbf{x}$).

In general, $\mathbf{x}$ is viewed as observations with nonrandom values whereas y is treated as having random components. Thus the essentials of regression without any assumptions are the conditional distribution of y on $\mathbf{x}$. The errors are calculated by subtracting the expected y from the observed y. From here, several questions arise. How do we get the expectation form of y conditioning on $\mathbf{x}$? And if we want separate errors from y, how do we treat them? If we just simply treat the unknown conditional distribution of y without considering any assumptions on errors, the model is nonparametric. If we treat errors as

independent identically distributed (*i.i.d.*) with arbitrary distribution *F*, the model is semi-parametric. If the distribution *F* is known, then the model is parametric.

4.1.2 *Gaussian* Linear Models

A *Gaussian* linear model is the most popular parametric model in regression, in which *F* has a *Gaussian* distribution, and the expectation of *y* is in the form of a linear combination of **x**. For example, we can write a simple linear model as

$$\mathrm{y} \mid \mathbf{x} \sim N(\eta_0 + \eta_1 x, \sigma^2) \tag{4.1.1}$$

In this simple model, **x** is just a scalar variable, and the linear combination of **x** is just itself plus a constant. The conditional distribution $y \mid \mathbf{x}$ has a *Gaussian* distribution with the mean $\eta_0 + \eta_1 x$ and the variance σ^2. If we want to introduce the error term in the model, we can rewrite the above model as

$$\mathrm{y} \mid \mathbf{x} = \eta_0 + \eta_1 x + e \tag{4.1.2}$$

$$e \sim N(0, \sigma^2)$$

Where *e* is a random error, and has a *Gaussian* distribution with the mean zero and the variance σ^2.

Many hypothesis tests of regression rely on the *Gaussian* model's normality assumptions. However, we are often able to reject *Gaussian* models in the real world due to their intensive assumptions.

4.1.3 Mean and Variance Functions

In order to focus on the study of how the conditional distribution of y changes as **x** is varied, we are mainly interested in the mean function $E(y \mid \mathbf{x})$ and the variance function $Var(y \mid \mathbf{x})$. Here $y \mid \mathbf{x}$ can be viewed as the subpopulation of the whole population y. Among these two, the mean function is usually over the variance function in the study interest.

From mathematical derivation, the mean function and variance function have the following properties:

$$E(y) = E(E(y \mid \mathbf{x})) \tag{4.1.3}$$

$$Var(y) = E(Var(y \mid \mathbf{x})) + Var(E(y \mid \mathbf{x})) \tag{4.1.4}$$

We can interpret the above properties as the mean response $E(y)$ is the weighted mean of $E(y \mid \mathbf{x})$ from subpopulations, and the variance of responses $Var(y)$ is equal to the expectation of variance function plus the variance of the mean function. The second equation is also recognized as ANOVA property, which states that the whole response's variance can be divided into the "within group" variance and "between group" variance.

When we study the mean function $E(y \mid \mathbf{x})$, the predictors **x** may take many forms, as long as they are known. For example, they could be in logarithm, polynomial, or power form; or several predictors might be combined in some way into an index. In order to distinguish between given predictors and predictors' forms in the mean function, we usually use "terms", which are built from the given predictors, to refer predictors' forms in the mean function. For example,

$$E(y \mid \mathbf{x}) = \eta_0 u_0 + \eta_1 u_1 + ... + \eta_{k-1} u_{k-1} \tag{4.1.5}$$

There are k terms in this mean function, from u_0 to u_{k-1}. Each term u_j is a function of given predictors, $u_j(\mathbf{x})$. Thus, while conditioning on **x**, every term u_j is also determined. Conventionally, the term u_0 is a constant and equal to one for the model including the intercept.

The simplest type of term is equal to one of the predictors, i.e. $u_1 = u_1(\mathbf{x}) = x_j$. For complicated forms, terms can be the power of the predictors, transformations of the predictors, polynomial functions of the predictors, or interaction of the predictors. In these complicated cases, since the coefficients, $\eta_0 ..... \eta_{k-1}$ are linear in the mean function, it is also said to be a "linear model". For categorical predictors, the corresponding terms could be their levels. For example, one categorical predictor with *l* levels needs *l*-1 terms in the mean function.

4.1.4 Estimates of Model's Parameters

The most common method of estimating a model's coefficients uses the "Least Squares" estimation. A least squares estimator manipulates parameter values to minimize a particular objective function. In regression, the particular objective function is the residual sum of squares function, *RSS* (**η**). Suppose we have a mean function

$$E(\mathbf{y} \mid \mathbf{X}) = \mathbf{U}\boldsymbol{\eta} \tag{4.1.6}$$

where $\mathbf{y}$ is an $n \times 1$ response vector, $\mathbf{X}$ is a $k \times 1$ predictor vector, $\mathbf{U}$ is an $n \times k$ model matrix, and $\boldsymbol{\eta}$ is a $k \times 1$ parameter vector. $RSS(\boldsymbol{\eta})$ can be defined as

$$RSS(\boldsymbol{\eta}) = (\mathbf{y} - \mathbf{U}\boldsymbol{\eta})^T (\mathbf{y} - \mathbf{U}\boldsymbol{\eta}) \tag{4.1.7}$$

It is called "Ordinary Least Squares" (OLS) estimates. Using calculus, minimizing (4.1.7) leads to the "normal equations", as follows

$$\mathbf{U}^T\mathbf{U}\boldsymbol{\eta} = \mathbf{U}^T\mathbf{y} \tag{4.1.8}$$

Thus, if the inverse of $(\mathbf{U}^T\mathbf{U})$ exists, the OLS estimates are given by

$$\hat{\boldsymbol{\eta}} = (\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\mathbf{y} \tag{4.1.9}$$

Suppose we have a weighted variance function

$$\text{var}(\mathbf{y} \mid \mathbf{X}) = \sigma^2\mathbf{W}^{-1} \tag{4.1.10}$$

where $\mathbf{W}$ is a diagonal matrix whose diagonal elements w_i are known positive weights. It is called "Weighted Least Squares" (WLS) estimates. The objective function becomes

$$RSS(\boldsymbol{\eta}) = (\mathbf{y} - \mathbf{U}\boldsymbol{\eta})^T \mathbf{W}(\mathbf{y} - \mathbf{U}\boldsymbol{\eta}) \tag{4.1.11}$$

We can convert a WLS regression into an OLS regression by converting $\mathbf{y}$ and $\mathbf{U}$ as follows

$$\mathbf{y}^* = \mathbf{W}^{1/2}\mathbf{y} \tag{4.1.12}$$

$$\mathbf{U}^* = \mathbf{W}^{1/2}\mathbf{U} \tag{4.1.13}$$

4.1.5 Statistical Inference

We have described regression as a way to study the conditional distribution $y \mid \mathbf{x}$, usually through the mean and variance functions. If we want to do more, we will probably apply statistical inference to address uncertainty in the data. However, where is this uncertainty from? It is from "sampling variability". When the data you have are a subset of a larger, full data set, and if you repeated the study, the data "on hand" (the term "on hand" refers to the specific sub dataset that was sampled in that run) would differ. Thus how does this sampling variability affect the result? That is what statistical inference is about. We usually use the term "population" for the full data set and "samples" for the subset.

We can use statistical inference to judge the model (the mean and variance function) through hypothesis tests such as t-test, F-test, and curvature-test, etc. In these tests, a p-value, which is a standardized quantity, will be used to make the judgment. However, statistical inference is only justified for the data that were actually generated by some random mechanism such as probability sampling, random assignments, or probability-equivalent sampling, etc. For the data that are a subset of a population, but were not selected by a random mechanism, we view them as "convenience samples".

Convenience samples might be biased, and not sufficient to represent the population in probability point of view. Thus statistical inference is not justified for convenience samples.

There is often a debate whether or not the statistical inference is justified. For most environmental data such as our highway experimental data, most statisticians think that they are convenience samples, and thus statistical inference is not justified. In this case, we restrict using regression to only describe the data "on hand".

4.1.6 Conclusions

Many modern regression techniques can help us to establish and evaluate the regression model, such as graphical regression and diagnostic plots. We will apply some of them in this study. Next, we do not intend to use the statistical inference in this study due to the convenience samples of our data set. In addition, there is always a tradeoff between model's fitting and clarity; in another words, simple models show a clearer relationship between response and predictors than complicated models, but may not fit the data as well as complicated models. Finally, the mean function will receive more research attention than the variance function.

4.2 The Site-Pooled COD Regression

The regression analysis involved in this chapter uses the computer program, *Arc* (Cook and Weisberg, 1999).

4.2.1 Introduction

In this section, we consider relating COD concentrations from grab samples to several potential predictors, including the corresponding cumulative rainfall, the corresponding rainfall intensity, the event's antecedent dry days, and the previous event's precipitation. All three sites' data (441 cases) will be used together in this regression, but some field data are missing in some cases (Appendix B shows this data set). The site's effect will be checked after building the site-pooled model. The abbreviations COD, AccR, RI, AtDry, and AtR will be used to represent these variables. Table 4.2.1 describes these variables.

We have examined the histogram and summary statistics for these variables in the previous chapter, and concluded that each should be logarithmically transformed. Thus the basic terms of predictors and response will appear in a log scale in the regression. In addition, before transforming, some of original variables need to be linearly rescaled in order to obtain better scale in the regression or to avoid singularity in a log scale. Table 4.2.2 describes the transformations for these basic terms and response.

Table 4.2.1 Descriptions of Responses and Potential Predictors

Variable	Description
COD	COD concentrations in mg/l
CumR	Cumulative rainfall corresponding to grab samples in inch
AtDry	Antecedent dry period before monitored events in days
AtR	Previous event's precipitation before monitored events in inch
RI	Rainfall intensity before grab samples in inch/15-mins

Table 4.2.2 Transformations of Responses and Potential Predictors

Original Variable	Transformed Variable
COD	log (COD)
CumR	log (CumRs); CumRs = CumR*100
AtDry	log (AtDry)
AtR	log (AtRs); AtRs = AtR*100
RI	log (RIs); RIs = RI *100 + 1

4.2.2 Building the Initial Model

4.2.2.1 The Scatterplot Matrix

A scatterplot matrix is a 2D array of 2D plots. The diagonal elements are variable names, and each off-diagonal element is a 2D plot describing the corresponding pair of variables. Scatterplot matrices actually tell us about "marginal" relationships between each pair of variables without reference to the other variables, or in another words, integrating over the other variables. We usually put the response variable on the bottom, and thus the last row of scatterplot matrices, which is called marginal response plots, displays the dependency of the response on each potential predictor. Therefore, the marginal response plots provide a visual lower bound for the goodness-of-fit that can be achieved with the full regression. However, without further qualification, the marginal response's dependency of the response is not the same as in the full regression. For example, if one marginal response plot suggests a non-linear trend between the response and some predictor, the trend is not necessarily non-linear in the regression.

Figure 4.2.1 shows the scatterplot matrix for the COD data with all potential predictors in this study. The COD and these predictors have been transformed to a natural log scale. In evaluating the regression model, we first would like to check all of the marginal response plots. The marginal response plot log [COD] versus log [CumRs] suggests that log [COD] linearly decreases with log [CumRs]. This marginal dependence between log [COD] and log [CumRs] is very strong, and CumRs is possibly a major component in the regression. The marginal response plot log [COD] versus log [AtDry] suggests that log [COD] increases with log [AtDry]. Next, the marginal response plot log [COD] versus log [RIs] suggests that log [COD] roughly decreases with log [RIs]. Finally, the marginal response plot log [COD] versus log [AtRs] suggests that there is no marginal dependency between log [COD] and log [AtRs].

Next, we would like to check the dependency between predictors. The only marginal plot showing a dependency is the plot log [CumRs] versus log [RIs], which suggests that

log [CumRs] increases with log [RIs]. This dependency may indicate that heavy rain comes later during the event.

4.2.2.2 The Initial Model

From the previous evaluating steps, we suggest the initial mean function as follows

$$E(y \mid \mathbf{x}) = \eta_0 + \eta_1 x_1 + \eta_2 x_2 + \eta_3 x_3 + \eta_4 x_4 \qquad (4.2.1)$$

where

$y = \log[COD]$

$\mathbf{x} = (x_1, x_2, x_3, x_4) = (\log[CumRs], \log[AtDry], \log[AtRs], \log[RIs])$

Table 4.2.3 shows the brief regression result for the above initial model. The R-squared vale is 0.67, which means that the above predictors can account for 67% of the responses' variation. The coefficients for log [CumRs], log [AtDry], log[AtRs], and log-[RIs] are -0.58, 0.39, -0.16, and -0.10 respectively

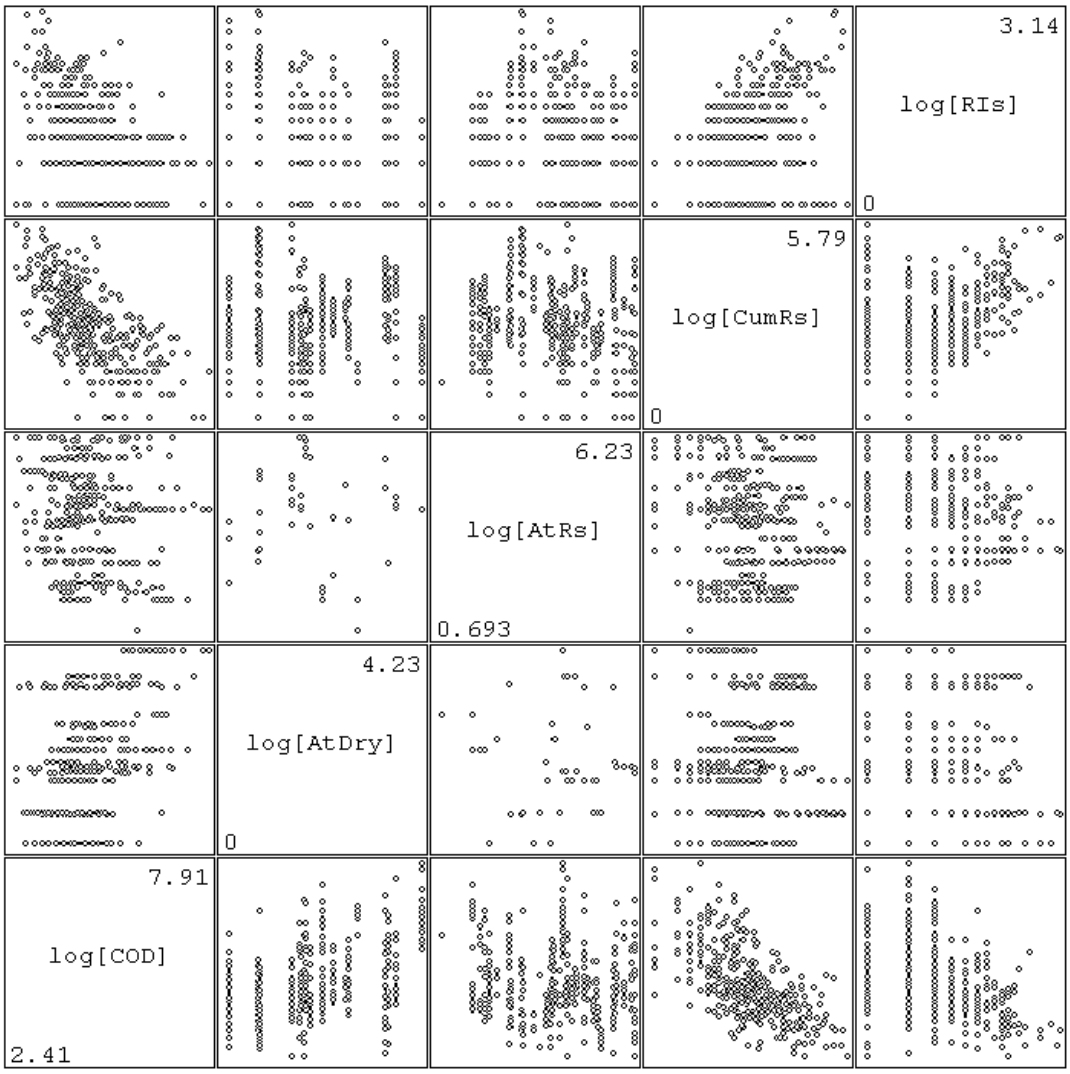


Figure 4.2.1 Scatterplot Matrix of Transformed Response and Potential Predictors

Table 4.2.3 Initial Regression Result for (4.2.1)

```
Data set = wq_pool, Name of Fit = COD Reg0
Kernel mean function = Identity
Response      = log[COD]
Terms         = (log[CumRs] log[AtDry]
                 log[AtRs] log[RIs])
Coefficient Estimates
Label        Estimate        Std. Error
Constant     6.14279         0.128495
log[CumRs]  -0.583173        0.0277896
log[AtDry]   0.390512        0.0259335
log[AtRs]   -0.161859        0.0216586
log[RIs]    -0.0980971       0.0439613
R Squared:                0.669661
Scale factor:             0.590761
Number of cases:              441
Number of cases used:         393
Degrees of freedom:           388
Summary Analysis of Variance Table
Source         df       SS            MS
Regression      4   274.505       68.6263
Residual      388   135.412      0.348999
```

4.2.3 Checking the Model

4.2.3.1 The 2D Added-Variable Plots

2D added-variable plots are scatter plots of the residualized response and term in the mean function. The properties of the plot can be derived as follows. Suppose we have a mean function

$$E(y \mid \mathbf{x}) = \eta^T \mathbf{u} = \eta_1^T \mathbf{u}_1 + \eta_2 u_2 \qquad (4.2.2)$$

where we have divided the k × 1 vector of terms into two pieces, $\mathbf{u}_1$ with k – 1 terms and u_2 with the remaining term. The 2D added-variable plot for u_2 is a plot of $\hat{e}(y \mid \mathbf{u}_1)$ versus $\hat{e}(u_2 \mid \mathbf{u}_1)$, where $\hat{e}(y \mid \mathbf{u}_1)$ and $\hat{e}(u_2 \mid \mathbf{u}_1)$ are the residuals of y and u_2 regressed on respectively. Fitting the simple linear mean function by OLS on the added-variable plot gives

$$\hat{E}(\hat{e}(y \mid \mathbf{u}_1) \mid \hat{e}(u_2 \mid \mathbf{u}_1)) = 0 + \hat{\eta}_2 \hat{e}(u_2 \mid \mathbf{u}_1) \qquad (4.2.3)$$

The estimated coefficient $\hat{\eta}_2$ will remain be the same in (4.2.3) as in the OLS fit of (4.2.2). In addition, the intercept is zero, and the residuals are also identical.

From (4.2.3), we know that 2D added-variable plots will provide visual information on the numerical calculation of the coefficient of a term. We can use them to assess the net effect of a predictor or term in model fitting. In practice, if points scatter about a horizontal line on an added variable plot, it indicates that adding this term to the model does not really help the fit. In addition, there is one condition we need to be aware. If points gather in a tiny range of $\hat{e}(u_2 \mid \mathbf{u}_1)$ compared to u_2, it means that u_2 is strongly collinear with other terms. In this case, the regression information from u_2 is already available from others, and thus u_2 works like a redundant term in the model.

Figures 4.2.2 (a) to (d) show the 2D added-variable plots for each term except the intercept in the initial mean function (4.2.1). We used a smooth function, *lowess* (Cook and Weisberg, 1999) with a smoothing parameter at 0.7, to help visualizing the trends. On the plot 4.2.2 (a), the smoothing line shows a large slope, and indicating that log [CumRs] is a useful additional term. The smoothing line also looks nearly straight.

Next, on the plot 4.2.2 (b), the smoothing line also shows a large slope, indicating the large net effect for the additional term log [AtDry]. The smoothing line, however, shows some curvature. Next, on the plot 4.2.2 (c), the smoothing line shows a small slope, and thus the net effect for the additional term log [AtRs] may not be large. Finally, on the plot 4.2.2 (d), the smoothing line is nearly flat, and indicates that the additional term log [RIs] may not be useful.

Table 4.2.4 shows the regression result if we delete the term log [RIs]. The R Squared and the coefficients for the remaining terms only changed a little: the R Squared is still the same as 0.67, the intercept changed from 6.14 to 6.08, the log [CumRs] coefficient changed from -0.58 to -0.60, the log [AtDry] coefficient changed from 0.38 to 0.40, and the log [AtR] coefficient is still the same as -0.16. Thus, the additional term log [RI] is not useful in the initial mean function (4.2.1).

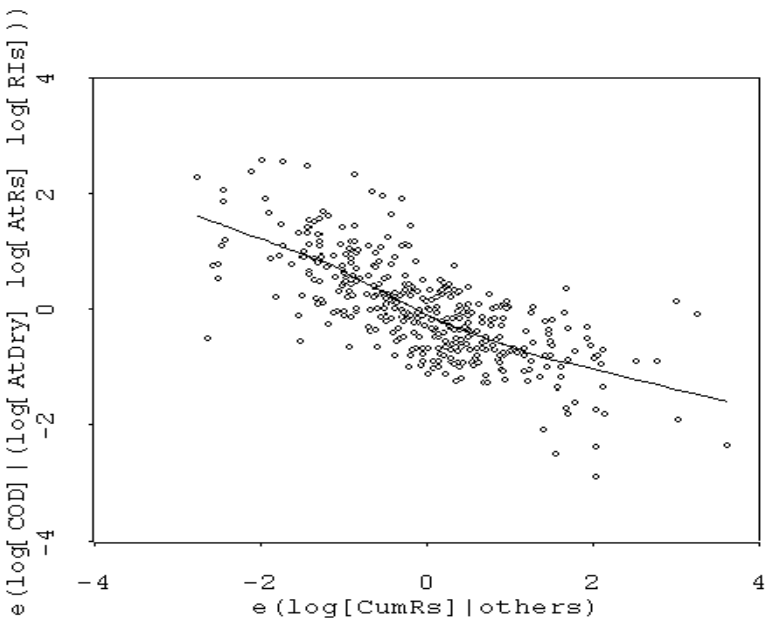
Table 4.2.5 shows the regression result if we continue to delete the term log [AtRs]. This time, the R Squared and the intercept changed more but the coefficients for the remaining terms only change slightly: The R Squared changed from 0.67 to 0.62; the intercept changed from 6.08 to 4.43; the log [CumRs] coefficient changed from -0.60 to 0.59, and the log [AtDry] coefficient changed from 0.40 to 0.38. We will still keep the log [AtR] in the initial mean function (4.2.1) due to the 8% decrease in fitting and the potentially physical meaning.

Table 4.2.4 Regression Result after Deleting log [RIs] in (4.2.1)

Data set = wq_pool, Name of Fit = COD Reg1		
Kernel mean function = Identity Response = log[COD] Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	6.08433	0.126437
log[CumRs]	-0.604614	0.0262086
log[AtDry]	0.402204	0.0255283
log[AtRs]	-0.164255	0.0217423
R Squared:	0.665422	
Scale factor:	0.593775	
Number of cases:	441	
Number of cases used:	393	
Degrees of freedom:	389	

Table 4.2.5 Regression Result after Deleting log [RIs] and log [AtRs] in (4.2.1)

Data set = wq_pool, Name of Fit = COD Reg2		
Kernel mean function = Identity Response = log[COD] Terms = (log[CumRs] log[AtDry])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	5.43378	0.0990119
log[CumRs]	-0.586660	0.0279139
log[AtDry]	0.383873	0.0271782
R Squared:	0.616333	
Scale factor:	0.635028	
Number of cases:	441	
Number of cases used:	393	
Degrees of freedom:	390	



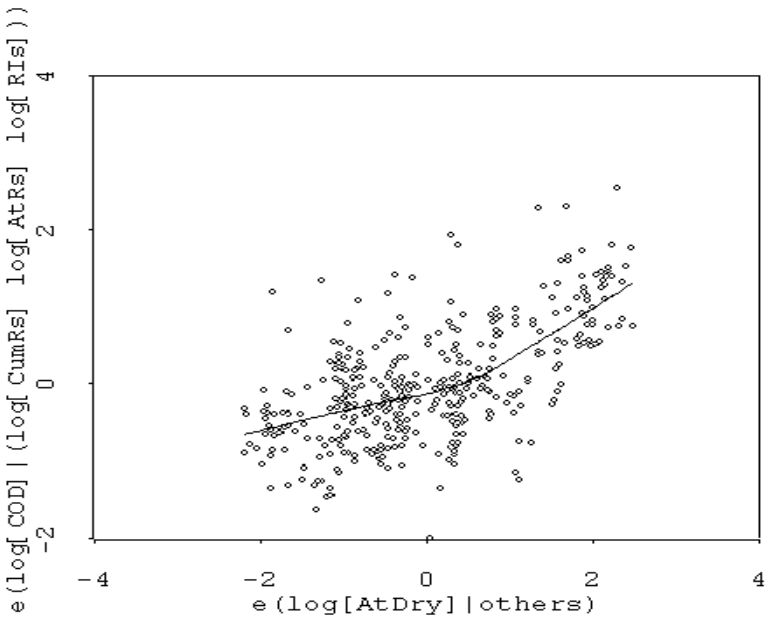
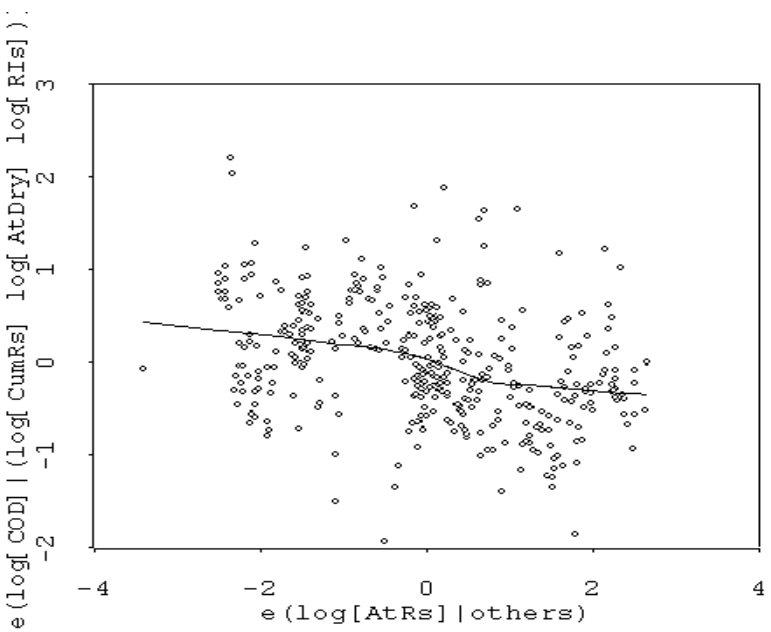


Figure 4.2.2(a) & (b) Added-Variable Plots for log [CumRs] and log [AtDry] in (4.2.1)



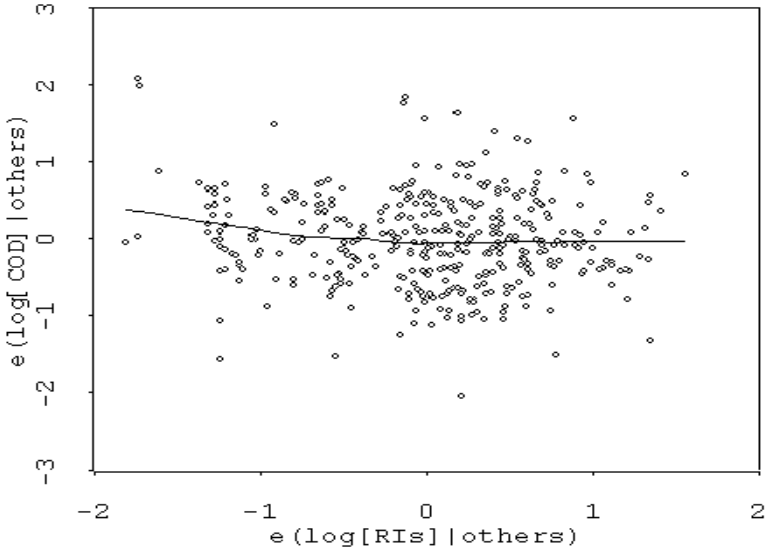


Figure 4.2.2(c) & (d) Added-Variable Plots for log [AtRs] and log [RIs] in (4.2.1)

4.2.3.2 The Model-Checking Plots

A common way to expose lack of fit in a regression is to observe residuals plots against predictors, terms, or any linear combination of predictors or terms. Since residuals should be independent of all the above variables, these residuals should have no pattern, such as curvature and have a zero mean. One way of checking the regression without using statistical inference is to use a smoothing function, which constructs the estimate of the conditional mean of residuals against any above horizontal variable in residual plots. If the curve produced by the smoothing function is close to a straight horizontal line located at zero, the mean function is validated.

Nevertheless, residual plots are able to indicate that the mean function may not be appropriate for the data. The residuals may sometimes be small relative to the range of response, and most of the variation in response can be explained by "incorrect" mean function. These incorrect mean functions judged by residual plots still provide useful information. In this case, the systematic pattern presented in residuals can be ignored. Therefore, in this case, model-checking plots, which are related to residual plots, provide additional information to let us determine the usefulness of the mean function.

Model-checking plots compare the estimated mean function computed from the model to the estimated mean function computed from a smoothing function. With response y on the vertical axis, the quantity h on the horizontal axis can be any function of the original predictors $\mathbf{x}$. The mathematical derivation is as follows

$$E(\mathrm{y} \mid h) = E[E(y \mid \mathbf{u}) \mid h)] \qquad (4.2.4)$$

$$= E[\eta^T \mathbf{u} \mid h)]$$

where **u** means the terms. If we substitute $\hat{y} = \hat{\eta}^T \mathbf{u}$ for $\eta^T \mathbf{u}$, then for any *h*, we can estimate $E(\mathrm{y} \mid h)$ from the model that produced $\hat{y}$ by smoothing the scatterplot of $\hat{y}$ versus *h*. If the model is correct, then the smoothed of y versus *h* and $\hat{y}$ versus *h* should agree with each other. In principle, we are required to examine model-checking plots for a variety of values of *h*.

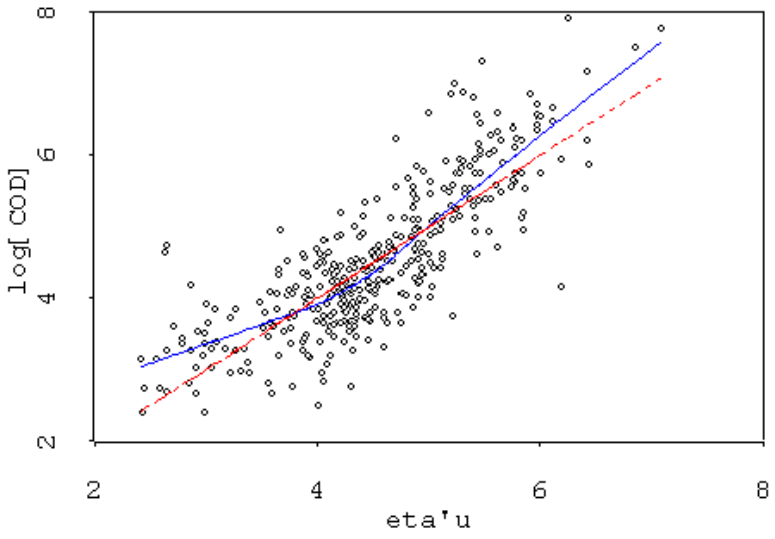
In practice, we usually check the model checking plots for the OLS fit and each individual term. There is no standardized quantification for the discrepancy between the data and the model on the plots. Thus, the judgment is made only through visualization. Additionally, the variance function can also be checked on model-checking plots. Similarly, we can compare the interval computed by the model to the interval computed by the smoothing function.

Figures 4.2.3(a) to 4.2.3(d) show the model-checking plots for the mean function. On plot 4.2.3(a), the data line is overall close to the model line under the horizontal variable as the OLS fitted values (shown as the symbol, eta'u, on the plot). The data line shows a little curvature on the two ends. Next, on the plot 4.2.3(b), the data line is very close to the model line under the horizontal variable as log [CumRs]. There is slight discrepancy between these two lines on the right end. Next, on the plot 4.2.3(c), the data line is very close to the model line under the horizontal variable as log [AtDry]. No obvious discrepancy was observed. Finally, on the plot 4.2.3(d), the data line is very close to the

model line under the horizontal variable as log [AtRs]. No obvious discrepancy was observed too.

Figures 4.2.4(a) to 4.2.4(d) show the model checking plots for checking the variance function. The data intervals are quite close to the model intervals on all of these four plots although some slight discrepancy happened somewhere. The variance function is a constant in this regression.

We can add some high-order polynomial terms in the mean function to correct the slight discrepancy as observed in Figure 4.2.3(a). This decision was not made due to the consideration of balancing the model's fit and complexity. Overall, we are satisfied with the current model.



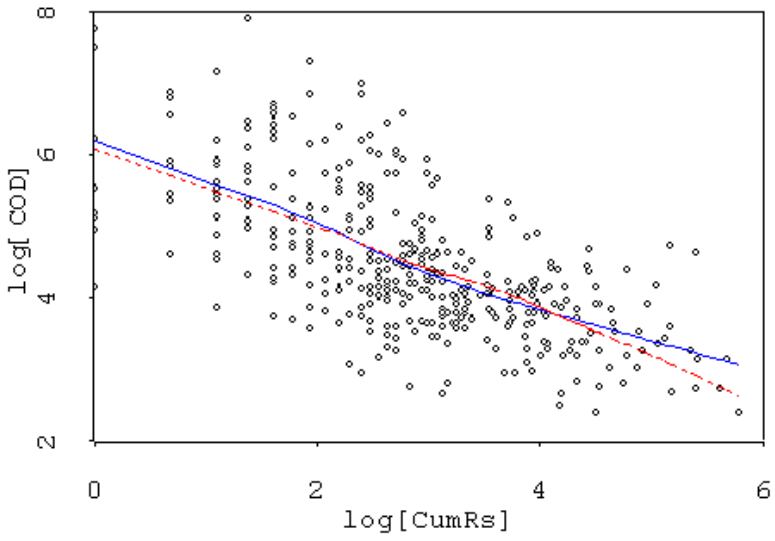
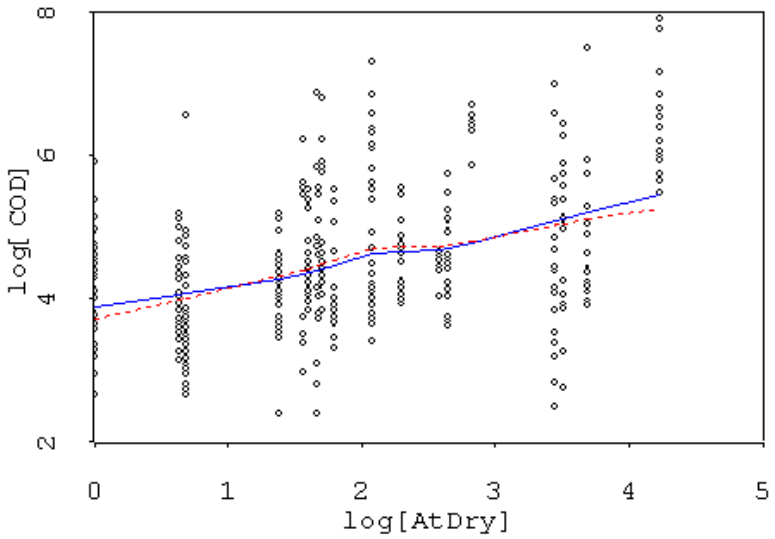


Figure 4.2.3(a) & (b) Model-Checking Plots for Checking OLS Fit (eta'u) and log [CumRs] in the Mean Function (the solid-blue line is for the data; the dashed-red line is for the model)



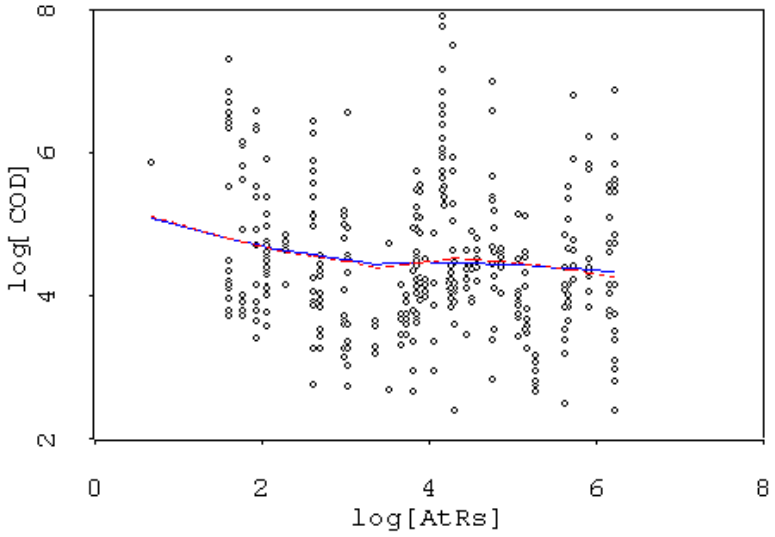


Figure 4.2.3(c) & (d) Model-Checking Plots for Checking log [AtDry] and log [AtRs] in the Mean Function (the solid-blue line is for the data; the dashed-red line is for the model)

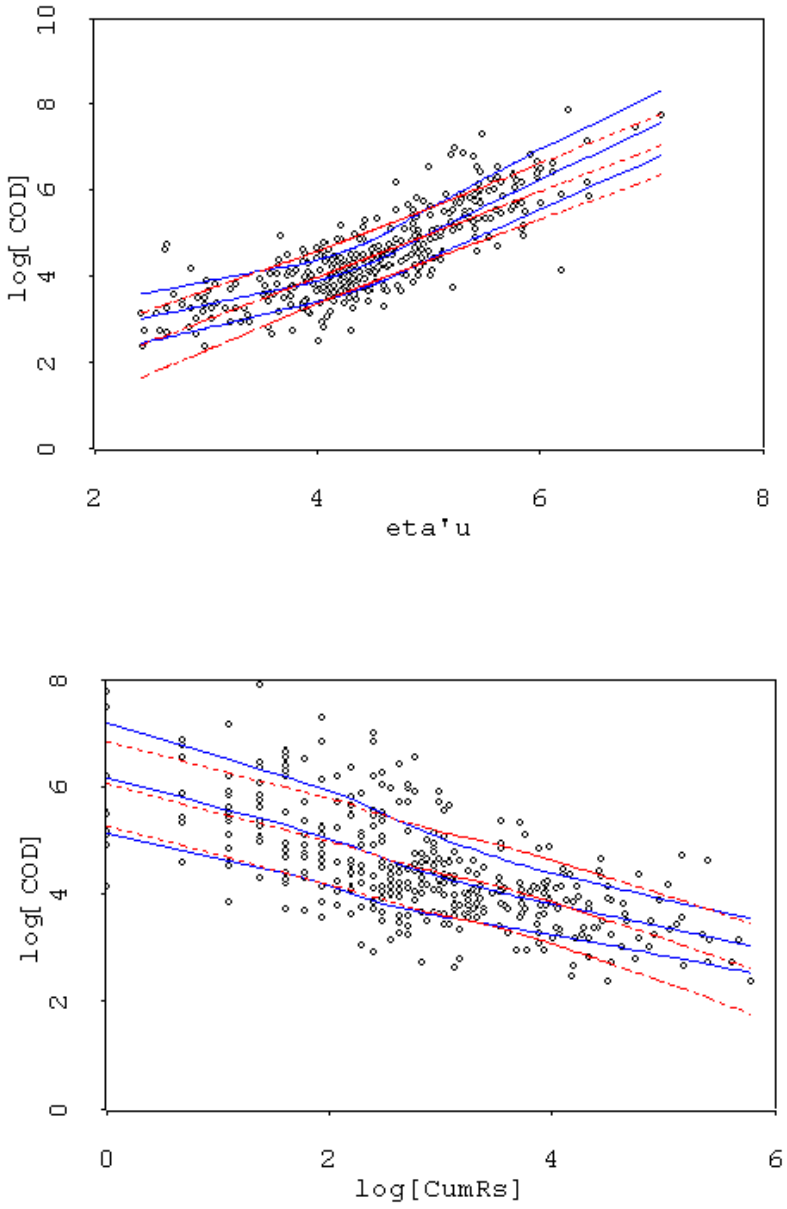
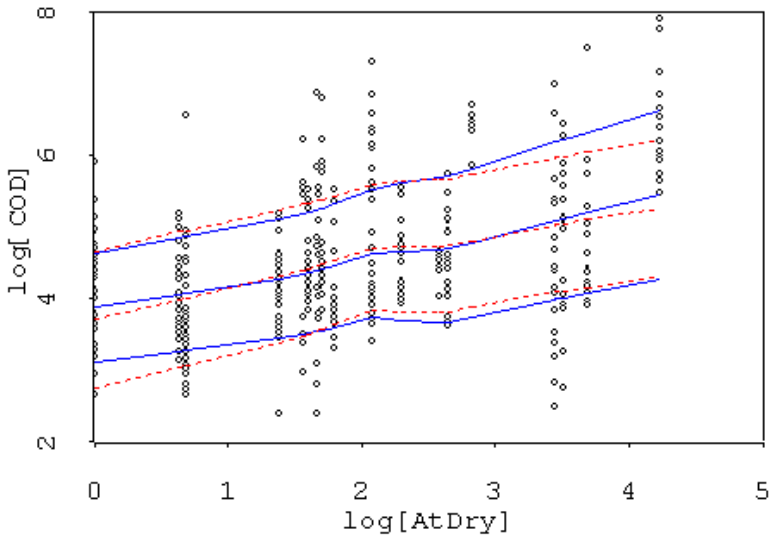


Figure 4.2.4(a) & (b) Model-Checking Plots for Checking OLS Fit (eta'u) and log [CumRs] in the Variance Function (the solid-blue line is for the data; the dashed-red line is for the model)



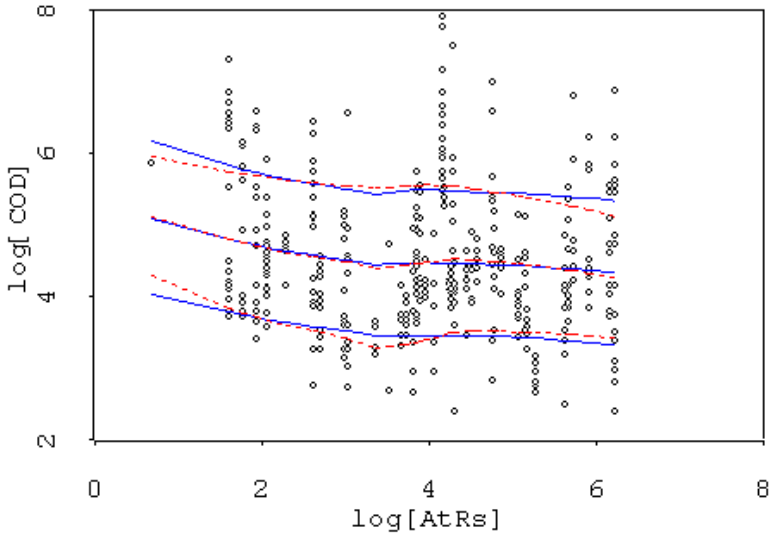


Figure 4.2.4(c) & (d) Model-Checking Plots for Checking log [AtDry] and log [AtRs] in the Variance Function (the solid blue line is for the data; the dashed-red line is for the model)

4.2.3.4 The Cook's Distance Plots

In a regression, some cases are considered to have a large influence on the results of the analysis if deleting them from the data set produces conclusions quite unlike those based on the full dataset. These influential cases are often far away from the rest of data. The explanation of these influential cases is a mistake, or a rare observations, or model shortcoming. Thus, we always would like to find the influential cases, and understand them.

