California Regional Water Quality (
Los Angeles Region
Recipient of the 2001 *Environmental Leadership Award* from Keep Calil

Alan C. Lloyd, Ph.D.
Agency Secretary

320 W. 4th Street, Suite 200, Los Angeles, California 90013
Phone (213) 576-6600 FAX (213) 576-6640 - Internet Address: http://www.waterboards.ca.gov/losangeles

SPECIAL HEARING
2/3/05
cc: BD, DI, DWQ
E-cys: BD, CC HMS, TH, CMW

3. larwqcb

TO: Bruce Fujimoto
Division of Water Quality, State Board

FROM: Xavier Swamikannu, D.Env.
Storm Water Program

DATE: January 31, 2005

SUBJECT: TRANSMITTAL OF (i) FINAL REPORT: INDUSTRIAL STORM WATER MONITORING, AND (ii) RESEARCH PAPER, FOR INCLUSION IN THE ADMINISTRATIVE RECORD FOR THE INDUSTRIAL GENERAL PERMIT

Please find enclosed the following two documents, (i) *Final Report: Industrial Storm Water Monitoring Program – Existing Statewide Permit Utility and Proposed Modifications, Stenstrom and Lee (2005)*, and (ii) *Utility of Stormwater Monitoring, Lee and Stenstrom, Water Environment Research, Vol. 77(1) (2005)*. The *Final Report: Industrial Storm Water Monitoring Program – Existing Statewide Permit Utility and Proposed Modifications*, was prepared pursuant to State Board Contract No. 02-172-140-0.

These documents will be of assistance to you and support changes to the Draft Industrial General Permit to improve its monitoring program to better assess the quality of storm water discharges, to collect data for Total Maximum Daily Load (TMDL) development, and to assess compliance. I request that they be included in the Administrative Record for the Industrial General Permit.

Attachments

Utility of Stormwater Monitoring

Haejin Lee, Michael K. Stenstrom

ABSTRACT: Stormwater runoff is now a major contributor to the pollution of coastal waters in the United States. Public agencies are responding by requiring stormwater monitoring to satisfy the National Pollutant Discharge Elimination System stormwater permit. However, studies to understand the utility of the current programs or to improve their usefulness have not yet been performed. In this paper, we evaluate the land-use-based program, the industrial stormwater permit program, and beach water-quality monitoring in the County of Los Angeles, California, to determine if the results will be helpful to planners and regulators in abating stormwater pollution. The utility of the program has been assessed based on the programs' ability to accurately estimate the emissions for different classes of land use. The land-use program appears successful, while the industrial monitoring program does not. Beach water-quality monitoring suffers from a lack of real-time monitoring techniques. We also provide suggested improvements, such as sampling method and time, and parameter selection. *Water Environ. Res.*, **77**, 000 (2005).

KEYWORDS: stormwater, monitoring, neural network, land use, coastal water.

Introduction

California coastal waters are important recreational and economic resources, which make the safety of coastal waters of concern to both state and county health departments and beachgoers (Jiang et al., 2001). The completion of wastewater treatment plants mandated by the Clean Water Act has reduced conventional water pollution to California's beaches and bays. As a result, non-point-source pollution, such as stormwater runoff, is now a major contributor to the pollution of the coastal water, including Santa Monica Bay, which is among the most severely polluted bays in the United States (Wong et al., 1997). The problem of stormwater pollution is becoming worse because of population growth, which results in increased impermeable area. Storm drains entering the ocean are a main cause of permanent beach postings at many California beaches (State Water Resources Control Board, 2001).

Public agencies are responding by requiring stormwater monitoring to satisfy the National Pollutant Discharge Elimination System (NPDES) stormwater permit as authorized by the Clean Water Act. For example, the Los Angeles County Department of Public Works (LACDPW) has been monitoring stormwater under the 1990 NPDES municipal permit (No. CA0061654), and later the 1996 municipal permit (No. CAS614001), since the 1994–1995 wet season. Additional sampling is required by other agencies, such as the City of Los Angeles and the California Department of Transportation. Similar programs are underway in other areas of California and the United States.

The existence of stormwater-monitoring programs should represent progress towards achieving clean-water goals; however, studies have not yet been performed to understand the utility of the current programs or to improve their usefulness. In this paper, we evaluate several monitoring programs to determine if the results will be helpful to planners and regulators in abating stormwater pollution. Datasets from a major municipal program, beach monitoring, a large self-monitoring program, and a research project were used. The results suggest that parts of the current monitoring programs will not be helpful to regulators and planners, and we make proposals for improvement, along with projected cost increases.

Background

The LACDPW has been monitoring stormwater since the early 1970s. In 1994, it began an improved program, which was designed to determine total pollutant emissions to Santa Monica Bay and determine land-use-specific discharges (Stenstrom and Strecker, 1993). Total emissions are estimated from flow-weighted composite samples that are collected at five sampling stations (four stations are required under the 1996 permit and one station remains from an earlier permit.). These stations are "mass emission" stations in that they were selected to sample the greatest runoff mass with the least number of stations. The stations are equipped with flow-monitoring equipment and operate unattended in secure facilities. Samples from specific land uses are also required by the 1996 municipal permit and are collected with composite samplers at engineered sampling stations. A large suite of water-quality parameters is measured, including indicator organisms, general minerals, nutrients, metals, semivolatile organic compounds, and pesticides.

Additional monitoring is being conducted by other agencies to satisfy regulations or for research. The California Department of Transportation (Caltrans) has a large monitoring program for their highways. Our laboratory has monitored three highway locations near the University of California at Los Angeles (UCLA) (adjacent to the 101 and 405 freeways) since 1999 (Stenstrom et al., 2000 and 2001). The study is also sponsored by Caltrans, and an extensive suite of parameters is measured, including indicator bacteria, general minerals, nutrients, metals, polycyclic aromatic hydrocarbons, and oil and grease.

The previous programs monitor discharges to the bay, but there are also programs that monitor coastal waters. The California Assembly passed Bill 411(chapter 765 of Statutes of 1997; http://www.swrcb.ca.gov/beach/bills/ab_411_bill_19971008_chaptered.pdf) to address the problem of declining beach water quality and restore confidence in beach swimming. Three types of indicators organisms are monitored, and retesting in the event of an exceedence is also required. The more restrictive procedures by the bill have increased the frequency of beach postings and closures. The closure of Huntington Beach in Orange County, California, during the summer of 1999, was the first example of beach closures caused by the new regulations (Grant et al., 2001; Orange County Sanitation District, 1999). Many organizations are monitoring the microbiological water quality of Southern California coastal waters (Noble et al., 2000).

An example of a new monitoring activity is the Industrial Activities Stormwater General Permit, which mandates all industrial stormwater

Table 1—Selected eight major industries and its case number according to the different sets of input parameters after clipping outliers for 1998–2001 seasons.

		Input parameters		
Major industries	SIC code	pH, TSS, SC, TOC, and O&G	pH, TSS, SC, and O&G	Lead, copper, and Zinc
Food and kindred products (FKP)	20	184	472	10
Chemical and allied products (CAP)	28	305	850	35
Primary metal industries (PMI)	33	144	773	100
Fabricated metal products, except machinery and transportation equipment (FMP)	34	417	1325	155
Transportation equipment (TE)	37	193	601	187
Motor freight transportation and warehousing (MFTW)	42	263	731	76
Electric, gas, and sanitary services (EGSS)	49	182	505	198
Wholesale trade-durable goods (WT)	50	120	723	471
Number of total cases		1808	5980	1232

permittees to analyze stormwater samples, twice per year, for at least four analytical parameters. The industries are classified by Standard Industrial Classification (SIC) code (www.swrcb.ca.gov/rwqcb4/html/programs/stormwater/sw_industrial.htmlreference). The monitored analytical parameters are pH, total suspended solids (TSS), specific conductance (SC), and total organic carbon (TOC). Oil and grease (O&G) may be substituted for TOC. In addition, the permittees must monitor any other pollutants, which they believe to be present in their stormwater discharge as a result of industrial activity (www.swrcb.ca.gov/stormwtr/docs/induspmt.doc). Permittees, in some cases, may be required to sample at more than one location.

It is natural to ask if the monitoring programs are valuable. Is the resulting water-quality database useful to planners and regulators to identify acute problems, improve long-term water quality, and understand land-use and water-quality relationships? An improved understanding of the relationship of land use to stormwater quality is an expected result, because land-use-specific sampling is required by the NPDES permit. The original purposes of the monitoring programs were to identify larger sources (e.g., "hot spots") and to create a database to help develop total mass daily loads (TMDLs) and other management tools. To answer this question, we reviewed the current industrial stormwater permit program. We also comment on other monitoring programs, and suggest improvements in sampling strategies and water-quality-parameter selection, with their anticipated cost increases.