Cook's distance can be used to summarize essential information about the influence of each case on the estimated regression coefficients. *Cook's* distance is a mathematical measurement for the impact of deleting a case. The measurement for accessing the influence of the i^{th} case is based on the vector difference among the coefficients estimates, $\hat{\eta}$ and $\hat{\eta}_{(i)}$. $\hat{\eta}$ is estimated from all of the data, and $\hat{\eta}_{(i)}$ is estimated from the data but without i^{th} case. *Cook's* distance is summarized by D_i, which is proportional to the squared distance between $\hat{\eta}$ and $\hat{\eta}_{(i)}$. If D_i is sufficiently large, then the case is influential for $\hat{\eta}$. For a linear model with mean function $E(y \mid \mathbf{x}) = \eta^T \mathbf{u}$, the D_i can be written as

$$D_i = \frac{(\hat{\eta}_{(i)} - \hat{\eta})^T (\mathbf{U}^T \mathbf{U})(\hat{\eta}_{(i)} - \hat{\eta})}{k\hat{\sigma}^2} \tag{4.2.5}$$

In (4.2.5), $\mathbf{U}$ denotes the $n \times k$ matrix defined by the terms in the mean function; n is the number of observations; k is the number of terms, and $\hat{\sigma}^2$ is the estimated variance.

Cook's distance plots show D_i values. In practice, we are interested in studying the cases that have $D_i > 0.5$, and always studying the cases with $D_i > 1$.

Figure 4.2.5 shows the *Cook's* distance plot for this regression and in no cases is D_i greater than 0.5. The three highest cases (marked by * on plot) are approximately 0.05. Table 4.2.6 shows the regression result if we delete these three cases. The R Squared increased from 0.67 to 0.70. The intercept increased from 6.08 to 6.17. The log [CumRs] coefficient decreased from -0.60 to -0.64. The log [AtDry] coefficient and the log [AtR]'s coefficient are approximately the same, from 0.40 to 0.41 and from -0.16 to -0.17 respectively.

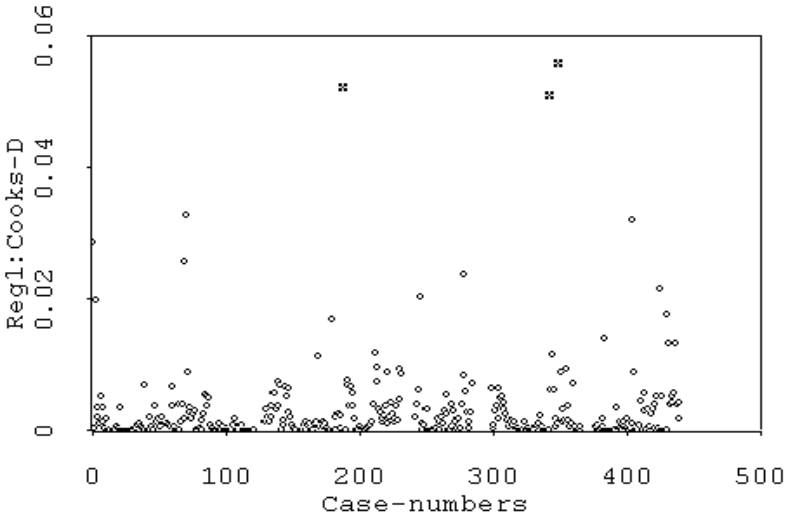


Figure 4.2.5 *Cook's* Distance Plot (three highest cases marked by *)

Table 4.2.6 Regression Result after Deleting Three Highest D_i Cases

Data set = wq_pool, Name of Fit = COD Reg1 Deleted cases are (S2-1999-09-C1 S3-1999-08-C2 S3-1999-10-01)		
Kernel mean function = Identity Response = log[COD] Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	6.17190	0.121658
log[CumRs]	-0.637184	0.0255451
log[AtDry]	0.409966	0.0244258
log[AtRs]	-0.168405	0.0207767
R Squared:	0.697159	
Scale factor:	0.56697	
Number of cases:	441	
Number of cases used:	390	
Degrees of freedom:	386	

4.2.4 Conclusions

One of the primary advantages of fitting a parametric mean function is that the regression can then be characterized by only a few numbers such as the estimates of the regression coefficients and $\hat{\sigma}$. Consequently, interpreting the regression coefficients is a necessary step to every regression analysis. Because measurements generally have units attached to them, and sometimes the mean function is very complicated, it will be difficult for us to quantify the effect on the response just from the values of the coefficients. In practice, the simple description of a regression coefficient is its net effect on response while keeping other terms fixed.

The result of the mean function for this site-pooled COD regression is

$$E(\log COD \mid \mathbf{x}) = 6.08 - 0.60 \log CumRs + 0.40 \log AtDry - 0.16 \log AtRs \qquad (4.2.6)$$

The intercept of the regression line is 6.08, and anchors the regression. Secondly, the term, log [CumRs]'s coefficient is -0.60. It means that COD concentration will decrease 0.6 % per 1% up of cumulative rainfall. Next, the term, log [AtDry]'s coefficient is 0.40. It means that COD concentration will increase 0.4 % per 1% up of antecedent dry period. Finally, the term, log [AtRs]'s coefficient is -0.16. It means that COD concentration will decrease 0.16 % per 1% up of previous rainfall.

In order to understand more about the above interpretation, we use the following examples to illustrate each term's effect on COD concentrations. In example one, one event is given with 7 days of dry period and 0.5 inch of previous rainfall. At 0.1-inch rainfall depth, the COD concentration is approximately 128 mg/l. At 0.3-inch rainfall

depth, the COD concentration is approximately 66 mg/l. The decrease is approximately 48 % compared to at 0.1-inch level. At 1.0-inch rainfall level, the COD concentration is approximately 32 mg/l. The decrease is approximately 75 % compared to at 0.1-inch depth. In fact, this relative ratio is fixed among rainfall levels for any event with given dry period and previous rainfall. Figure 4.2.6 shows this decreasing trend for the range of rainfall depth from 0.1 to 3 inches.

In example two, one event is given with 0.5 inch of previous rainfall and three dry period conditions, 7, 20, and 40 days respectively. At 0.1-inch rainfall depth, the COD concentration is approximately 128 mg/l for 7 days of dry period, 195 mg/l for 20 days of dry period, and 257 mg/l for 40 days of dry period. The relative concentration ratio among 7, 14, and 40 days is 1:1.5:2. At 0.5-inch rainfall depth, the COD concentration is approximately 49 mg/l, 74 mg/, and 98 mg/l for 7 days, 20 days, and 40 days of dry period respectively. The concentration ratio among 7, 14, and 40 days is still the same. In fact, this ratio is fixed too at any rainfall level for any event with given previous rainfall. Figure 4.2.7 shows this relative trend for the range of dry period from 3 days to 50 days.

Similarly to example 2 with given, different previous rainfall, the concentration ratio is also fixed at any rainfall level for any event with given dry period. Figure 4.2.8 shows this relative trend for the range of previous rainfall from 0.1 to 3 inches.

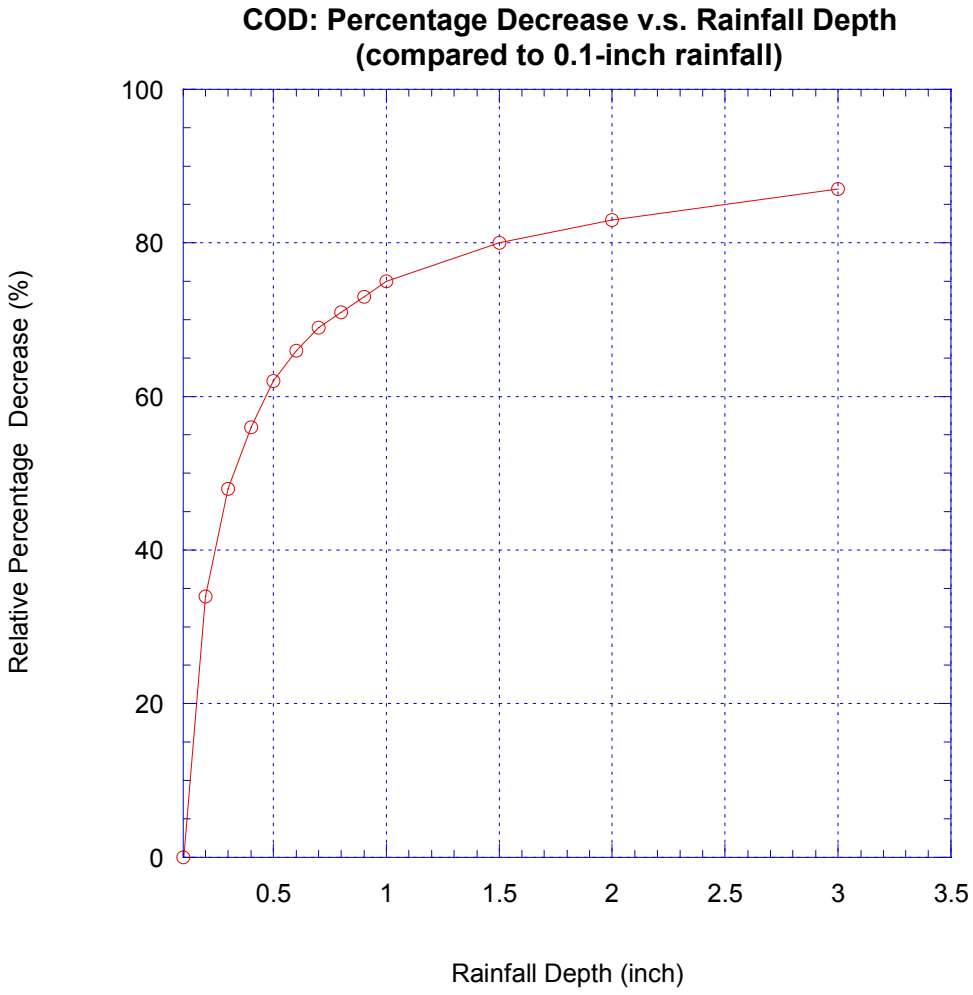


Figure 4.2.6 Relative Percentage Decrease in COD Concentration versus Cumulative Rainfall with Fixed Antecedent Dry Days and Previous Rainfall (reference level: 0.1-inch cumulative rainfall)

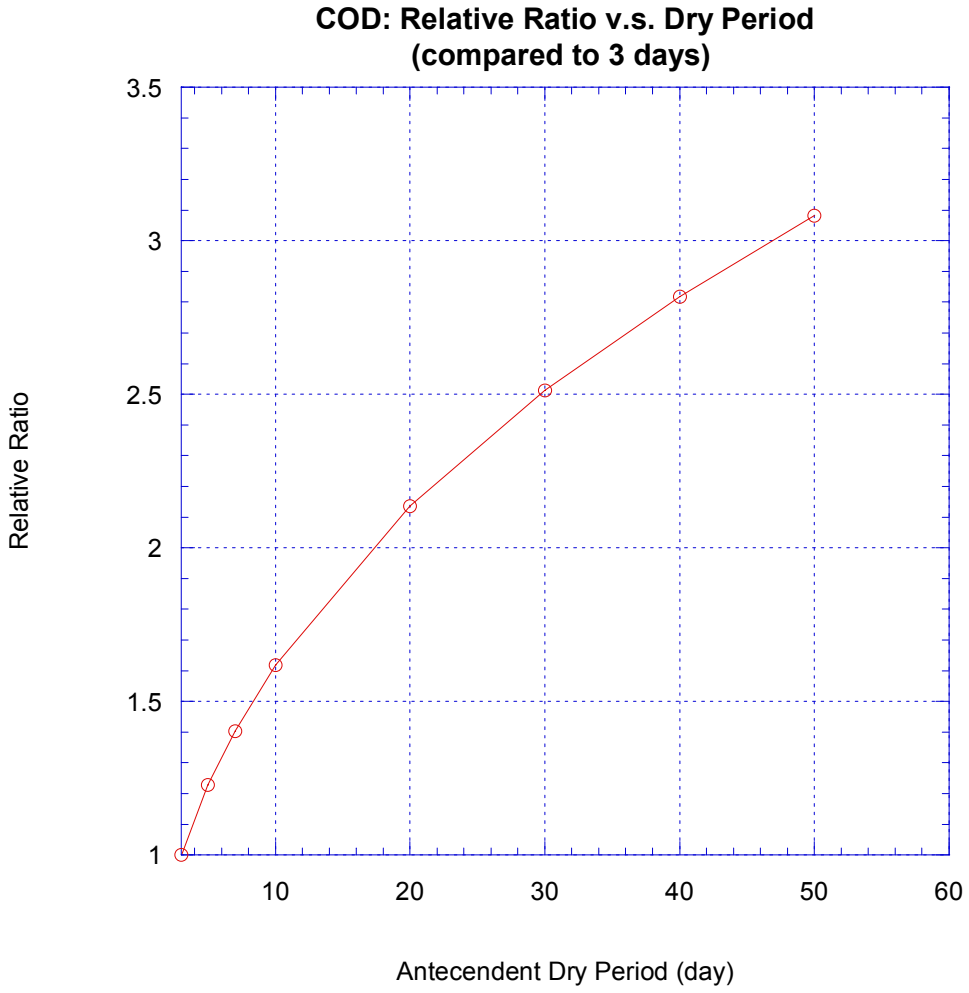


Figure 4.2.7 Relative Ratio in COD Concentration versus Antecedent Dry Period with Fixed Rainfall Depth and Previous Event Precipitation (reference level: three days of antecedent dry period)

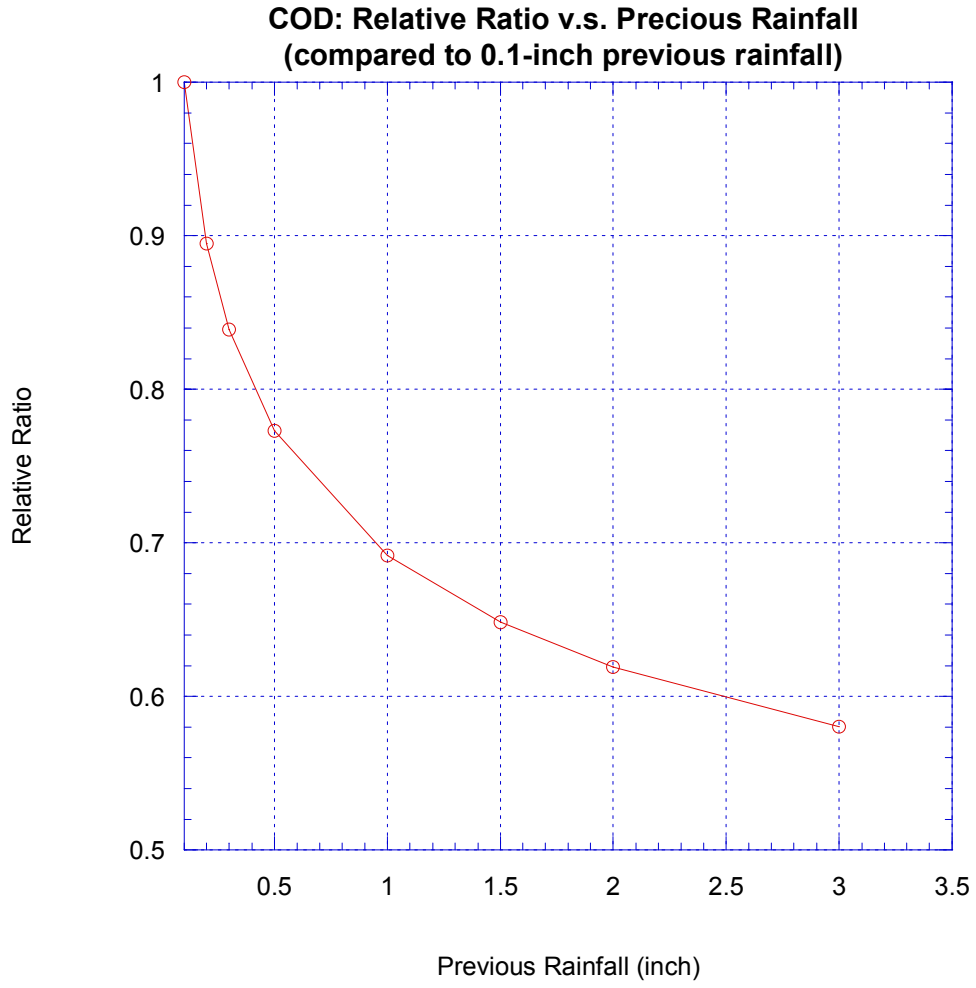


Figure 4.2.8 Relative Ratio in COD Concentration versus Previous Event Precipitation with Fixed Rainfall Depth and Antecedent Dry Period (reference level: 0.1-inch of previous event precipitation)

4.3 The Site Effect on the COD Regression

Continuing with the COD regression, we now consider the site effect on the mean function. In order to distinguish the data from different sites, we have to use a categorical predictor that represents the site. Usually a categorical predictor with *k* levels requires (*k*--1) indicator variables. Each indicator variable is assigned the value one or zero. And for any case, no more than one of these indicator variables is assigned to one, and all the rest are equal to zero. If they are all zeros, then this categorical predictor belongs to its base level. In the COD regression, the site categorical predictor (using symbol *S*) has 3 levels. If we set Site 1 as the base level, thus we need two indicator variables to represent Site 2 and 3.

When considering categorical predictor *S* in the COD regression, we have to think of how *S* can affect the mean function. We will illustrate the role of a categorical predictor from a simple case: a simple linear function with an intercept, one continuous predictor *X*, and one categorical predictor *S* with two levels (1 or 2). Suppose that for each fixed level of *S*, the mean function can be expressed as

$$E(Y \mid X, S = s) = \eta_{0s} + \eta_{1s} X \tag{4.3.1}$$

The additional subscript *s* on coefficients ηs implies the possibility that the effect of *S* could be reflected in the intercept, the slope, or both. Thus for a simple mean function such as (4.3.1), there are four possibilities about the effect of a categorical predictor. We will distinguish these four cases as below:

1. Unrelated Regression Lines: This is the most general case. The slope and intercept parameters for each level of *S* are totally different. Even the trends can be opposite. From the regression point of view, we would say that there is an interaction between the categorical predictor *S* and the continuous predictor *X*, and each indicator variable should also be included in the mean function. The mean function is as follows

$$E(Y \mid X, S) = \eta_0 + \eta_{02} l_2 + \eta_1 X + \eta_{12} l_2 X \tag{4.3.2}$$

where l_2 is the indicator variable, so $l_2 = 1$ when $S = 2$.

2. Parallel Regression Lines: In this case, the slopes are equal. That is $\eta_{11} = \eta_{12}$, but the intercepts may differ. Thus the effect of the categorical predictor *S* does not depend on *X*. In regression point of view, only each indicator variable should be included in the mean function. The mean function becomes

$$E(Y \mid X, S) = \eta_0 + \eta_{02} l_2 + \eta_1 X \tag{4.3.3}$$

3. Equal Intercept Regression Lines: In this case, the intercepts are equal ($\eta_{01} = \eta_{02}$) but the slopes may differ. Thus, the ordering of the *S* levels are always the same, but the size of differences changes with *X*. From the regression point of view, the interaction between *S* and *X* exists, but indicator variables should not be included in the mean function. The mean function is as follows

$$E(Y \mid X, S) = \eta_0 + \eta_1 X + \eta_{12} l_2 X \tag{4.3.4}$$

4. Coincident Regression Lines: In this case, all the intercepts and slopes are the same. Thus, the categorical predictor *S* has no effect on responses. From the regression point of view, no additional terms about categorical predictor *S* should be included in the mean function. The mean function is still the same

$$E(Y \mid X, S) = E(Y \mid X) = \eta_0 + \eta_1 X \tag{4.3.5}$$

The COD regression becomes more complicated due to *S* with three levels and more than one continuous predictor in the mean function. For the full model (Case 1), we will have two more intercepts and six (2×3) more interaction terms based on the mean function (4.2.8). Our goal was to examine how data from different site respond differently to the mean function. The answer is among the four possibilities. If we achieve a result that COD concentrations have totally opposite trends from different sites corresponding to the same predictor, it might imply that we missed very important information from sites, and that the missing site-related variable is more important than the current predictors for determining the COD concentrations. Therefore, it seems unlikely because we know that the pooled COD regression fits well.

In order to avoid complexity of using *S* in the beginning, we would like to use the dimension reduction technique that let the fitted values from the pooled COD regression become the only continuous predictor in the COD regression. This technique is as follows

$$Fit - Values = 6.08 + 0.60 \log CumRs + 0.40 \log AtDry - 0.16 \log AtRs \tag{4.3.6}$$

Thus the case is now simplified into one categorical predictor and one continuous predictor. We can then get a visual comparison for different levels of *S* in this simplified case by using *S* as a marking variable and plotting log [COD] versus *Fit-Values*. Figure 4.3.1 shows the 2D plot from different sites with different symbols and colors to indicate the marking variable *S*. The points from different sites are distributed evenly around the regression line.

Next we try to fit the simplified model with those different possibilities, by fitting the restricted model with the equal intercept but different slopes. Figure 4.3.2 shows the result. These three regression lines look close to each other although Site 2 has the highest slope and Site 3 has the lowest. Next, we fit the restricted model with the equal slope but different intercepts. Figure 4.3.3 shows the fitting result. These three regression lines also look close to each other, and Site 2 is above Site 1 and Site 1 is above Site 3. Finally, we fit the most general case with different slopes and different intercepts. Figure 4.3.4 shows the fitting result. These three regression lines also look close to each other and similar to the above two cases.

We conclude that there is no site effect on the COD regression. Even we think that it shows minor site effects on the plots (from Figure 4.3.2 to 4.3.4), the site effects are negligible compared to the other major predictors.

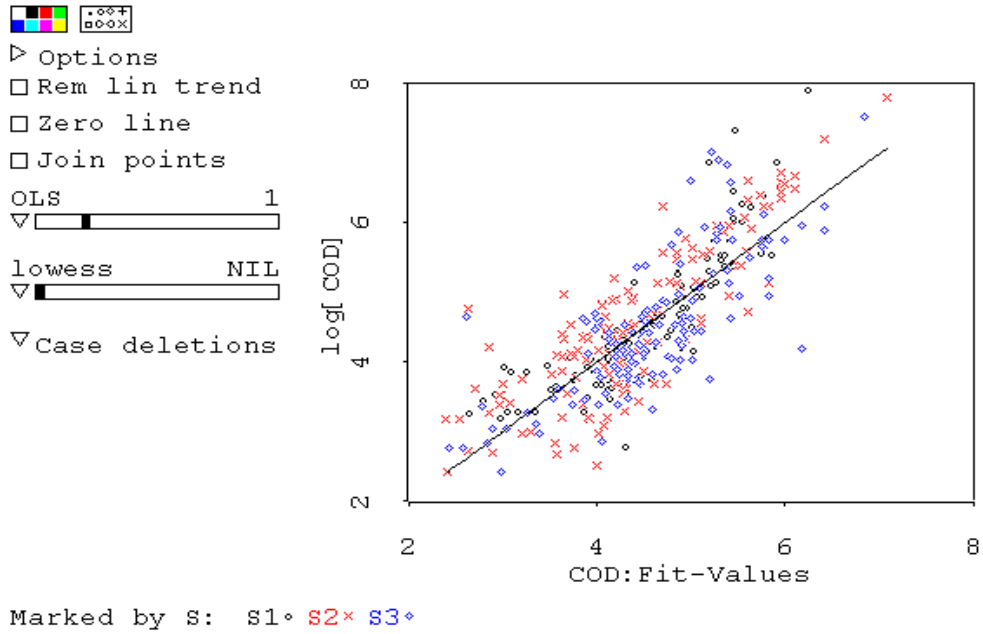


Figure 4.3.1 log [COD] versus Fit-Values Using a Site-Marking Variable S

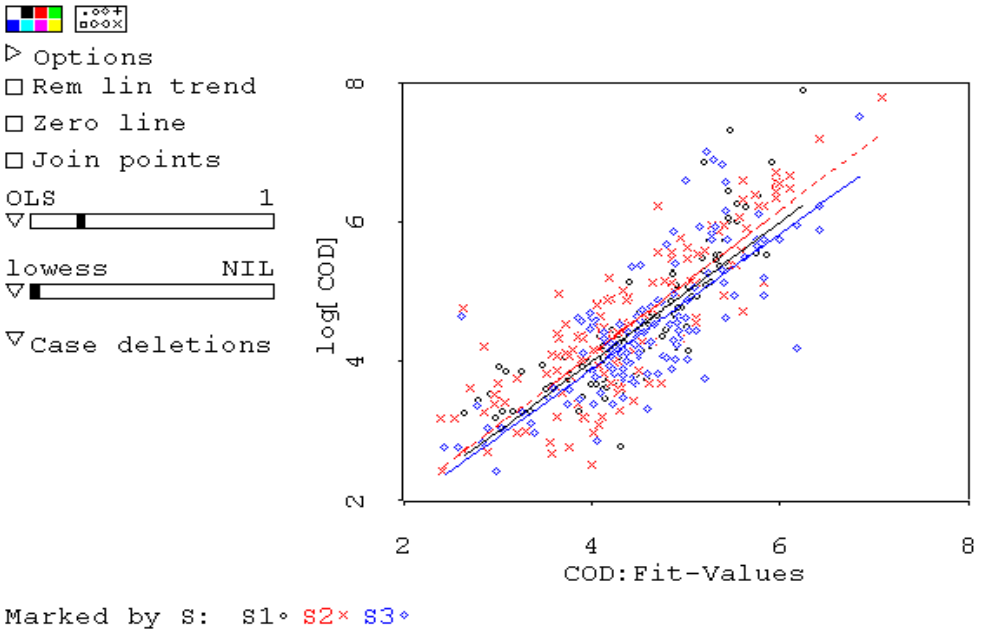


Figure 4.3.2 Fitting Case 3: Equal Intercept but Different Slope Regression Lines

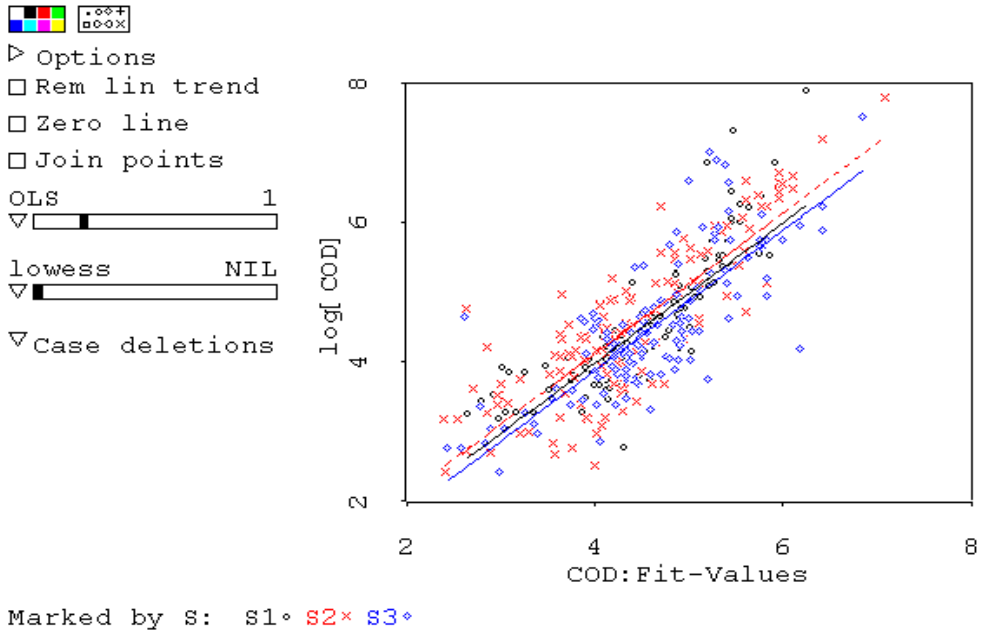


Figure 4.3.3 Fitting Case 2: Parallel Regression Lines

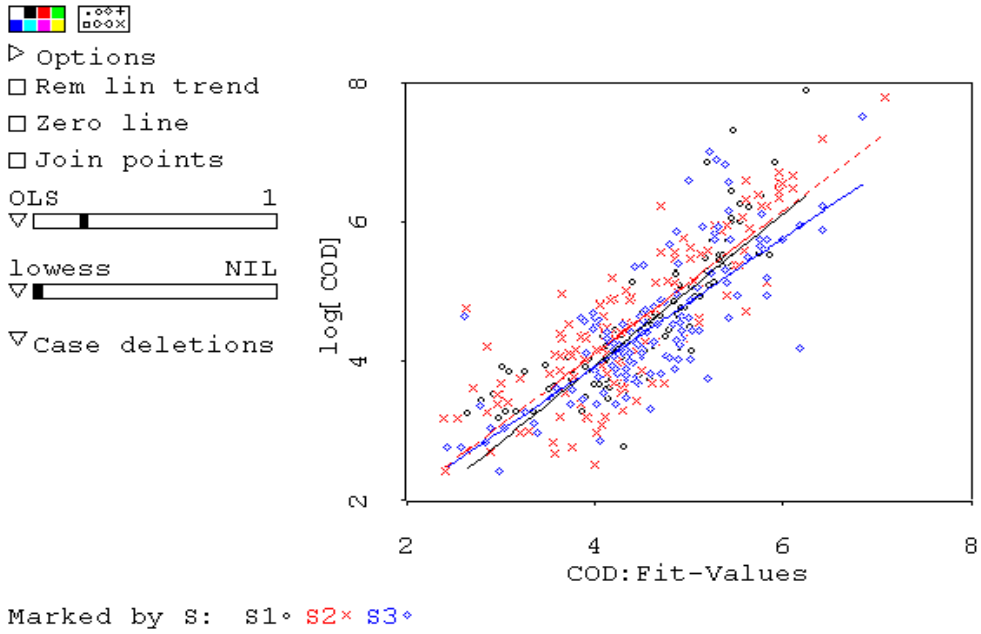


Figure 4.3.4 Fitting Case 1: Different Intercept and Slope Regression Lines

4.4 Regressions for COD-Correlated Parameters

We know that if a parameter is correlated well with COD on the log scale, a similar regression can be achieved using the same model structure:

$$E(\log y \mid \mathbf{x}) = \eta_0 + \eta_1 \log CumRs + \eta_2 \log AtDry + \eta_3 \log AtRs \qquad (4.4.1)$$

Figure 4.4.1 shows the scatterplot matrix for COD and parameters that correlates well with COD using log scale, including Oil & Grease (O&G), Dissolved Organic Carbon (DOC), Total Kjeldahl Nitrogen (TKN), dissolved Phosphorus (P_dis), dissolved Copper (Cu_dis), and dissolved Nickel (Ni_dis). The sample correlation coefficients are 0.91 for O&G, 0.94 for DOC, 0.93 for TKN, 0.85 for P_dis, 0.91 for Cu_dis, and 0.82 for Ni_dis. The model structure (4.4.1) works well for these parameters, showing the R Squared range from 0.65 to 0.77 except Ni_dis as 0.55. No serious discrepancies were found on model-checking plots. Appendix C shows these regression results.

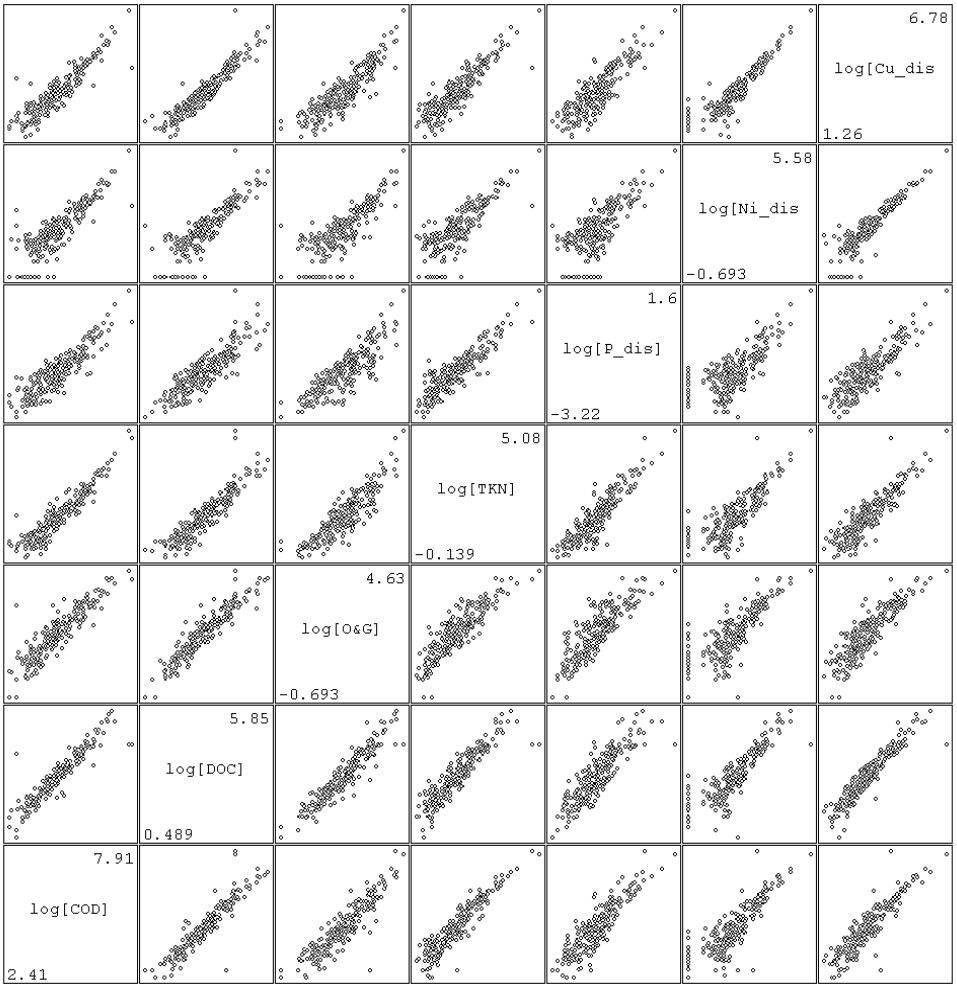


Figure 4.4.1 Scatterplot Matrix of COD and COD-Correlated Parameters

5. EVENT MEAN CONCENTRATIONS

5.1 Introduction

Event mean concentrations (EMCs) have been extensively used in the past to calculate pollutant loads. The EMC, by its name, represents the average concentration of the pollutant throughout the storm event. Some studies show that EMCs might be very different from case to case. Driscoll (1986) used probability to describe the variation of EMCs, and concluded that EMCs vary from one storm to another following a log-normal distribution. James et al. (1999) compared the mean and median of EMCs for 10 parameters from NURP (Nationwide Urban Runoff Program) and the three-pooled data sources (NURP, USGS, NPDES), and concluded that the differences could be large enough to merit re-evaluating selected management strategies in existing control programs. Generally, the land use type in catchments is believed to be the most important factor that causes EMCs to be different (Wong et al, 1997). Although differences of EMC data have been noticed, the differences have never been addressed from the aspect of their estimation.

Practically, an EMC is estimated from either an automated composite sampler or a series of grab samples taken during a storm event. When estimating EMCs from grab samples, each grab sample, physically, represents an instant concentration of pollutants within a storm event, and the EMC is calculated from these instant concentration values. A reasonable calculation method, used by many authors (Charbeneau and Barrett, 1998; Wu et al., 1998, etc) is to use a discharge-weighted average of these instant

concentrations. An automatic sampler collects a large number of individual samples, and is a series of instant concentration samples. The EMC is equal to the result of analyzing the single, large sample. However, in this case, EMCs can still mathematically be viewed as a result from instant concentration measurements.

The goal of this study is to investigate the reliability of EMC data from the aspect of their estimation. A mathematical definition and its related calculation forms will be introduced first. Then, a stochastic approach will be used through theories and computer simulations. In the theory part, the objective is toward general cases, not parameter or site specific. In the computer simulation part, a pre-described concentration model will be used for a particular case, in which the field data were collected from three highway-monitoring sites for two seasons (1999-2001). Therefore, the simulation results could still provide useful information that helps most storm water monitoring programs.

5.2 Methodology

5.2.1 Definition and Calculation Forms of EMCs

Mathematically, EMCs can be defined as total pollutant mass (M) discharged during an event divided by total volume (V) discharge of the storm event.

$$EMC = \frac{M}{V} = \frac{\int C(t)Q(t)dt}{\int Q(t)dt} \tag{5.2.1}$$

In (5.2.1), $C(t)$ is a smooth real-valued function of time that represents the pollutant concentration curve, and $Q(t)$ is also a smooth real-valued function of time that represents the stormwater discharge flow rate curve. However, in practice, we estimate the integrals in (5.2.1) not by the functions of $Q(t)$ and $C(t)$ but by the measurements of $Q(t)$ and $C(t)$. We estimate the EMC from discrete values. If we assume we measure the concentration and the discharge rate based on equal time-interval in a storm event, the EMC can be estimated as

$$EMC = \frac{\sum_i c_i q_i}{\sum_i q_i} \tag{5.2.2}$$

where q_i and c_i are the measurements for the discharge rate and pollutant concentration in the i^{th} interval. From the point of view of approximating the continuous functions in (5.2.1), the more measurements we take, the more accurate approximation we can obtain by (5.2.2).

When we view the measurements of the discharge rate as the weights, (5.2.2) becomes the discharge-weighted average throughout the storm event.

$$EMC = \sum_i w_i c_i \tag{5.2.3}$$

$$w_i = \frac{q_i}{\sum_i q_i} \tag{5.2.4}$$

where w_i is the flow weight, and $\sum_{i=1}^{n} W_i = 1$. In practice, one common situation is the number of concentration measurements does not match the number of discharge measurements. Generally there are many fewer concentration measurements. This occurs because concentration measurements are much more expensive and time consuming; discharge measurements can be easily and automatically obtained by the instrument. For most situations we have to adjust the weights for each concentration measurement in (5.2.3). One of the reasonable ways to adjust the weights is to use the discharge volume. One approach (Charbeneau and Barrett, 1998) splits the discharge volume from the mid-point between two consecutive concentration measurements. Figure 5.2.1 shows this approach. The adjusted weight can be written as:

$$w_i = \frac{V_i}{\sum_i V_i} \tag{5.2.5}$$

where V_i is the corresponding discharge volume for the i^{th} concentration measurement. This mid-discharge splitting method can also be applied for measurements at unequal time-interval bases.

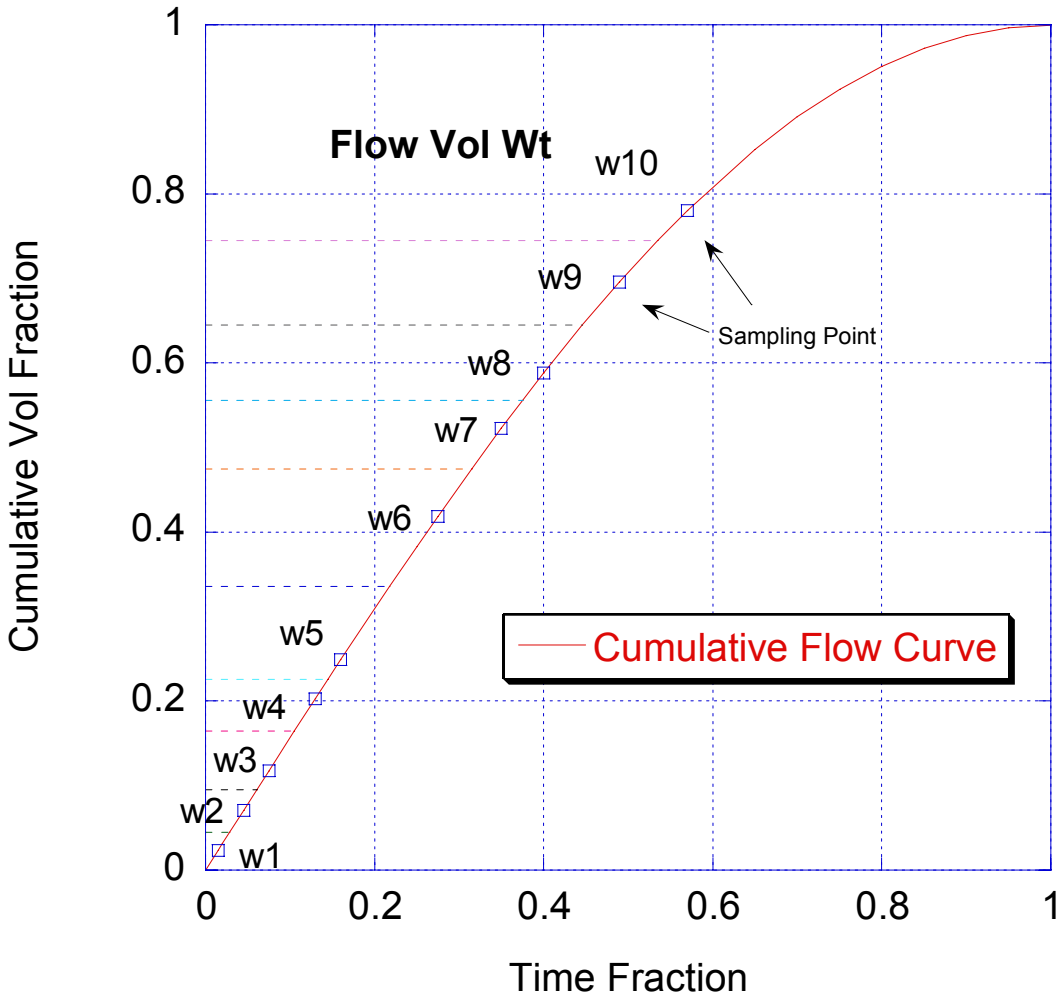


Figure 5.2.1 Determination of Flow Weights (w1 to w10) for Grab Samples

When we take concentration measurements based on constant discharge volume, the weighted average of C_is from (5.2.3) is reduced to the arithmetic average

$$EMC = \overline{C}_n \tag{5.2.6}$$

Ideally, automated samplers can collect samples in proportion to discharge volume. Additionally there are always slight errors (noise) in sample volume and pace that change the equal weights. Thus, EMC is still an inherent weighted average of concentration measurements.

5.2.2 Asymptotic Distributions of EMCs

In this section, we will check the asymptotic distributions of EMCs for several cases under different assumptions. We can view these asymptotic distributions as theoretical treatment for obtaining reliable EMC estimates under these circumstances. In principle, concentration measurement (C_i) will be treated as a stochastic property, usually having an identical independent distribution (*i.i.d.*) with finite mean and variance, or having a trend plus *i.i.d.* errors (with mean zero and finite variance). Weights (W_i) will be considered as constants, except in one case. The theoretical background in this section is the central limit theorem (CLT) and its extension from "Large Sample Theory".

5.2.2.1 Case 1: Sampling from Equal Discharge-Volume

Assume concentration measurements (C_i) are *i.i.d.* with mean μ and finite variance σ^2 during a given storm event. The concentration samples are collected using equal discharge-volume, which assumed perfect (without noise or error). This is the case for (5.2.6). Thus, from (5.2.6), we know

$$EMC_n^* = \bar{C}_n \tag{5.2.7}$$

where EMC_n^* means an estimate of the EMC based on n samples.

By CLT, we obtain

$$\sqrt{n}(\bar{C}_n - \mu) \xrightarrow{L} N(0, \sigma^2) \tag{5.2.8}$$

(5.2.8) means that $\sqrt{n}(\bar{C}_n - \mu)$ will converge (by law) to a normal distribution with mean 0 and variance σ^2.

When C_i has a trend (μ_i) plus *i.i.d.* error (ε) with mean zero and finite variance σ^2, then (5.2.8) becomes

$$\sqrt{n}(\bar{C}_n - \frac{\sum_{i=1}^{n} \mu_i}{n}) \xrightarrow{L} N(0, \sigma^2) \tag{5.2.9}$$

5.2.2.2 Case 2: Sampling from Equal Time-Interval

Assume concentration measurements (C_j) are *i.i.d.* with mean μ and finite variance σ^2 during a given storm event. The concentration samples are collected based on equal

time-interval. It is the case for (5.2.5) that we could adjust the weights based on the mid-discharge volume method, and the EMC estimate is a weighted average of C_js. Since the storm event is given, the weights can be deduced from the number of samples under equal-time sampling. For this case, the Lindeberg-Feller Theorem (Ferguson, 1996), which is a type of generalized CLT in one dimension, can be applied to obtain the asymptotic distribution for EMCs.

For applying the Lindeberg-Feller theorem, we have to use a new random variable, Z_{nj} such that $E(Z_{nj}) = 0$. Let $Z_{nj} = W_{nj}(C_j - u)$, then $E(Z_{nj}) = 0$, and $\text{var}(Z_{nj}) = W_{nj}{}^2\sigma^2$. W_{nj} is determined by the mid-discharge volume method for n samples using equal-time sampling, and $\sum_{j=1}^{n} W_{nj} = 1$. By the Lindeberg-Feller Theorem, let $S_n = \sum_{j=1}^{n} Z_{nj}$ and $B_n^2 = Var(S_n) = \sigma^2 \sum_{j=1}^{n} W_{nj}^2$, then

$$\frac{S_n}{B_n} \xrightarrow{L} N(0,1) \tag{5.2.10}$$

provided the Lindeberg Condition holds. (5.2.10) can be written in the form of a weighted average of C_js,

$$\frac{(\sum_{j=1}^{n} W_{nj} C_j - \mu)}{B_n} \xrightarrow{L} N(0,1) \tag{5.2.11}$$

When C_i has a trend (μ_i) plus *i.i.d.* error (ε) with mean zero and finite variance σ^2, then (5.2.11) becomes

$$\frac{(\sum_{j=1}^{n} W_{nj} C_j - \sum_{j=1}^{n} W_{nj} \mu_j)}{B_n} \xrightarrow{L} N(0,1) \qquad (5.2.12)$$

5.2.2.3 Case 3: Sampling from Equal Discharge-Volume with Weighting Noises

Next we try to evaluate the actual working of an automated sampler, which is operated to collect equal discharge-volume samples, but coupled with weighting noise. The weighting noise may result from collecting individual samples with unequal sample volume or errors in flow pacing signals. Even if we do not have actual experimental data for these weights, we still think that they are probably field-specific.

In order to process this case in spite of the lack of experimental data, we need to make reasonable assumptions about these weighting noises. The weighting noises are assumed to be *i.i.d.* under normal operations, although they are likely to be correlated to form an *m*-dependent sequence of data. In addition, they are also assumed to be independent of concentration measurements. A Beta distribution might be one reasonable guess because the weights will be centered somewhere in the range from 0 to 1.

Assume that concentration measurement (C_j) is *i.i.d.* with mean μ and finite variance σ^2 during a given storm event. The weighting noise (W_i) is also *i.i.d*, and has a beta distribution, *Beta* (a, b). For determining the asymptotic distribution of EMCs, we need to apply the Lindeberg-Fetter Theorem again.

Let $Z_{nj} = W_j(C_j - u)$, such that

$$E(Z_{nj}) = 0, \text{ and } Var(Z_{nj}) = E(Z_{nj})^2 = E(W_j^2(X_j - \mu)^2) = \sigma^2 E(W_j^2)$$

W_i has a *Beta* (a, b) distribution, so that

$$E(W_j) = \frac{a}{a+b} \text{ and } Var(W_j) = \frac{ab}{(a+b)^2(a+b+1)}$$

Thus,

$$E(W_j^2) = Var(W_j) + (E(W_j))^2 = \left(\frac{ab}{(a+b)^2(a+b+1)} + \frac{a^2}{(a+b)^2} \right)$$

Let $S_n = \sum_{j=1}^{n} Z_{nj}$, and let $B_n^2 = \text{var}(S_n)$, so that

$$B_n^2 = \text{var}(S_n) = \sum_{j=1}^{n} \text{var}(Z_{nj}) = n\sigma^2 \left[\frac{ab}{(a+b)^2(a+b+1)} + \frac{a^2}{(a+b)^2} \right]$$

By the Lindeberg-Feller Theorem,

$$\frac{S_n}{B_n} \xrightarrow{L} N(0,1) \tag{5.2.13}$$

provided the Lindeberg Condition holds. Equation (5.2.13) can be written in the form of the weighted average of C_js as

$$\frac{(\sum_{j=1}^{n} W_j C_j - \mu \sum_{j=1}^{n} W_j)}{B_n} \xrightarrow{L} N(0,1) \tag{5.2.14}$$

When C_i has a trend (μ_i) plus *i.i.d.* error (ε) with mean zero and finite variance σ^2, then (5.2.14) becomes

$$\frac{(\sum_{j=1}^{n} W_j C_j - \sum_{j=1}^{n} W_j \mu_j)}{B_n} \xrightarrow{L} N(0,1) \tag{5.2.15}$$

5.2.3 Computer Simulations of EMCs

Computer simulations, compared to theoretical treatments, provide more flexible solutions to understand the EMCs, although the theoretical solutions are more general. For example, we can explore the relationship between known pollutant concentration trends and different flow patterns (which decide the weights). We can also evaluate the performances of different "sampling protocols". For some cases, such as randomly collected samples during a storm event, it is really difficult to obtain the theoretical solution. However, in order to precede the computer simulation, a model must be used as our knowledge to describe the washed-out concentration of pollutant during a storm event.

5.2.3.1 Simulation Descriptions

A previously developed COD regression model will be used in this simulation. In principle, the regression's mean function indicates the concentration trend, and the error term indicates the variability around the trend. The COD regression model is describes as:

$$E(\log COD \mid \mathbf{x}) = 6.08 - 0.60 \log CumRs + 0.40 \log AtDry - 0.16 \log AtRs \qquad (5.2.16)$$

$$\varepsilon_i \sim N(0, 0.59^2) \qquad (5.2.17)$$

Table 5.2.1 describes the above variables, and Figure 5.2.2 shows the model's fitted values vs. the observations.

Table 5.2.1 Description of Variables Used in Regression

Variable	Description
COD	COD concentrations in mg/l
CumRs	Cumulative rainfall corresponding to grab samples in 0.01 inch
AtDry	Antecedent dry period before monitored events in days
AtRs	Previous event's precipitation before monitored events in 0.01 inch

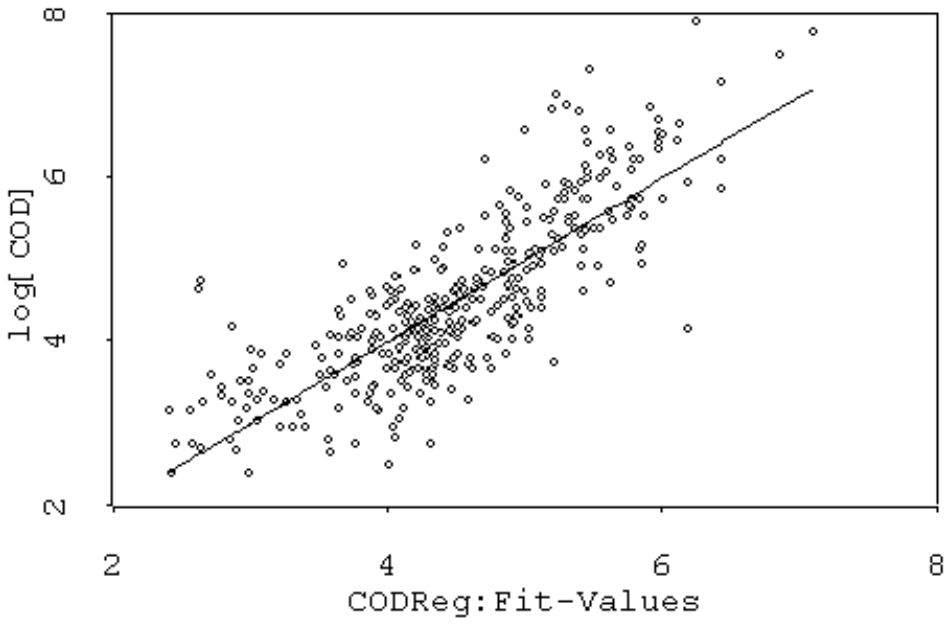


Figure 5.2.2 Regression's Fitted Values vs. Observations

This model uses logarithm scale. When we convert it back to normal scale, the responses will have a lognormal distribution around the regression line. For a lognormal variable X, where log X has a normal distribution with mean μ and variance σ^2, it is not true that $\mu_X = E(X) = e^{\mu}$. The mean and median of X are respectively as

$$\mu_X = e^{\mu+\sigma^2/2} \tag{5.2.18}$$

$$m_X = e^{\mu} \tag{5.2.19}$$

The variance, Var (X), is derived as

$$Var(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \tag{5.2.20}$$

From (5.2.18), we know that the COD mean function actually represents the median concentration response on the normal scale. From the Lindeberg-Feller Theorem, the weighted averages will converge to the weighted mean instead of the weighted median. Thus, the EMC estimates will not converge to $\sum_i w_i e^{y_i}$, and in fact, they will be larger.

There is a special simulation, in which COD concentrations are generated by (5.2.16) and (5.2.17) using one-minute interval. This special simulation will be used as the benchmark in simulation tasks, because one-minute is the actual time interval in our rainfall and flow data. The EMC is then calculated using (5.2.2), where the weights are the discharge rates. Under this one-minute simulation, each EMC result, symbolized as EMC^o, is not affected by the sampling strategy. After multiple simulations, EMC^os will

form a normal distribution, according to the Lindeberg-Feller Theorem. From (5.2.18), the population mean of EMC^o, symbolized as μ^o, is expressed as

$$\mu^o = e^{\sigma^2/2} \cdot \sum_i w_i e^{\log c_i} \tag{5.2.21}$$

In (5.2.21), w_i is the weight, and $w_i = q_i \Big/ \sum_i q_i$ (q_i is the discharge rate); log c_i is the generated COD response, and σ^2 is the error variance

In order to illustrate this one-minute simulation, one real event (dated on 01/25/99, at Site 1) is used for demonstration. Figure 5.2.3 shows the original and the smoothed event hydrograph. The smoothed hydrograph will be used in simulation to correct fluctuations in original data. Figure 5.2.4 shows the histogram of EMC^{o}s after 1000 runs. The sample mean is 116.36 (mg/l). The population mean (μ^o), calculated by (5.2.21), is 116.25 (mg/l). The population median ($=\sum_i w_i e^{\log c_i}$) is 97.69 (mg/l). μ^o will be used as the true value of EMC in later error evaluation.

In contrast to the one-minute simulation, we conduct other simulations by taking fewer samples during a storm event. Each type of simulation will generate a distribution of EMCs after multiple runs. The EMC calculated under simulations with larger intervals is called EMC*. The more samples we collect, the closer the distribution of EMC*s to EMC^o's. Figure 5.2.5 illustrates EMC*s by showing the histograms from different numbers of samples (n = 10, 20, 40, and 60) for the same tested event, where samples are collected using equal time sampling.

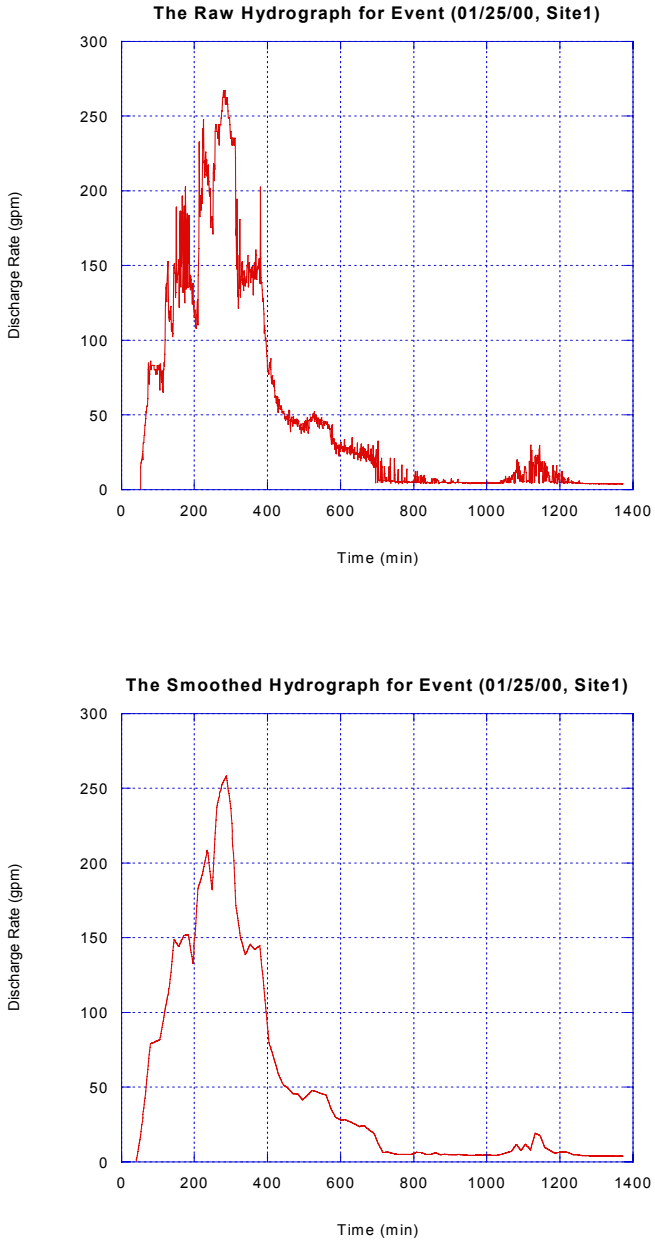


Figure 5.2.3 Original and Smoothed Hydrographs (event recorded on 01/25/00, Site 1)

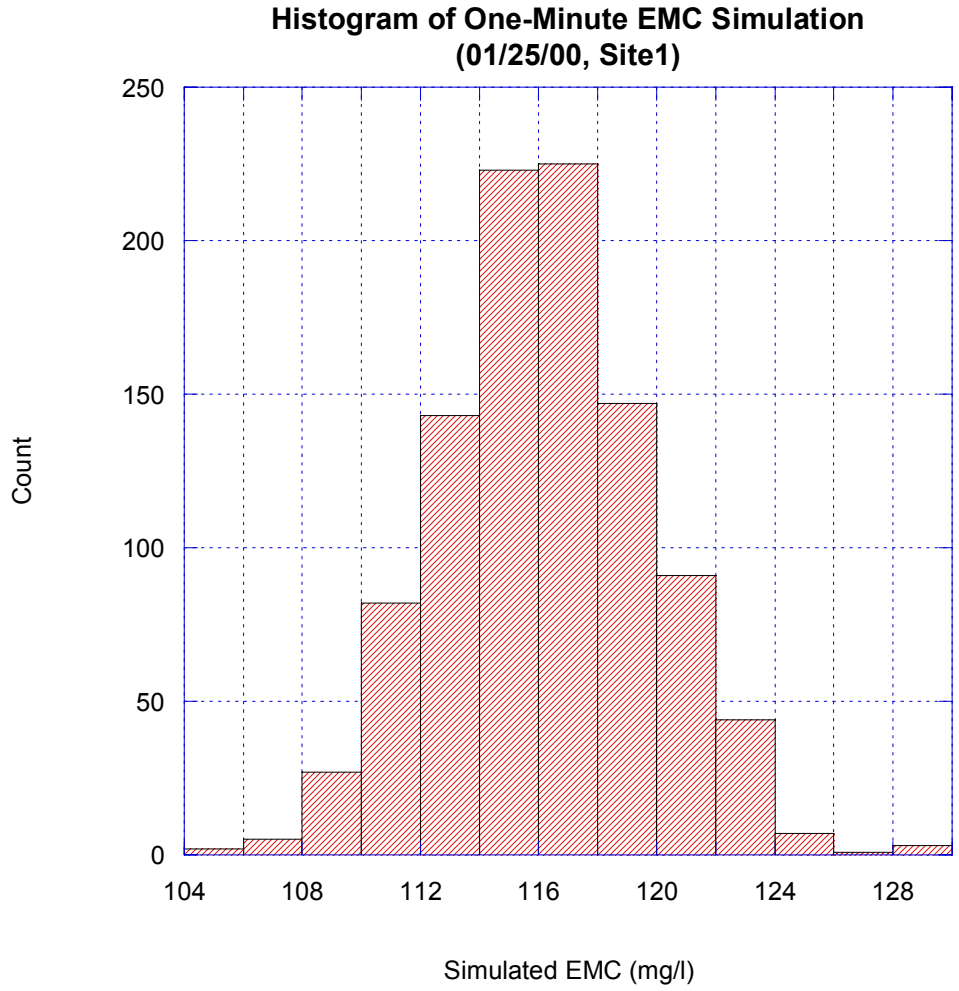


Figure 5.2.4 Sampling Distribution for One-Minute EMC Simulation (event recorded on 01/25/00, Site 1)

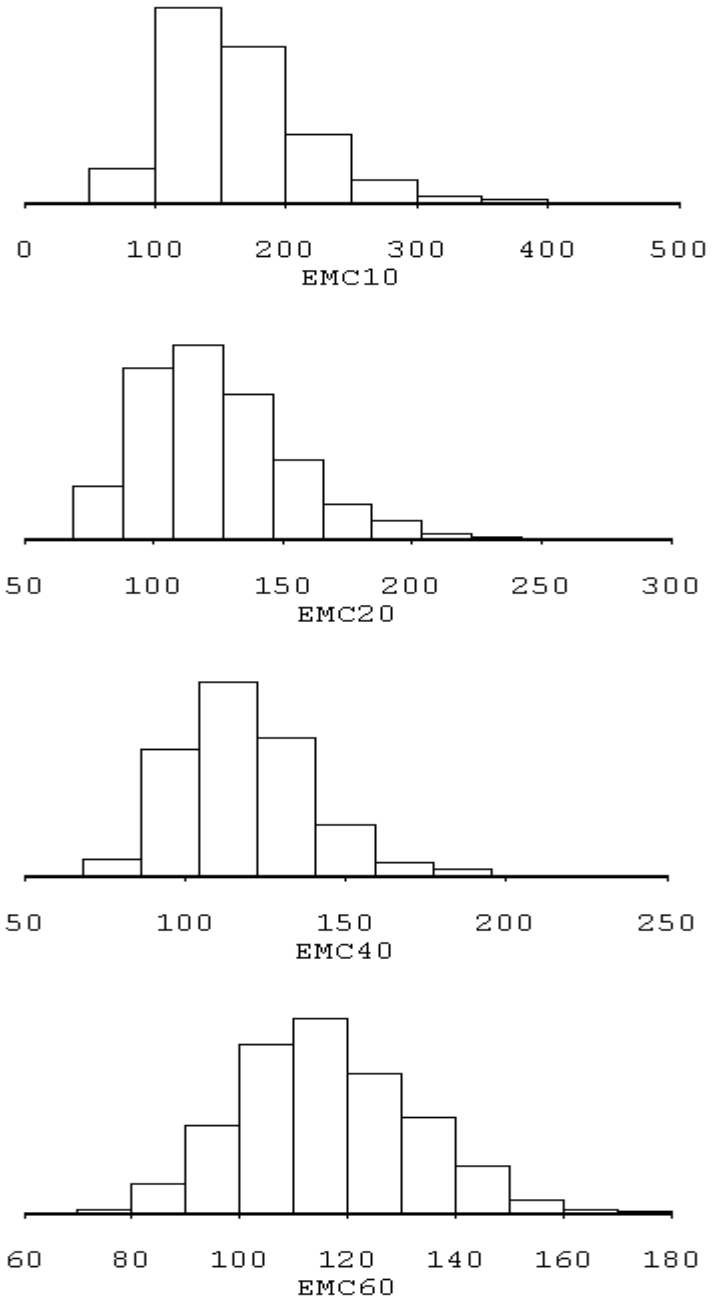


Figure 5.2.5 Sampling Distributions for n = 10, 20, 40, and 60 Using Equal-Time Sampling (event recorded on 01/25/00, Site 1)

To analyze the difference between EMC^{o}s and EMC*s, the *mean squared error* (MSE) will be used. The MSE of EMC*, which depends on the variance and the bias of EMC*, is given by

$$MSE(EMC^*) = Bias(EMC^*)^2 + Var(EMC^*) \tag{5.2.22}$$

The bias of EMC*, which can be thought as the "long-run average error", is given by

$$Bias(EMC^*) = \overline{EMC_n^*} - \mu^o \tag{5.2.23}$$

where $\overline{EMC_n^*}$ is the sample mean of EMC* simulated from n samples, and μ^o is the population mean from the one-minute simulation. Therefore, we probably prefer using the expression of error percentages, given by

$$Err\%(EMC^*) = \frac{\sqrt{MSE(EMC^*)} - \mu^o}{\mu^o} \tag{5.2.24}$$

The error percentage for the one-minute simulation (EMC^o) is caused by its variance. Since no other simulation can achieve a lower error bound, we will use this error percentage as to evaluate EMC*s. The error percentages for the case (Figure 5.2.5) are 64%, 26%, 18%, and 14% for n=10, 20, 40, and 60 respectively. The error percentage for the one-minute simulation is 3%.

5.2.3.2 Simulation Tasks

Several specific simulations were performed to check the influence of sample size on estimating EMCs. Each simulation task used 35 monitored event patterns, in which the flow and rainfall data were well measured at one-minute intervals with automatic electronic logging. The COD regression was also partially derived from these events. In each simulation task, the error percentage is calculated for each event after 1000 runs. In addition, some special sampling protocols are tested to find the best sampling approach for this particular highway case. Table 5.2.2 summarizes the hydrologic characteristics for used events.

Table 5.2.2 Hydrologic Characteristics for 35 Monitored Events

Hydrologic Property	Average	StdDev	Minimum	Median	Maximum
Total Rainfall (in)	1.17	1.54	0.08	0.67	6.14
Max Rain Intensity (in/hr)	0.31	0.33	0.02	0.19	1.28
Discharge Volume (gal)	75022	99293	1799.5	36808	374217
Max Discharge Rate (gpm)	340	304	17	258	1465
Rain Duration (min)	660.51	512.66	93	610	2376

Task 1: Random Sampling

The simulation assumes a sample set with specified size (n) that is randomly collected from all possible time elements during each tested event. It is a random permutation of size n for a sequence. Theoretically, this is the most general case for a sample set with fixed size. The influence of sample size on EMC results is evaluated by setting n = 10, 20, 40, 60, and 100.

Task 2: Equal-Time Sampling

The simulation assumes a sample set with specified size (n) that is equally spaced in tome during each tested event. To avoid the extreme result of a sample sequence, each selected sample sequence will be randomly shifted forward or backward in a range (10 minutes) in simulations. The influence of sample size on EMC results is evaluated by setting n = 10, 20, 40, 60, and 100.

Task 3: Equal-Rainfall Interval Sampling

The simulation assumes a sample set with specified size (n) that is equally spaced in rainfall depth during each tested event. The influence of sample size on EMC results is evaluated by setting n = 10, 20, 40, 60, and 100.

Task 4: Equal Discharge-Volume Sampling

The simulation assumes a sample set with specified size (n) that is equally spaced in the discharge volume during each tested event. No weighting noise is assumed in this task. The influence of sample size on EMC results is evaluated by setting n = 10, 20, 40, 60, and 100.

Task 5: Equal Discharge-Volume Sampling with Weighting Noise

Task 5 has almost the same setup as Task 4 but with independent weighting nose. The weighting noise is assumed to have a Beta (5, 2) distribution. Figure 3.2.6 shows the probability density for Beta (5, 2). The influence of sample size on EMC results is evaluated by setting n = 10, 20, 40, 60, and 100.

Task 6: First Flush vs. Non-First Flush Sampling

Task 6 evaluates the performance of "first flush" and "non-first flush sampling protocols. A first flush protocol assumes a more frequent sampling pace in the beginning of a storm event, and then gradually reduces to a slower constant pace near the end of the event. In contrast to first flush sampling, a non-first flush protocol always keeps the same pace throughout the storm event. The paces for first flush protocols can depend on the time or rainfall. Two proposed first flush protocols are FFR (based on rainfall) and FFT (based on time). Similarly, two proposed non-first flush protocols are CSR (based on rainfall) and CST (based on time).

These protocols are described as follows:

FFR: Collect five samples during the first 0.1-inch rainfall depth of an event. Next, collect four samples during the rainfall depth from 0.1 to 0.3 inch. Afterwards, collect one sample for every 0.1-inch rainfall depth.

CSR: In contrast to FFR, collect one sample for every 0.05-inch rainfall depth.

FFT: Collect five samples during the first hour of an event, each separated by 15 minutes. Next, collect four samples in the next two hours, each separated by 30 minute. Afterwards, collect one sample for every hour.

CST: In contrast to FFT, collect one sample every 45 minutes.

There are two approaches to evaluate these four protocols. One is to set the maximum number of samples as 15 in each protocol; the other is to keep sampling according to each protocol until the end of an event.

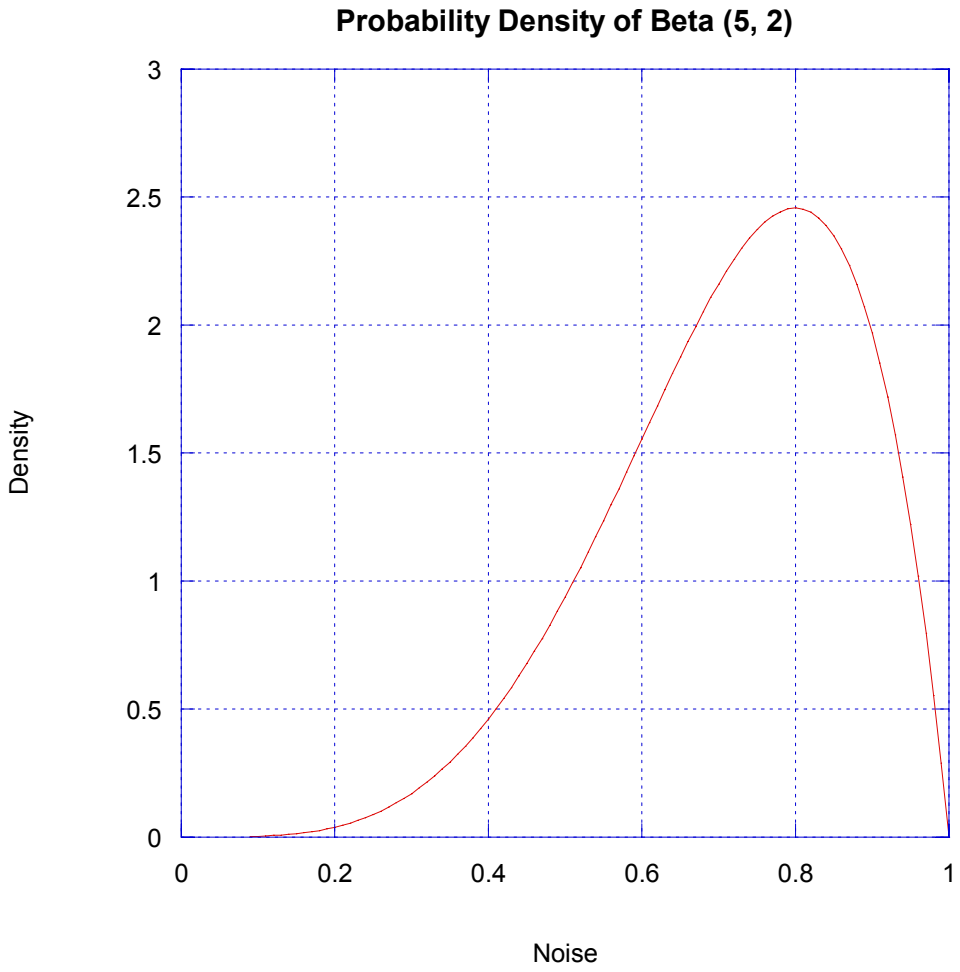


Figure 5.2.6 Probability Density of Beta (5, 2)

5.3 Results and Discussion

5.3.1 Results of Simulation Tasks

Figure 5.3.1 shows the sample distributions from Task 1. The worst error percentage can be up to 80% for $n = 10$. Table 5.3.1 shows the summary statistics for Task 1's result. The average error percentages for $n = 10$, 20, 40, 60, and 100, are 47.0%, 30.2%, 19.5%, 15.3%, and 11.6% respectively. The medians of errors are slightly lower than the averages. The corresponding standard deviations are 13.9%, 7.2%, 4.1%, 2.9% and 2.2%. Task 1's result provides a benchmark on the influence of sample size for estimating EMCs, and is the most general sample set in Task 1.

Figure 5.3.2 shows the sample distributions from Task 2. Only one outlier was found for each n. The worst case is for $n = 10$ and is approximately 66%, which is much improved from Task 1. Table 5.3.2 shows the summary statistics for Task 2's result. The average error percentages for $n = 10$, 20, 40, 60, and 100, are 37.2%, 21.7%, 15.2%, 12.4%, and 9.2% respectively. The medians of errors are generally the same as the averages. The corresponding standard deviations are 11.1%, 4.4%, 2.7%, 2.8% and 1.7%. These statistics show an improvement over random sampling.

Figure 5.3.3 shows the sample distributions from Task 3. Although several outliers were found for $n = 10$, the worst case is only around 30%, which is much improved over Task 2. Table 5.3.3 shows the summary statistics for Task 3's result. The average error percentages for $n = 10$, 20, 40, 60, and 100, are 23.9%, 17.5%, 13.5%, 11.9%, and 10.5% respectively. The medians of errors are generally the same as the averages. The

corresponding standard deviations are 2.2%, 2.2%, 2.6%, 3.2% and 3.7%. These standard deviations show an improvement over equal time sampling.

Figure 5.3.4 shows the sample distributions from Task 4. It is obvious on plot that this is the best result from the aspect of outliers, averages, or variances. Table 5.3.4 shows the summary statistics for Task 4's result. The average error percentages for n = 10, 20, 40, 60, and 100, are 23%, 16.6%, 12.0%, 9.7%, and 7.5% respectively. The medians are generally the same as the averages. The corresponding standard deviations are 2.5%, 1.6%, 1.2%, 1.0% and 0.7%.

Figure 5.3.5 shows the sample distributions from Task 5. Even though we added independent noise, the noise has little effect. Table 5.3.5 shows the summary statistics for Task 5's result. The average error percentages for n = 10, 20, 40, 60, and 100, are 23.5%, 17.1%, 12.3%, 10.1%, and 7.8% respectively. The corresponding standard deviations are 2.1%, 1.6%, 1.3%, 0.9% and 0.8%.

Figure 5.3.6 shows the sample distributions from Task 6's limited case (the maximum number of samples is set as 15). The average numbers of samples for CSR, FFR, CST, and FFT are 11.5, 11.7, 12.5, and 12.8 respectively. Several outliers were found for each group. Without considering outliers, FFT and FFR appear a little better than CST and CSR on plot due to fewer worse cases and smaller variances. Table 5.3.6 shows the summary statistics. The average error percentages for CSR, FFR, CST, and FFT are 37.6%, 32.7%, 35.7%, and 34.9% respectively. The medians are smaller than the averages in each group. The corresponding standard deviations are 29.5%, 25.2%, 28.9%, and 38.6%.

Figure 5.3.7 shows the sample distributions for Task 6's unlimited case (keep sampling until the end of an event). The average numbers of samples for CSR, FFR, CST, and FFT are 22.7, 16.9, 17.9, and 18.5 respectively. No outliers occurred. FFT and FFR are obviously better than CST and CSR on plot due to better worse cases and smaller variances. Table 5.3.7 shows the summary statistics. The average error percentages for CSR, FFR, CST, and FFT are 23.8%, 21.5%, 27.0%, and 21.3% respectively. The medians are a little smaller than the averages in each group. The corresponding standard deviations are 11.7%, 6.7%, 9.9%, and 3.8%.