Monitoring Program Utility

It is generally recognized that different human activities will create different types and varying concentrations of stormwater contaminants (Stenstrom et al., 1984). For example, runoff from transportation-associated land use is a primary source of metals and hydrocarbons (LACDPW and Woodward-Clyde, 1998). Vehicles release hydrocarbons from leaks, engine byproducts and unburned fuel, and various metals from corrosion, fuel combustion, and wearing surfaces, such as brake pads (Rogge et al. 1993; Sansalone and Buchberger, 1997). Differences in land-use patterns will likely result in different pollutant concentrations, and, therefore, land-use-related control strategies are essential to control stormwater pollution effectively.

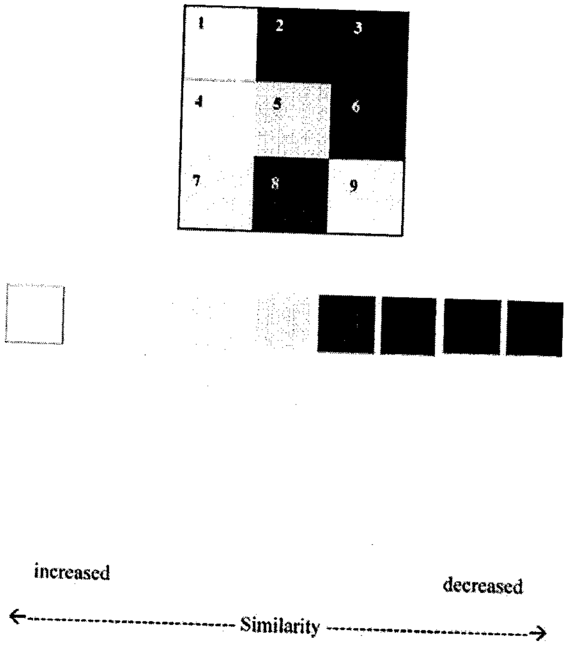


Figure 1—Activation map having 3 × 3 neurons obtained by a Kohonen neural model, which was trained with four input parameters (pH, TSS, SC, and O&G). The shading intensity indicates the degree of similarity to their neighbor nodes. Numbers indicate node in the Kohonen layer.

Land-Use-Monitoring Data. The land-use-based program administered by the LACDPW is a useful example. The land-use-monitoring program, required by the 1996 Municipal Permit, was examined to determine if different land uses produce different stormwater quality. If the monitoring program is successful, land uses should be identifiable from the collected data. We developed a neural network approach to identify the various types of land use (commercial, residential, industrial, transportation, and vacant) as a function of stormwater-quality data (Ha and Stenstrom, 2003). The neural model uses a Bayesian network, and was trained using LACDPW data collected during 1996–2001 wet seasons. Among approximately 90 water-quality parameters that were measured, 42 candidate parameters were initially selected because they were detected in more than 25% of the storm events. We then selected the top 10 most useful parameters for classifying the target land-use classes using a discriminant analysis. The 10 water-quality parameters are potassium, sulfate, alkalinity, dissolved phosphorus, nitrite-N, total dissolved solids, volatile suspended solids, TSS, dissolved copper, and dissolved zinc. The model was successful at classifying 92% of the cases. The model was useful in that a data set

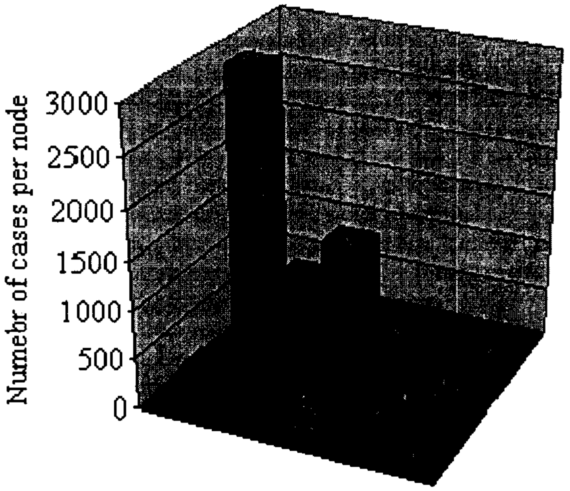


Figure 2—Number of cases per node obtained by the Kohonen neural model explained in Figure 1.

could be manipulated by changing various water-quality parameters, and the changes in classifications could be observed. It is also possible to determine which parameters are most sensitive for the classification and which are most active in a particular case. The model will eventually be useful to automatically examine many datasets to identify abnormally high or low parameters for a particular land use and label these as opportunities for investigation or improvement.

Industrial Stormwater Monitoring. Based on this experience, a similar approach was applied to the industrial stormwater discharge data for the 1998–2001 wet seasons. This dataset contains approximately 14 000 cases. Neural networks were trained to differentiate between several industrial categories, based on SIC code and water-quality data. It was hoped that the trained model would be help to identify industrial "hot sources" or outliers. Eight industrial categories were selected, based their prevalence in Los Angeles County, which means some SIC codes have many more cases than others. The selected eight industrial categories, and each category's case number for the three years, are shown in Table 1. The data cases that contain both the mandatory water-quality parameters (pH, TSS, SC, TOC, and O&G) and metals are limited. Because of this reason, a neural model trained separately with the water-quality data and metal data. Outliers in this study were defined as the upper 2% of the whole range of the data set for each parameter, and these cases were removed.

In this study, Neural Connection 2.1 (SPSS, Inc. and Recognition Systems, Inc, Chicago, Illinois) was used to build the neural models. Three supervised, feed-forward neural networks, namely, Multi-Layer Perceptron, Radial Basis Function, and Bayesian Network were used to differentiate the various types of industries. The neural models were extensively trained with various architectures; however, the performance of all models was very poor. This indicates a weak or almost no relationship between the industrial categories based on the SIC code and the available water-quality data.

To further seek a relationship between water-quality data and various land uses of industries, an unsupervised Kohonen neural network was used. The goal of Kohonen network is to map the spatial relationships among clusters of data points into hyperdimensional space (Aguilera et al., 2001). Once trained successfully, it may be used to identify unknown data patterns, and it was hoped that useful patterns between water quality and industrial categories would be identified.

A Kohonen neural model with two dimensions in the Kohonen layer was trained with different node sizes of 3 × 3, 5 × 5, and 7 × 7. The method performs square normalization, which normalizes the original input data patterns to zero mean and unit variance. The results were generally unsatisfactory, and it was difficult to make a decision to cluster from the activation maps by the neural model.

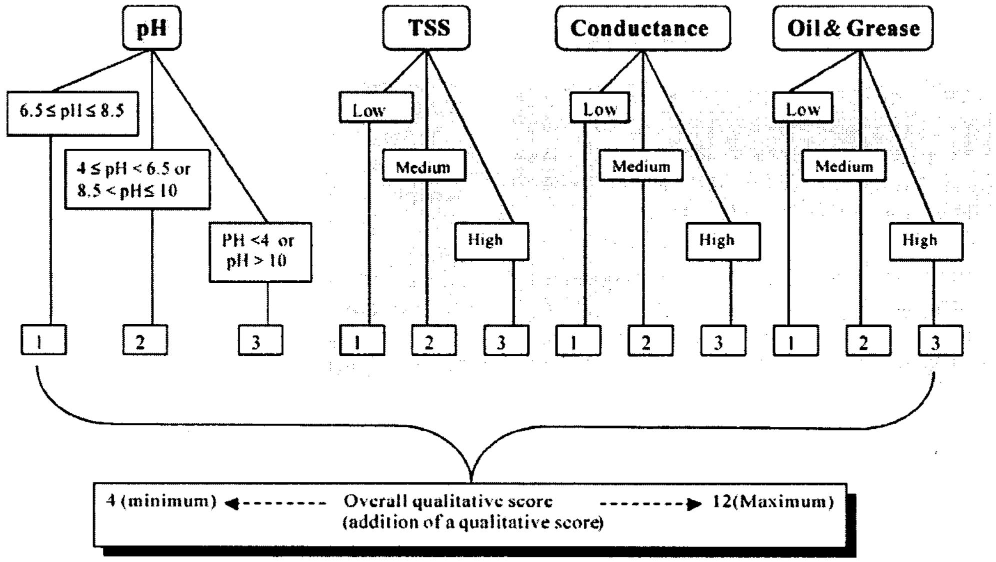


Figure 3—Overall process of producing an overall qualitative score with four parameters (shade area: a Kohonen network having three nodes in the Kohonen layer was trained for each parameter).

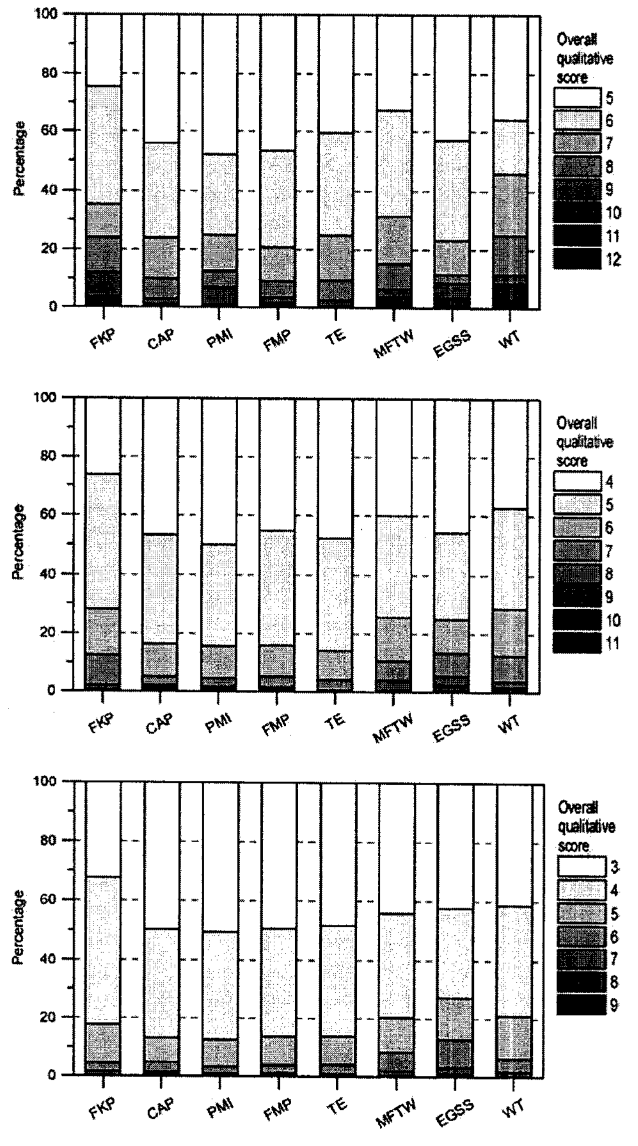


Figure 4—Distribution of the overall qualitative score for each industrial category. Higher scores indicate low water quality. Number of cases per category with different sets of parameters was shown in Table 1 (top: five parameters [pH, TSS, SC, TOC, and O&G] were used; middle: four parameters [pH, TSS, SC, and O&G] were used; bottom: three parameters [Pb, Cu, and Zn] were used).

Figure 1 shows an activation map having 3 × 3 neurons obtained by a Kohonen model that was trained with four parameters: pH, TSS, SC, and O&G. The shading intensity indicates the degree of similarity to their neighbor nodes; lighter shades indicate similar characteristics as the neighboring node, and darker patterns indicate greater differences. There are two possible clusters. The first contains nodes 1, 4, 5, and 7, and the second contains nodes 2, 3 and 6. Figure 2 shows the number of cases assigned to the various nodes. Nodes 4 and 5 contain most of the cases, and the majority (82%) would be assigned to the first cluster. A classification system that assigns such a large fraction to a single cluster is not useful; basically, the classification system is saying that it can find no difference in the available water-quality parameters among the majority of the SIC codes. Nodes 4 and 5 tend to have the lowest pollutant concentrations, but the members are not distinguished by SIC codes. Similar results were obtained using 5 × 5 or 7 × 7 neurons and with different set of input parameters. The conclusion from this analysis is that stormwater quality is not distinguishable by SIC code using the current water-quality parameters.

To further investigate possible relationships, the water-quality data were transformed into a three-member fuzzy set with categories of low, medium, and high. Each of the resulting data sets (except for pH) was examined using a Kohonen neural model. Each model had three nodes in the one-dimensional Kohonen layer and was trained for each parameter separately. The output result of the model was assigned a specific node number, from 1 to 3, for every case. The nodes were reordered so that higher node numbers always indicated greater pollutant concentration, with the number 3 representing the highest pollutant concentration for each parameter. For the case of pH, a qualitative score was assigned manually, based on deviation from neutrality. An overall qualitative score was created by summing the fuzzy states. Figure 3 shows the overall process. For example, when we used four parameters, the possible minimum and maximum overall qualitative score is 4 and 12, respectively, with 12 representing the worst water quality.

Figure 4 shows the distribution of the overall qualitative score for the various types of industries with different sets of input parameters. The upper figure used all five water-quality parameters (pH, TSS, SC, TOC, and O&G). The middle part shows the classifications when TOC is left out. The bottom shows the classification using the metal analysis. In general, no distinguishing differences were found among industrial categories and with different sets of input parameters. The food and kindred product facilities have the least abundance of low scores for the three different sets of input parameters, suggesting that it has worse stormwater quality. When five water-quality parameters were used (Figure 4, top), the wholesale trade-durable goods category had least abundance of small scores, if scores up to 6 are considered.

In general, no distinguishing differences were found among industrial categories for metals (Figure 4, bottom), if scores up to 4 are considered. However, it was necessary to remove more outliers in lead and zinc concentrations for primary metal industries and copper for transportation equipment industries than for other categories. This suggests that these two industries have the worst stormwater quality, with respect to metals. The statistical significance of these findings has not been evaluated, and it all likelihood, a new or different method would need to be used.

The industrial data set was also examined to determine if a seasonal first flush could be identified. Los Angeles has two distinct rainfall seasons. The late spring to late fall or early winter is generally dry. Most rainfall occurs in winter and early spring. This rainfall pattern creates a long period for pollutant buildup, and the first storm of the season generally has abnormally high pollutant concentrations, which is called a seasonal first flush. The industrial permit requires the first storm to be sampled and one later storm to be sampled, which was required to identify the seasonal first flush.

To determine if the industrial stormwater monitoring program was successful in identifying the seasonal first flush, the data (for 2000–2001 season only) were divided into first and second sample datasets. In some cases, the first sample does not represent the first rainfall event. In cases when there were more than two samples collected, the later samples were ignored. Cases with only one sample were also ignored. The first to second samples for the 2000–2001 season are compared in Figure 5 using notched bar plots.

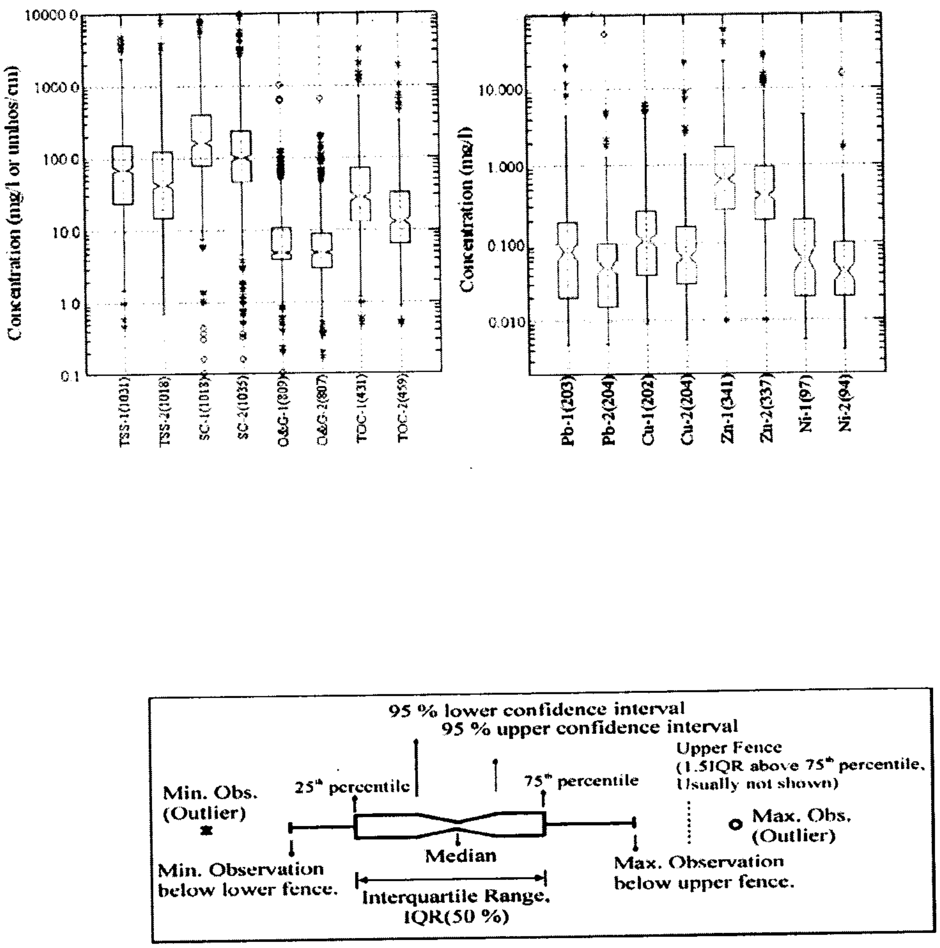


Figure 5—Comparison of first to second sample for the 2000–2001 wet season. All outliers are now shown. The "1" in the x-axis indicates the first sample, and "2" indicates the second sample. Number in a parenthesis in x-axis indicates number of cases.

Concentrations for all parameters were higher in the first sample than the second sample by 0 to 120% for the median and 20 to 85% for the mean. The TOC showed the greatest difference between first and second samples; O&G showed the smallest difference. Statistically significant differences can be observed in the notched bar plot.