Table 5.3.1 Simulation Summary for Random Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	4.0	1.3	2.2	3.6	7.3
EMC10	35	47.0	13.9	29.5	43.2	80.8
EMC20	35	30.2	7.2	20.8	28.0	49.5
EMC40	35	19.5	4.1	13.5	18.8	31.6
EMC60	35	15.3	2.9	11.2	14.8	24.3
EMC100	35	11.6	2.2	8.0	11.3	17.8

Table 5.3.2 Simulation Summary for Equal-Time Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	4.0	1.3	2.2	3.6	7.1
EMC10	35	37.2	11.1	22.8	33.2	66.1
EMC20	35	21.7	4.4	15.8	21.2	41.1
EMC40	35	15.2	2.7	10.8	15.2	22.2
EMC60	35	12.4	2.8	8.8	12.0	22.9
EMC100	35	9.2	1.7	6.9	8.9	14.3

Table 5.3.3 Simulation Summary for Equal-Rainfall Depth Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	4.0	1.3	2.2	3.6	7.4
EMC10	35	23.9	2.2	20.5	23.4	29.2
EMC20	35	17.5	2.2	14.2	17.2	23.2
EMC40	35	13.5	2.6	10.8	12.7	23.2
EMC60	35	11.9	3.2	9.0	10.8	24.2
EMC100	35	10.5	3.7	7.2	9.1	23.5

Table 5.3.4 Simulation Summary for Perfect Equal-Discharge Volume Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	4.0	1.3	2.1	3.5	7.2
EMC10	35	23.0	2.5	20.1	22.3	30.5
EMC20	35	16.6	1.6	13.9	16.3	21.5
EMC40	35	12.0	1.2	10.2	11.6	15.2
EMC60	35	9.7	1.0	8.4	9.5	12.9
EMC100	35	7.5	0.7	6.5	7.4	9.5

Table 5.3.5 Simulation Summary for Noised Equal-Discharge Volume Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	3.9	1.3	2.2	3.5	6.9
EMC10	35	23.5	2.1	20.6	23.1	29.8
EMC20	35	17.1	1.6	15.0	16.5	21.7
EMC40	35	12.3	1.3	10.5	12.0	15.8
EMC60	35	10.1	0.9	8.7	9.8	12.7
EMC100	35	7.8	0.8	6.7	7.6	10.5

Table 5.3.6 Simulation Summary for the Limited Case in Task 6 (maximum: 15 samples) plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	3.9	1.3	2.2	3.5	7.0
CSR	35	37.6	29.5	17.2	25.9	134.1
FFR	35	32.7	25.2	17.4	24.2	124.0
CST	35	35.7	28.9	18.1	25.4	144.5
FFT	35	34.9	38.6	19.7	24.0	210.7

Table 5.3.7 Simulation Summary for the Unlimited Case in Task 6 plus One-Minute Simulation

Case	N	Average Err%	StdDev Err%	Minimum Err%	Median Err%	Maximum Err%
1-Min	35	4.0	1.3	2.2	3.5	7.2
CSR	35	23.8	11.7	8.0	21.6	57.2
FFR	35	21.5	6.1	9.3	21.7	33.3
CST	35	27.0	9.9	13.8	23.6	54.3
FFT	35	21.3	3.8	14.5	21.4	28.5

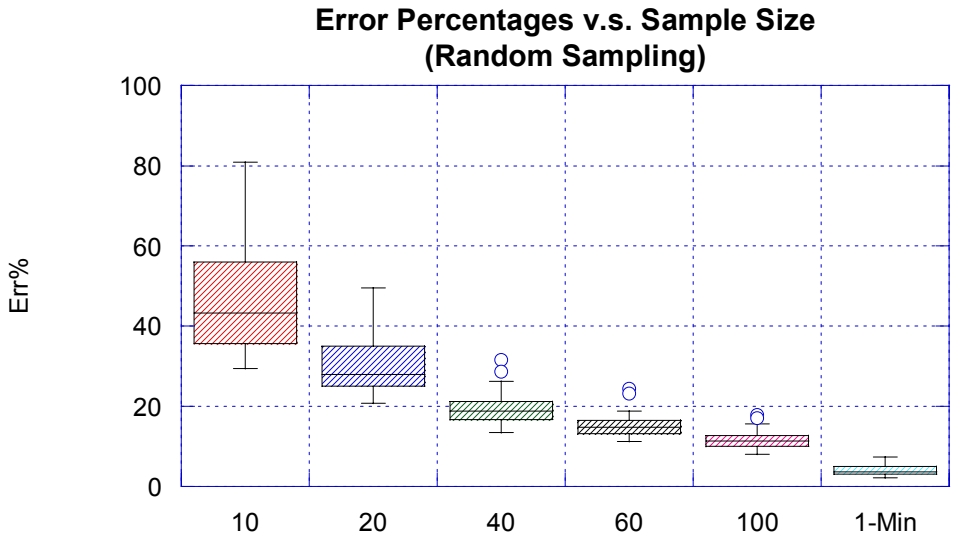


Figure 5.3.1 Sampling Distributions for Random Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

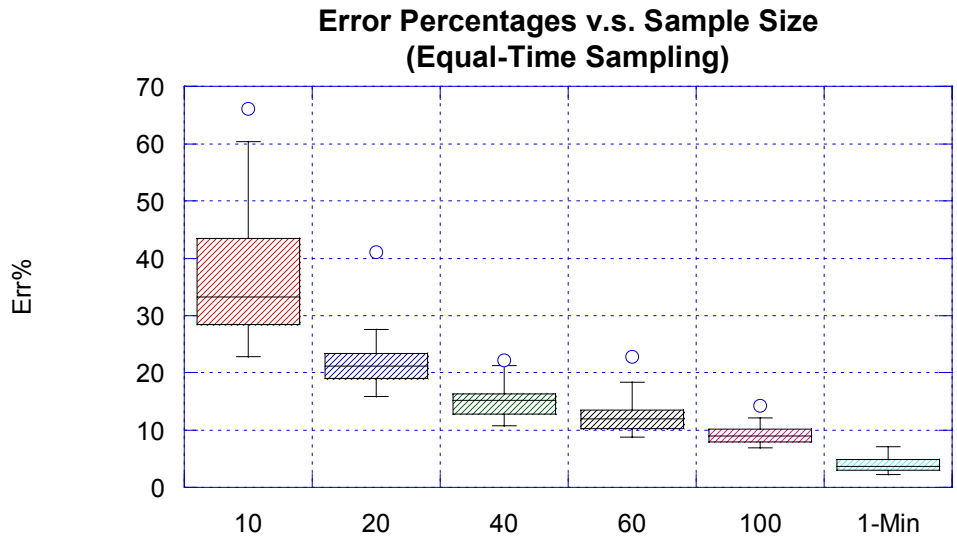


Figure 5.3.2 Sampling Distributions for Equal-Time Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

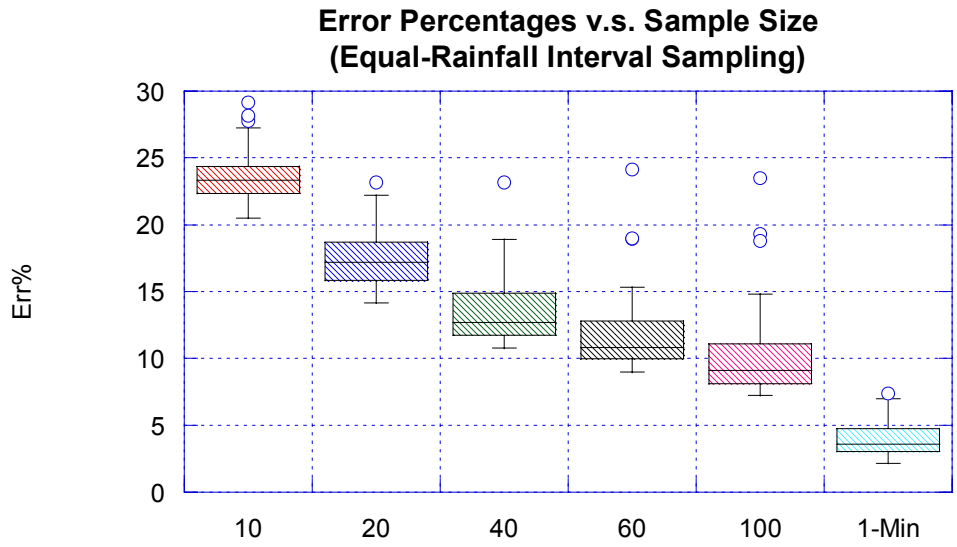


Figure 5.3.3 Sampling Distributions for Equal-Rainfall Interval Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

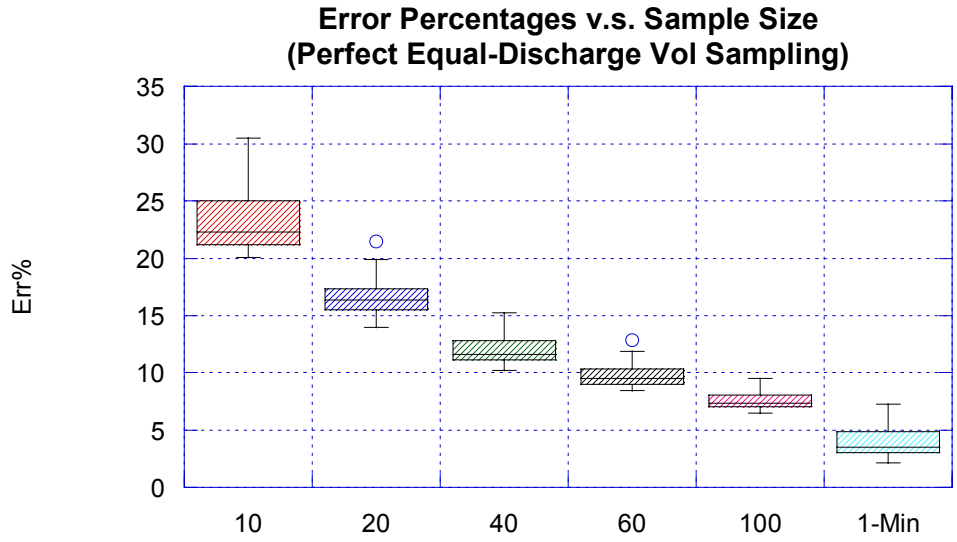


Figure 5.3.4 Sampling Distributions for Perfect Equal-Discharge Volume Sampling (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

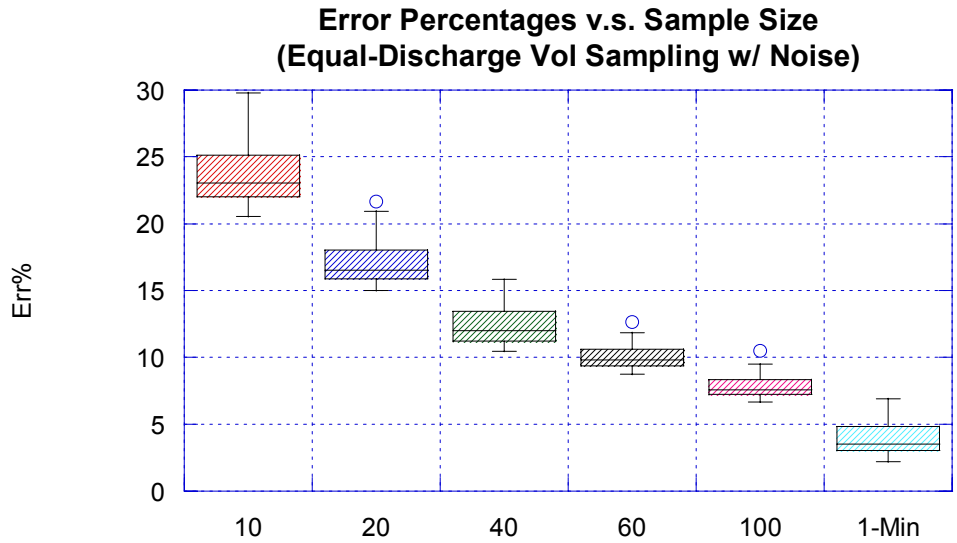


Figure 5.3.5 Sampling Distributions for Equal-Discharge Volume Sampling with Noise (as n = 10, 20, 40, 60, and 100) plus One-Minute Simulation

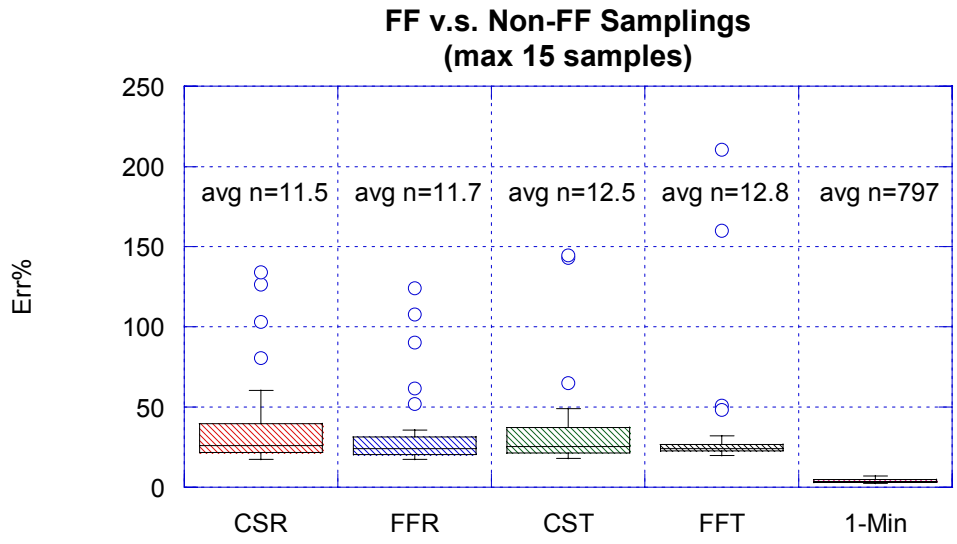


Figure 5.3.6 Sampling Distributions for First Flush vs. Non-First Flush Sampling (maximum 15 samples) plus One-Minute Simulation

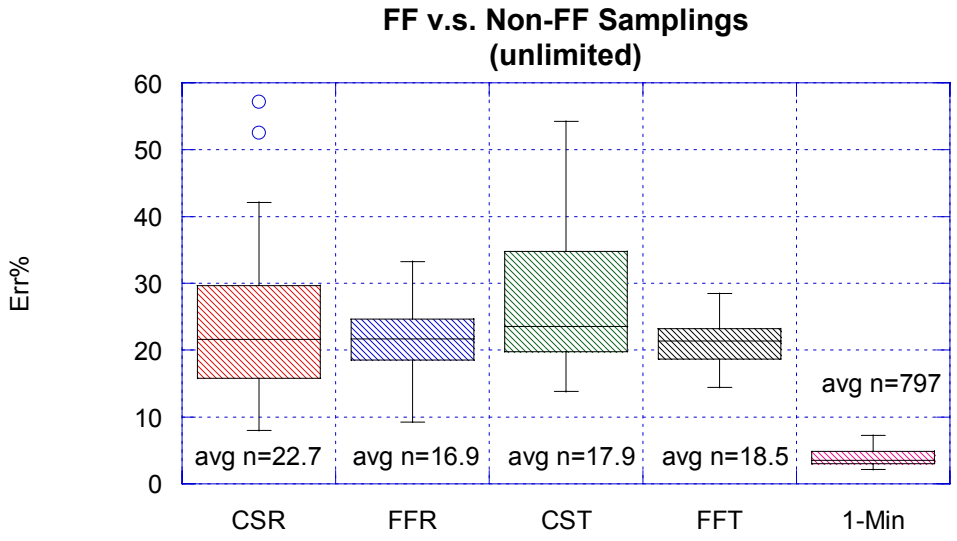


Figure 5.3.7 Sampling Distributions for First Flush vs. Non-First Flush Sampling (the unlimited case) plus One-Minute Simulation

5.3.2 Reliability Checking of Existing EMC Data

As shown in the simulation and theory, the reliability of an EMC result is highly controlled by the number of its composite instant concentrations and their related weights. When this number is large, the EMC result is stable and the influence of their related weights is not significant. On the other hand, when this number is small, the EMC result is not stable and the influence of their related weights is significant. Generally, it is better to have at least 40 composed samples for reliable EMCs while the average error could reasonably limit to around 10%.

EMC results obtained from composite samples collected by automated samplers are generally more reliable due to the large number of their composite instant concentrations. It is almost cost-and effort-free to achieve this large number by programming. The only remaining concerns for automated samplers are mainly field problems, such as sensors or tubing clogging during an event. Our experience was 30% failure of automatic samplers.

EMC results obtained from grab samples contain more uncertainty due to fewer composite concentrations. It is very expensive and unfeasible to manually collect more than 20 grab samples for a storm event, and it is rare to find EMC results obtained from over 20 grab samples. Therefore, EMC results obtained from less than 10 samples should be used cautiously because the errors might be over 50%. For an existing EMC result obtained from 10 to 20 samples, the weight distribution provides additional information to check the reliability. For example, as shown earlier, the average error is around 46% under random sampling for the sample size as 10. While all of weights are less than 0.2, the average error was reduced to 36%; while at least one weight is greater than 0.4, the

average error increased to 56%. Thus, the value of "2/n" might be used as the maximum weight to check the weight distribution.

5.4 Conclusions

In brief, the following conclusions can be made:

1. Realistically, an EMC result is a weighted average of its composite instant concentration values, even if they were obtained from an automated sampler. When obtained from grab samples, the corresponding discharge distribution determines the weights; when obtained from composite samples, the inherent equipment noise determines the weights. At least 40 composite samples might be needed for obtaining reliable EMCs.
2. In spite of performance concerns for equipment and operational problems, composite samples collected by automated samplers are a better way to obtain EMCs, because of the large number of composite concentrations achieved by programming. Unfortunately equipment failures detract from this result, and our experience was 30% failure of automatic samplers.
3. EMC results obtained from less than 10 grab samples should be cautiously used because the error might be above 50%. The 2/n principle in the weight distribution could be used for further reliability checking when the number of grab samples ranges from10 to 20. It is very rare to find existing EMC results obtained from more than 20 grab samples.
4. Stochastic approaches are proper way to handle natural data due to the complexity of the system. The derived asymptotic distributions provide general ideas about the reliability of EMC data. However, to precisely estimate the reliability, knowledge describing the behavior of washed-out concentrations is needed.

5. For this particular highway COD case, stratifying grab samples as much equal weights as possible still assures obtaining better results. Practically, it is possible to distribute the weights more equally by using equal-rainfall interval sampling. Additionally, it is worth collecting more samples earlier during an event by applying “first flush sampling” due to the rapid change in COD concentration in that period. Therefore it is valuable to sample the entire event to achieve better results.

6. STUDY OF FIRST FLUSH NOTATIONS AND CRITERIA

This chapter contains existing and new first flush notations and criteria. Each section includes background, methodologies, results and discussion. The introduction includes a statement of the problem and objectives. The conclusions summarize both notations and include recommendations.

6.1. Introduction

In the past 15 years, the first flush phenomenon has been the subject of research and debate. Although interest in the subject has been high, the understanding of this phenomenon is still in a very early stage, and even this phenomenon's existence is debated. Previous researchers (Geiger, 1987; Bertrand-Krajewski et al., 1998) have proposed mass emission based definitions, and many events do not have a sufficiently large initial mass discharge to qualify as a first flush. Large watersheds may not exhibit a first flush because of the large transportation lag (Ma et al, 2002). In spite of the debate, a neutral and consensus upon qualitative description of the first flush phenomenon is "the first part of runoff in a storm event is the most polluted". Based upon this understanding, many facilities and devices have been invented and installed to treat or capture the first flush. For example, one frequently used Best Management Practice (BMP) is a sedimentation basin, which functions as suspended solids removal device as well as a detention basin. These sedimentation basins or storage tanks are usually designed to store the first few centimeters of runoff (Novotny and Olem, 1994).

Why is there no consensus of the first flush phenomenon? It is useful to review what other authors have written about the first flush. "A strongly distinctive first flush effect of suspended solids was not recorded at either catchment"; "The first flush of conductivity was not a regular feature at either catchment" (Deletic, 1998). "The characteristics of the M (V) curves (used to compare pollutant discharge) depend on the pollutant, on the site, on the rainfall event and the overall operation of the sewer system"; "No clear and general multi-regression relationships can be established to explain their shapes and their variability" (Bertrand-Krajewski et al., 1998). "The first flush load was shown to correlate well with the peak rainfall intensity, the storm duration and the antecedent dry weather period" (Gupta and Saul, 1996). "For all of the events of our records, the first flush is observed only once" (Saget et al., 1996). "The run-off from this specific catchment has shown significant first flush effects"; "The total run off per event of SS, COD, and BOD clearly depend on the preceding dry weather period" (Larsen et al., 1998). "The first flush for the particulate-bound fractions of these metals was not well defined"; "A first flush occurred for all events for all solids fractions" (Sansalone and Buchberger, 1997).

From the above citations, it is obvious that no consensus exists. Generally, several factors might affect the existence of the first flush phenomenon, such as the rainfall's characteristics, the catchment's characteristics, the nature of the pollutant, and the observer's definition of a first flush. , The lack of consensus of a definition and notation for first flush is a large aspect of the debate.

The objective of this chapter is to examine current and widely used first flush notations and criteria, and through better notation and criteria, clarifies the understanding of the first flush phenomenon. In addition, we show the impact of first flush on BMP selection. Finally, we develop a new first flush notation to overcome some existing weaknesses of current notations.

6.2. Existing First Flush Notations and Criteria

6.2.1 Background

Most existing first flush concepts are established around the belief that the first portion of the discharge volume contains a higher fraction of pollutant mass load that is transported later in the storm event. From the above concept, a quantitative criterion has been developed. For example, many researchers have used plots of the cumulative fraction of total pollutant mass load versus the cumulative fraction of total runoff volume for the event to define the first flush phenomenon. The type of curves, which we call FF (First Flush) Mass-Load Curves in this study, can be used as a dimensionless representation for every storm event. Geiger (1987) defined a first flush notation when such curves have an initial slope greater than 45 percent. Gupta and Saul (1996) used Geiger's criteria, as did Larsen et al (1998). Sansalone and Buchberger (1997) used a more liberal notation than Geiger's that first flush is perceived when a mass cumulative curve is above the runoff volume curve. Saget et al (1995) suggested a very strict first flush notation that at least 80 percent of the pollutant mass is transferred in the first 30 percent of the runoff volume. Some researchers preferred to compare the fraction of the mass load at the same point in the first part of the runoff cumulative volume like Saget et al (1995). For example, Vorreiter and Hickey (1994) chose 25 percent of the runoff cumulative volume; Deletic (1998) chose 20 percent of the runoff cumulative volume. In a summary, researchers have frequently used FF mass-load curves to create various first flush criteria.

6.2.2 Methodology

6.2.2.1 Mathematical Derivations

The mathematical expression for defining FF Mass-Load curves is as follows

$$(x, y) = \left(\frac{\int_0^{t_1} Q(t)dt}{\mathbf{V}}, \frac{\int_0^{t_1} C(t)Q(t)dt}{\mathbf{M}}\right) \tag{6.2.1}$$

where **M** is total mass load, **V** is total discharge volume for a storm event, and t_1 is event's running time since beginning. From (6.2.1), y/x can be expressed as

$$\frac{y}{x} = \frac{\frac{\int_0^{t_1} C(t)Q(t)dt}{\mathbf{M}}}{\frac{\int_0^{t_1} Q(t)dt}{\mathbf{V}}} = \frac{\int_0^{t_1} C(t)Q(t)dt}{\int_0^{t_1} Q(t)dt} \times \frac{\mathbf{V}}{\mathbf{M}} \tag{6.2.2}$$

In fact, M/V is Event Mean Concentration (EMC) by definition (Ref). Thus, V/M in (6.2.2) is the inverse of EMC. Therefore, $\int_0^{t_1} C(t)Q(t)dt \Big/ \int_0^{t_1} Q(t)dt$ in (6.2.2) is the EMC for the storm event up to time t_1, which we will call the "running EMC" in this study. From (6.2.2), y/x can be rewritten as

$$\frac{y}{x} = \frac{EMC_{t_1}^r}{EMC} \tag{6.2.3}$$

When t_1 is zero, (x, y) is defined as (0, 0) because $EMC_t^r = 0$. When t_1 is the entire storm time, (x, y) is defined as (1, 1) because $EMC_t^r = EMC$. Thus (0, 0) and (1, 1) are the two fixed points on the FF mass-load curve.

From (6.2.3), we know that FF Mass-Load curves actually record the change of running EMC with cumulative flow volume or time. They also record the relative relationship between running EMC and EMC. For example, when $EMC_i^r > EMC$, the curve will fall above the diagonal line; when $EMC_i^r < EMC$, the curve will fall below the diagonal line. In addition, when EMC_i^r is increasing, the curve's slope will be greater than 45 degrees; when EMC_i^r is decreasing, the curve's slope will be less than 45 degrees.

To further understand FF Mass-Load curves, we need to study discrete cases of running EMCs. For example, based on a fine and equal discretization on discharge volume, a running EMC sequence is as follows

$$0, EMC_1^r, EMC_2^r, EMC_3^r, ..., EMC_{n-1}^r, EMC_n^r \tag{6.2.4}$$

The last element in the running EMC sequence, EMC_n^r, is the EMC for the event. To define y/x on the FF Mass-Load curve, the above running EMC sequence is divided by EMC_n^r = EMC. Equation (6.2.4) becomes

$$0, \frac{EMC_1^r}{EMC_n^r}, \frac{EMC_2^r}{EMC_n^r}, \frac{EMC_3^r}{EMC_n^r}, ..., \frac{EMC_{n-1}^r}{EMC_n^r}, 1 \tag{6.2.5}$$

The sequence defined by (6.2.5) actually sketches the profile of a FF Mass-Load curve.

From (6.2.5), we know that both running EMC and EMC determine each point on the FF Mass-Load curve. Thus, even though we may obtain several accurate estimates of running EMCs, we will distort the curve with an inaccurate estimate of the EMC. For

example, if we overestimate the EMC by 30 percent, the curve will be underestimated by 23 percent. In addition, discretization plays a role in defining FF Mass-Load curves. For example, for poorly representative discretization, the points that define the curve will have gaps, which are filled with a straight line. This introduces error in FF Mass-Load curve.

6.2.2.2 Constructing FF Mass-Load Curves

Theoretically, we can construct each FF Mass-Load curve from normalizing the running EMC sequence defined in (6.2.4) and the corresponding cumulative flow sequence. Practically, these two sequences need to be constructed from discrete measurements. For most cases, we are able to measure flow accurately and frequently. Thus, the cumulative flow sequence will be determined without much difficulty. Constructing the pollutant mass sequence is more difficulty because samples are less frequently collected. Composite samples could be collected up to defined points in the runoff event; for example, ten automated samplers could be programmed to collect samples from the storm beginning to defined volumes, which can be compared to total runoff volume. This method is unlikely since composite samplers are expensive. It is more likely that the EMC will be determined from grab samples collected at defined times.

When analyzing grab samples, calculating the EMCs or running EMCs is not as straight forward as analyzing composite samples, because each concentration measurement represents an instantaneous value during an event. The EMCs or running

EMCs must be estimated from the flow-weighted average of these instantaneous concentrations. A common way to decide the weights is to split the discharge volume between two consecutive measurements at the mid-volume point. The process of constructing a FF Mass-Load curve from grab samples is illustrated as follows:

Suppose we collect 10 grab samples in time order during an event, and the measured concentrations are 20.1, 14.2, 11.5, 11.7, 8.0, 9.5, 7.0, 7.2, 6.1, and 4.9. The corresponding normalized cumulative flow volumes are 0.02, 0.07, 0.12, 0.20, 0.25, 0.42, 0.52, 0.59, 0.70, and 0.78. The determination of the flow weights for each sample is illustrated in Figure (6.2.1). The result of the weighting sequence is 0.045, 0.05, 0.065, 0.065, 0.11, 0.135, 0.085, 0.09, 0.095 and 0.26. The cumulative flow sequence is just the cumulative sequence of the weight sequence as

(0.045, 0.095, 0.16, 0.225, 0.335, 0.47, 0.555, 0.645, 0.74, 1)

Then, the running EMC sequence is calculated by

$$EMC_i^r = \frac{\sum_{j=1}^{i} C_j * W_j}{\sum_{j=1}^{i} W_j} \tag{6.2.6}$$

The result of the running EMC sequence is

(20.1, 17.0, 14.8, 13.9, 11.9, 11.2, 10.6, 10.1, 9.6, 8.4)

From the cumulative flow sequence and the running EMC sequence, the (x, y) pairs of the FF Mass-Load curve can be defined as

{ (0, 0), (0.045, 0.108), (0.095, 0.193), (0.16, 0.282), (0.225, 0.373), (0.335, 0.478), (0.47, 0.631), (0.555, 0.702), (0.645, 0.779), (0.74, 0.845), (1, 1) }

Figure 6.2.2 shows the above FF Mass-Load curve.

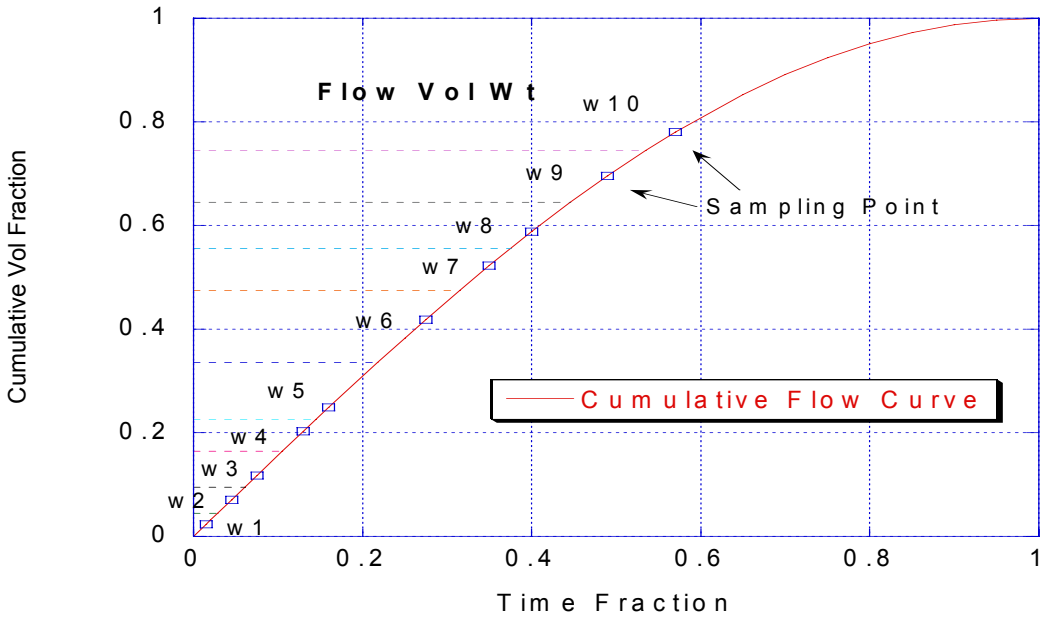


Figure 6.2.1 Determination of Flow Weights (w1 to w10) from Grab Samples

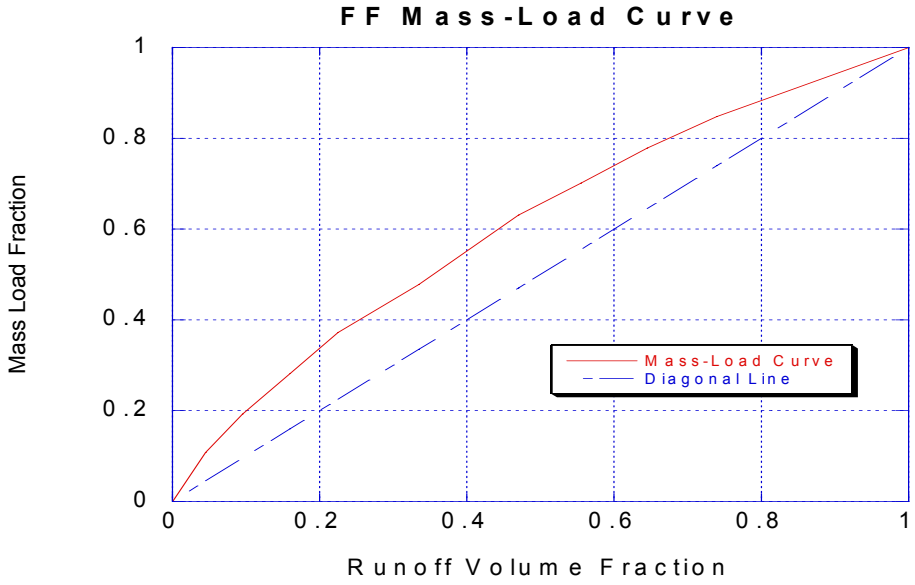


Figure 6.2.2 FF Mass-Load Curve

6.2.2.3 Error Analysis for FF Mass-Load Curves

In this section, we try to evaluate the errors associated with constructed FF Mass-Load curves. Two factors can affect the accuracy of estimated FF Mass-Load curves. One is from the intervals of a running EMC sequence. The other is from the estimation of running EMCs themselves. For example, to accurately describe an x-y curve from discrete points in a specified range, the adequacy of number of points and their distribution belong to the interval problem. The accuracy of each data point belongs to the estimation problem. The interval problem is clearly related to the discretization. However, discretization also exists in the estimation problem. This results because the EMCs or running EMCs are estimated from discretizing the continuous concentration quantity during an event, irrespective of the source of concentration data (grabs or composite samples). When EMC or running EMC is obtained from grab samples, the EMC or running EMC is clearly a flow-weighted average of analytical results of grab samples. When the EMC or running EMCs is obtained from composite samples, the EMC or running EMC is an inherently a flow-weighted average of an analytical measurement of a collection of samples taken by automated samplers. Both methods have discretization-related errors, and to evaluate their impact it is necessary to use simulation.

To perform this error analysis, we need a model to describe the pollutant's concentration variation. The COD regression model suggested below is used.

$$E(\log COD \mid \mathbf{x}) = 6.08 - 0.60 \log CumRs + 0.40 \log AtDry - 0.16 \log AtRs \qquad (6.2.7)$$

$$\varepsilon_i \sim N(0, 0.59^2)$$

Table 5.2.1 previously describes the variables in (6.2.7). There are 391 cases for the regression analysis. These data were collected over a period of two years (from 1999 to 2001) from three highway runoff-monitoring stations. Figure 5.2.2 previously shows the fitted values versus observations. The R-squared is 0.66.

In the computer simulations, we used 35 monitored events, in which the flow data were well measured at one-minute interval and the concentration data were used for the regression. To evaluate the errors on a continuous curve, several specific locations on the curve are examined. These locations are defined as First Flush ratios (FF ratios). A FF ratio is the ratio as y/x on First Flush Mass-Load curves. For each test event, we simulate COD concentration using (6.2.7) and random noise defined by the standard error, for every minute. One minute is the finest interval based on the source, and is consider the most precise. Multiple simulations were performed (1000 runs) to obtain a representative curve (mean and variance) for each event. This case is represents a benchmark for other cases, where the COD sampling interval is increased. Each sampling strategy (equal-time, equal-discharge volume, random) was evaluated with sample sizes varying from 10, 20, 40, 60, and 100. For equal time, and equal discharge volume strategies, the starting point was also randomly varied to avoid any situation that might cause error associates with a specific sampling sequence. Multiple simulations were performed (1000 runs) to obtain a representative curve (mean and variance) for each event. The square root of MSE (Mean

Squared Error) was used for comparing cases. The MSE is defined as the square of the estimator's bias plus the estimator's variance. In this case, the bias is the difference between the benchmark mean and the cases mean, and the variance is just the case variance. Through thirty-five test events, first flush ratios were recorded at 10, 20, and 30 %of the runoff volume, symbolized as "FF10", "FF20", and "FF30" respectively in this dissertation.

6.2.3 Results and Discussion

6.2.3.1 Results of Error Analysis

Figure 6.2.3 shows the boxplots of $\sqrt{MSE}$, evaluated at 10%, 20%, and 30% of total runoff volume, for different number of samples using a random sampling strategy. Figures 6.2.4 and 6.2.5 show similar results, except for equal-time and equal-discharge volume sampling strategies. Figures 6.2.4 to 6.2.6 show that the decline in error as a 2^{nd}-order concave curve while the number of samples increases from 10 to 100. And as expected, the variability decreases as the volume fraction increases.