Beach Monitoring. Assembly Bill 411 created improved beach water-quality-monitoring requirements. The improved monitoring was mandated after an epidemiological study of Santa Monica Bay swimmers suggested increased health risk associated with swimming near storm drains (Haile et al., 1999). Daily samples for total and fecal coliforms and enterococcus were mandated with new, lower levels that trigger a beach posting or closure. Leecaster and Weisberg (2001) examined sampling data from 24 sites in Los Angeles County between 1995 and 1999. They report that over 70% of the water-quality exceedences were for only one day.

The time required to analyze indicator organism data is generally more than 24 hours. This created a chronology as follows:

(1) Day 1—a sample is collected and analysis begins.
(2) Day 2—the sample is analyzed and an exceedence is noted, with a beach posting; a new sample is collected and analysis begins.
(3) Day 3—the second sample result is negative 70% of the time, and the beach posting is removed.

The chronology creates a situation that beaches are posted when the samples do not exceed standards and open when they do. Clearly, the problem is a monitoring program that cannot be implemented with current technology. Rapid indicators are needed. Furthermore, the utility of conventional indicator organisms for fecal contamination in beach waters is in question.

Discussion and Recommendations

Three stormwater monitoring programs were discussed. The land-use-monitoring program was generally successful and showed the anticipated differences in water quality, based on land use. This program used automatic, flow-weighted composite samplers with trained personnel. The second program, the industrial-monitoring program, used grab samples collected at various times for two storms with SIC codes as land-use or industrial-use descriptors. The program is generally unsuccessful in identifying relationships between water quality and land use. It was successful in showing a seasonal first flush, and its utility for identifying acute problems is questionable, based on outliers, to be discussed later. The third is

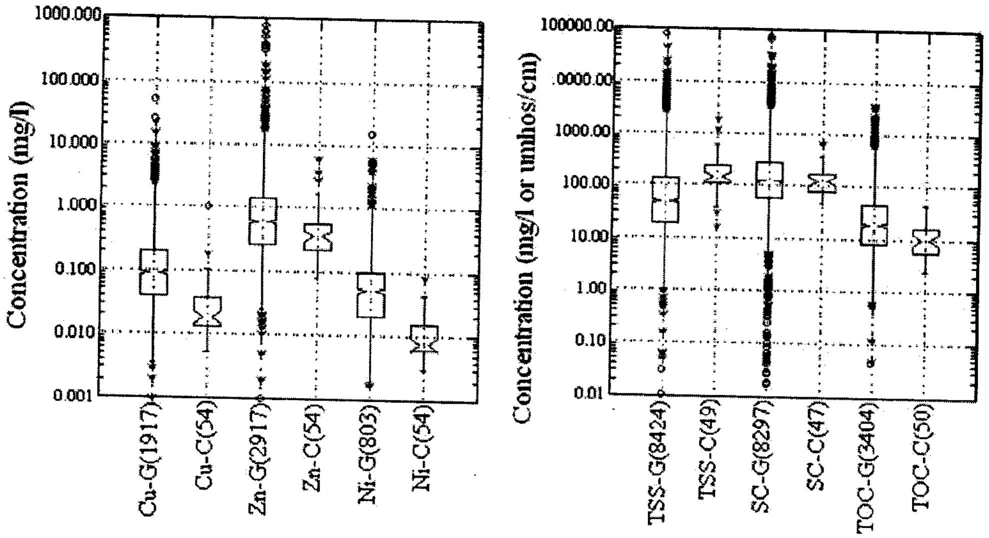


Figure 6—Comparison of grab sample from the industrial stormwater discharge data during 1998–2001 to flow-weighted composite sample from the land-use-monitoring data (industrial land use alone) during 1996–2001. The "G" in the x-axis indicates a grab sample, and "C" indicates a composite sample. The number in parenthesis in the x-axis indicates the number of cases.

the beach water-quality-monitoring program, which uses grab samples and analyses that are not real-time. This creates problems of beach postings, which are out of phase with exceedences.

In this section, we discuss possible reasons for less successful program and suggest ways to improve monitoring. Some suggestions will require new technology.

Sampling Method. For the industrial stormwater permit monitoring, grab samples are allowed, and facility operators are instructed to collect the sample during the first hour of discharge from the first storm event for the wet season (October to May) and at least one other storm event in the wet season. A grab sample is a discrete sample taken within a short period of time, typically less than 15 minutes. Flow-weighted composite samples were collected in the land-use program, which requires instrumentation and, perhaps, site preparation to create a channel for flow measurement and security for equipment. The flow-weighted samplers collect a composite by combining a series of discrete samples of specific volume, collected at specific flow-weighted intervals over the duration of a storm event (LACDPW and Woodward-Clyde, 1998).

It is useful to compare the results of the two programs. They are analogous in that both programs attempt to measure the emissions from a particular human activity, although the industrial program also attempts to identify high dischargers. The results of the 1998–2001 industrial permit using grab samples were compared to the 1996–2001 land-use monitoring that used flow-weighted composites. Figure 6 is a notched bar plot that shows the differences.

There are a large number of outliers among the grab samples and only a few outliers among the composite samples. The number of outliers suggests the need for a quality assurance program, and is helpful in understanding why the neural networks could not identify significant differences in stormwater from SIC codes.

The standard deviations of the concentrations are much lower among the composite samples (Table 2). For example, the standard deviation for TOC is 174 for grab samples and 9.7 for composite samples, or a ratio of 18. The other parameters have ratios of standard deviations from 2.3 (pH) to 66 (zinc). With this large range of differences, one has to question to the utility of such a monitoring program for any purpose. The application of any normalization method of the original data is not useful to generalize for use with a neural model. In addition, there are too many upper and lower outliers in the data set, which results in excessive clipping.

A flow-weighted composite sample for a storm event generally better represents the storm event than a single grab sample, which may be biased because of the collection time. The results of the flow-weighted composite sample can be considered as an event mean concentration (EMC), which can be multiplied by the flowrate to calculate overall mass emissions. This is useful for spreadsheet load models (Wong et al., 1997), which are finding widespread use for planners and TMDL development.

A grab sample suffers from a variety of errors and biases, but one that has not been fully explored is the effect of first flush. Many parameters exhibit a first flush, which is typified by a declining concentration from storm beginning to storm end (Ma et al., 2002a). When the grab sample is collected early in the storm, it will be higher than the EMC; conversely, if collected too late, it will be lower than the EMC. The industrial-monitoring program suggests collecting a sample within the first hour. To improve the results from grab samples, it is necessary to find the best time to sample. For example, the best time for sampling O&G from highway land use is between two and three hours and is related to cumulative rainfall and duration (Ma et al., 2002b). There might be some improvement in the existing program with better definition of collection times.

It is almost universally recognized that composite samplers are better for stormwater monitoring; however, to collect a flow-weighted composite sample, an automatic sampler must be installed and operated properly before a storm event. It would be a burden to all industrial permittees to construct composite sampling facilities. Additionally, several water-quality parameters, such as O&G and indicator bacteria, are not easily measured by a composite sample.

Table 2—Comparison of grab to composite sample (0 indicates level below detection limit).

Water-quality parameters	Grab Industrial stormwater permit sample 1998–2001	Flow-weighted composite Landuse monitoring sample 1996–2001 (industrial site alone)
pH		
Number	8584	51
Minimum	0.1	6.04
Maximum	12.7	8.32
Median	6.88	6.82
Mean	6.91	6.83
Standard deviation	0.96	0.41
Total suspended solids (mg/L)		
Number	8424	49
Minimum	0	16
Maximum	101 000	1865
Median	48	140
Mean	219.11	232.55
Standard deviation	1693	298
Specific conductance (µmhos/cm)		
Number	8297	47
Minimum	0.017	48.9
Maximum	71 000	691
Median	121	126
Mean	365.17	150.06
Standard deviation	1555	111
Oil and grease (mg/L)		
Number	6685	
Minimum	0	
Maximum	6640	
Median	5	Not analyzed
Mean	13.63	
Standard deviation	95	
Total organic carbon (mg/L)		
Number	3404	50
Minimum	0	2.4
Maximum	3700	45.62
Median	18	9.85
Mean	56.01	12.67
Standard deviation	174	9.7
Lead (mg/L)		
Number	171	Low detection frequency
Minimum	0	
Maximum	90	
Median	0.06	
Mean	0.402	
Standard deviation	3.5	
Copper (mg/L)		
Number	1917	54
Minimum	0	0.0053
Maximum	49.5	0.99
Median	0.084	0.0185
Mean	0.337	0.047
Standard deviation	1.6	0.13
Zinc (mg/L)		
Number	2917	54
Minimum	0	0.079
Maximum	2200	5.97
Median	0.6	0.36
Mean	4.86	0.63
Standard deviation	64.4	0.97
Nickel (mg/L)		
Number	803	54
Minimum	0	0
Maximum	15.1	0.0804
Median	0.05	0.005995
Mean	0.196	0.0082
Standard deviation	0.76	0.013

To improve sampling, it might be reasonable to randomly select a small subset of industrial users for composite sampling. This might be funded by fee permittees or by allowing a reduced number of grab samples to be collected. A trained team would also increase quality assurance to eliminate outliers. Such an approach might be a better or less-expensive method of determining stormwater emissions on receiving waters.