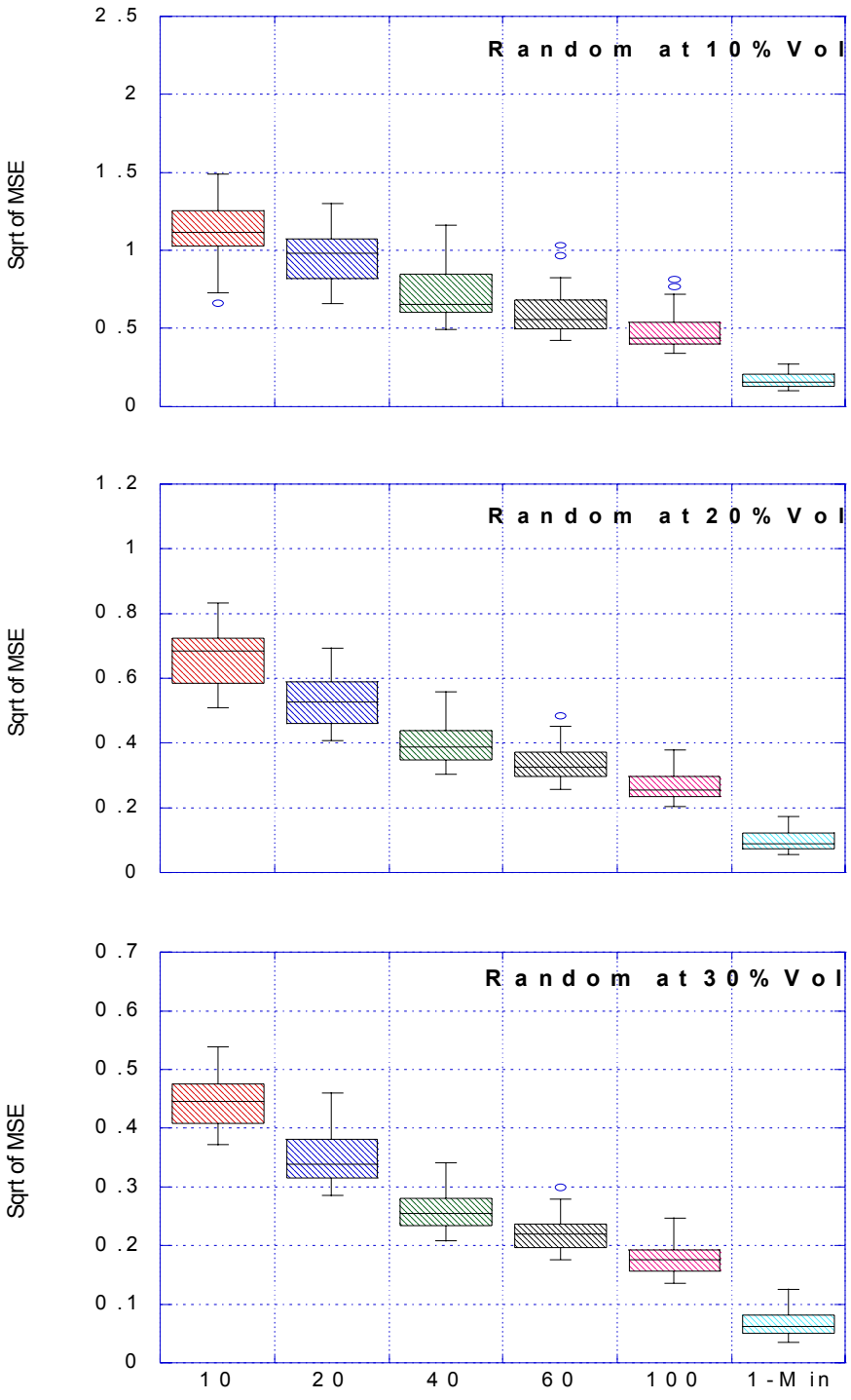


Figure 6.2.3 Boxplots of $\sqrt{MSE}$ for Random Sampling with Sample Size as 10, 20, 40, 60, and 100 plus One-Minute Sampling

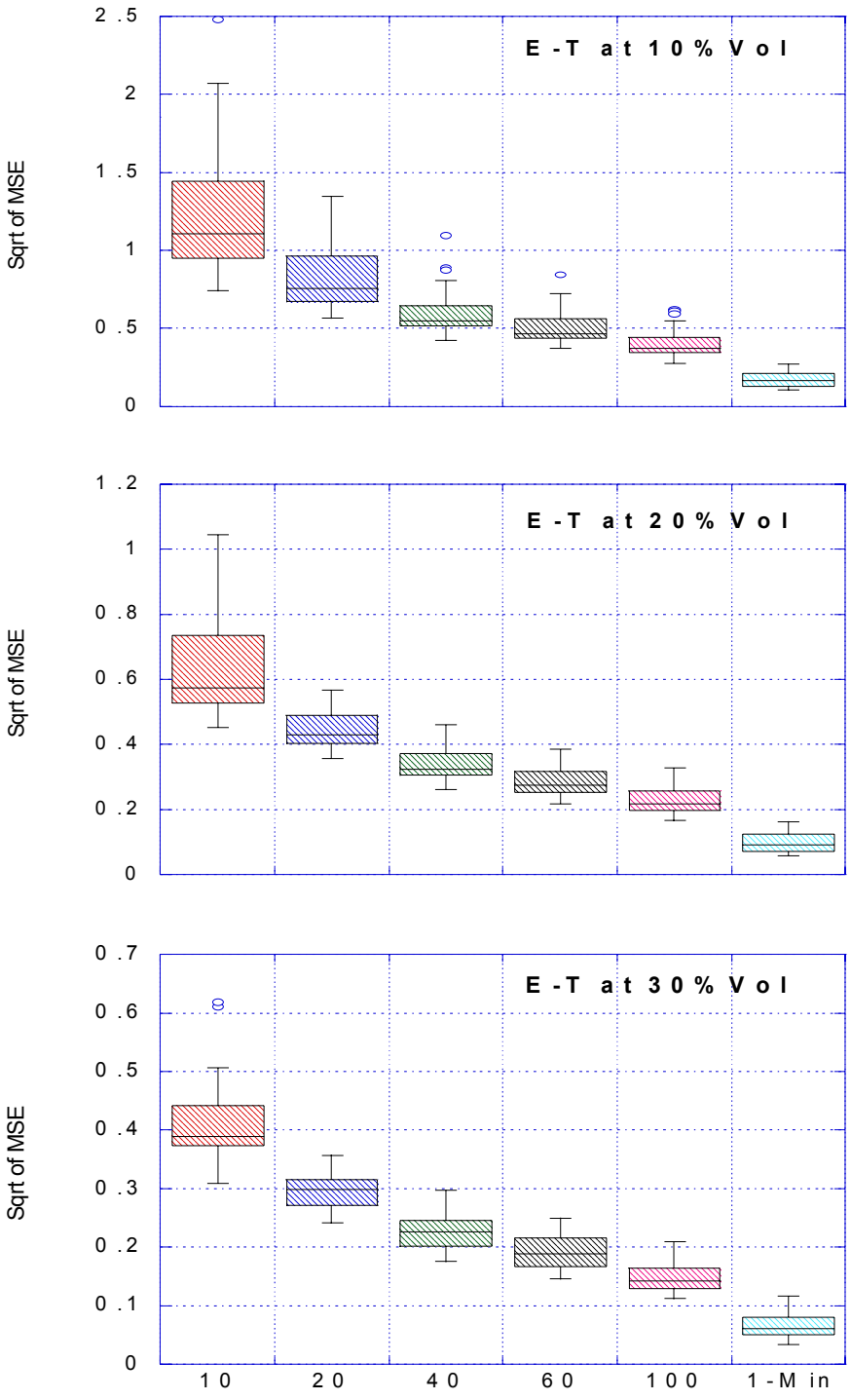


Figure 6.2.4 Boxplots of $\sqrt{MSE}$ for Equal-Time Sampling with Sample Size as 10, 20, 40, 60, and 100 plus One-Minute Sampling

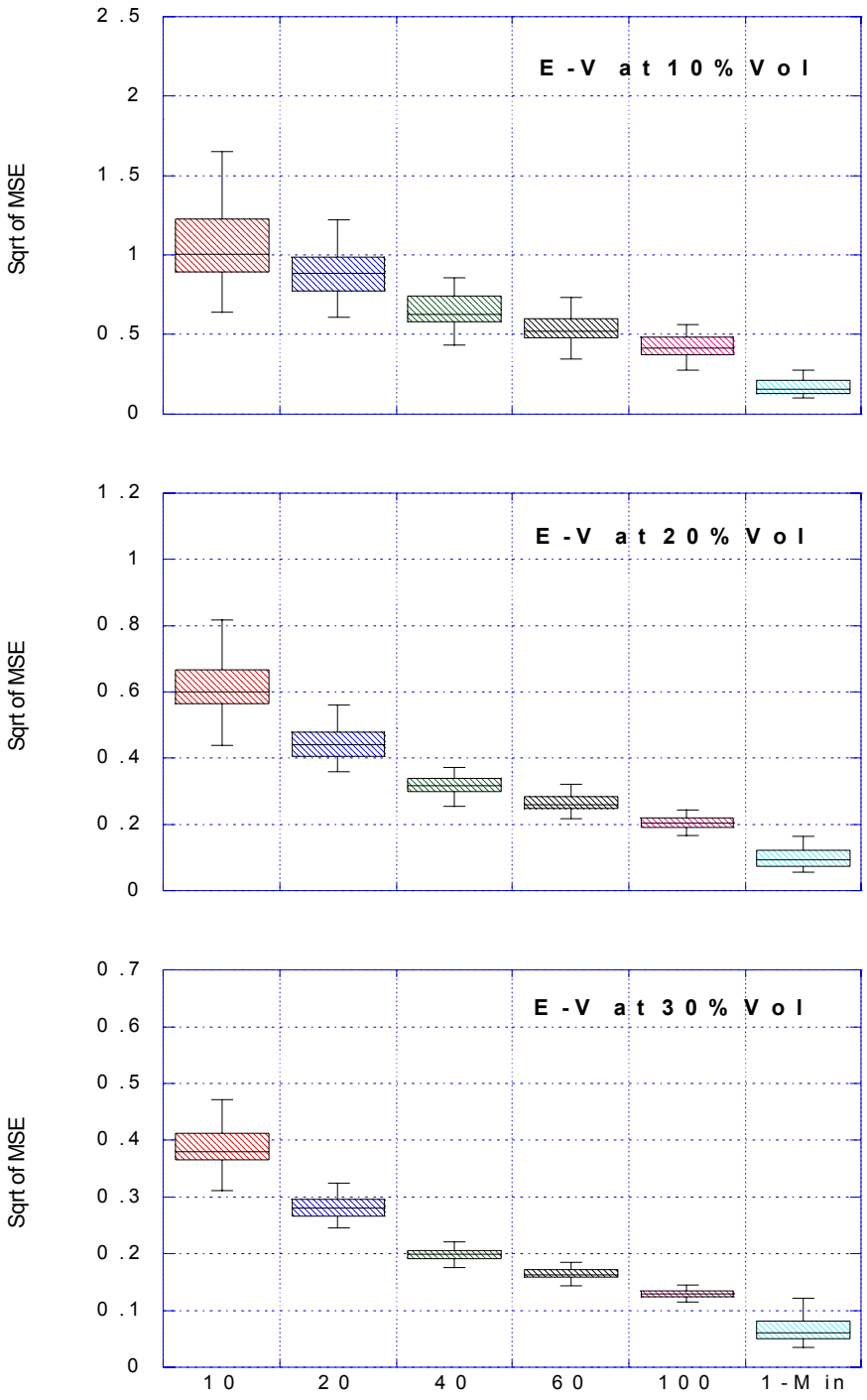


Figure 6.2.5 Boxplots of $\sqrt{MSE}$ for Equal-Discharge Volume Sampling with Sample Size as 10, 20, 40, 60, and 100 plus One-Minute Sampling

Figure 6.2.6 compares the event's mean FF ratios among the benchmark (one-minute sampling) and different samplings with sample size as 10, 20, and 40. For low number of samples, equal-time and equal-discharge volume samplings are biased. For high number of samples, every sampling strategy is unbiased.

For each sampling strategy, the benchmark's MSE reflects model's inherent randomness in (6.2.7). Based on R-squared as 0.66, the inherent randomness contributes one third of concentration variation in samples. Other samplings' MSE, therefore, includes model's inherent randomness and discretization errors. As a result, the benchmark's MSE is quite small compared to other simulations'. It implies that discretization errors strongly impact the accuracy of FF Mass-Load curves.

The equal-discharge volume strategy has large negative bias with fewer samples. This may result because the strategy misses the first flush. If only 10 samples are used, the first 10% of the storm is not sampled. The average FF10 from the benchmark cases is 2.2, which means that an equal volume strategy will miss 22% of the discharge mass. The random and equal time strategies show less bias, and this occurs because the first part of the storm is not always missed.

The simulation results show that it is necessary to collect approximately 40 samples to reduce discretization error to the benchmark case. 20 samples' sampling is generally unbiased, but with larger variances.

Figure 6.2.7 shows the comparison among regression FF ratios (calculated from regression) and field FF ratios (calculated directly from grab samples). The average number of samples is approximately 10 for 35 used events. As expected, there is more

variability in FF Mass-Load curves from field than from regression. There is little difference in the bias.

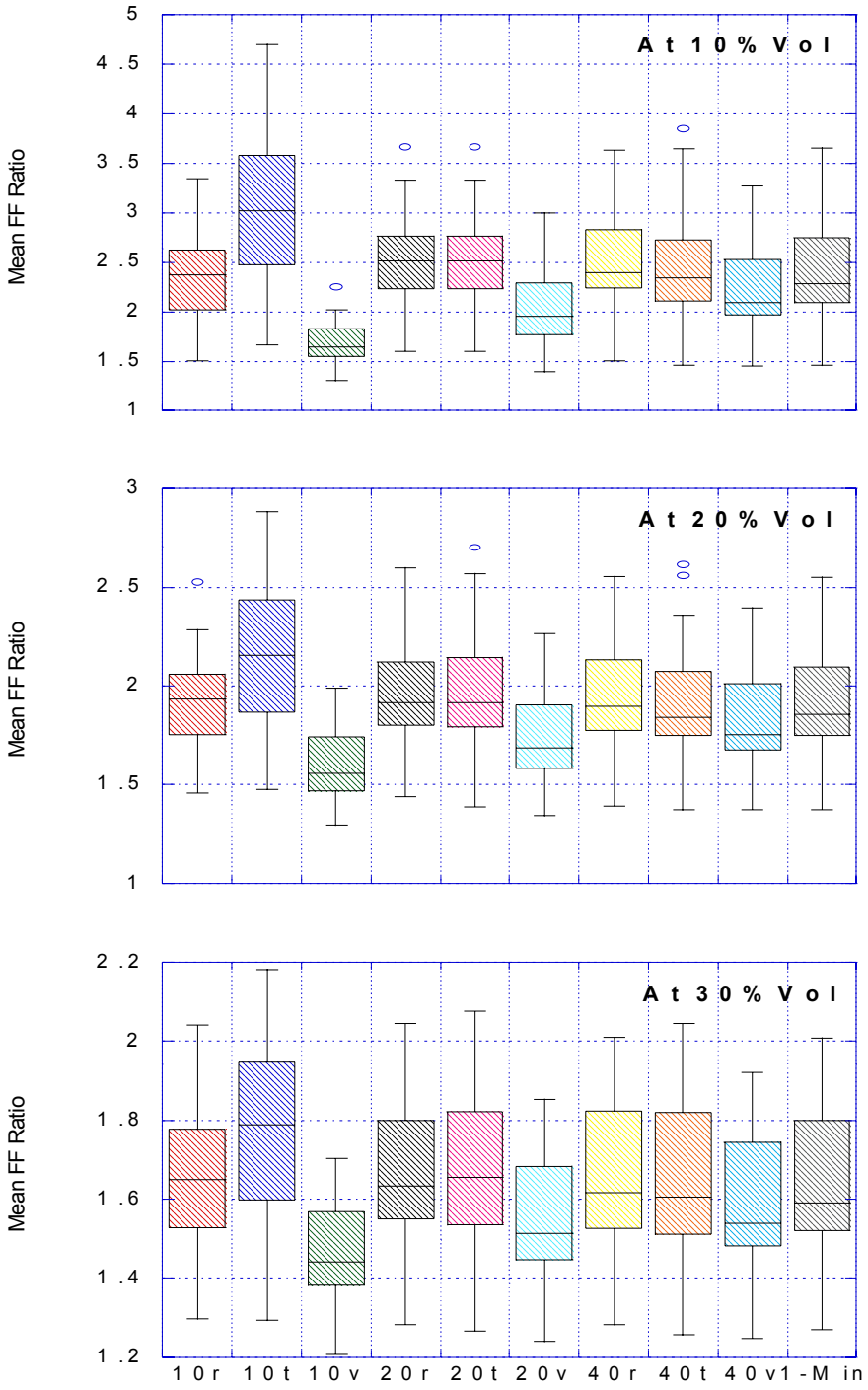


Figure 6.2.6 Comparison of Mean FF Ratios from Benchmark and Different Samplings (r, t, v) as 10, 20, and 40 Samples

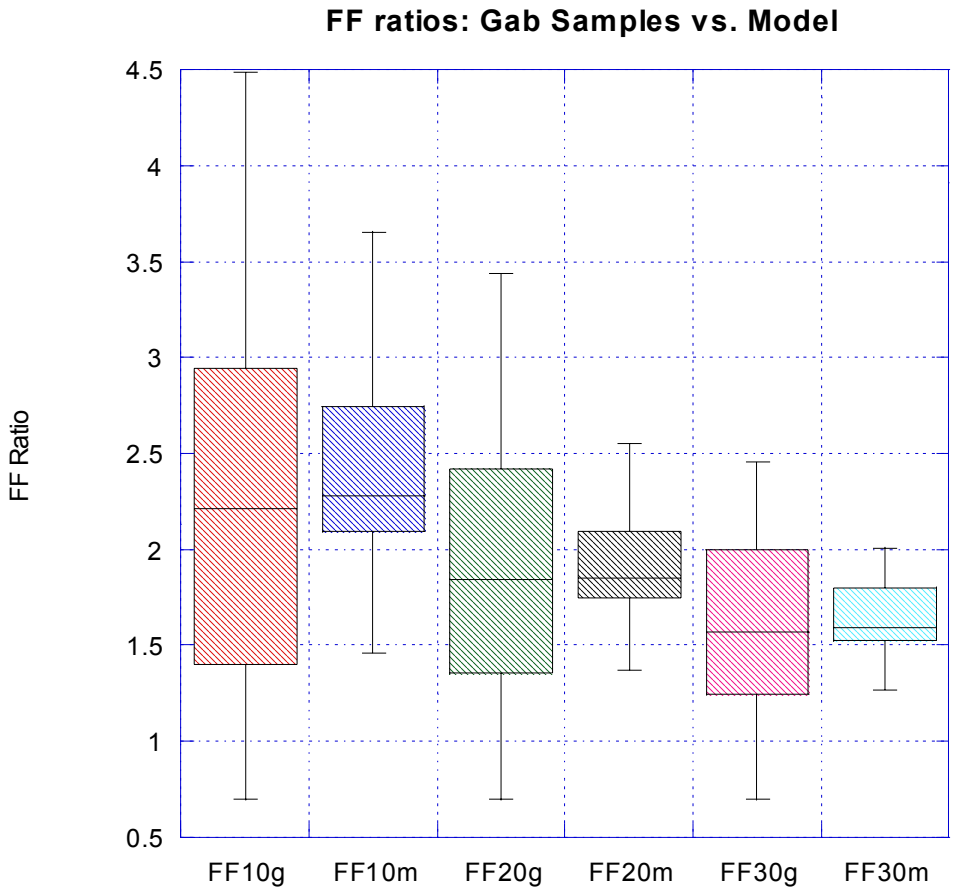


Figure 6.2.7 Comparison among Regression's and Field's FF Ratios (g: from grab samples; m: from regression)

6.2.3.2 Implications for BMP Selection

A very important reason to use FF Mass-Load curves for defining first flush phenomenon is to obtain information for pollutant loading and potential mass removal efficiency. It is necessary to estimate removal efficiencies and mass loads to comply with Total Maximum Daily Load (TMDL) provisions issued by the US EPA regulations in 1985 and 1992 that implement section 303(d) of the Clean Water Act. A TMDL specifies the maximum amount of a pollutant that a water body can receive and still meet water quality standards, and allocates pollutant loadings among point and nonpoint pollutant sources. By law, states, territories, and authorized tribes are required to develop lists of impaired waters, and establish priority rankings for waters on the lists to develop TMDLs. The TMDL makes pollutant loading an important aspect for stormwater management. Therefore, FF Mass-Load curves provide related mass loading information.

Here we suggest a hypothetical scenario to illustrate how to use FF Mass-Load curves in a BMP for a stormwater treatment plan. This treatment plan is designed to remove potential first flush pollution for COD from highway runoff by filtration. The filtration capacity is designed to be able to treat 5-year flow. Assume the treatment efficiency fixed at 80%. The only cost consideration is related to the treated flow volume, and the cost constraint limits treatment to no more than 15% of annual runoff volume. The corresponding decision to make is to determine the filter operation for each storm event. To simplify the decision, there are only four operation options:

(d0) do not treat;

(d1) treat first 10% of flow volume;

(d2) treat first 20% of flow volume, or

(d3) treat first 30% of flow volume.

The objective is to remove as much pollutant mass as possible within the cost constraint. In implementing this BMP, we used the 35 previously events and divide them into four groups based on rainfall size: S (Small), MS (Medium Small), ML (Medium Large), and L (Large). The group size is selected so that each group has approximately the same frequency. Thus, each group ranges as: S <= 0.3, 0.3 < MS <= 0.9, 0.9 < ML <= 2.4, and 2.4 < L. Therefore, the decision now can be simplified to choose the fraction option for treatment for each group.

The representative values of FF10, FF20, and FF30 for each group need to be used for calculation. The COD regression (6.2.7) is used to calculate the FF ratios. The relationship between FF ratios and event rainfall is used. Figure 6.2.8 shows the second order OLS (Ordinary Least Square) fit and the *lowess* smoothing fit for the FF ratios as a function of log *EventRain*. As seen earlier, the fitted line of FF10 is always higher than FF20 and then FF30. Based on the OLS fit in Figure 6.2.7, the representative FF10, FF20, and FF30 are selected for each group based on the median rainfall in each group.

Several assumptions are used to facilitate the calculations: the runoff coefficient is equal to 1.0; the catchment area is equal to 1.0 and the EMC is equal to 1.0. These assumptions make the calculations dimensionless. Thus, based on the above assumptions,

representative FF ratios, and representative rainfall (median rainfall), the representative mass removal can be calculated. Table 6.2.1 shows the mass removal calculations for each group under each operation option.

There are total $4^4 = 256$ decision types. After searching for all qualified decisions (satisfying the cost constraint), the BMP decision is that d3 for S, d3 for MS, d2 for ML, and d1 for L. As shown in Table 6.2.2, the mass removal is approximately 32 percent, and the treated volume is approximately 15%.

This hypothetical case shows that accounting for first flush effects can improve treatment. It can be extended to multiple pollutants cases, where the optimal decision needs to be searched among multiple decision tables. In addition, more realistic and complicated conditions can be considered, such as variable treatment efficiency, variable event mean concentration, variable runoff coefficients, etc. The data requirements will be greater for these cases.

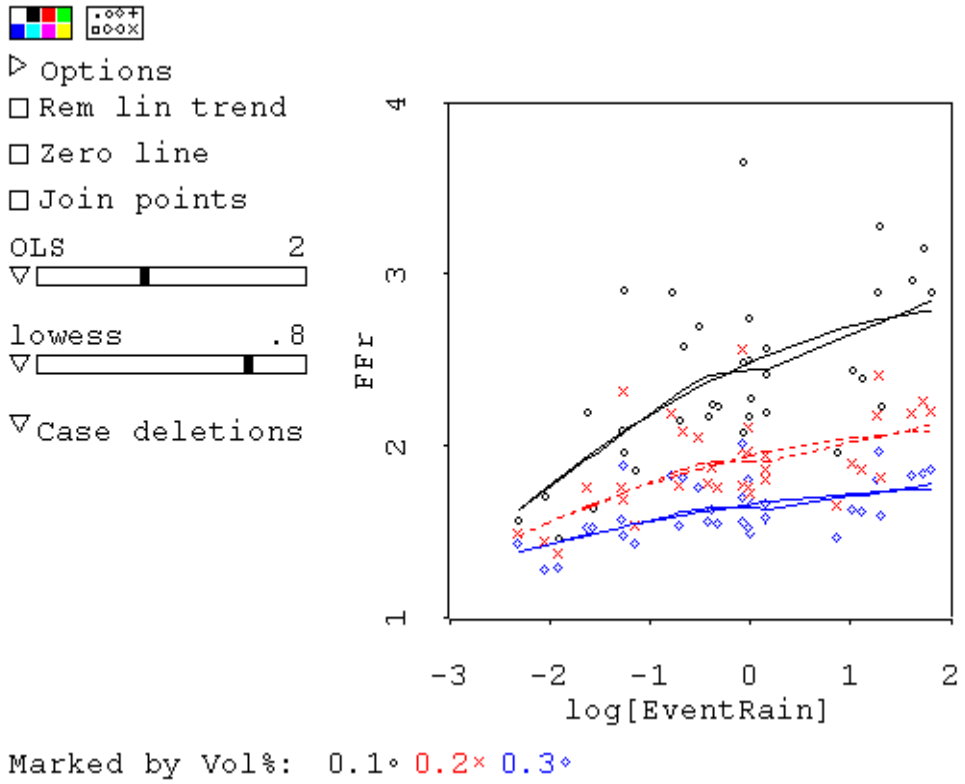


Figure 6.2.8 2nd-Order OLS and *lowess* Smoothing Fits for FF10 (black), FF20 (red), and FF30 (blue)

Table 6.2.1 Mass Removal Calculation

Option	Event Group	**S**	**MS**	**ML**	**L**
	Unit Event Vol	0.21	0.57	1.00	3.64
d0	Unit Vol Treated	0.00	0.00	0.00	0.00
	Unit Mass Rem	0.00	0.00	0.00	0.00
d1	FF10	**1.96**	**2.33**	**2.49**	**2.74**
	Unit Vol Treated	0.02	0.06	0.10	0.36
	Unit Mass Rem	0.04	0.13	0.25	1.00
d2	FF20	**1.66**	**1.86**	**1.95**	**2.07**
	Unit Vol Treated	0.04	0.11	0.20	0.73
	Unit Mass Rem	0.07	0.21	0.39	1.51
d3	FF30	**1.50**	**1.61**	**1.66**	**1.74**
	Unit Vol Treated	0.06	0.17	0.30	1.09
	Unit Mass Rem	0.09	0.27	0.50	1.90

Table 6.2.2 Efficiency Calculation for Selecting BMPs

Event Group	S	MS	ML	L	
BMP Decision	**d3**	**d3**	**d2**	**d1**	Efficiency %
Mass Removal	0.09	0.27	0.39	1.00	1.75/5.41 =32%
Volume Treated	0.06	0.17	0.20	0.36	0.79/5.41 =15%

6.3 New First Flush Notations and Criteria

6.3.1 Background

6.3.1.1 Observations of First Flush Phenomenon

The most persuasive evidence that the first flush "believers" provided to document the first flush phenomenon is simple observations. For example we can confirm their observations, especially for the first event of a season, because of the very dark and turbid runoff during the beginning of an event. The disbelievers may discredit these observations because they cannot verify the first flush from the collected data using their criteria. Both positions are acknowledged, but there must be some unexplored criteria that can unify their conclusions. The criteria or definition of a first flush is a candidate. The color or turbidity in the runoff reflects pollutant concentration, which suggests that the believers observe a decreasing trend in pollutant concentration. For most unbelievers, their first flush criteria are probably based on Mass-Loading concepts. The interpretation of this type of criteria: "the main portion of total pollutant mass load is transported during the early part of a storm event". This type of criteria relates to mass load fraction. The mass fraction can never be evaluated by simple observation and requires mathematical analysis. So we realize that the believers and the disbelievers focus on different aspects on the first flush phenomenon: the concentration aspect versus the mass load aspect.

6.3.1.2 Concentration Aspects vs. Mass Load Fraction Aspects

Using concentration criteria could show results for the first flush effect than using mass load fractions. For example, when the concentration declines rapidly in a certain absolute amount of runoff volume, and stabilizes at a lower value, storm event size will affect the FF mass fraction. Large events will have lower FF ratios for any given volume fraction, even though the declining concentration mechanism is independent of storm size. The only situation where FF ratios could be the same for different storm sizes would occur if the concentration decline were proportional to runoff volume.

When we recall the qualitative description about the first flush phenomenon - "the first part of runoff in a storm event is the most polluted", concentration aspects actually are more meaningful than mass load fraction aspects because we generally define the pollution from concentration.

Concentration aspects could also be better than mass load fraction aspects from the data requirement point of view. We have shown that it is not easy to get reliable FF mass load curves using only a few grab samples; however, concentration aspects can be focused on just the partial event profile instead of the whole profile, in contrast to mass load fraction aspects. Thus, taking a few grab samples could efficiently describe an event's partial information for studying first flush.

When considering BMPs, concentration aspects have many advantages over mass aspects. Theses advantages lost when performing first flush characterizations using dimensionless analysis. For example, consider two equal rainfall events, A and B. Event A discharges 50% of the pollutant mass in the first 20% of the runoff volume, and Event

B discharges 30% of the pollutant mass in the same volume fraction. Using mass loading criteria, Event A is a more significant first flush effect than Event B. The actual pollutant masses are not known, since the masses are normalized. Suppose that the pollutant concentration in event B is five times higher than in event A. In this case, if a BMP were used that was limited by runoff volume, it would be more beneficial to treat Event B even though it has a lower mass FF ratio that Event A.

Another example shows the difficulty of making BMP decisions from mass load aspects. Assume there is an event that produces a significant first flush; say 80 percent of pollutant mass in first 30 percent of runoff volume. However, if the washed-off pollutant concentration is always below a critical or regulatory limit, there is little value in treating the runoff. In this case it would be better to treat a more harmful pollutant, even if it does not exhibit a first flush.

In this section, we develop the concentration-based first flush criteria and notation. The purpose of developing the new first flush notation is to provide a less controversial and more applicable notation for quantifying first flush phenomenon.

6.3.2 Methodology

The concentration-based first flush notation could characterize first flush phenomenon as "the concentration decreases during the event". From our collected data, we have found that several pollutants routinely show this decreasing concentration trend, such as COD, oil & grease, DOC, TKN, dissolved phosphorus, and dissolved copper, etc. Figure 6.3.1 shows the decreasing trends. The data for this graph is the pooled grab sample data from all events, divided into groups based on corresponding cumulative rainfall. Although some concentrations increased during an event, the magnitude and number of theses trends is very small compared to the general decreasing trends. We could conclude that a concentration-decreasing trend exists for these parameters.

There are several considerations for developing a new first flush notation. We could create a new first flush criterion to indicate whether or not first flush phenomenon occurs. This notation would be less useful that continuous index which is able to indicate much more than a binary result (yes or no). Secondly, this index should be standardized and flexible for all parameters and events so that we have a standard way to compare different cases. Next, the data requirement cannot be too intensive or hard to satisfy. Finally, it should be applicable for BMPs.

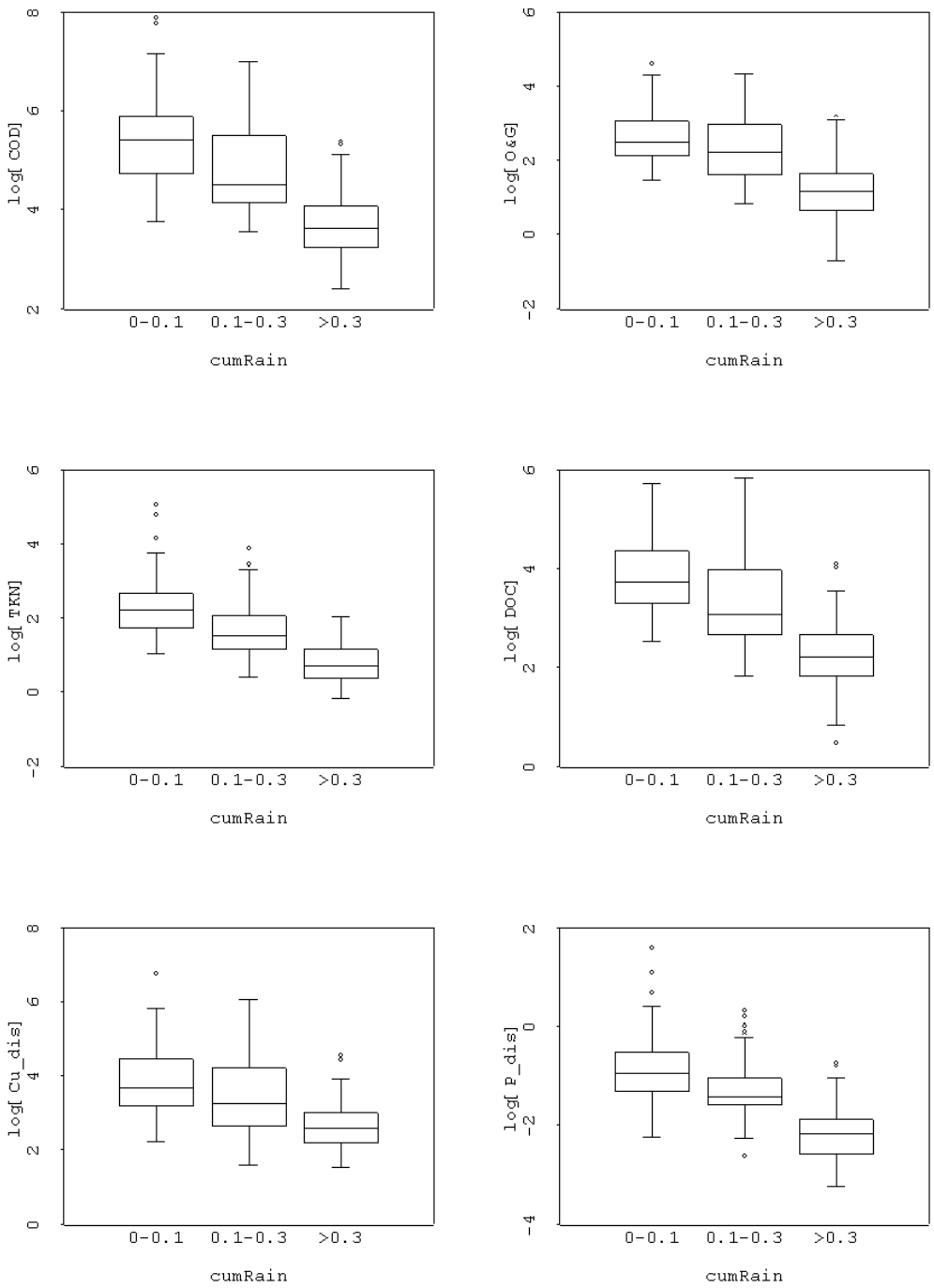


Figure 6.3.1 Concentration Downside Trends (log scale) Described by Cumulative Rainfall Depth (symbolized as cumRain on plot) in Inches

Based on the above principles, a new concentration-based first flush index is proposed from measured pollutant concentrations at a base condition and other reference conditions. The cumulative rainfall is selected to represent the reference condition because it can represent the time series, and it is normalized to the catchment area. The base level, where we measure the concentration for representing the initial condition for the event, is selected at 0.1-inch cumulative rainfall. The 0.1-inch cumulative rainfall is selected because this amount of rainfall should produce runoff for impervious surfaces. In order to characterize different size storms, two reference cumulative rainfall depths are selected. These two reference levels are 0.3-inch and 0.5-inch cumulative rainfall. The 0.3-and 0.5-inch cumulative rainfalls are selected because for BMP selection and rainfall frequency. Most rainfall events in the study area are less than 0.5 inch, and regulators are likely to require BMPs for this range of rainfall. To reduce error, replicate samples are suggest at each reference rainfall, and the actual rainfall might be flexibly defined (e.g., +/- 0.03 inch). Then, we transform each measured concentration into nature logarithm scale. The logarithm transformation is used to create a standardized way to accommodate different concentration ranges. Finally, a new concentration-based first flush notation (*cFF*) is suggested as below:

$$cFF_n^{0.3} = \frac{\sum_{i=1}^{n} \log(C_i^{0.1}) - \sum_{i=1}^{n} \log(C_i^{0.3})}{n}, \qquad cFF_n^{0.5} = \frac{\sum_{i=1}^{n} \log(C_i^{0.1}) - \sum_{i=1}^{n} \log(C_i^{0.5})}{n} \tag{6.3.1}$$

In *cFF* notation, the super script indicates the reference level at either 0.3-or 0.5-inch cumulative rainfall; *n* is the number of replicate samples taken at each rainfall level. This index is a simple subtraction of the average log-transformed concentration at the reference level from the average at the base level. The interpretation of this first flush index is analogous to the percentage drops for pollutants at reference levels.

6.3.3 Results & Analysis

6.3.3.1 Field Data Analysis

From our collected data in two seasons, there are 14 events that can be used for calculating *cFF*s. Among these events, nine events can be used for calculating $cFF^{0.3}$; seven events can be used for calculating $cFF^{0.5}$; and two events can be used for both. The number of samples for the base level (at 0.1-inch cumulative rainfall) is usually two or three, and the number for the reference level (at 0.3-or 0.5-inch cumulative rainfall) is usually one or two. Since the number of samples is not consistent for calculating *cFF*s, we will drop the subscript *n* for the moment. The above sample counts considered the tolerance range (within +/– 0.03-inch rainfall at each sample-taken level).

Figure 6.3.2 shows the $cFF^{0.3}$s for COD, dissolved copper, DOC, oil & grease, and dissolved phosphorous, where the event sequence is sorted by their $cFF^{0.3}$s of COD. For convenience of this study, we set $cFF^{0.3} >= 0.7$ (which equals the concentration decline of approximately 50 percent) as indicating "first flush confirmed" and $cFF^{0.3} >= 2$ (which equals the concentration decline of approximately 90 %) as indicating "significant first flush". There are two events showing "significant first flush" for COD, and four events

showing "first flush confirmed" for almost every parameter except for dissolved phosphorous ($cFF^{0.3} \sim 0.5$). The remaining five events do not show any obvious decreasing trends. Figure 6.3.3 shows the $cFF^{0.5}$s for the same parameters. All $cFF^{0.5}$s of nine events are greater than 0.7 (most are greater than one). The results of $cFF^{0.5}$s are higher than the previous results of $cFF^{0.3}$s as we expect.

Figure 6.3.4 shows the $cFF^{0.3}$s for solid related parameters, such as TSS, VSS, and turbidity, where the event sequence is sorted by their $cFF^{0.3}$s of TSS. There are four events for TSS and VSS and two events for turbidity showing "first flush confirmed". No event shows "significant first flush". There are three events showing negative values for each parameter. Figure 6.3.5 shows the $cFF^{0.5}$s for the same parameters. Basically, these $cFF^{0.5}$s show approximately the same range as the $cFF^{0.3}$s. Thus, it seems either that solid related parameters do not show a strong decreasing tendency in concentration, or that there are unconsidered factors, which control their washoff concentrations.

Figure 6.3.6 shows the comparison between two notations, $cFF^{0.3}$s and FF20, for the parameter COD. The $cFF^{0.3}$s are from the same events as above. The FF20s are the mean values that are calculated by the regression model (6.2.7), using the real data of the same events as calculating the $cFF^{0.3}$s. The event sequence in Figure 6.3.6 is sorted by their $cFF^{0.3}$s. We found that all FF20s are about the same (~ 1.8), and the corresponding $cFF^{0.3}$s show a wide value range of from 0.04 to 2.8. Thus, this concentration based criteria is more sensitive.