Parameter Selection. A variety of metal-related industries are included among the SIC codes in the industrial monitoring program. Many industries should be sources of metals, such as chromium, copper, lead, nickel, and zinc (Woodward Clyde, 1992). Figure 7 shows the mean concentration of the basic analytical parameters and metals as a function of their industrial categories for the 1998–2001 seasons. Outliers, defined earlier, have been removed. The numbers of cases for all parameters vary with as many as 800 for conventional parameters and only approximately 80 for metals. The conventional water-quality parameters show much less relation to industrial category than metals. The mean concentrations of lead, zinc, and nickel were highest for the primary metal industries category, and copper was highest at the transportation equipment facilities category. Mean concentrations of O&G and TOC were highest for whole trade-durable good industries and mean concentration of conductivity and suspended solids were highest for electric, gas, and sanitary-service facilities.

In addition, concentrations of metals exceeded the stormwater benchmark values suggested by the U.S. Environmental Protection Agency (U.S. EPA) more frequently than the basic water-quality parameters (Table 3). The concentrations of metals (except nickel) are mostly above the benchmark levels. For the cases of zinc, approximately 90% of observations exceeded the benchmark.

The addition of metals to the basic permit's requirement for basic water-quality parameters would be a useful way of adding information to the dataset. A neural model trained with both metals and basic parameters will perform better than that trained with existing water-quality parameters or metals alone. The addition of metals will increase the monitoring cost. Table 4 shows the current costs for laboratory analysis. The addition of metals to the permit will approximately double or triple the laboratory costs. The cost of collecting the samples should be quite similar. Cost increases are probably inevitable, but this approach may be less expensive than other approaches.

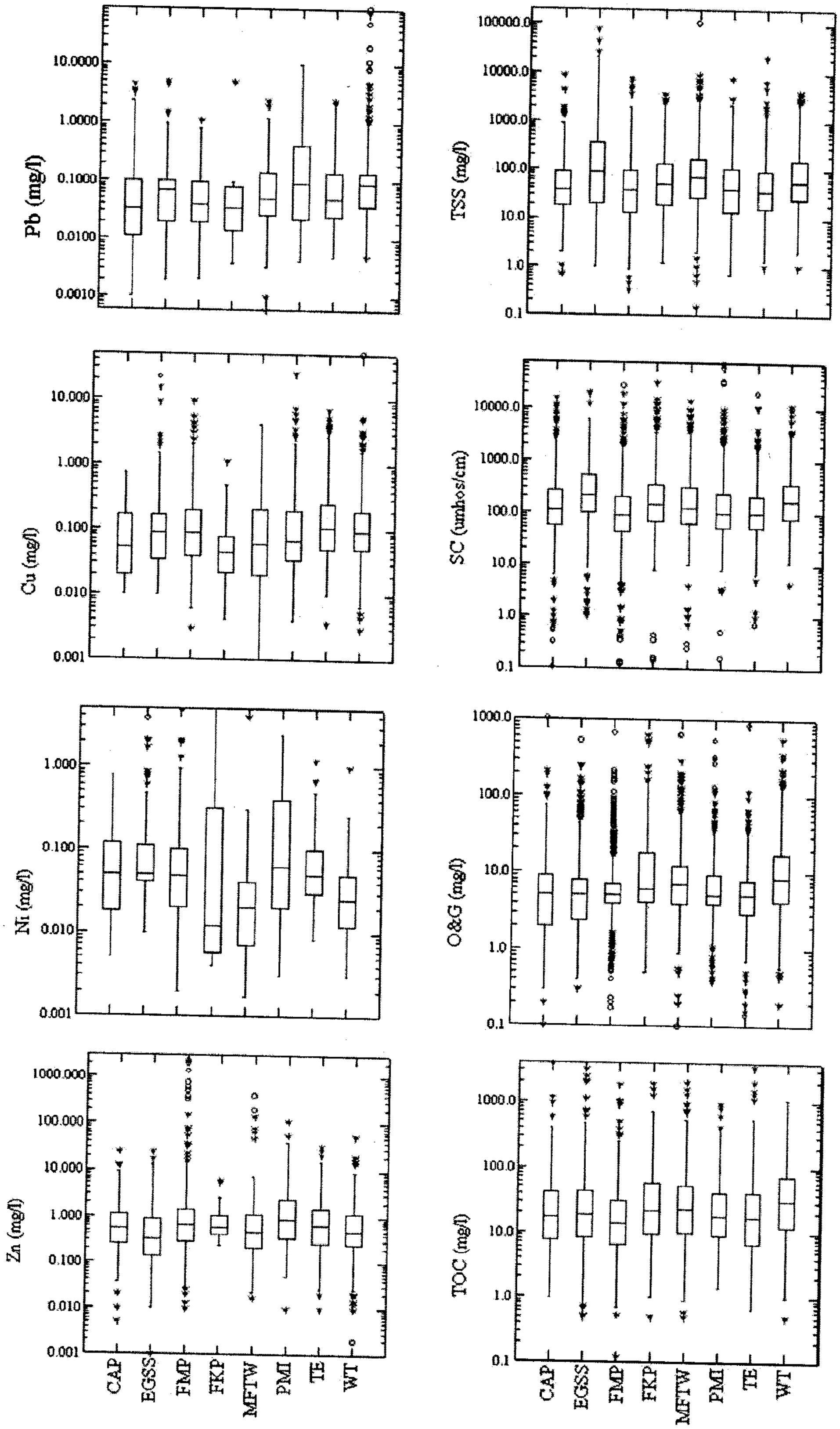


Figure 7—Distribution of the analytical parameters and metals for each industrial category. The number of cases is varied.

Table 3—Comparison with the U.S. EPA benchmark levels for parameters for the 1998–2001 wet seasons.

Water-quality parameters	No. of observations	No. of outside benchmark values	Percentage of outside benchmark values
pH	13 770	1571	11.4
Total suspended solids	13 527	4266	31.5
Specific conductance	13 155	4617	35.1
Oil and grease	10 780	1560	14.5
Total organic carbon	5406	499	9.2
Aluminum	1487	898	60.4
Copper	2505	1441	57.5
Iron	1762	1205	68.4
Nickel	1122	20	1.8
Lead	2230	906	40.6
Zinc	3615	3227	89.3

Conclusions

This paper has examined three stormwater monitoring programs. The utility of the programs have been assessed based on the programs' ability to accurately estimate the emissions for different classes of land uses and other obvious benefits. The following conclusions are made:

(1) Data collected by grab samples had much higher variability than composite samplers. The coefficients of variation (standard deviation divided by the mean) for the same parameters were generally 2 to 9 times higher for the grab samples. The variability suggests that composite samples should be collected, even if it means a reduction in the total number of samples or facilities that can be monitored.

(2) The time required to analyze a sample must be commensurate with the intended use of the results. Beach water-quality monitoring suffers from analysis time for indicator organisms. The data suggests that 70% of the beach postings are out of phase with the water-quality parameter exceedence.

(3) Metals are major pollutants in industrial land use and potentially more useful to distinguish industrial categories or land-use patterns. Metal concentrations frequently exceeded U.S. EPA benchmark concentrations. Adding them to existing permits might increase cost, but will add value to the resulting monitoring database.

Managing stormwater is a developing technology and much remains to be done. This paper has shown that, even with the limited experience we have thus far, there are improvements that can be made.

Acknowledgments

Credits. This research was sponsored in part by a U.S. EPA-sponsored watershed project. We are grateful to the County of Los Angeles, Department of Public Works (California) and Kosta Kaporis, Bureau of Sanitation, City of Los Angeles, for his assistance in obtaining datasets.

Authors. At the time of this work, Haejin Lee was a postdoctoral researcher in the Department of Civil and Environmental Engineering at University of California at Los Angeles (UCLA). Michael K. Stenstrom is a professor in the Department of Civil and Environmental Engineering at UCLA. Correspondence should be addressed to Michael K. Stenstrom, Prof., Department of Civil and Environmental Engineering, 5714 Boelter Hall, University of California, Los Angeles, CA 90005; e-mail: stenstro@seas.ucla.edu.

Submitted for publication September 30, 2002; revised manuscript submitted November 6, 2003; accepted for publication November 6, 2003.

The deadline to submit Discussions of this paper is (??).

References

Aguilera, P. A.; Garrido Frenich, A.; Torres, J. A.; Castro, H.; Martinez Vidal, J. L.; Canton, M. (2001) Application of the Kohonen Neural Networks in Coastal Water Management: Methodological Development for the Assessment and Prediction of Water Quality. *Water Res.*, **35,** 4053.