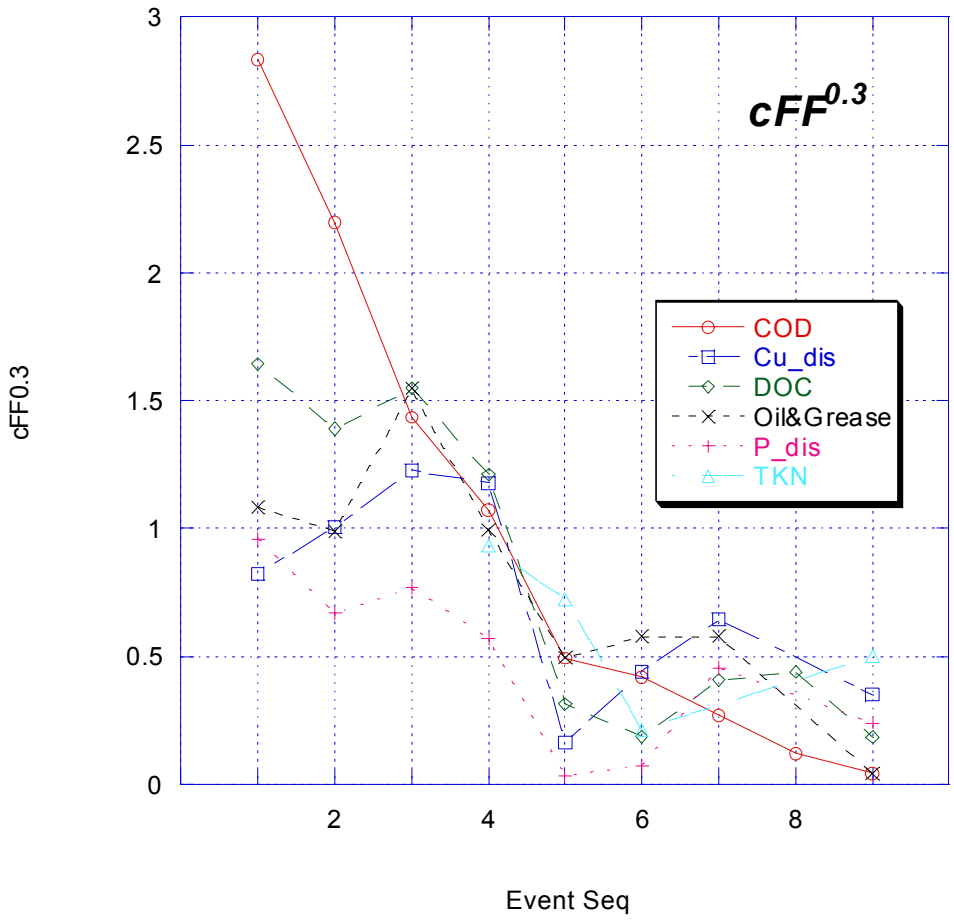


Figure 6.3.2 Field Observations for $cFF^{0.3}$s Showing First Flush

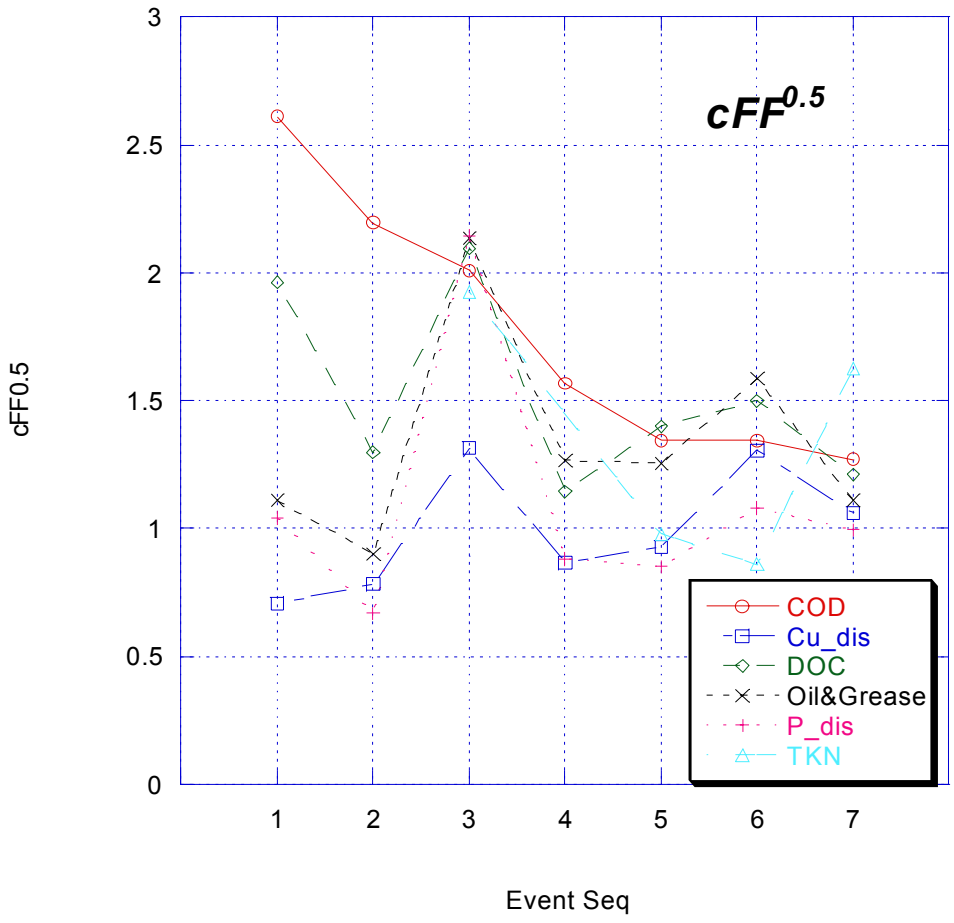


Figure 6.3.3 Field Observations for $cFF^{0.5}$s Showing First Flush

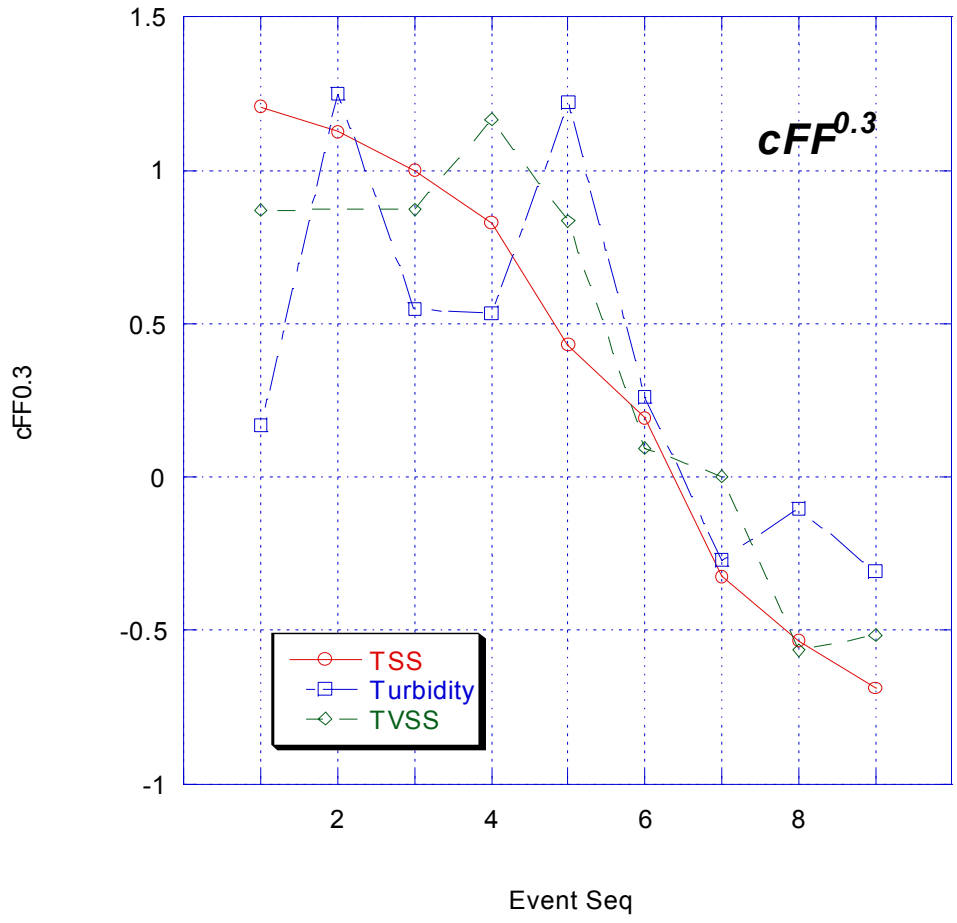


Figure 6.3.4 Field Observations for $cFF^{0.3}$s for Solid-Related Parameters

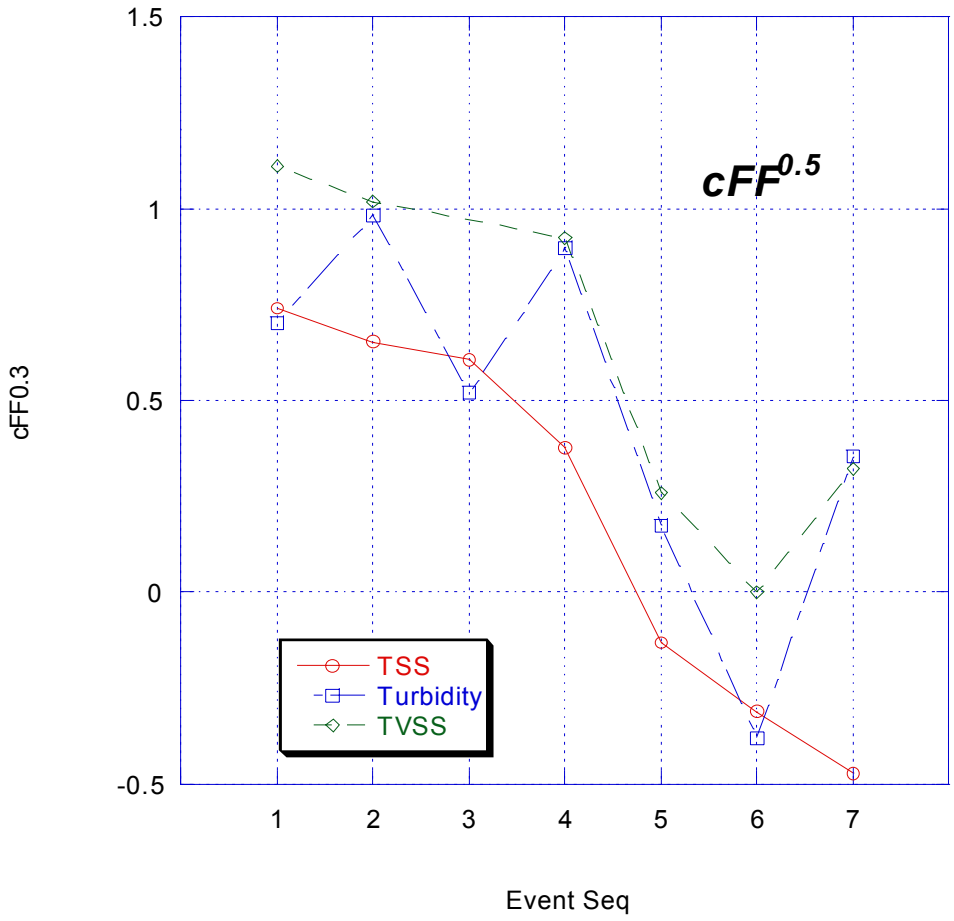


Figure 6.3.5 Field Observations for $cFF^{0.5}$s for Solid-Related Parameters

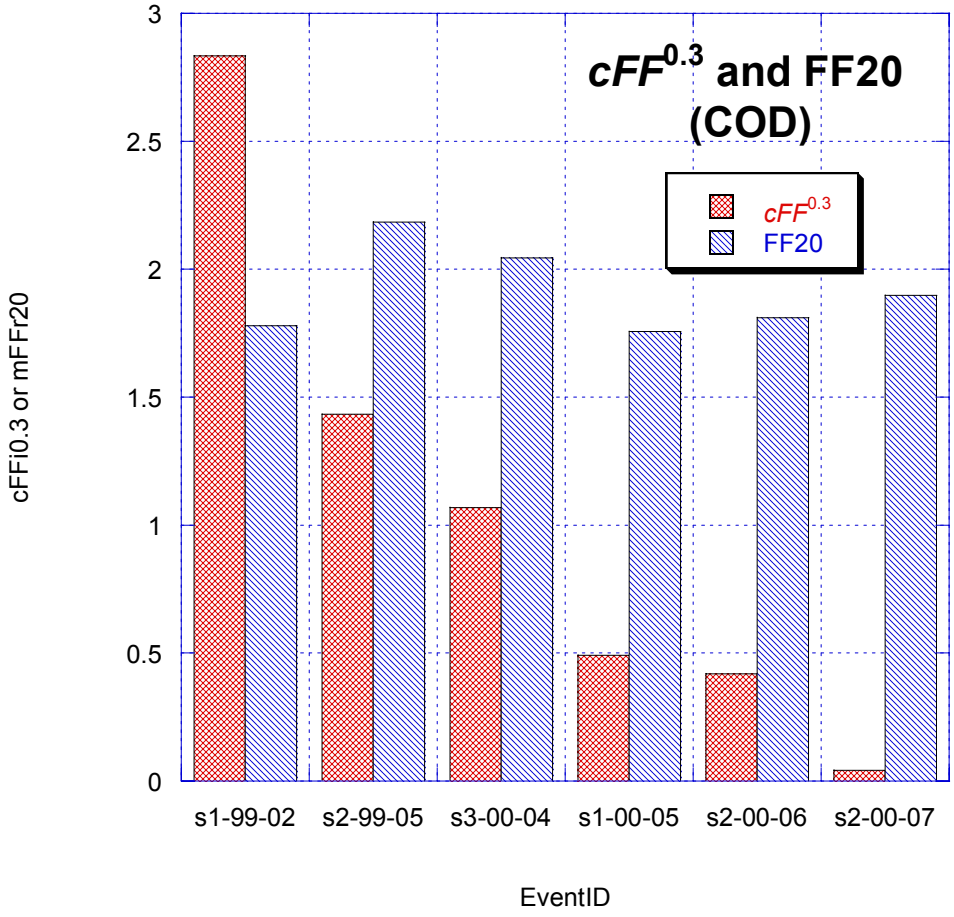


Figure 6.3.6 Comparison between FF20 and $cFF^{0.3}$

6.3.3.2 Implications for cFFs from Regression

The COD regression (6.2.7) accounts for antecedent dry days and antecedent rainfall, which implies that the mean of $cFF_n^{0.3}s$ or $cFF_n^{0.5}s$ should be independent of these two parameters. Based on (6.2.7), the mean of $cFF_n^{0.3}s$ is calculated as

$$\mu_{0.3} = -0.6 (\log 10 - \log 30) = 0.7$$

Similarly, the mean of $cFF^{0.5}$s is calculated as

$$\mu_{0.5} = -0.6 (\log 10 - \log 50) = 1.0$$

The variance of *cFF*s includes two random errors in (6.2.7). So let e_i and e_j be these two independent random errors. Then, from

$$\sigma^2 = Var(e_i - e_j) = Var(e_i) + Var(e_j) - 2Cov(e_i, e_j)$$

Since $Cov(e_i, e_j) = 0$ and $Var(e_i) = Var(e_j)$,

$$\sigma^2 = 2Var(e_i) = 2 \times 0.59^2 \cong 0.7 \text{, and } \sigma \cong 0.8$$

Based on the central limit theorem,

$$cFF_n^{0.3} \to N(0.7, 0.7/n) \text{, and } cFF_n^{0.5} \to N(1.0, 0.7/n) \qquad (6.3.2)$$

Equation (6.3.2) means that $cFF_n^{0.3}s$ will converge to a normal distribution with mean 0.7 and variance 0.7/n. Similarly, $cFF_n^{0.5}s$ will converge to a normal distribution with mean 1.0 and variance 0.7/n.

We evaluate the field results by using (6.3.2). In most cases, there is just one sample (n = 1) used for calculating cFFs. Thus, the 95% confidence interval for $cFF_1^{0.3}$ is from 1 to 2.3, and for $cFF_1^{0.5}$ is from -0.7 to 2.6. After checking the field results, only one from 16 cases is not within 95% confidence interval.

6.3.4 Discussion

Using cFF_n^x, we can judge whether there exists a decreasing concentration trend in pollutant concentrations, and we can also know exactly which part of a storm event is under examination. $cFF_n^{0.3}$ is preferred because of the large difference in rainfall in defining $cFF_n^{0.5}$. The reason is that there is more possibility that the concentration trend is not a monotone in the interval covered by $cFF_n^{0.5}$. Thus, $cFF_n^{0.5}$ is should be used as an additional tool to extend the rainfall interval for examination.

From the stormwater management point of view, cFF_n^x can provide simple answers for decision makers. For example, if cFF_n^x shows a first flush, decision makers choose a criterion among first 0.3 and 0.5-inch rainfall depth, and send the corresponding flow for treatment based on the efficiency. Other rainfall values could be used by interpolation. This type of answer is clear and easy to follow. If cFF_n^x does not show first flush,

decision makers at least know that the first part of a storm event is not special in emitting a pollutant, and they can use the average concept to consider the whole profile of stormwater pollution.

From the watershed management point of view, decision makers need to use both information derived from cFF_n^x and information collected from point sources. The treatment cost efficiency (mass removal/cost) needs to be explicitly evaluated. Our field data suggests that the decreasing trend in pollutant concentration can be used to derive the cost savings presented by a first flush.

One hypothetical scenario is used to illustrate the watershed management regarding of a receiving water body. In this drainage basin, there is one point source (**P1** that might represent a wastewater treatment plant) and two non-point sources (**NP1** and **NP2** that might represent two areas with different land use types). The pollutant oil & grease is being evaluated and load reductions are required. The point source **P1** shows a constant loading contribution. The non-point sources, **NP1** shows significant first flush phenomenon as indicated by $cFF_{10}^{0.3}$ and $cFF_{10}^{0.5}$, but **NP2** does not. The decision makers' job is to set a priority list to treat or abate the sources, based on the cost of treatment. Since **NP2** does not show a first flush, and there is no further information to reveal the concentration trend, **NP2** is treated the same as **P1**, which has constant oil & grease contribution. Therefore they just have one average unit cost estimate. For **NP1**, there are four unit cost estimates corresponding to two first flush costs, one residual cost, and one average cost. The assumed first treatment unit cost for **NP1** are 0.6 kg/$1 for first 0.3-inch rainfall, 0.4 kg/$1 for first 0.5-inch rainfall, and $1 for residual (> 0.5-inch rainfall).

The average unit costs for both **NP1** and NP2 are 0.15 kg/$1. The average unit cost for **P1** is 0.2 kg/$1. Table 6.3.1 shows the above cost structures. After comparing the efficiencies, the order of treatment priority should be **NP1** (first 0.5-inch), **P1**, **NP2**, and **NP1** (residual).

Table 6.3.1 Hypothetical Treatment Unit Costs

Target Pollutant	Oil & Grease		
Source Type	Non-point Source		Point Source
Source ID	**NP1**	**NP2**	**P1**
First Flush?			
$cFF^{0.3}$	Yes	No	
$cFF^{0.5}$	Yes	No	
Mass Removal Efficiency:			
First 0.3 inch	0.6 (kg/$1)		
First 0.5 inch	0.4 (kg/$1)		
Residual (>0.5 in)	0.1 (kg/$1)		
Average	0.15 (kg/$1)	0.15 (kg/$1)	0.2 (kg/$1)

6.4. Conclusions

Based upon the field data and analysis, the first flush phenomenon probably occurs for most pollutants. The pollutant source is greater at the beginning of a storm and if the washed off mass is proportional to the accumulated mass in the catchment, there must be a declining concentration. The lack of consensus over the existence of a first flush probably exists because of the difficulty in measuring and documenting the first flush, as well as differences in definitions. For example the equal volume sampling strategy missed the first flush for a small number of samples (~10).

Thus, a standard and easy-to-follow measuring protocol is needed for studying first flush. This measuring protocol needs to be capable of verifying and quantifying the first flush. In addition, it needs to produce useful information for engineers who are abating stormwater pollution. Otherwise, the first flush is a topic of only academic interest.

First flush Mass-Load curves and related first flush ratios, which are based on the concept of mass load, have been frequently used in quantifying first flush. However, they produce some confusion about understanding first flush. This confusion is partially caused by improper discretization of a continuous Mass-Load curve. In generally xx discrete samples are need to minimize the error in creating a Mass-Load curve, which is probably much higher than used in most monitoring programs. Additionally, the first flush Mass-Load curve is dimensionless, which is sometimes difficult for practicing engineers to use.

The concentration-based first flush notation, cFF_n^x, was developed to provide a less confusing, more reliable, and less expensive way to quantify first flush. The cFF_n^x ratio

quantifies a pollutant's change in concentration from the beginning of a storm event to a specific point later in the event. The cFF_n^x ratio can just be calculated from grab samples. One advantage for cFF_n^x is that it does not need intensive field samples for evaluation; the entire runoff period need not be sampled. Five samples for each rainfall depth are proposed. The cFF_n^x ratio can also assist decision makers determine if the first portion of a storm event is a promising candidate for treatment. The cFF_n^x ratio does not provide information to examine the whole profile of a storm event.

For future research, the mechanism of first flush can be studied after the definition of first flush is clarified. It may be helpful to reexamine previously published data to clarify the existence of a first flush.

7. CONCLUSIONS

This dissertation has showed the procedures and results of highway runoff monitoring, the exploration of a regression model for water quality in highway runoff, and the evaluation of methods for estimating EMCs and first flush phenomenon. Although highway was under investigation in this dissertation, any catch-basin with high imperviousness and short transportation lag could respond similarly to the highway surface.

The overall research framework takes advantage of stochastic methods to reduce the complexity of problems. This approach was successful in complicated runoff systems where the pure deterministic approach is difficult because of its intensive data requirement. In addition, a stochastic approach was also used to estimate the errors from the discretization of time series variables.

Stormwater monitoring is very expensive and labor consuming, but produces large quantities of data. To efficiently use the monitoring results, it was necessary to compile the existing data into database that facilitated its storage and retrieval. A monitoring database could be established at the watershed level to facilitate comprehensive watershed master planning. Recording the corresponding field conditions, such as the rainfall and flow, will extend the usage of monitoring results for further study.

The flow-weighted average concentration, called the event mean concentration (EMC) is currently used for estimating stormwater mass loading, and total maximum daily loads (TMDLs), are established using EMCs. This dissertation has discussed the approaches to properly measure and estimate EMCs. EMCs can be estimated from a series of grab samples or by flow-weighted composite samplers. Simulations show that a composite sample is a better way to obtain the EMCs due to the large number of individual samples.

Understanding the first flush phenomenon will benefit BMP selection for stormwater pollution control, because the best mass removal efficiency can be potentially achieved by treating the first flush volume. However, many factors affect the evaluation of these first flush BMPs. For example, there is a large data requirement for properly analyzing the first flush phenomenon using the current notations. Controversy exists in documenting the first flush phenomenon, which may be caused by incomplete data sets. Additionally, runoff systems vary a lot, and it is difficult develop protocols which adequately address all situations. An improved, concentration-based notation, which was designed to require less data, is provided, and shows applicability for BMP selection. Future BMP selection should utilize the improved understanding of the first flush phenomenon; otherwise, the phenomenon is only of academic interest.

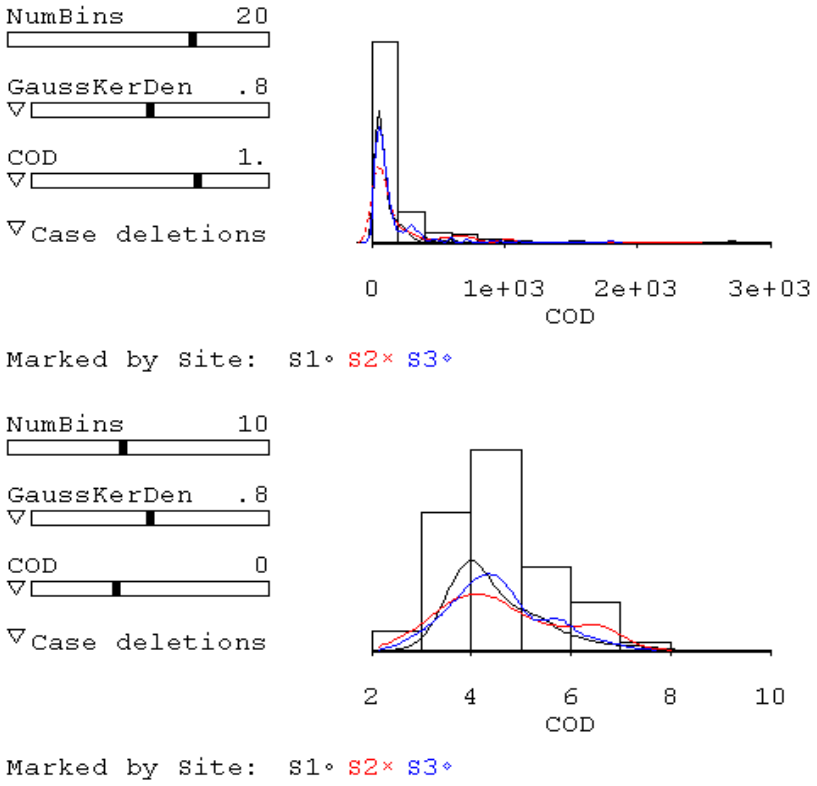
APPENDIX A. Water Quality Monitoring Results

A.1: COD

Summary Statistics:

COD (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	11.1	15.8	11.1	11.1
Maximum	2714.3	2714.3	2381.0	1800.0
Median	80.6	66.7	80.6	83.3
Mean	187.6	174.1	222.7	161.5
Standard Dev	295.9	332.4	322.6	225.9

Histograms plus Site's Density Curves Using Normal and Log Scales:

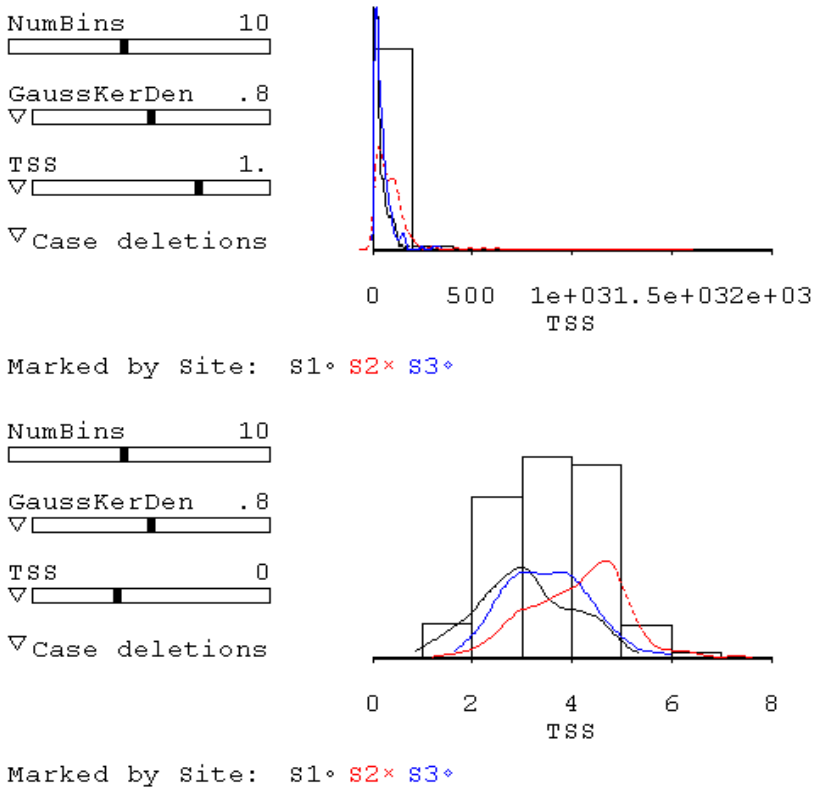


A.2: TSS

Summary Statistics:

TSS (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	2.9	2.9	4.5	6.3
Maximum	1534.7	174.7	1534.7	331.6
Median	37.4	20.6	68.3	31.8
Mean	62.3	33.4	97.9	47.7
Standard Dev	98.2	33.2	144.8	49.1

Histograms plus Site's Density Curves Using Normal and Log Scales:

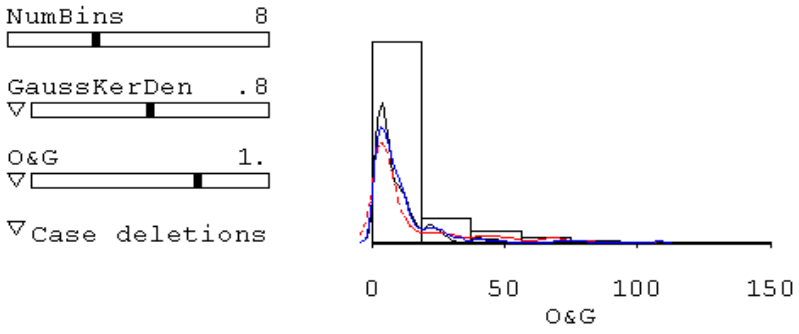


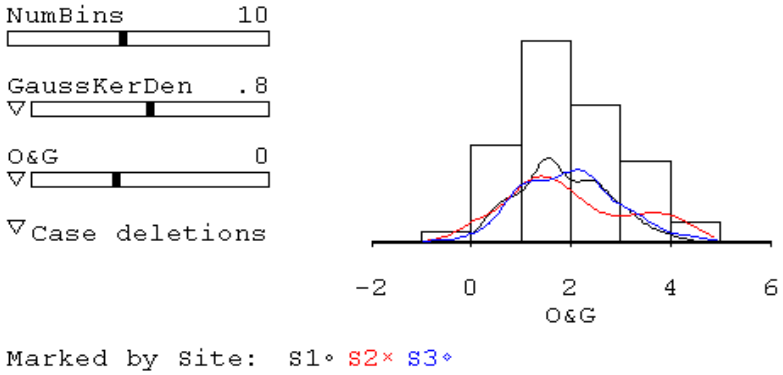
A.3: O & G

Summary Statistics:

O&G (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	0.5*	1.3	0.5	0.5
Maximum	108.0	73.0	102.3	108.0
Median	6.2	5.6	5.9	7.3
Mean	12.6	9.8	15.5	11.7
Standard Dev	16.7	10.6	21.0	15.1

Histograms plus Site's Density Curves Using Normal and Log Scales:



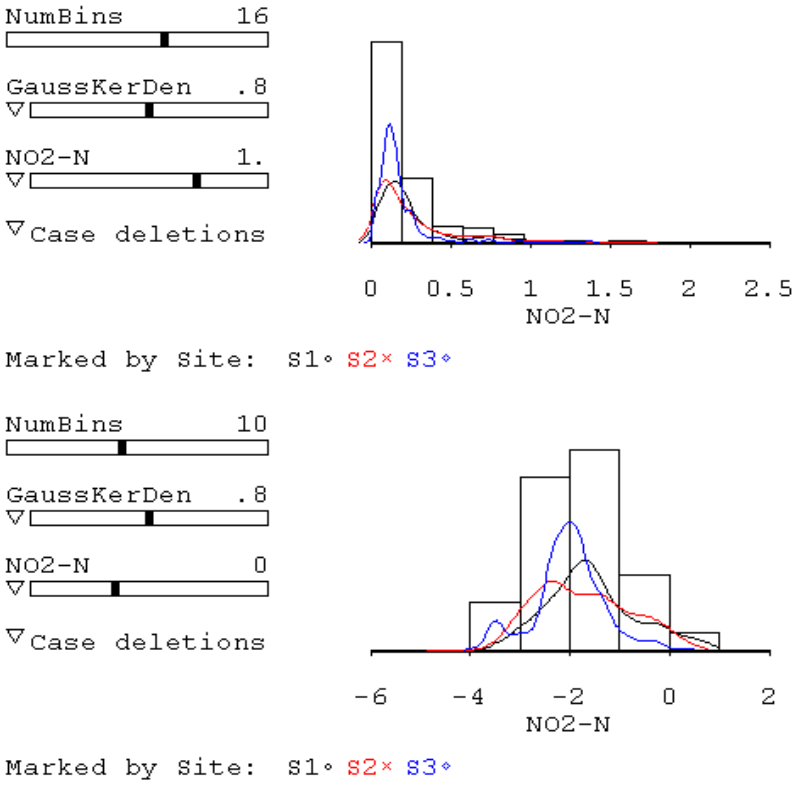


A.4: Nitrite

Summary Statistics:

NO2-N (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	0.01	0.03	0.01	0.02
Maximum	2.22	2.22	1.72	1.37
Median	0.16	0.19	0.16	0.14
Mean	0.26	0.33	0.28	0.18
Standard Dev	0.30	0.40	0.30	0.17

Histograms plus Site's Density Curves Using Normal and Log Scales:

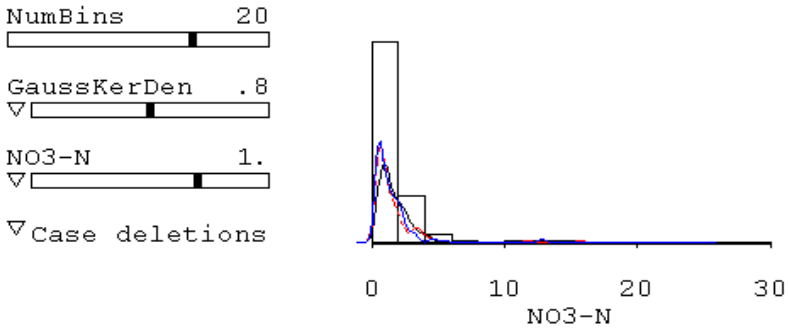


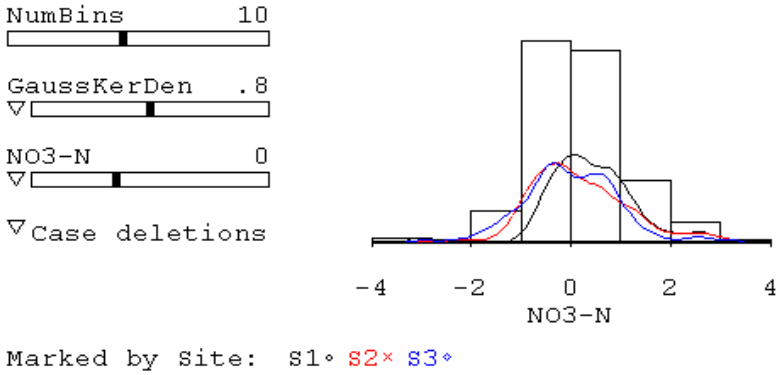
A.5: Nitrate

Summary Statistics:

NO3-N (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	0.05*	0.05	0.05	0.05
Maximum	24.74	15.73	20.89	24.74
Median	1.13	1.37	0.97	0.94
Mean	1.97	2.26	2.08	1.63
Standard Dev	2.86	2.71	3.14	2.64

Histograms plus Site's Density Curves Using Normal and Log Scales:



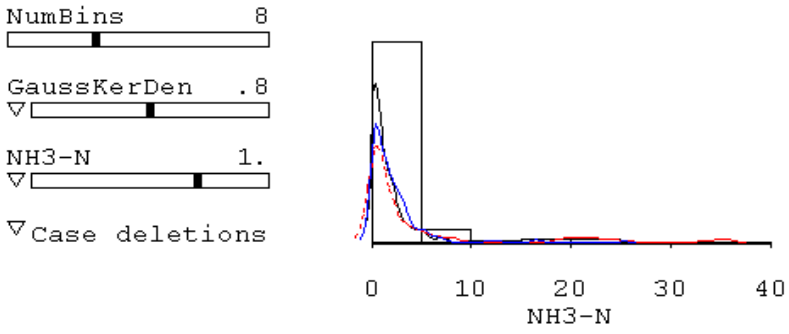


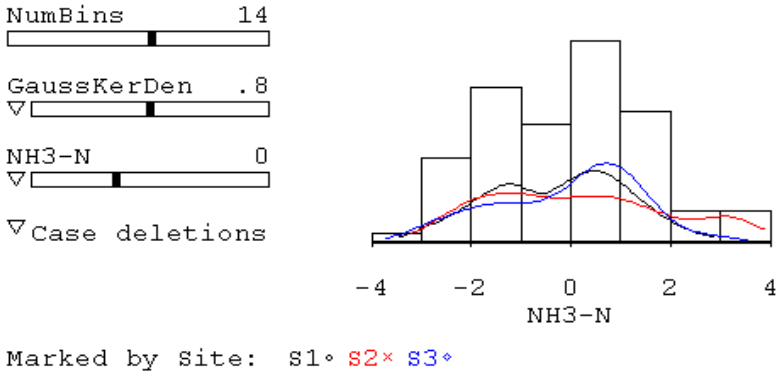
A.6: Ammonia

Summary Statistics:

NH3-N (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	441	123	163	155
Minimum	0.03	0.04	0.06	0.03
Maximum	35.96	13.69	35.96	25.41
Median	1.07	0.94	1.04	1.31
Mean	3.04	1.69	4.94	2.13
Standard Dev	5.93	2.27	8.72	3.21

Histograms plus Site's Density Curves Using Normal and Log Scales:



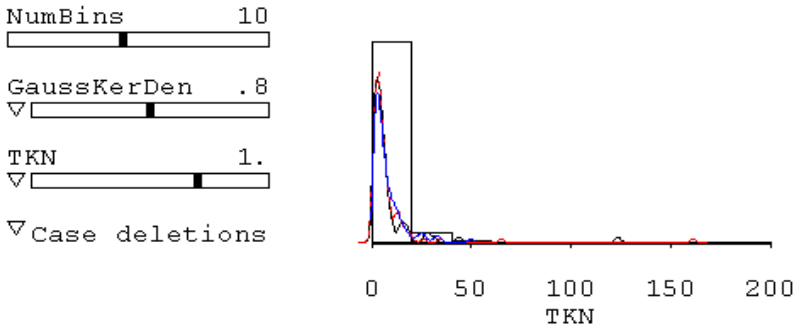


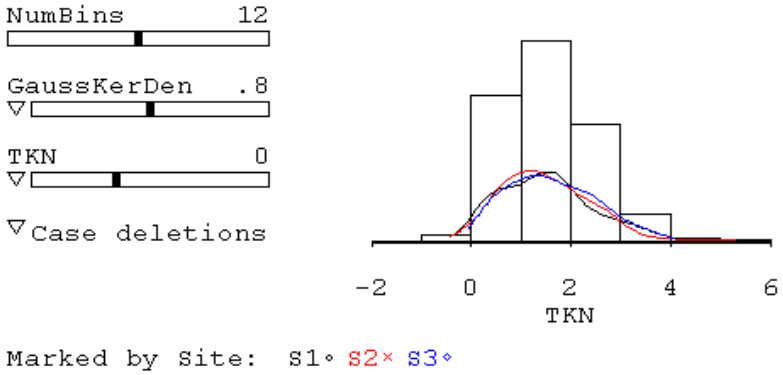
A.7: TKN

Summary Statistics:

TKN (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	219	63	82	74
Minimum	0.87	0.97	0.87	1.14
Maximum	161.30	123.60	161.30	50.02
Median	4.23	4.43	3.97	4.67
Mean	8.47	8.88	8.75	7.82
Standard Dev	15.54	16.77	19.16	8.63

Histograms plus Site's Density Curves Using Normal and Log Scales:



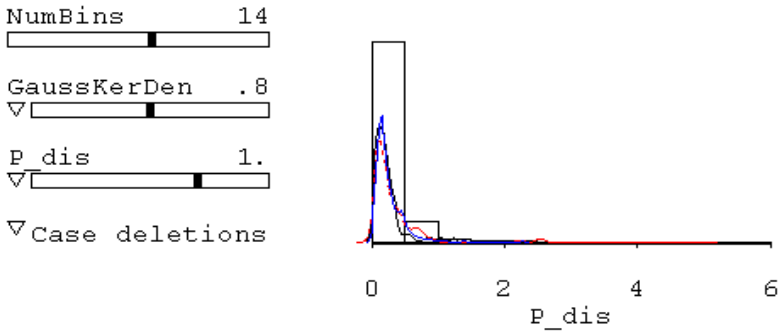


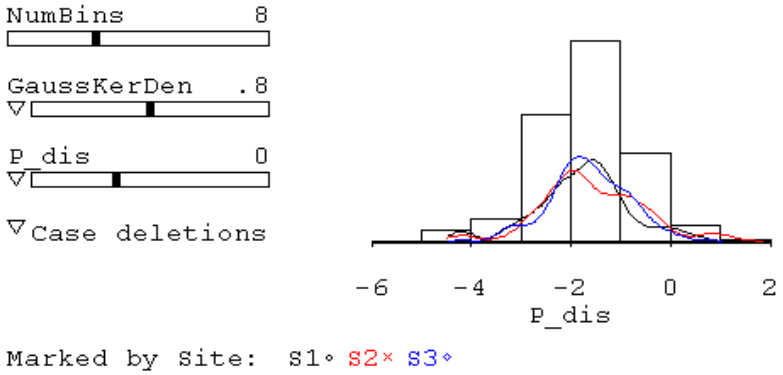
A.8: Dissolved Phosphorous

Summary Statistics:

P_dis (mg/l)	Pool	Site 1	Site 2	Site 3
N of cases	434	121	162	151
Minimum	0.02*	0.02	0.02	0.02
Maximum	4.97	2.02	4.97	2.28
Median	0.18	0.19	0.18	0.18
Mean	0.31	0.27	0.39	0.27
Standard Dev	0.44	0.31	0.62	0.27

Histograms plus Site's Density Curves Using Normal and Log Scales:



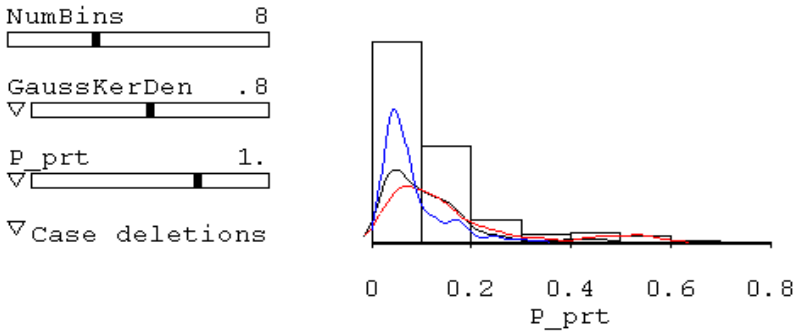


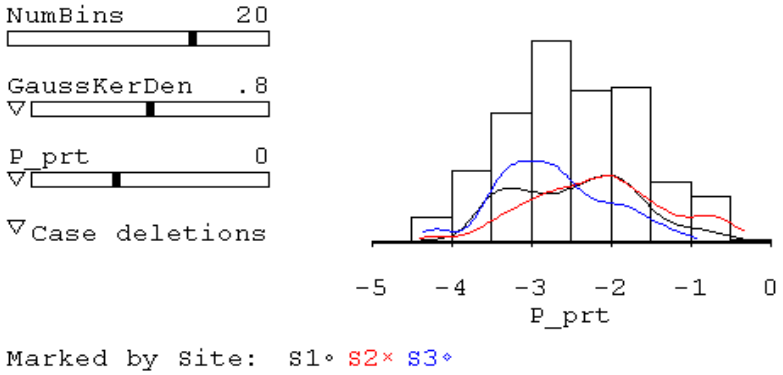
A.9: Particulate Phosphorous

Summary Statistics:

P_prt (mg/l)	Pool	Site 1	Site 2	Site 3
hN of cases	417	121	156	140
Minimum	0.02*	0.02	0.02	0.02
Maximum	0.68	0.68	0.61	0.34
Median	0.08	0.10	0.12	0.06
Mean	0.12	0.12	0.16	0.08
Standard Dev	0.12	0.11	0.15	0.06

Histograms plus Site's Density Curves Using Normal and Log Scales:



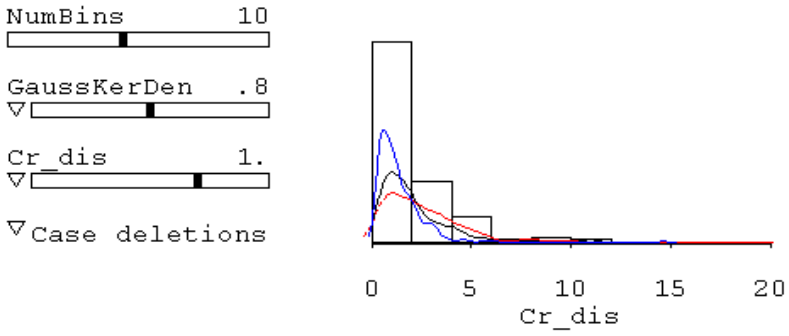


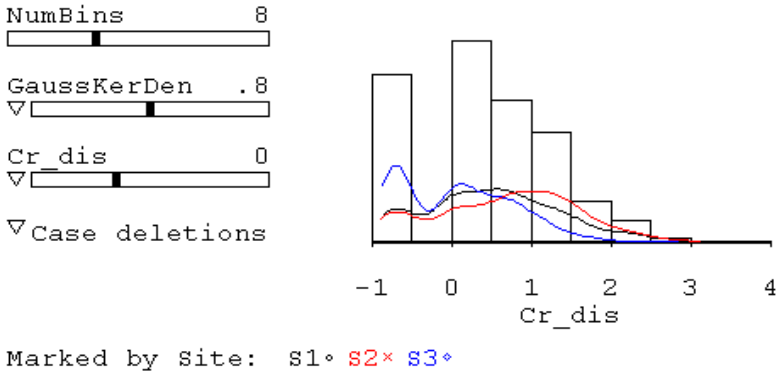
A.10: Dissolved Chromium

Summary Statistics:

Cr_dis (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	0.5*	0.5	0.5	0.5
Maximum	19.3	11.1	19.3	14.6
Median	1.5	1.6	2.0	1.0
Mean	2.2	2.3	2.8	1.4
Standard Dev	2.2	2.2	2.6	1.5

Histograms plus Site's Density Curves Using Normal and Log Scales:



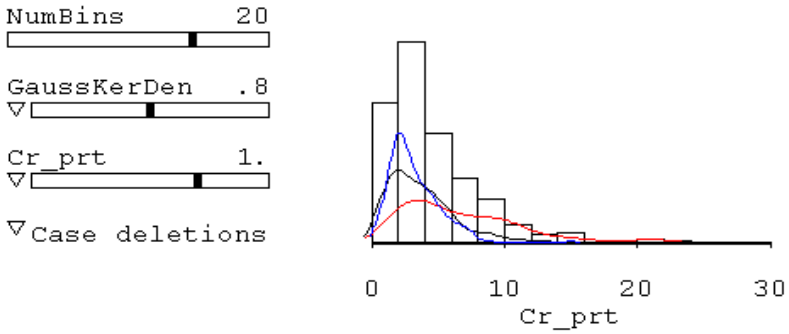


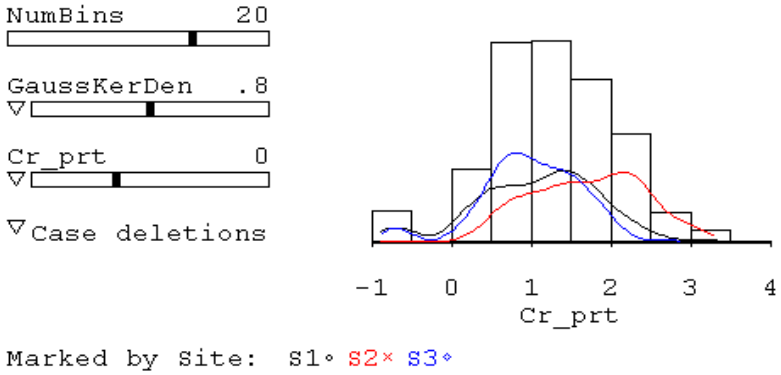
A.11: Particulate Chromium

Summary Statistics:

Cr_prt (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	417	121	156	140
Minimum	0.5*	0.5	0.5	0.5
Maximum	24.3	24.3	22.4	15.0
Median	3.7	3.3	5.6	2.7
Mean	4.8	3.9	6.9	3.2
Standard Dev	3.9	3.1	4.7	2.0

Histograms plus Site's Density Curves Using Normal and Log Scales:



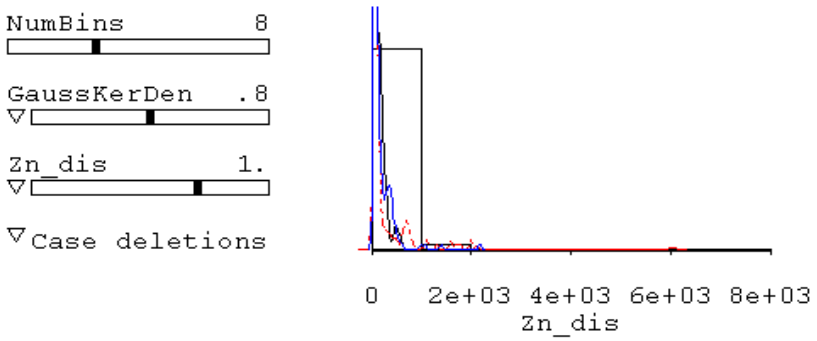


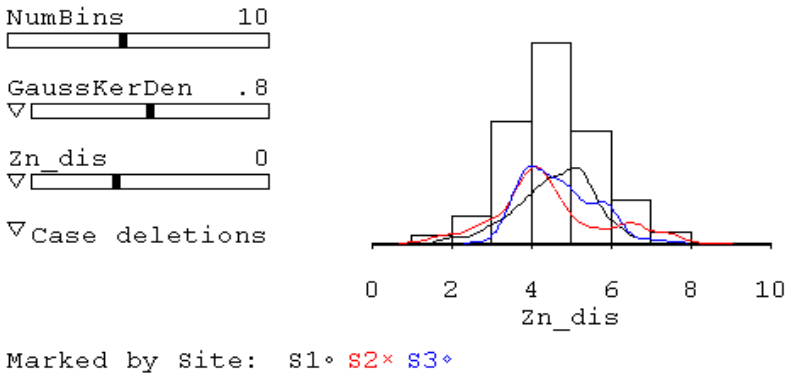
A.12: Dissolved Zinc

Summary Statistics:

Zn_dis (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	3.0	6.0	3.0	13.2
Maximum	6041.2	590.9	6041.2	2180.7
Median	86.0	105.5	69.0	102.2
Mean	202.8	142.4	268.9	180.5
Standard Dev	412.6	121.5	616.5	252.5

Histograms plus Site's Density Curves Using Normal and Log Scales:



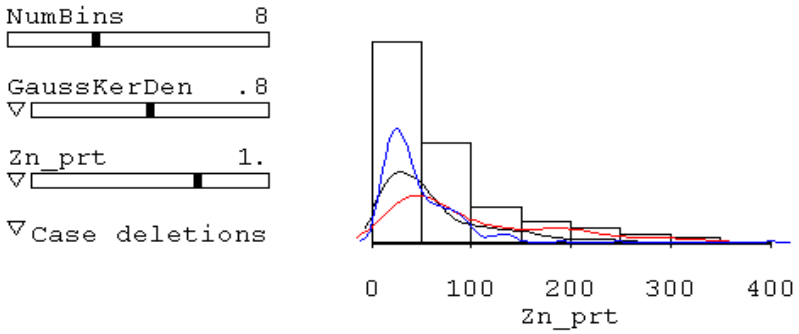


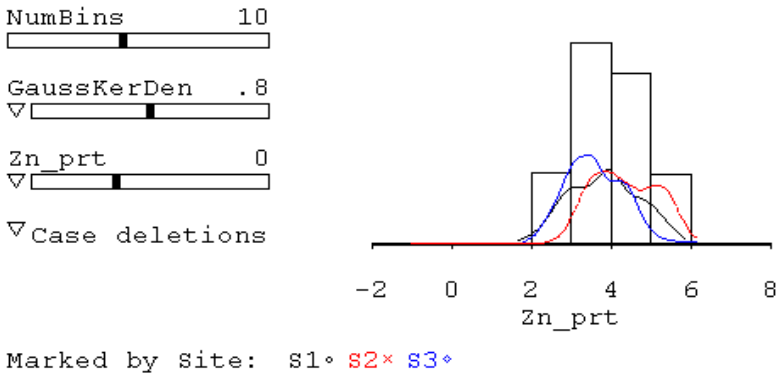
A.13: Particulate Zinc

Summary Statistics:

Zn_prt (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	417	121	156	140
Minimum	0.5*	6.5	0.5	7.5
Maximum	400.8	294.9	346.2	400.8
Median	48.4	47.2	73.6	34.7
Mean	74.9	65.8	104.6	49.5
Standard Dev	69.7	58.7	82.5	47.6

Histograms plus Site's Density Curves Using Normal and Log Scales:



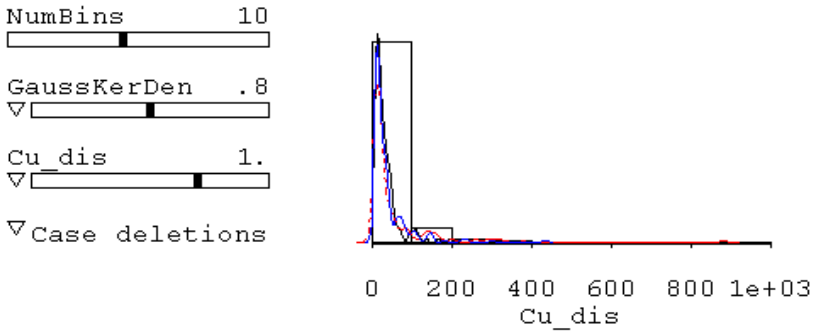


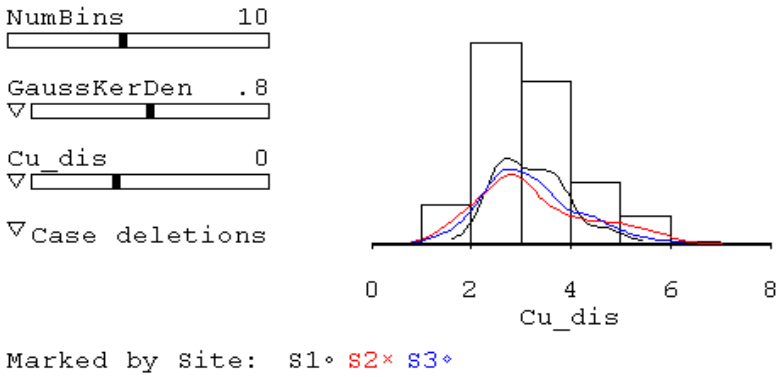
A.14: Dissolved Copper

Summary Statistics:

Cu_dis (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	3.0	6.0	3.0	3.0
Maximum	882.4	198.6	882.4	435.1
Median	21.0	23.0	20.2	21.0
Mean	44.4	33.6	56.2	40.5
Standard Dev	71.0	31.8	98.1	56.1

Histograms plus Site's Density Curves Using Normal and Log Scales:



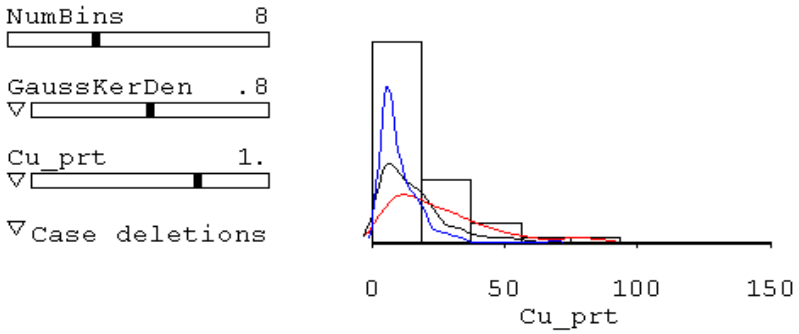


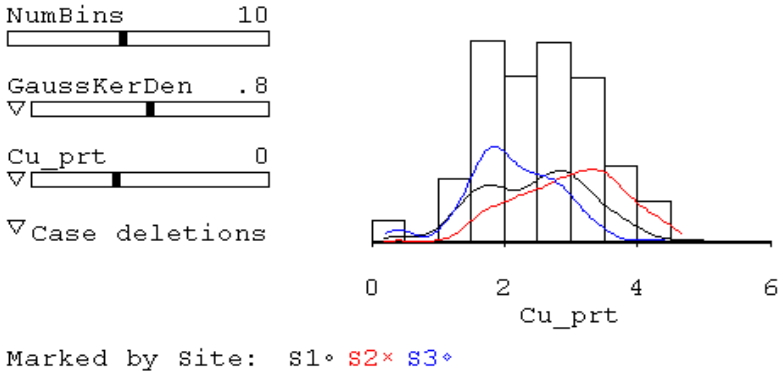
A.15: Particulate Copper

Summary Statistics:

Cu_prt (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	417	121	156	140
Minimum	1.5*	1.5	1.5	1.5
Maximum	107.8	107.8	88.1	68.9
Median	12.9	12.5	21.0	8.0
Mean	18.1	16.3	26.2	10.7
Standard Dev	16.7	14.8	19.9	8.4

Histograms plus Site's Density Curves Using Normal and Log Scales:



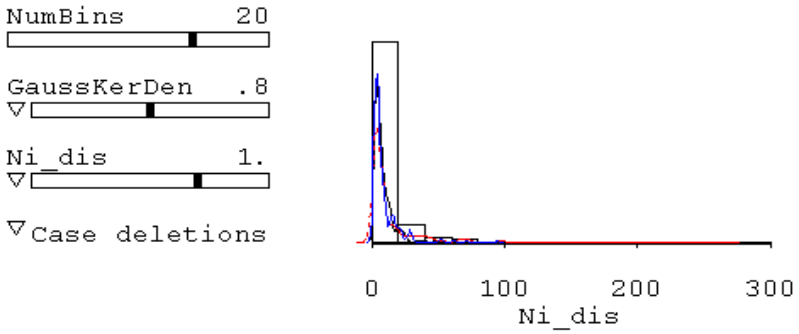


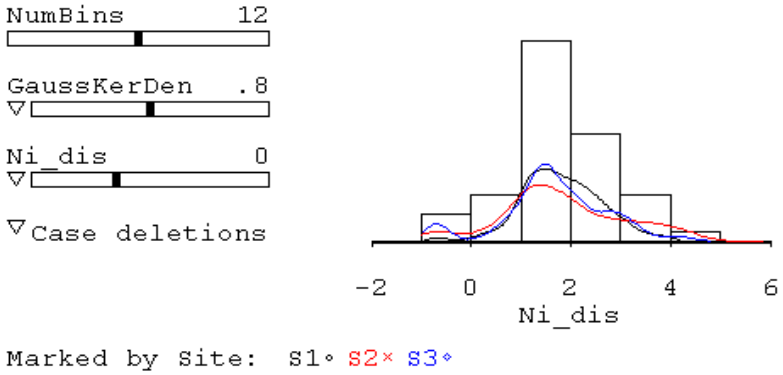
A.16: Dissolved Nickel

Summary Statistics:

Ni_dis	Pool	Site 1	Site 2	Site 3
N of cases	437	122	163	152
Minimum	0.5*	0.5	0.5	0.5
Maximum	264.2	52.6	264.2	94.9
Median	5.0	6.0	5.0	5.0
Mean	11.2	8.7	14.7	9.5
Standard Dev	18.9	8.6	27.1	12.9

Histograms plus Site's Density Curves Using Normal and Log Scales:



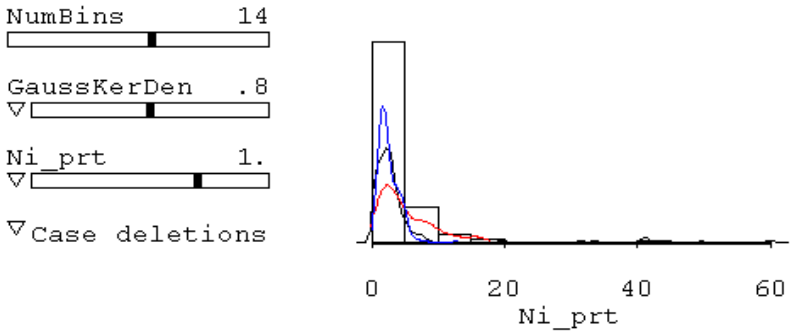


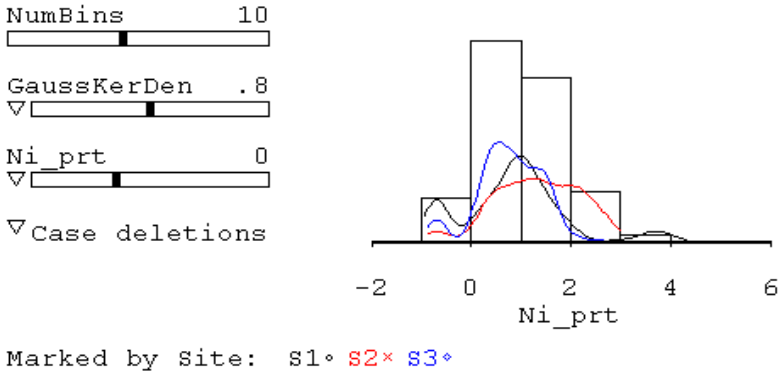
A.17: Particulate Nickel

Summary Statistics:

Ni_prt (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	417	121	156	140
Minimum	0.5*	0.5	0.5	0.5
Maximum	59.9	59.9	17.1	12.6
Median	2.7	2.4	3.8	2.2
Mean	4.3	5.0	5.2	2.6
Standard Dev	6.1	9.9	4.1	1.7

Histograms plus Site's Density Curves Using Normal and Log Scales:



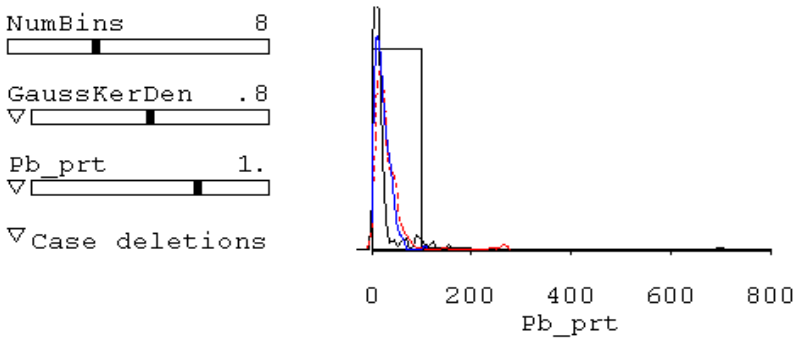


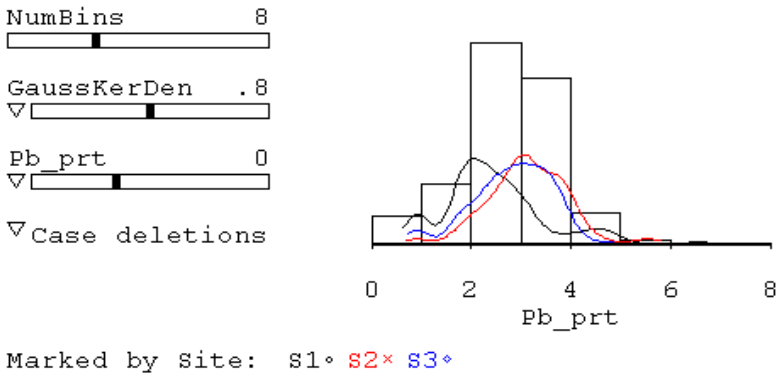
A.18: Particulate Lead

Summary Statistics:

Pb_prt (ug/l)	Pool	Site 1	Site 2	Site 3
N of cases	417	121	156	140
Minimum	2.5*	2.5	2.5	2.5
Maximum	700.7	700.7	268.0	107.8
Median	17.8	10.1	22.4	18.6
Mean	26.8	25.0	32.5	21.9
Standard Dev	44.1	67.6	37.5	15.5

Histograms plus Site's Density Curves Using Normal and Log Scales:





Appendix B. Dataset for COD Regression

SN	Site	SampleID	COD (mg/l)	cumR (inch)	RI (inch/15min)	AtDry (day)	AtR (inch)
1	s1	s1-1999-02-01	1509.98	0.07	0.01	8	0.05
2	s1	s1-1999-02-02	250	0.09	0.02	8	0.05
3	s1	s1-1999-02-03	944.44	0.11	0.02	8	0.05
4	s1	s1-1999-02-04	88.89	0.15	0.02	8	0.05
5	s1	s1-1999-02-05	66.67	0.18	0.03	8	0.05
6	s1	s1-1999-02-06	77.78	0.23	0.02	8	0.05
7	s1	s1-1999-02-07	41.67	0.3	0.04	8	0.05
8	s1	s1-1999-02-08	44.44	0.36	0.03	8	0.05
9	s1	s1-1999-02-09	52.78	0.42	0.04	8	0.05
10	s1	s1-1999-02-10	61.11	0.48	0.03	8	0.05
11	s1	s1-1999-02-11	44.44	0.52	0.02	8	0.05
12	s1	s1-1999-02-12	61.11	0.57	0.02	8	0.05
13	s1	s1-1999-02-13	63.89	0.6	0.02	8	0.05
14	s1	s1-1999-03-03	253.52	0.03	0.01	5	0.67
15	s1	s1-1999-03-04	216.67	0.03	0	5	0.67
16	s1	s1-1999-03-05	200	0.04	0.01	5	0.67
17	s1	s1-1999-03-06	197.22	0.07	0.01	5	0.67
18	s1	s1-1999-03-07	191.67	0.07	0	5	0.67
19	s1	s1-1999-04-01	64.12	0.14	0	10	0.1
20	s1	s1-1999-04-02	117.65	0.16	0.02	10	0.1
21	s1	s1-1999-04-03	117.65	0.17	0	10	0.1
22	s1	s1-1999-04-04	105.4	0.18	0.01	10	0.1
23	s1	s1-1999-04-05	107.97	0.18	0	10	0.1
24	s1	s1-1999-04-06	128.53	0.19	0	10	0.1
25	s1	s1-1999-04-07	102.83	0.27	0.01	10	0.1
26	s1	s1-1999-05-01	38.55	0.25	0.05	1	0.29
27	s1	s1-1999-05-02	36.14	0.28	0.04	1	0.29
28	s1	s1-1999-05-03	26.51	0.37	0.09	1	0.29
29	s1	s1-1999-05-04	26.51	0.49	0.12	1	0.29
30	s1	s1-1999-05-05	26.51	0.59	0.08	1	0.29
31	s1	s1-1999-05-06	24.1	0.68	0.08	1	0.29
32	s1	s1-1999-06-01	39.47	0.17	0.02	4	0.42
33	s1	s1-1999-06-02	39.47	0.19	0.02	4	0.42
34	s1	s1-1999-06-03	31.58	0.22	0.02	4	0.42
35	s1	s1-1999-06-04	36.84	0.23	0.01	4	0.42
36	s1	s1-1999-06-05	42.11	0.23	0	4	0.42
37	s1	s1-1999-06-c1	55.26	0.23	0	4	0.42
38	s1	s1-1999-06-c2	50	1.45	0	4	0.42
39	s1	s1-1999-07-01	91.57	0.07	0.03	4	1.6
40	s1	s1-1999-07-02	171.08	0.1	0.03	4	1.6
41	s1	s1-1999-07-03	57.83	0.12	0.02	4	1.6
42	s1	s1-1999-07-04	50.6	0.12	0	4	1.6
43	s1	s1-1999-07-c1	55.42	0.13	0.01	4	1.6
44	s1	s1-1999-09-01	35.71	0.07	0.02	2	1.8
45	s1	s1-1999-09-02	45.24	0.08	0.02	2	1.8
46	s1	s1-1999-09-03	38.1	0.09	0	2	1.8
47	s1	s1-1999-09-04	39.52	0.12	0.03	2	1.8
48	s1	s1-1999-09-05	32.38	0.14	0.01	2	1.8
49	s1	s1-1999-09-c1	26.19	0.15	0.01	2	1.8
50	s1	s1-1999-10-01	68.35	0.41	0.09	40	0.7
51	s1	s1-1999-10-02	63.29	0.47	0.09	40	0.7

52	s1	s1-1999-10-03	60.76	0.52	0.08	40	0.7
53	s1	s1-1999-10-04	63.29	0.52	0	40	0.7
54	s1	s1-1999-10-05	63.29	0.52	0	40	0.7
55	s1	s1-1999-10-c1	65.82	0.52	0	40	0.7
56	s1	s1-2000-01-01	530.12	0.12	0.01	33.6	0.14
57	s1	s1-2000-01-02	626.51	0.14	0.02	33.6	0.14
58	s1	s1-2000-01-03	313.25	0.15	0.01	33.6	0.14
59	s1	s1-2000-01-04	171.05	0.19	0.04	33.6	0.14
60	s1	s1-2000-01-05	163.16	0.21	0.02	33.6	0.14
61	s1	s1-2000-01-06	145.27	0.35	0.06	33.6	0.14
62	s1	s1-2000-01-07	57.89	0.45	0.02	33.6	0.14
63	s1	s1-2000-01-08	71.05	0.53	0.02	33.6	0.14
64	s1	s1-2000-01-09	47.37	0.67	0.03	33.6	0.14
65	s1	s1-2000-01-10	50	0.93	0	33.6	0.14
66	s1	s1-2000-01-11	15.79	0.94	0	33.6	0.14
67	s1	s1-2000-02-01	2714.29	0.04	0.01	69	0.65
68	s1	s1-2000-02-02	952.38	0.07	0.01	69	0.65
69	s1	s1-2000-02-03	595.24	0.09	0.01	69	0.65
70	s1	s1-2000-02-04	500	0.11	0.01	69	0.65
71	s1	s1-2000-02-05	404.76	0.13	0.01	69	0.65
72	s1	s1-2000-02-06	404.76	0.15	0	69	0.65
73	s1	s1-2000-02-07	428.57	0.15	0	69	0.65
74	s1	s1-2000-03-01	95.24	0.09	0.04	1.9	0.15
75	s1	s1-2000-03-02	71.43	0.13	0.04	1.9	0.15
76	s1	s1-2000-03-03	85.71	0.16	0.03	1.9	0.15
77	s1	s1-2000-03-04	77.92	0.2	0.05	1.9	0.15
78	s1	s1-2000-03-05	57.14	0.25	0.04	1.9	0.15
79	s1	s1-2000-03-06	51.95	0.54	0.05	1.9	0.15
80	s1	s1-2000-03-07	46.75	0.77	0.07	1.9	0.15
81	s1	s1-2000-03-08	46.75	1.04	0.07	1.9	0.15
82	s1	s1-2000-03-09	33.77	1.34	0.09	1.9	0.15
83	s1	s1-2000-03-10	31.17	1.69	0.08	1.9	0.15
84	s1	s1-2000-03-11	25.97	2.11	0.08	1.9	0.15
85	s1	s1-2000-04-01	309.52	0.08	0.02	14.2	0.48
86	s1	s1-2000-04-02	238.1	0.09	0.01	14.2	0.48
87	s1	s1-2000-04-03	162.86	0.1	0.01	14.2	0.48
88	s1	s1-2000-04-04	140	0.1	0	14.2	0.48
89	s1	s1-2000-04-05	137.14	0.11	0.01	14.2	0.48
90	s1	s1-2000-04-06	85.71	0.17	0.01	14.2	0.48
91	s1	s1-2000-04-07	37.14	0.49	0.05	14.2	0.48
92	s1	s1-2000-04-08	57.14	0.51	0	14.2	0.48
93	s1	s1-2000-04-09	68.57	0.51	0	14.2	0.48
94	s1	s1-2000-05-01	255.81	0.01	0.01	5.4	4.8
95	s1	s1-2000-05-02	232.56	0.02	0.01	5.4	4.8
96	s1	s1-2000-05-03	163.89	0.04	0.02	5.4	4.8
97	s1	s1-2000-05-04	113.89	0.06	0.01	5.4	4.8
98	s1	s1-2000-05-05	100	0.08	0.02	5.4	4.8
99	s1	s1-2000-05-06	63.89	0.13	0.02	5.4	4.8
100	s1	s1-2000-05-07	44.44	0.26	0.02	5.4	4.8
101	s1	s1-2000-05-08	41.67	0.27	0	5.4	4.8
102	s1	s1-2000-05-09	57.22	0.28	0	5.4	4.8
103	s1	s1-2000-07-01	250.67	0.01	0.01	6	2.91
104	s1	s1-2000-07-02	213.33	0.02	0.01	6	2.91
105	s1	s1-2000-07-03	158.33	0.04	0.02	6	2.91
106	s1	s1-2000-07-04	105.56	0.07	0.03	6	2.91
107	s1	s1-2000-07-05	63.89	0.09	0.02	6	2.91
108	s1	s1-2000-07-06	38.89	0.2	0.02	6	2.91

109	s1	s1-2000-07-07	47.22	0.23	0.01	6	2.91
110	s1	s1-2000-07-08	50	0.26	0	6	2.91
111	s1	s1-2000-07-09	50	0.26	0	6	2.91
112	s1	s1-2000-07-10	55.56	0.26	0	6	2.91
113	s1	s1-2000-08-01	98.7	0.19	0	13.2	1.32
114	s1	s1-2000-08-02	93.51	0.19	0	13.2	1.32
115	s1	s1-2000-08-03	106.49	0.2	0.01	13.2	1.32
116	s1	s1-2000-08-04	85.71	0.26	0.06	13.2	1.32
117	s1	s1-2000-08-05	80.52	0.28	0.02	13.2	1.32
118	s1	s1-2000-08-06	57.14	0.32	0	13.2	1.32
119	s2	s2-1999-02-01	641.98	0.04	0	17	0.05
120	s2	s2-1999-02-02	691.36	0.05	0.01	17	0.05
121	s2	s2-1999-02-03	567.9	0.05	0.01	17	0.05
122	s2	s2-1999-02-04	617.28	0.05	0.01	17	0.05
123	s2	s2-1999-02-05	567.9	0.05	0	17	0.05
124	s2	s2-1999-02-06	617.28	0.05	0	17	0.05
125	s2	s2-1999-02-07	706.17	0.05	0	17	0.05
126	s2	s2-1999-02-08	814.81	0.05	0	17	0.05
127	s2	s2-1999-02-c1	814.81	0.05	0	17	0.05
128	s2	s2-1999-03-01	583.33	0.04	0.01	8	0.07
129	s2	s2-1999-03-02	555.56	0.05	0.02	8	0.07
130	s2	s2-1999-03-03	722.22	0.05	0.01	8	0.07
131	s2	s2-1999-03-04	111.11	0.05	0.01	8	0.07
132	s2	s2-1999-03-05	216.67	0.06	0.01	8	0.07
133	s2	s2-1999-03-06	138.87	0.07	0.01	8	0.07
134	s2	s2-1999-03-07	250	0.11	0.02	8	0.07
135	s2	s2-1999-03-08	38.89	0.21	0.02	8	0.07
136	s2	s2-1999-03-09	38.89	0.26	0.03	8	0.07
137	s2	s2-1999-03-10	30.56	0.34	0.03	8	0.07
138	s2	s2-1999-03-11	38.89	0.42	0.04	8	0.07
139	s2	s2-1999-03-12	38.89	0.51	0.04	8	0.07
140	s2	s2-1999-03-13	44.44	0.57	0.03	8	0.07
141	s2	s2-1999-03-14	50	0.63	0.03	8	0.07
142	s2	s2-1999-03-c1	63.89	0.69	0.01	8	0.07
143	s2	s2-1999-04-04	91.67	0.13	0.01	5	0.99
144	s2	s2-1999-04-05	66.67	0.17	0.03	5	0.99
145	s2	s2-1999-04-06	80.56	0.21	0.01	5	0.99
146	s2	s2-1999-04-07	102.78	0.22	0	5	0.99
147	s2	s2-1999-04-c1	122.22	0.23	0	5	0.99
148	s2	s2-1999-05-01	164.71	0.12	0.01	10	0.5
149	s2	s2-1999-05-02	235.29	0.12	0	10	0.5
150	s2	s2-1999-05-03	258.82	0.12	0	10	0.5
151	s2	s2-1999-05-04	70.59	0.2	0.08	10	0.5
152	s2	s2-1999-05-05	51.41	0.29	0.08	10	0.5
153	s2	s2-1999-05-06	65.81	0.34	0.03	10	0.5
154	s2	s2-1999-05-c1	131.11	0.35	0.01	10	0.5
155	s2	s2-1999-06-01	77.11	0.13	0.06	1	0.46
156	s2	s2-1999-06-02	142.17	0.19	0.06	1	0.46
157	s2	s2-1999-06-03	44.82	0.24	0.05	1	0.46
158	s2	s2-1999-06-04	19.28	0.4	0.16	1	0.46
159	s2	s2-1999-06-05	28.92	0.59	0.19	1	0.46
160	s2	s2-1999-06-c1	14.46	0.67	0.08	1	0.46
161	s2	s2-1999-07-01	36.84	0.14	0.02	4	0.75
162	s2	s2-1999-07-02	36.84	0.15	0.02	4	0.75
163	s2	s2-1999-07-03	50	0.16	0.01	4	0.75
164	s2	s2-1999-07-04	57.89	0.16	0	4	0.75
165	s2	s2-1999-07-05	60.53	0.19	0.01	4	0.75

166	s2	s2-1999-07-c1	86.84	0.19	0	4	0.75
167	s2	s2-1999-07-c2	65.79	1.56	0.08	4	0.75
168	s2	s2-1999-07-c3	11.08	3.26	0	4	0.75
169	s2	s2-1999-08-01	166.27	0.03	0	4	1.76
170	s2	s2-1999-08-02	101.2	0.03	0	4	1.76
171	s2	s2-1999-08-03	93.98	0.03	0	4	1.76
172	s2	s2-1999-08-04	74.7	0.04	0.01	4	1.76
173	s2	s2-1999-08-05	62.65	0.09	0.05	4	1.76
174	s2	s2-1999-08-c1	33.73	0.12	0.03	4	1.76
175	s2	s2-1999-09-c1	114.57	1.77	0	2	0.34
176	s2	s2-1999-09-c2	14.81	1.78	0	2	0.34
177	s2	s2-1999-10-01	21.43	0.1	0.04	2	2
178	s2	s2-1999-10-02	19.05	0.11	0.02	2	2
179	s2	s2-1999-10-03	23.81	0.13	0.02	2	2
180	s2	s2-1999-10-04	15.71	0.17	0.04	2	2
181	s2	s2-1999-10-05	14.29	0.23	0.05	2	2
182	s2	s2-1999-10-c1	16.67	0.24	0.03	2	2
183	s2	s2-1999-11-01	134.18	0.54	0.11	40	0.92
184	s2	s2-1999-11-02	83.54	0.61	0.1	40	0.92
185	s2	s2-1999-11-03	81.01	0.65	0.07	40	0.92
186	s2	s2-1999-11-04	53.16	0.7	0.06	40	0.92
187	s2	s2-1999-11-05	78.48	0.85	0.09	40	0.92
188	s2	s2-1999-11-06	50.63	0.91	0.06	40	0.92
189	s2	s2-2000-01-01	361.45	0.1	0	33.6	0.14
190	s2	s2-2000-01-02	265.06	0.11	0.01	33.6	0.14
191	s2	s2-2000-01-03	265.06	0.11	0	33.6	0.14
192	s2	s2-2000-01-04	216.87	0.12	0.01	33.6	0.14
193	s2	s2-2000-01-05	313.25	0.14	0.02	33.6	0.14
194	s2	s2-2000-01-06	265.06	0.21	0	33.6	0.14
195	s2	s2-2000-01-07	168.42	0.43	0.05	33.6	0.14
196	s2	s2-2000-01-08	71.05	0.53	0.02	33.6	0.14
197	s2	s2-2000-01-09	47.37	0.67	0.03	33.6	0.14
198	s2	s2-2000-01-10	26.32	0.93	0	33.6	0.14
199	s2	s2-2000-02-01	2380.95	0.01	0	69	0.65
200	s2	s2-2000-02-02	1309.52	0.03	0.02	69	0.65
201	s2	s2-2000-02-03	785.71	0.05	0.02	69	0.65
202	s2	s2-2000-02-04	690.48	0.06	0.01	69	0.65
203	s2	s2-2000-02-05	500	0.08	0.02	69	0.65
204	s2	s2-2000-02-06	428.57	0.12	0	69	0.65
205	s2	s2-2000-02-07	380.95	0.16	0.01	69	0.65
206	s2	s2-2000-02-08	380.95	0.2	0.01	69	0.65
207	s2	s2-2000-03-01	166.67	0.01	0.01	1.9	0.2
208	s2	s2-2000-03-02	166.67	0.07	0.06	1.9	0.2
209	s2	s2-2000-03-03	148.05	0.12	0.05	1.9	0.2
210	s2	s2-2000-03-04	179.22	0.15	0.03	1.9	0.2
211	s2	s2-2000-03-05	122.07	0.19	0.04	1.9	0.2
212	s2	s2-2000-03-06	59.74	0.42	0.07	1.9	0.2
213	s2	s2-2000-03-07	41.56	0.76	0.09	1.9	0.2
214	s2	s2-2000-03-08	38.96	1.06	0.08	1.9	0.2
215	s2	s2-2000-03-09	25.97	1.37	0.09	1.9	0.2
216	s2	s2-2000-03-10	36.36	1.77	0.1	1.9	0.2
217	s2	s2-2000-03-11	23.38	2.27	0.22	1.9	0.2
218	s2	s2-2000-03-12	23.38	2.92	0.16	1.9	0.2
219	s2	s2-2000-05-01	232.56	0.03	0.02	4.8	5.07
220	s2	s2-2000-05-02	279.07	0.03	0	4.8	5.07
221	s2	s2-2000-05-03	279.07	0.03	0	4.8	5.07
222	s2	s2-2000-05-04	255.81	0.05	0.02	4.8	5.07

223	s2	s2-2000-05-05	502.33	0.05	0	4.8	5.07
224	s2	s2-2000-05-06	42.78	0.23	0.06	4.8	5.07
225	s2	s2-2000-05-07	19.44	0.51	0.08	4.8	5.07
226	s2	s2-2000-05-08	29.44	0.73	0.06	4.8	5.07
227	s2	s2-2000-05-09	33.33	0.86	0.01	4.8	5.07
228	s2	s2-2000-05-10	33.33	0.87	0.01	4.8	5.07
229	s2	s2-2000-06-01	170.73	0.03	0.01	1	0.08
230	s2	s2-2000-06-02	146.34	0.04	0.01	1	0.08
231	s2	s2-2000-06-03	117.81	0.06	0.02	1	0.08
232	s2	s2-2000-06-04	112.33	0.06	0	1	0.08
233	s2	s2-2000-06-05	104.11	0.07	0.01	1	0.08
234	s2	s2-2000-06-06	93.15	0.1	0.01	1	0.08
235	s2	s2-2000-06-07	79.45	0.13	0	1	0.08
236	s2	s2-2000-06-08	90.41	0.16	0.01	1	0.08
237	s2	s2-2000-06-09	73.97	0.2	0.01	1	0.08
238	s2	s2-2000-06-10	54.79	0.21	0	1	0.08
239	s2	s2-2000-06-11	76.71	0.22	0.01	1	0.08
240	s2	s2-2000-06-12	60.27	0.27	0.02	1	0.08
241	s2	s2-2000-07-01	506.67	0.01	0.01	5.5	3.74
242	s2	s2-2000-07-02	346.67	0.02	0.01	5.5	3.74
243	s2	s2-2000-07-03	320	0.04	0.02	5.5	3.74
244	s2	s2-2000-07-04	80.56	0.09	0.05	5.5	3.74
245	s2	s2-2000-07-05	72.22	0.13	0.04	5.5	3.74
246	s2	s2-2000-07-06	58.33	0.28	0.02	5.5	3.74
247	s2	s2-2000-07-07	91.67	0.3	0.01	5.5	3.74
248	s2	s2-2000-07-08	75	0.34	0	5.5	3.74
249	s2	s2-2000-07-09	58.33	0.35	0	5.5	3.74
250	s2	s2-2000-07-10	47.22	0.35	0	5.5	3.74
251	s2	s2-2000-08-01	129.27	0.35	0.05	31.5	2.8
252	s2	s2-2000-08-02	46.34	0.42	0.07	31.5	2.8
253	s2	s2-2000-08-03	59.1	0.45	0.03	31.5	2.8
254	s2	s2-2000-08-04	129.27	0.49	0.04	31.5	2.8
255	s2	s2-2000-08-05	24.39	0.55	0.06	31.5	2.8
256	s2	s2-2000-08-06	12.2	0.66	0.01	31.5	2.8
257	s2	s2-2000-08-07	24.39	0.76	0.01	31.5	2.8
258	s2	s2-2000-08-08	29.27	0.83	0.01	31.5	2.8
259	s2	s2-2000-08-09	63.41	0.91	0.03	31.5	2.8
260	s2	s2-2000-08-10	34.15	1.08	0.05	31.5	2.8
261	s2	s2-2000-08-11	24.39	1.19	0.01	31.5	2.8
262	s2	s2-2000-08-12	80.49	1.19	0	31.5	2.8
263	s3	s3-1999-03-01	355.56	0.03	0	17	0.02
264	s3	s3-1999-04-01	277.78	0.04	0.01	8	0.06
265	s3	s3-1999-04-02	444.44	0.04	0.01	8	0.06
266	s3	s3-1999-04-03	138.89	0.06	0.02	8	0.06
267	s3	s3-1999-04-04	472.22	0.07	0.01	8	0.06
268	s3	s3-1999-04-05	338.03	0.09	0.02	8	0.06
269	s3	s3-1999-04-06	55.56	0.14	0.02	8	0.06
270	s3	s3-1999-04-07	55.56	0.17	0.02	8	0.06
271	s3	s3-1999-04-08	44.44	0.23	0.03	8	0.06
272	s3	s3-1999-04-09	44.44	0.28	0.03	8	0.06
273	s3	s3-1999-04-10	41.67	0.37	0.05	8	0.06
274	s3	s3-1999-04-11	41.67	0.43	0.03	8	0.06
275	s3	s3-1999-04-12	44.44	0.5	0.04	8	0.06
276	s3	s3-1999-04-13	50	0.54	0.01	8	0.06
277	s3	s3-1999-05-01	83.33	0.05	0.03	5	0.72
278	s3	s3-1999-05-02	75	0.06	0.02	5	0.72
279	s3	s3-1999-05-03	83.33	0.06	0	5	0.72