Grant S. B.; Sanders, B. F.; Boehm, A. B.; Redman, J. A.; Kim, J. H.; Mrse, R. D.; Chu, A. K.; Gouldin, M.; McGee, C. D.; Gardiner, N. A.; Jones, B. H.; Svejkovsky, J.; Leipzig, G. V. (2001) Generation of Enterococci

Table 4—Comparison of the current requirement parameters to the complementary parameters on the costs of laboratory analysis (data source: Los Angeles County Department of Agricultural Commissioner/Weights and Measures Environmental Toxicology Laboratory, Arcadia, California).

		Current requirement parameters		Complementary parameters	
Water-quality parameters	Cost per sample ($)	pH, TSS, SC, and TOC	pH, TSS, SC, and O&G	pH, TSS, lead (Pb), copper (Cu), and zinc (Zn)	pH, TSS, SC, O&G, Pb, Cu, Zn, aluminum (Al), and iron (Fe)
pH	3.50	3.50	3.50	3.50	3.50
Total suspended solids (TSS)	7.68	7.68	7.68	7.68	7.68
Specific conductance (SC)	6.40	6.40	6.40	6.40	6.40
Total organic carbon (TOC)	23.46	23.46		23.46	23.46
Oil and grease (O&G)	36.25		36.25		
Copper (Cu)	20.29			20.29	20.29
Lead (Pb)	20.29			20.29	20.29
Zinc (Zn)	20.29			20.29	20.29
Aluminum (Al)	20.29				20.29
Iron (Fe)	20.29				20.29
Total cost per sample ($)		41.04	53.83	101.91	142.49

bacteria in a Coastal Saltwater Marsh and its Impact on Surf Zone Water Quality, *Environ. Sci. Technol.*, **35,** 2407.

Ha, H.; Stenstrom, M. K. (2003) Identification of Land Use with Water Quality Data in Stormwater using a Neural Network. *Water Res.*, **37,** 4222.

Haile, R. W.; Witte, J. S.; Gold, M.; Cressey, R.; Mcgee, C.; Millikan, R. C.; Glasser, A.; Harawa, N.; Ervin, C.; Harmon, P.; Harper, J.; Dermand, J.; Alamillo, J.; Barrett, K.; Nides, M.; Wang, G. (1999) The Health Effects of Swimming in Ocean Water Contaminated by Storm Drain Runoff. *Epidemiology,* **10,** 355.

Jiang, S.; Noble, R.; Chu, W. (2001) Human Adenoviruses and Coliphages in Urban Runoff Impacted Coastal Waters of Southern California. *Appl. Environ. Microbiol.*, **67,** 179.

LACDPW Woodward-Clyde (1998) Los Angeles County 1997–1998 Stormwater Monitoring Report. Los Angeles County Department of Public Works, Alhambra, California, Woodward-Clyde, Oakland, California, and Southern California Coastal Water Research Project, Westminster, California.

Leecaster, M.; Weisberg, S. (2001) Effect of Temporal Sample Frequency on Shoreline Microbiology Assessments. Southern California Coastal Water Research Project, Westminster, California.

Ma, J-S.; Khan, S.; Li, Y-X.; Kim, L-H.; Ha, S.; Lau, S-L.; Kayhanian, M.; Stenstrom, M. K. (2002a) First Flush Phenomena for Highways: How it Can Be Meaningfully Defined. *Proceedings of the 9th International Conference on Urban Drainage*, Portland, Oregon, September 8–13.

Ma, J-S.; Khan, S.; Li, Y-X.; Kim, L-H.; Ha, H.; Lau, S-L.; Kayhanian, M.; Stenstrom, M. K. (2002b) Implication of Oil and Grease Measurement in Stormwater Management Systems. *Proceedings of the 9th International Conference on Urban Drainage*, Portland, Oregon, September 8–13.

Noble, R. T.; Dorsey, J. H.; Leecaster, M.; Orozco-Borbon, V.; Reid, D.; Schiff, K.; Weisberg, S. B. (2000) A Regional Survey of the Microbiological Water Quality along the Shoreline of the Southern California Bight. *Environ. Monitor. Assess.*, **64,** 435.

Orange County Sanitation District (1999) Huntington Beach Closure Investigation—Final Report. Orange County Sanitation District: Fountain Valley, California.

Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R. (1993) Sources of Fine Organic Aerosol. 3. Road Dust, Tire Debris, and Organometallic Break Lining Dust: Roads as Sources and Sinks. *Environ. Sci. Technol.*, **27,** 1892.

Sansalone, J. J.; Buchberger, S. G. (1997) Partitioning and First Flush of Metals in Urban Roadway Stormwater. *J. Environ. Eng.*, **123,** 134.

State Water Resources Control Board (2001) California Beach Closure Report 2000. Division of Water Quality, State Water Resources Control Board, California Environmental Protection Agency, Sacramento, California.

Stenstrom, M. K.; Strecker, E. (1993) Assessment of Storm Drain Sources of Contaminants to Santa Monica Bay, Vol. III, Surface Drainage Water Quality Monitoring Program Plan. UCLA-ENG-93-64, May 1993, Vol. III, 1–59.

Stenstrom, M. K.; Silverman, G. S.; Bursztynsky, T. A. (1984) Oil and Grease in Urban Stormwaters. *ASCE J. Environ. Eng. Div.*, **110,** 58.

Stenstrom, M. K.; Lau, S-L.; Lee, H-H.; Ma, J-S.; Ha, H.; Kim, L-H.; Khan, S.; Kayhanian, M. (2000) First Flush Stormwater Runoff from Highways. Research Report for California Department of Transportation, University of California, Los Angeles.

Stenstrom, M. K.; Lau, S-L.; Ma, J-S.; Ha, H.; Lim, L-H.; Lee, S.; Khan, S.; Kayhanian, M. (2001) First Flush Stormwater Runoff from Highways. Research Report for Department of California Transportation, University of California, Los Angeles.

Wong, K.; Strecker, E. W.; Stenstrom, M. K. (1997) A Geographic Information System to Estimate Stormwater Pollutant Mass Loadings. *ASCE J. Environ. Eng.*, **123,** 737.

Woodward Clyde (1992) Source Identification and Control Report. Prepared for Santa Clara Valley Urban Pollution Program, December 1992. Woodward-Clyde, Oakland, California.

Final Report
Industrial Storm Water Monitoring Program
Existing Statewide Permit Utility and Proposed Modifications

Michael K. Stenstrom and Haejin Lee
Civil and Environmental Engineering Department, UCLA
Los Angeles, California

January 2005

Final Report
Industrial Storm Water Monitoring Program
Existing Statewide Permit Utility and Proposed Modifications

Michael K. Stenstrom and Haejin Lee
Civil and Environmental Engineering Department, UCLA
Los Angeles, California

January 2005

ABSTRACT

Data collected from the California statewide General Industrial Storm Water Permit (GISP) was examined over the nine-year period from 1992 to 2001. The data were evaluated to determine if it could be used to identify permittees with high emissions and if the storm water loads from various classes of industries could be characterized in order to create rankings and typical emission rates. It was hoped that the GISP would provide information for regulators and others to better implement future storm water management programs.

The data collected by the permittees is highly variable, with coefficients of variation as high as 15. This compares to coefficients of variation, generally less than 0.5 for other environmental monitoring programs, such as water and wastewater treatment plant influents. There are several sources for the variability and the use of grab samples, untrained sampling personnel, and a limited selection of monitored parameters are among the largest sources.

The requirements of a new monitoring program are proposed. They include a broadened suite of parameters, use of composite samples and certified laboratories, joint sampling programs and a web-based reporting system.

ACKNOWLEDGEMENTS

This work was supported in part by funds from the US EPA through cooperative agreement CP-82969201, from the California State Water Resources Control Board, contract number 02-172-140-0. The authors are grateful for their support.

The authors are also grateful for the assistance of the review team and Xavier Swamikannu and Dan Radulescu of the California Regional Water Control Board, Los Angeles Region.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The completion of wastewater treatment plants mandated by the Clean Water Act has reduced pollution from point sources to the waters of the United States. As a result, non-point sources pollution such as storm water runoff are now major contributors to pollution of receiving waters. Storm water pollution control has not enjoyed the technology benefits of secondary treatment and remains a difficult task for agencies and permittees. The problem of storm water pollution is growing worse because of continuing development, which results in increased impervious surface area. In order to reduce storm water pollution, regulatory agencies are requiring storm water monitoring programs. The programs are implemented under the National Pollutant Discharge Elimination System (NPDES) storm water permits.

The overall goal of storm water monitoring program includes the identification of high-risk polluters, also in addition to the development of a better understanding of the mechanisms and sources of storm water pollution, with the long-term goal of reducing pollutants to less harmful levels. Most storm water monitoring programs are relatively new, and evaluations of their usefulness for satisfying these goals are only now possible (Duke et al., 1998; Lee and Stenstrom, 2004; Pitt et al., 2003). Lee and Stenstrom (2004) recently evaluated facility-monitoring in Los Angeles County under the statewide General Industrial Storm Water Permit (GISP) and permittee monitoring in the Los Angeles County Municipal Storm Water Permit (LA County Municipal Storm Water Permit) for three wet seasons (1998 - 2001) and five wet seasons (1996 - 2001), respectively. The results of our previous study suggest that parts of the current GISP monitoring program will not be helpful to identify high dischargers, nor will they be useful in developing Total Maximum Daily Loads (TMDLs). The design and requirements of the monitoring program do not produce data with sufficient precision for decision-making.