280	s3	s3-1999-05-04	58.33	0.08	0.02	5	0.72
281	s3	s3-1999-05-05	69.44	0.1	0.02	5	0.72
282	s3	s3-1999-05-06	58.33	0.12	0.01	5	0.72
283	s3	s3-1999-05-07	52.78	0.14	0.02	5	0.72
284	s3	s3-1999-05-08	61.11	0.15	0.01	5	0.72
285	s3	s3-1999-05-09	69.44	0.21	0.03	5	0.72
286	s3	s3-1999-05-10	47.22	0.28	0.04	5	0.72
287	s3	s3-1999-06-01	92.54	0.12	0.07	10	0.53
288	s3	s3-1999-06-02	56.56	0.13	0.08	10	0.53
289	s3	s3-1999-06-03	92.54	0.16	0.04	10	0.53
290	s3	s3-1999-06-04	69.41	0.23	0.1	10	0.53
291	s3	s3-1999-06-05	53.98	0.26	0.07	10	0.53
292	s3	s3-1999-06-06	64.27	0.29	0.03	10	0.53
293	s3	s3-1999-06-07	58.1	0.34	0.05	10	0.53
294	s3	s3-1999-07-01	130.12	0.03	0.02	2	0.59
295	s3	s3-1999-07-02	130.12	0.05	0.03	2	0.59
296	s3	s3-1999-07-03	65.06	0.07	0.04	2	0.59
297	s3	s3-1999-07-04	48.19	0.11	0.04	2	0.59
298	s3	s3-1999-07-05	28.92	0.24	0.11	2	0.59
299	s3	s3-1999-07-06	19.28	0.44	0.07	2	0.59
300	s3	s3-1999-08-01	68.42	0.05	0.03	4	0.87
301	s3	s3-1999-08-02	71.05	0.08	0.04	4	0.87
302	s3	s3-1999-08-03	81.58	0.09	0.03	4	0.87
303	s3	s3-1999-08-04	78.95	0.11	0.02	4	0.87
304	s3	s3-1999-08-05	60.53	0.14	0.02	4	0.87
305	s3	s3-1999-08-06	68.42	0.15	0.01	4	0.87
306	s3	s3-1999-08-c1	31.58	0.48	0.09	4	0.87
307	s3	s3-1999-08-c2	103.61	2.23	0	4	0.87
308	s3	s3-1999-09-01	178.31	0.01	0.01	4	1.21
309	s3	s3-1999-09-02	139.76	0.01	0.01	4	1.21
310	s3	s3-1999-09-03	101.2	0.02	0.01	4	1.21
311	s3	s3-1999-09-04	72.29	0.05	0.03	4	1.21
312	s3	s3-1999-09-05	60.24	0.09	0.03	4	1.21
313	s3	s3-1999-09-c1	33.73	0.18	0.09	4	1.21
314	s3	s3-1999-10-01	64.2	0.01	0.01	6	0.4
315	s3	s3-1999-10-02	41.98	0.05	0.04	6	0.4
316	s3	s3-1999-10-03	27.16	0.14	0.1	6	0.4
317	s3	s3-1999-10-04	39.51	0.17	0.03	6	0.4
318	s3	s3-1999-10-c1	32.1	0.21	0.01	6	0.4
319	s3	s3-1999-11-01	47.62	0.03	0.02	2	1.61
320	s3	s3-1999-11-02	40.48	0.06	0.04	2	1.61
321	s3	s3-1999-11-03	47.62	0.07	0.03	2	1.61
322	s3	s3-1999-11-04	42.86	0.11	0.04	2	1.61
323	s3	s3-1999-11-05	30.95	0.15	0.01	2	1.61
324	s3	s3-1999-12-01	1800	0.01	0.01	40	0.74
325	s3	s3-1999-12-02	377.78	0.03	0.02	40	0.74
326	s3	s3-1999-12-03	311.11	0.06	0.04	40	0.74
327	s3	s3-1999-12-04	200	0.11	0.05	40	0.74
328	s3	s3-1999-12-05	154.43	0.18	0.05	40	0.74
329	s3	s3-1999-12-c1	103.79	0.24	0.06	40	0.74
330	s3	s3-2000-02-01	500	0.03	0	69	0.65
331	s3	s3-2000-02-02	309.52	0.06	0.03	69	0.65
332	s3	s3-2000-02-03	309.52	0.08	0.02	69	0.65
333	s3	s3-2000-02-04	285.71	0.09	0.01	69	0.65
334	s3	s3-2000-02-05	238.1	0.11	0.02	69	0.65
335	s3	s3-2000-02-06	309.52	0.15	0.01	69	0.65
336	s3	s3-2000-02-07	309.52	0.2	0.01	69	0.65

337	s3	s3-2000-03-01	714.29	0.02	0.02	2	0.21
338	s3	s3-2000-03-02	142.86	0.05	0.03	2	0.21
339	s3	s3-2000-03-03	96.1	0.09	0.04	2	0.21
340	s3	s3-2000-03-04	90.91	0.13	0.04	2	0.21
341	s3	s3-2000-03-05	75.32	0.17	0.04	2	0.21
342	s3	s3-2000-03-06	36.36	0.41	0.08	2	0.21
343	s3	s3-2000-03-07	25.97	0.72	0.08	2	0.21
344	s3	s3-2000-03-08	20.78	1.03	0.09	2	0.21
345	s3	s3-2000-03-09	20.78	1.31	0.05	2	0.21
346	s3	s3-2000-03-10	28.57	1.58	0.04	2	0.21
347	s3	s3-2000-03-11	15.58	2.23	0.21	2	0.21
348	s3	s3-2000-03-12	15.58	2.78	0.12	2	0.21
349	s3	s3-2000-04-01	166.67	0.06	0.03	14.2	0.48
350	s3	s3-2000-04-02	190.48	0.08	0.02	14.2	0.48
351	s3	s3-2000-04-03	82.86	0.1	0.02	14.2	0.48
352	s3	s3-2000-04-04	100	0.12	0.02	14.2	0.48
353	s3	s3-2000-04-05	94.29	0.14	0.02	14.2	0.48
354	s3	s3-2000-04-06	40	0.3	0.03	14.2	0.48
355	s3	s3-2000-04-07	62.86	0.36	0.01	14.2	0.48
356	s3	s3-2000-04-08	42.86	0.58	0.01	14.2	0.48
357	s3	s3-2000-04-09	88.57	0.61	0	14.2	0.48
358	s3	s3-2000-05-01	979.74	0.02	0.01	5.3	5.07
359	s3	s3-2000-05-02	348.84	0.04	0.02	5.3	5.07
360	s3	s3-2000-05-03	127.78	0.05	0.01	5.3	5.07
361	s3	s3-2000-05-04	113.89	0.07	0.02	5.3	5.07
362	s3	s3-2000-05-05	63.89	0.12	0.05	5.3	5.07
363	s3	s3-2000-05-06	22.22	0.49	0.11	5.3	5.07
364	s3	s3-2000-05-07	11.11	0.91	0.09	5.3	5.07
365	s3	s3-2000-05-08	16.67	1.16	0.02	5.3	5.07
366	s3	s3-2000-06-01	368.85	0.02	0.01	1	0.08
367	s3	s3-2000-06-02	219.51	0.04	0.02	1	0.08
368	s3	s3-2000-06-03	82.19	0.09	0.05	1	0.08
369	s3	s3-2000-06-04	43.84	0.1	0.01	1	0.08
370	s3	s3-2000-06-05	63.01	0.11	0.01	1	0.08
371	s3	s3-2000-06-06	82.19	0.14	0.01	1	0.08
372	s3	s3-2000-06-07	98.63	0.16	0.01	1	0.08
373	s3	s3-2000-06-08	101.37	0.17	0	1	0.08
374	s3	s3-2000-06-09	60.27	0.2	0.01	1	0.08
375	s3	s3-2000-06-10	95.89	0.21	0	1	0.08
376	s3	s3-2000-06-11	101.37	0.22	0.01	1	0.08
377	s3	s3-2000-06-12	35.62	0.26	0.01	1	0.08
378	s3	s3-2000-07-01	906.67	0.02	0.02	5.5	3.09
379	s3	s3-2000-07-02	373.33	0.03	0.01	5.5	3.09
380	s3	s3-2000-07-03	118.33	0.07	0.04	5.5	3.09
381	s3	s3-2000-07-04	83.33	0.09	0.02	5.5	3.09
382	s3	s3-2000-07-05	102.78	0.09	0	5.5	3.09
383	s3	s3-2000-07-06	69.44	0.14	0.04	5.5	3.09
384	s3	s3-2000-08-01	1108.43	0.11	0.01	31.6	1.19
385	s3	s3-2000-08-02	722.89	0.16	0.05	31.6	1.19
386	s3	s3-2000-08-03	289.16	0.22	0.06	31.6	1.19
387	s3	s3-2000-08-04	216.87	0.35	0.13	31.6	1.19
388	s3	s3-2000-08-05	207.23	0.41	0.06	31.6	1.19
389	s3	s3-2000-08-06	29.27	0.57	0.05	31.6	1.19
390	s3	s3-2000-08-07	17.07	0.76	0.02	31.6	1.19
391	s3	s3-2000-08-08	29.27	0.8	0.01	31.6	1.19
392	s3	s3-2000-08-09	85.37	0.84	0	31.6	1.19
393	s3	s3-2000-08-10	107.32	0.85	0	31.6	1.19

APPENDIX C. Regression Results for COD-Correlated Parameters

C.1 Oil & Grease (O&G)

The mean function for log [O&G] is as

$$E(\log O\&G \mid \mathbf{x}) = 3.31 - 0.56\log CumRs + 0.43\log AtDry - 0.18\log AtRs \qquad \text{(C.1)}$$

The total eligible cases for the O&G regression are 389. The R-squared is about 0.66. Table C.1 shows the brief regression result. Figure C.1 (a) and (b) show the model-checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows a very little curvature, and the data interval is quite consistent. Consequently, we think that (C.1) properly reflects the data.

Table C.1 Regression Results for (C.1)

Data set = wq_pool, Name of Fit = O&G Reg		
Kernel mean function = Identity Response = log[O&G] Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	3.30686	0.126547
log[CumRs]	-0.558414	0.0261632
log[AtDry]	0.430439	0.0254400
log[AtRs]	-0.176893	0.0217630
R Squared:	0.657808	
Scale factor:	0.591404	
Number of cases:	441	
Number of cases used:	389	
Degrees of freedom:	385	

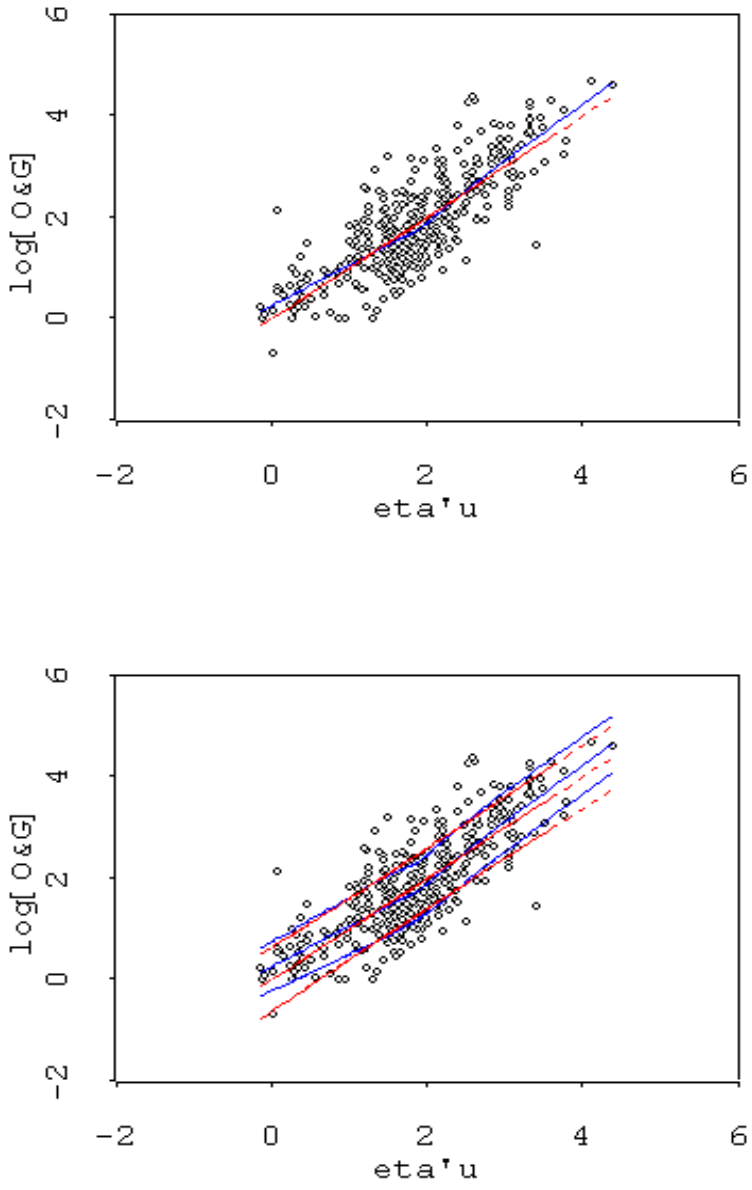


Figure C.1 (a) & (b) Model-Checking Plots for Checking (C.1) (the solid-blue line is for the data; the dashed-red line is for the model)

C.2 TKN

The mean function for log [TKN] is as

$$E(\log TKN \mid \mathbf{x}) = 3.38 - 0.60 \log CumRs + 0.33 \log AtDry - 0.16 \log AtRs \qquad \text{(C.2)}$$

The total eligible cases for the TKN regression are 200 only from the period of 2000 to 2001. The R-squared is about 0.77. Table C.2 shows the brief regression result. Figure C.2 (a) and (b) show the model-checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows mild curvature, and the data interval gets narrower on the left lower end. The inconsistent data interval might be caused by limited cases. Therefore, we think that (C.2) is still useful to reflect the data.

Table C.2 Regression Results for (C.2)

Data set = wq_pool, Name of Fit = TKN Reg		
Kernel mean function = Identity		
Response = log[TKN]		
Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	3.38856	0.136406
log[CumRs]	-0.608409	0.0269510
log[AtDry]	0.330106	0.0251255
log[AtRs]	-0.165478	0.0239235
R Squared:	0.768131	
Scale factor:	0.464762	
Number of cases:	441	
Number of cases used:	200	
Degrees of freedom:	196	

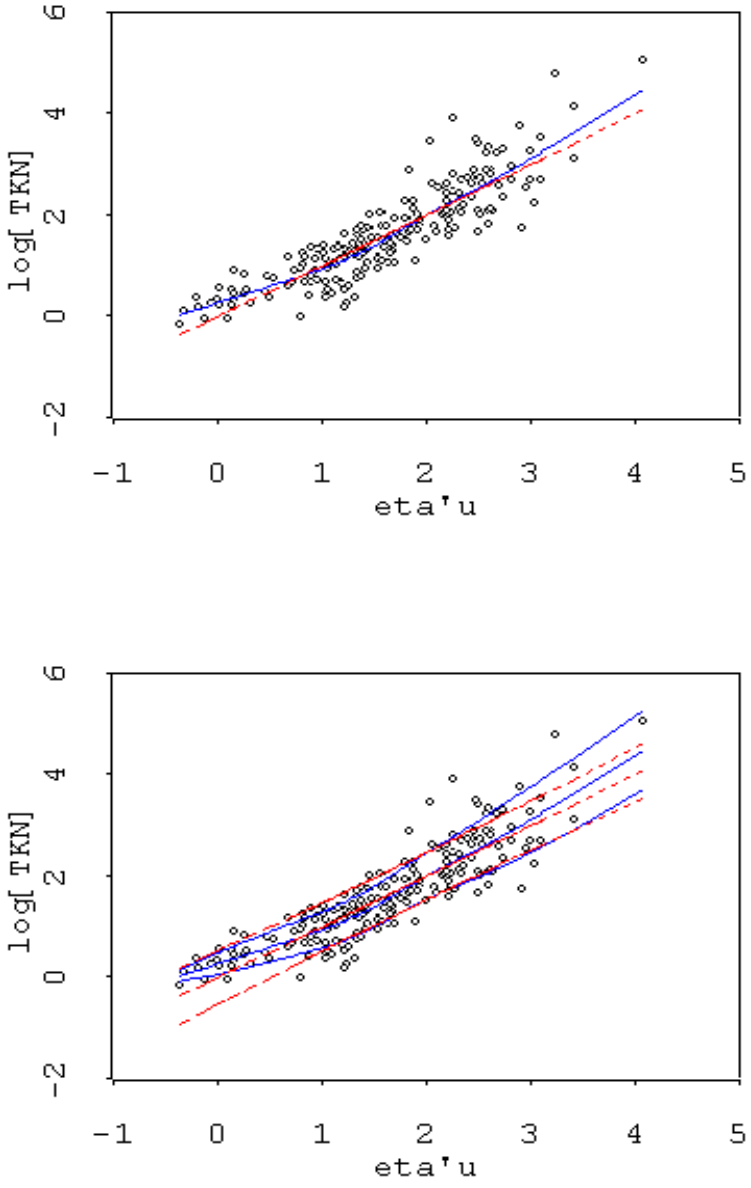


Figure C.2 (a) & (b) Model-Checking Plots for Checking (C.2) (the solid-blue line is for the data; the dashed-red line is for the model)

C.3 Dissolved Organic Carbon (DOC)

The mean function for log [DOC] is as

$$E(\log DOC \mid \mathbf{x}) = 4.44 - 0.57 \log CumRs + 0.42 \log AtDry - 0.17 \log AtRs \qquad \text{(C.3)}$$

The total eligible cases for the DOC regression are 393. The R-squared is about 0.67. Table C.3 shows the brief regression result. Figure C.3 (a) and (b) show the model checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows a little curvature on the left end, and the data interval is quite consistent. Consequently, we think that (C.3) properly reflects the data.

Table C.3 Regression Results for DOC

Data set = wq_pool, Name of Fit = DOC Reg		
Kernel mean function = Identity		
Response = log[DOC]		
Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	4.43776	0.124312
log[CumRs]	-0.569784	0.0257681
log[AtDry]	0.424184	0.0250992
log[AtRs]	-0.174477	0.0213769
R Squared:	0.665034	
Scale factor:	0.583796	
Number of cases:	441	
Number of cases used:	393	
Degrees of freedom:	389	

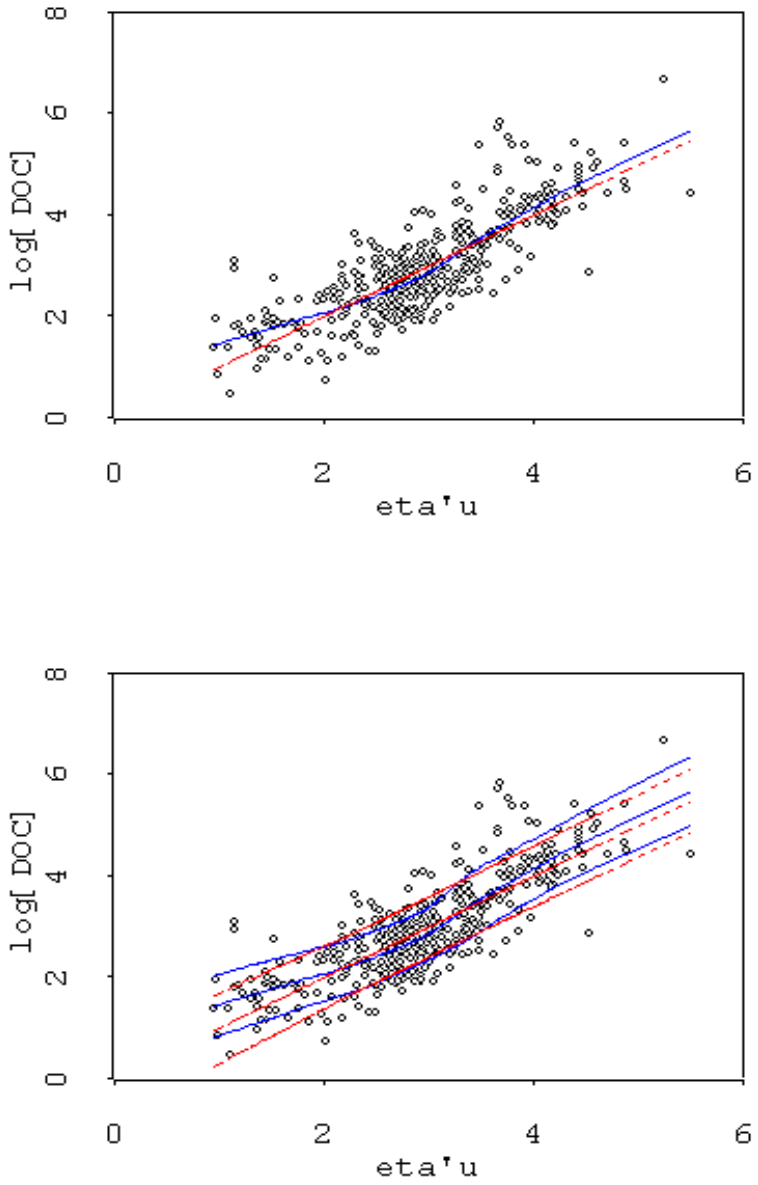


Figure C.3 (a) & (b) Model-Checking Plots for Checking (C.3) (the solid-blue line is for the data; the dashed-red line is for the model)

C.4 Dissolved Phosphorus (P_dis)

The mean function for log [P_dis] is as

$$E(\log P\_dis \mid \mathbf{x}) = -0.59 - 0.43 \log CumRs + 0.35 \log AtDry - 0.16 \log AtRs \qquad \text{(C.4)}$$

The total eligible cases for the P_dis regression are 387. The R-squared is about 0.49. Table C.4 shows the brief regression result. Figure C.4 (a) and (b) show the model checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows mild curvature, and the data interval is overall consistent except for the left end. Therefore, we think that (C.4) is still useful to reflect the data.

Table C.4 Regression Results for P_dis

Data set = wq_pool, Name of Fit = P_dis Reg1		
Kernel mean function = Identity		
Response = log[P_dis]		
Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	-0.591316	0.140533
log[CumRs]	-0.425109	0.0295018
log[AtDry]	0.355199	0.0281248
log[AtRs]	-0.158143	0.0240224
R Squared:	0.490238	
Scale factor:	0.653122	
Number of cases:	441	
Number of cases used:	387	
Degrees of freedom:	383	

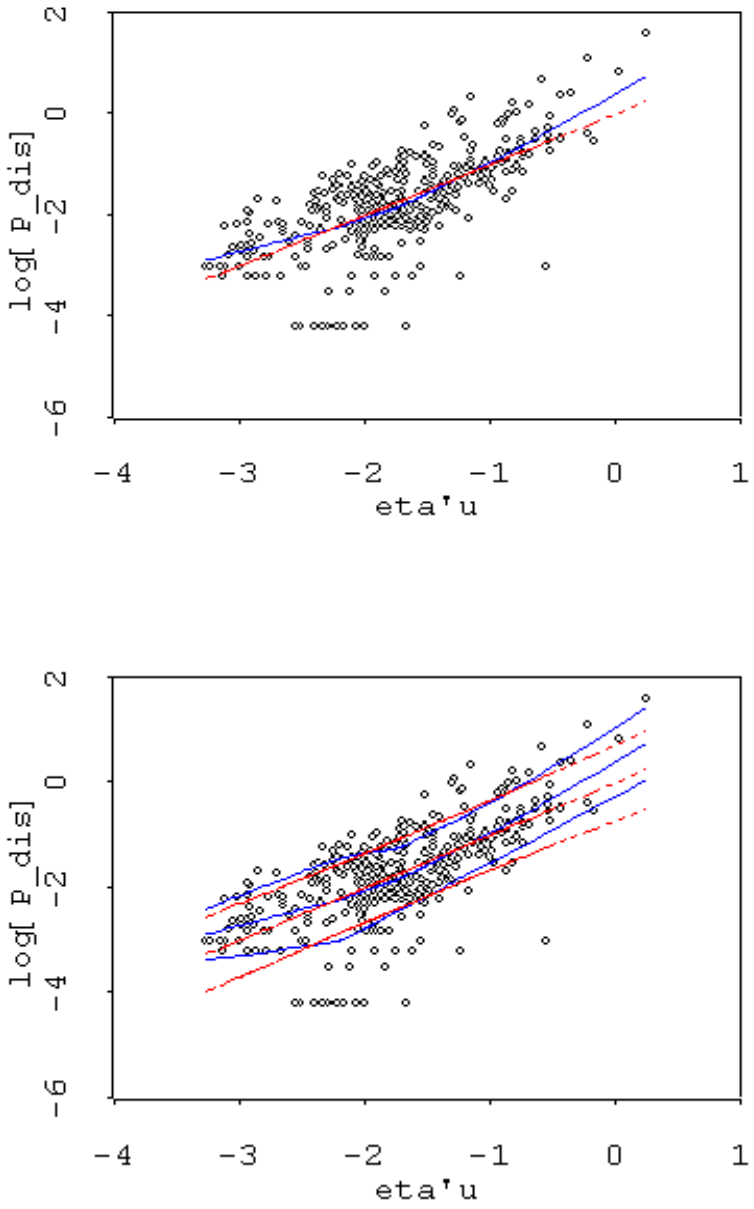


Figure C.4 (a) & (b) Model-Checking Plots for Checking (C.4) (the solid-blue line is for the data; the dashed-red line is for the model)

C.5 Dissolved Nickel (Cu_dis)

The mean function for log [Cu_dis] is as

$$E(\log Cu\_dis \mid \mathbf{x}) = 4.38 - 0.48 \log CumRs + 0.41 \log AtDry - 0.16 \log AtRs \qquad \text{(C.5)}$$

The total eligible cases for the Cu_dis regression are 390. The R-squared is about 0.65. Table C.5 shows the brief regression result. Figure C.5 (a) and (b) show the model checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows mild curvature, and the data interval is quite consistent. Consequently, we think that (4.4.6) properly reflects the data.

Table C.5 Regression Results for Cu_dis

Data set = wq_pool, Name of Fit = Cu_dis Reg		
Kernel mean function = Identity		
Response = log[Cu_dis]		
Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	4.38654	0.116773
log[CumRs]	-0.482444	0.0241533
log[AtDry]	0.408707	0.0234832
log[AtRs]	-0.164237	0.0200603
R Squared:	0.645253	
Scale factor:	0.546032	
Number of cases:	441	
Number of cases used:	390	
Degrees of freedom:	386	

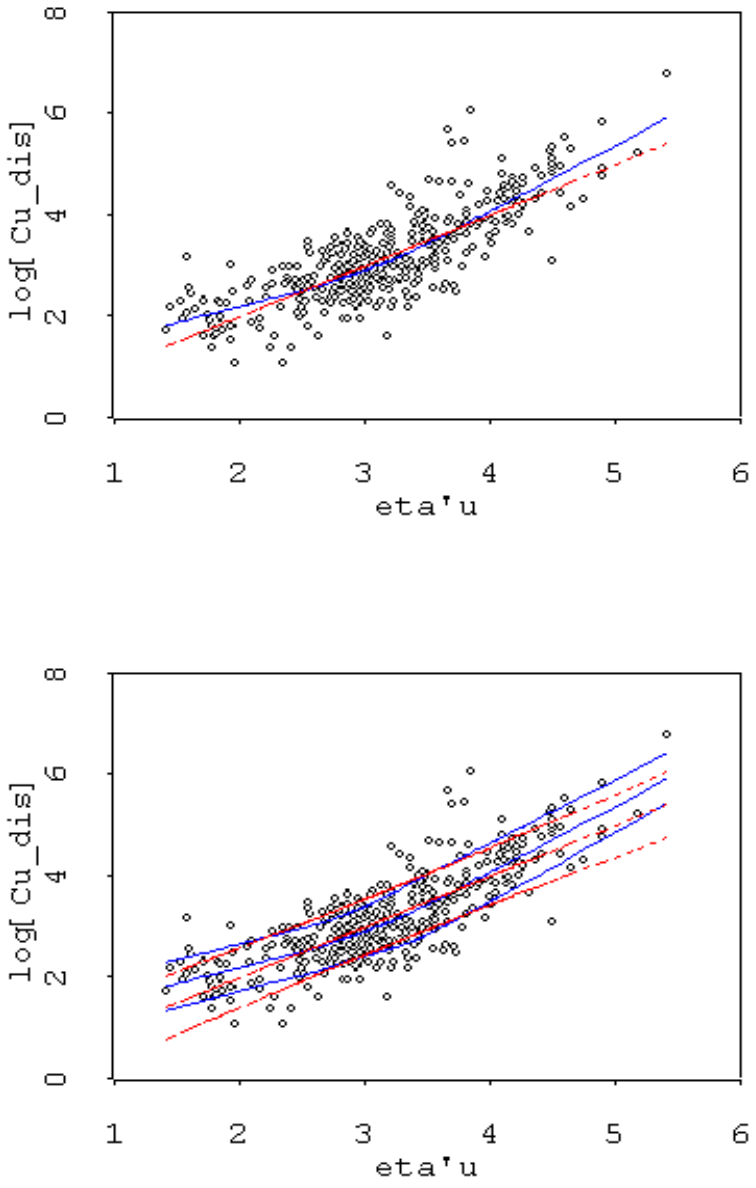


Figure C.5 (a) & (b) Model-Checking Plots for Checking (C.5) (the solid-blue line is for the data; the dashed-red line is for the model)

C.6 Dissolved Nickel (Ni_dis)

The mean function for log [Ni_dis] is as

$$E(\log Ni\_dis \mid \mathbf{x}) = 3.06 - 0.51 \log CumRs + 0.42 \log AtDry - 0.18 \log AtRs \qquad \text{(C.6)}$$

The total eligible cases for the Ni_dis regression are 390. The R-squared is about 0.55. Table C.6 shows the brief regression result. Figure C.6 (a) and (b) show the model checking plots for checking the mean function and the variance function, where the horizontal variable is the OLS fitted values (eta'u). The data line shows some curvature, and the data interval gets bigger on the left end. Consequently, we think that (C.6) is useful to moderately reflect the data.

Table C.6 Regression Results for Ni_dis

Data set = wq_pool, Name of Fit = Ni_dis Reg		
Kernel mean function = Identity		
Response = log[Ni_dis]		
Terms = (log[CumRs] log[AtDry] log[AtRs])		
Coefficient Estimates		
Label	Estimate	Std. Error
Constant	3.05913	0.149023
log[CumRs]	-0.513983	0.0308240
log[AtDry]	0.419359	0.0299689
log[AtRs]	-0.177199	0.0256006
R Squared:	0.552058	
Scale factor:	0.696836	
Number of cases:	441	
Number of cases used:	390	
Degrees of freedom:	386	

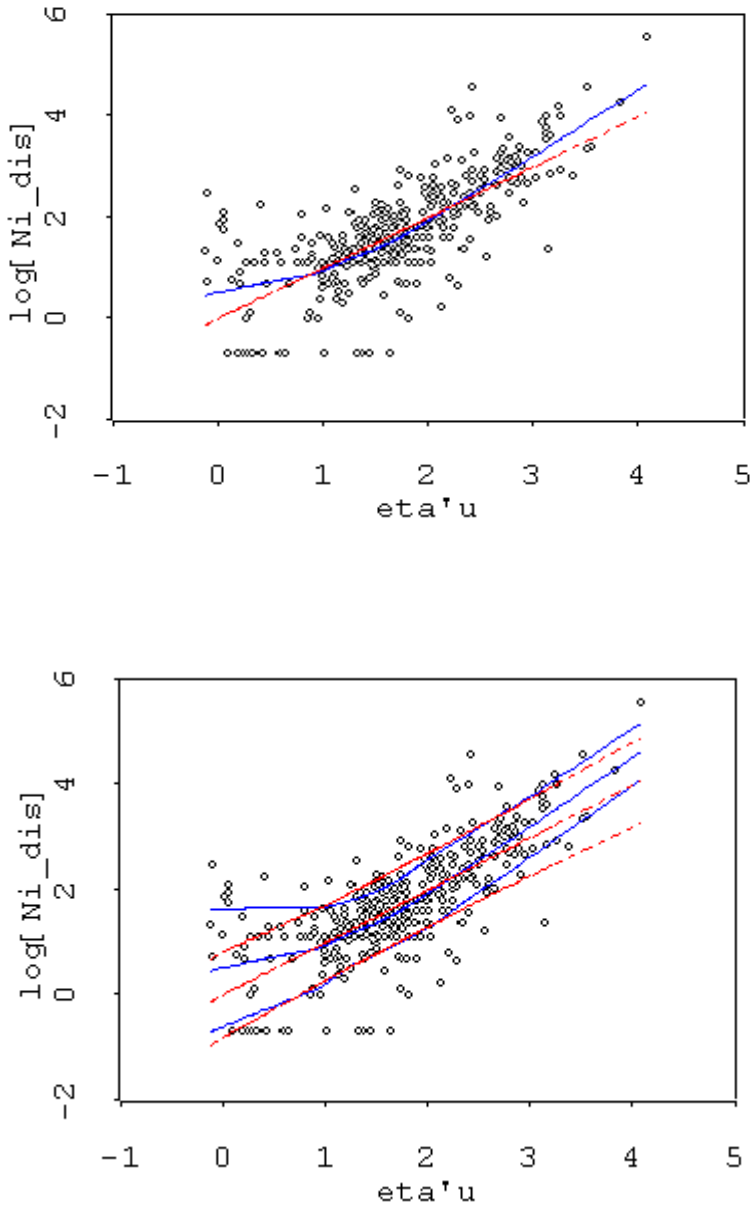


Figure C.6 (a) & (b) Model-Checking Plots for Checking (C.6) (the solid-blue line is for the data; the dashed-red line is for the model)

REFERENCES

American Sigma (2002). *950 Flow Meter Instrument Manual*, 2nd Edition, Oakville, Ontario, Canada.

American Sigma (2002). *900 Max Portable Sampler Instrument Manual*, 4th Edition, Oakville, Ontario, Canada.

Bertrand-Krajewski J., Chebbo G. and Saget A. (1998). "Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon", *Wat. Res.*, 32(8), 2341-2356.

Caltrans (2000). *Guidance Manual: Stormwater Monitoring Protocol*, 2nd Edition, California Department of Transportation, Sacramento, CA.

Charbeneau R. J. and Barrett M. E. (1998). "Evaluation of methods for estimating stormwater pollutant loads", *Water Environment Research*, 70, 1295-1302.

Cook R.D. and Weisberg S. (1999). *Applied Regression Including Computing and Graphics*, John Wiley & Sons, Inc., New York, N.Y.

Corbitt R. A. (1989). *Standard handbook of environmental engineering*, McGraw-Hill, Inc., New York, N.Y.

Deletic A. (1998). "The first flush load of urban surface runoff", *Wat. Res.*, 32(8), 2462-2470.

Driscoll E. D., P. E. Shelley, and Strecker, E. W. (1990). "Pollutant loadings and impacts from stormwater runoff, Vol. III", *Analytical Investigation and Res. Rep*. FHWA-RD-88-008, Fed. Highway Admin., Washington, D.C.

Ferguson T.S. (1996). *A Course in Large Sample Theory*, Chapman & Hall, New York, N.Y.

Geiger W. (1987). "Flushing effects in combined sewer systems", *Proceedings of the 4th Int. Conf. on Urban Drainage*. 40-46, Lausanne, Switzerland.

Gilliland M.W. and Baxter-Potter W. (1987). "A geographic information system to predict non-point source pollution potential", *Water Resour. Bull.*, 23(2). 281-291.

Gupta K. and Saul A.J. (1996). "Specific relationships for the first flush load in combined sewer flows", *Wat. Res.*, 30(5), 1244-1252.

Larsen T., Broch K., and Andersen M.R. (1998). "First flush effects in an urban catchment area in Aalborg", *Wat. Sci. & Tec*., 37(1), 251-257.

Lau S. L., Bay S. and Stenstrom M. K. (1993). " Contamination in urban runoff and their impact on receiving waters", *Proc., Asian WaterQual's 93, Fourth IAWQ Asian Regional Conf. on Water Conservation and Pollution Control*, Vol I-2.

Lau S.L., Ma J.-S., Kayhanian K., and Stenstrom M. K., "First flush of organics in highway runoff", *Proceeding of Global Solutions for Urban Drainage*, *9th International Conference on Urban Drainage*, September, 2002, Portland, Oregon.

National Research Concil. (1990). "Monitoring Southern California's coastal waters", *Rep.,* National Academy Press, Washington, D.C.

Novotny V. and Olem H. (1994). *Water Quality, Prevention, Identification and Management of Diffuse Pollution,* Van Nostrand Reinhold, New York.

Ma J.-S., Khan S., Li Y.-X. Li, Kim L.-H., Ha H., Lau S.L., Kayhanian K., and Stenstrom M. K. (2002). "First flush phenomena for highways: how it can be meaningfully defined", *Proceeding of Global Solutions for Urban Drainage*, *9th International Conference on Urban Drainage*, September, 2002, Portland, Oregon.

Ma J.-S., Khan S., Li Y.-X. Li, Kim L.-H., Ha H., Lau S.L., Kayhanian K., and Stenstrom M. K. (2002). "Implication of Oil and Grease Measurement in Stormwater Management Systems", *Proceeding of Global Solutions for Urban Drainage*, *9th International Conference on Urban Drainage*, September, 2002, Portland, Oregon.

Saget A., Chebbo G. and Bertrand-Krajewski J. (1995). "The first flush in sewer system", *Proceeding of the 4th Int. Conf. on Sewer Solids-Characteristics, Movement, Effects and Control*, 58-65, Dundee, UK.

Sansalone J.J. and Buchberger S.G. (1997). "Partitioning and first flush of metals in urban roadway storm water", *J. of Environ. Engineering*, 123(2), 134-143.

Smullen, J. T., Shallcross, A. L., Cave, K. A. (1999). "Updating the U.S. nationwide urban runoff quality data base". *Wat. Sci. Tech.*, 39 (12), 9-16.

U.S. EPA (1983). "Results of nationwide urban runoff program: Executive summary", *WH-554*, Water Plng. Div., U.S. Envir. Protection Agency, Washington, D.C.

U.S. EPA (1990). "National water quality inventory – 1988 report to Congress", *EPA 440-4-90-003,* U.S. Envir. Protection Agency, Washington, D.C.

Vorreiter L. and Hickey C. (1994). "Incidence of the first flush phenomenon in catchments of the Sydney region", *National Conf. Publication-Institution of Engineers*, Vol. 3, 359-364, Australia.

Wong, K. M., Strecker, E. W., and Stenstrom, M. K. (1997). "GIS to estimate storm-water pollution mass loadings", *J. Environ. Eng.*, 123 (8), 737-745.

Wu, J. S., Allan, C. J., Saunders, W. L., and Evett, J. B. (1998). "Characterization and pollutant loading estimation for highway runoff", *J. Environ. Eng.*, 584-592.