In this report, we summarize the results of our evaluation of several storm water monitoring programs to determine their usefulness in achieving their dual goals, as well as making recommendations for improvement. Our analysis included more than 20 monitoring programs and datasets. The appendix includes a draft of a submitted paper that more thoroughly discusses the statistical results.

Based upon these findings, the recommended requirements for a new GISP that will achieve the stated objectives are proposed. The proposed new requirements include additional parameters, trained sampling teams with QA/QC requirements, more sophisticated sampling techniques which are mindful of the first flush phenomenon, real-time reporting requirements and possible reorganization of industry group categorizations.

2. PROJECT BACKGROUND

2.1 Study objectives

The goal of this study is to evaluate the existing statewide General Industrial Storm Water Permit (GISP) and to develop a new program that better characterizes storm water discharges and better serves the regulated community to make informed decisions to reduce storm water pollution. The following steps were envisioned:

- compile the industrial wet weather monitoring information
- conduct crucial QA/QC evaluation of these data
- conduct informative statistical analyses
- evaluate the current monitoring program
- recommend a new monitoring plan enabling management decision-making

As a part of this process, a committee of experts and representatives from the various stakeholders in the monitoring plan was assembled to review our findings and comment on proposed modifications to the monitoring program.

The original goal of the monitoring program associated with the GISP was to identify polluters and modify their behavior to deter or reduce future storm water pollution. A second and equally important goal was to collect information to better understand storm water pollution, develop Total Maximum Daily Loads (TMDLs) and formulate pollution reduction plans. The overall goal is to eventually reduce pollutants in storm water discharges from various industries at reasonable cost.

2.2 Report Organization

The report is divided into five sections. The basic statistic results for data from nine years of GISP monitoring for the Los Angeles Region (1992 –2001) are presented in section 3.1. The last three years of this data set were analyzed previously (Ha and Stenstrom, 2002). The first six years of data became available afterwards. Section 3.2 includes the analysis of the early data and confirms previously cited conclusions.

Identifying a review committee is described in section 4. Developing the requirements for an improved monitoring program are discussed in section 5. The findings and suggestions based on the present results are summarized in section 6.

2.3 Background

On November 19, 1991, the State Water Resource Control Board (the State Board) issued the statewide General Permit for Discharges of Storm Water Associated with Industrial Activities Excluding Construction Activities (the GISP). The GISP requires facilities that discharge storm water associated with the industrial activities directly or indirectly into the State waters must apply and obtain coverage under the permit. In 1997, the California State Water Resources Control Board (State Board) renewed the GISP as State Board Order No. 97-03-DWQ.

Currently, there are approximately 2962 permittees within Los Angeles County that are covered under the GISP. Under the monitoring requirements, permittees must collect water quality samples from two storms per year and analyze for four basic parameters: pH, specific conductance, total suspended solids, and oil and grease. Total organic carbon can be substituted for oil and grease. Certain facilities must analyze for specific additional pollutants, and permittees must analyze for pollutants that they believe are pervasive in their storm water. Permittees in some cases may be required to sample at more than one location. Industries are categorized by Standard Industrial Classification Codes (SIC codes).

The original goal of the monitoring program was to identify high-risk polluters, influence their pollution prevention behavior, and to create a database to help in the development of Total Maximum Daily Loads (TMDLs). It is natural to ask if the monitoring program is valuable. To answer this question we reviewed the current GISP program with the latest three years of monitoring data. We determined that the monitoring data was unrelated to industry type based on SIC code. We therefore propose the selection of appropriate pollutant parameters and sampling methods to improve upon the current monitoring program.

3. ANALYSIS OF EXISTING DATA

Table 3.1 shows case numbers for each of the nine years (92-93, 93-94, 94-95,95-96,96-97, 97-98, 98-99, 99-00, 00-01). A case number is a set of reported data for a single sample, and includes all the analysis for that particular sample. Table 3.2 shows the cases by year and reports the number of observations by water quality parameter. For some years and parameters, no information was provided. In general SIC codes, imperviousness and other site-specific information were provided.

3.1 Basic statistic analysis

This report uses many figures and tables that have been generated from various software packages. Systat 10.2 (SPSS Inc., Chicago, IL) was used for most cases. It generates "box plots," and the general key to the box plots is shown in Figure 3.1 The Figure shows the box, which extends from the 25 to 75% of the data, as ordered by the parameter being examined. The horizontal line in the box plot shows the median (50% point) of the data. The "whiskers" or upper and lower fences represent the 25 or 75 percentile values, plus or minus 1.5 times the difference of the 25 and 75 percentile values. Values beyond the two fences are termed outliers and may be deleted in further analysis, depending on the nature of the analysis. This procedure is very commonly applied and further information is available (SYSTAT 10 user's guide).

Outliers in the data set were first eliminated. Outliers are defined as any data point that lies below lower inner fence or above upper inner fence (Figure 3.1). Histograms are most useful for large data set to display the shape of data distribution. Histograms for TSS, Specific Conductance, Oil and Grease, and TOC are shown in Figure 3.2.The histogram shows the number of samples that fall within specific concentration intervals. Histograms for BOD, COD, aluminum, and copper are shown in Figure 3.3. Histograms for iron, lead, nickel, and zinc are shown in Figure 3.4. The left height of each bar represents the number of data points that fall inside the interval covered by the bar. The right of each bar indicates the proportion of data points. In general, the distribution is

Table 3.1 Case number of each season for nine years (1992-2001)

Season	Case number	Comment
92-93	632	
93-94	2021	
94-95	2555	
95-96	2812	
96-97	2393	
97-98	863	
98-99	4477[a]	Previously analyzed
99-00	4719[a]	Previously analyzed
00-01	4949[a]	Previously analyzed
00-02	NA	Not received yet

[a] After deletion of cases that have no value.

Table 3.2. Contents of each season and case number according to its parameter for nine years (1992-2001)

Season	92-93	93-94	94-95	95-96	96-97	97-98	98-99	99-00	00-01
Sequence	y	y	y	y	y	y	y	y	y
Date of Storm	n	n	y	y	n	y	y	y	y
Date Analyzed	n	n	n	n	n	n	y	y	y
pH	619	2004	2477	2749	2347	824	4374	4588	4870
TSS (mg/L)	556	1928	2374	2680	2182	827	4371	4543	4637
SC (μmhos/cm)	553	1864	2316	2612	2212	802	4219	4383	4588
O&G (mg/L)	367	1242	1831	2013	1670	690	3615	3602	3554
TOC (mg/L)	228	727	896	1037	1045	346	1784	1866	1777
BOD (mg/L)	–	–	99	120	115	20	–	214	157
COD (mg/L)	–	–	183	124	86	87	414	482	453
Al (mg/L)	–	25	44	10	–	49	442	461	581
Cr (mg/L)	–	127	–	–	91	–	–	–	–
Cd (mg/L)	–	62	–	–	–	–	–	–	–
Cu (mg/L)	–	–	366	313	16	145	758	846	899
Fe (mg/L)	–	–	18	19	–	41	458	600	701
Pb (mg/L)	–	245	405	342	166	135	604	800	816
Ni (mg/L)	–	–	291	237	130	77	288	424	403
Zn (mg/L)	–	471	440	414	5	208	1046	1222	1338
Toluene	–	–	–	–	354	–	–	–	–
Xylene	–	–	–	–	6	–	–	–	–
TDS	–	–	–	–	2	–	–	–	–
Facility_SiteSize	y	y	y	y	y	y	y	y	y
FacilitySizeUnits	y	y	y	y	y	y	y	y	y
FacilityPercentImpervious	y	y	y	y	y	y	y	y	y
SIC1	y	y	y	y	y	y	y	y	y
SIC2	y	y	y	y	y	y	n	n	n
SIC3	y	y	y	y	y	y	n	n	n

1) – indicates sample was not analyzed
2) y indicates the information was provided but n indicates the information was not provided

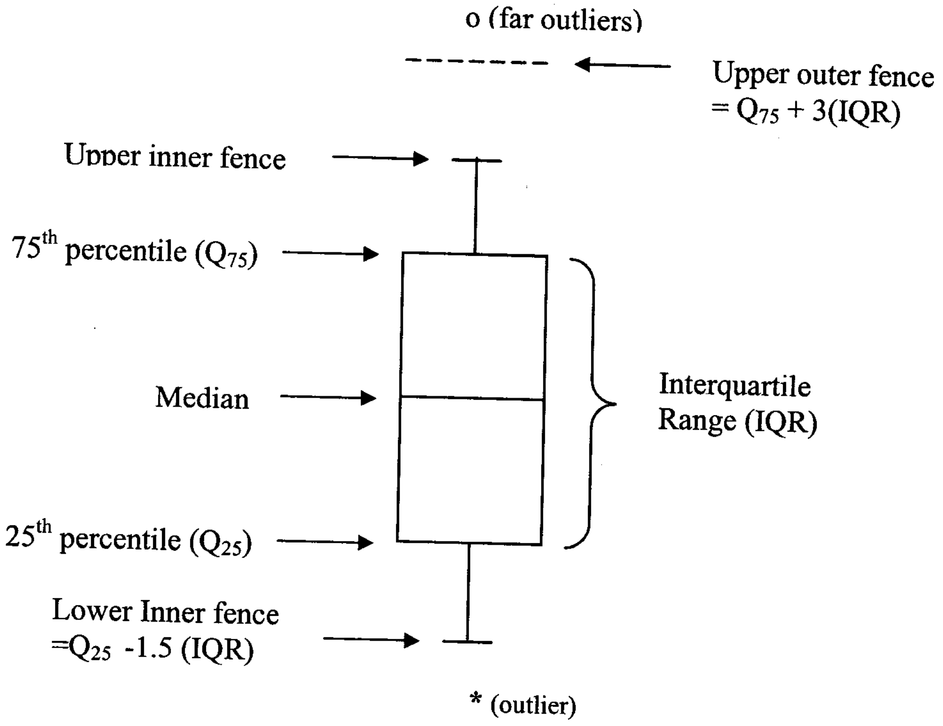


Figure 3.1. Anatomy of a boxplot.

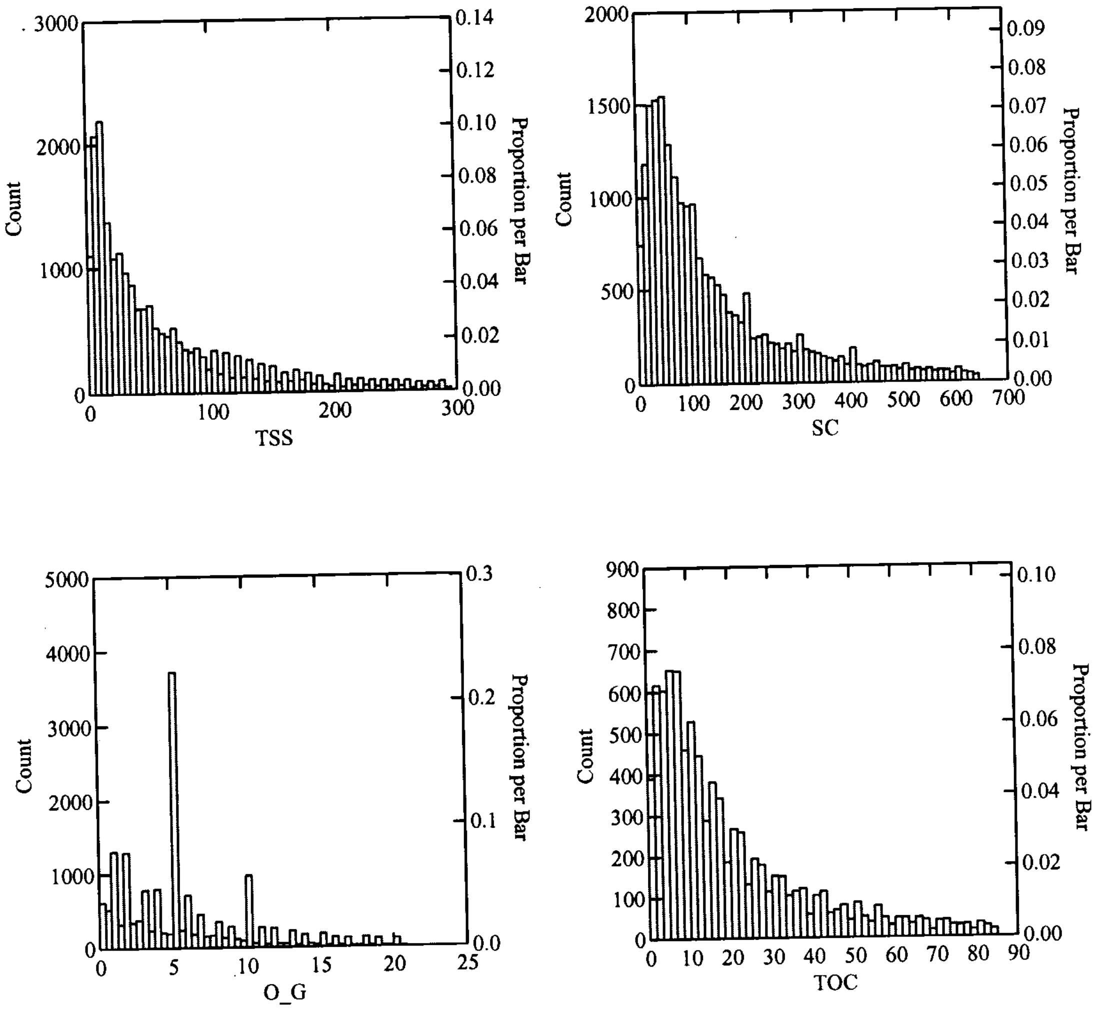


Figure 3.2. Histogram for TSS, SC (Specific Conductivity), O_G(Oil and Grease) and TOC(Unit of the parameters are mg/L except SC. The unit for SC is μmhos/cm).

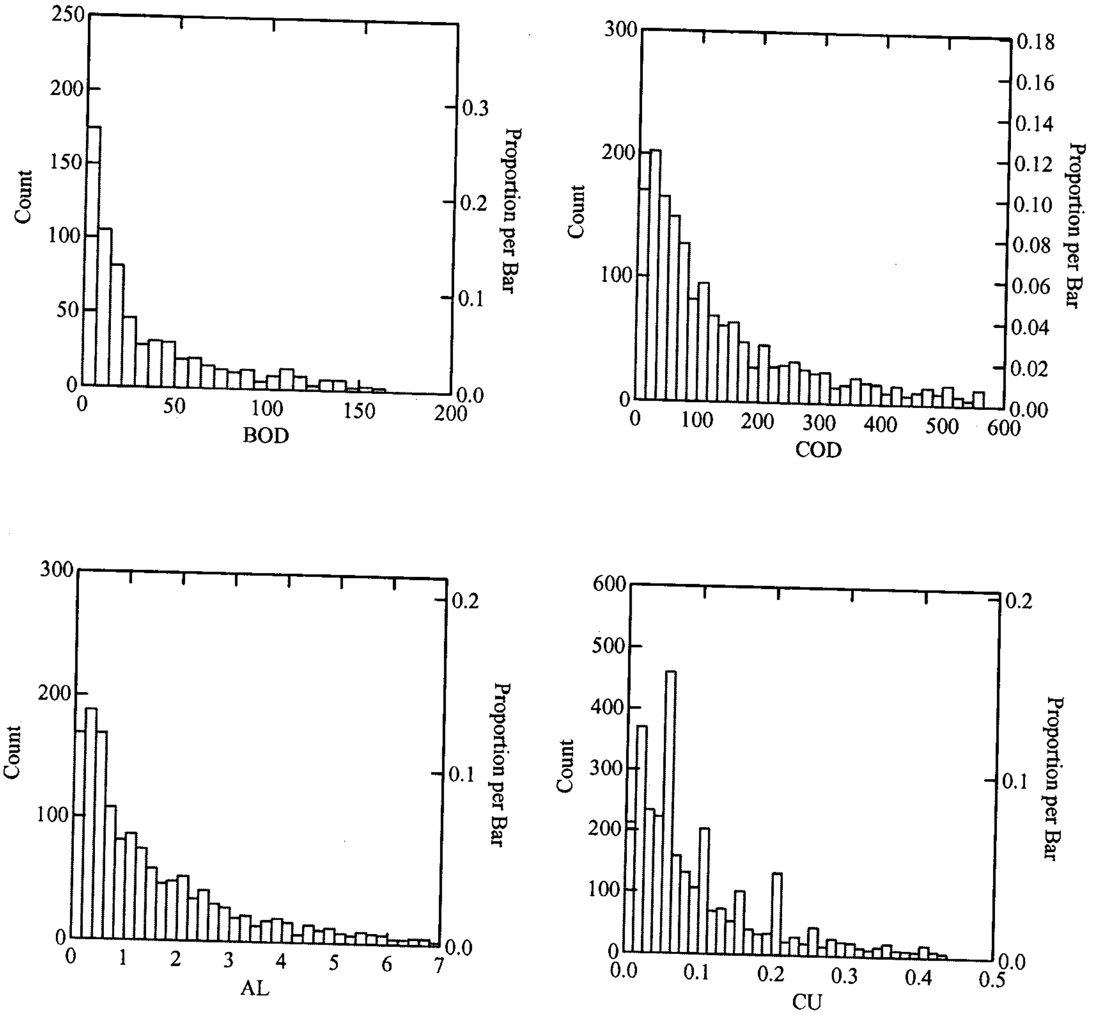


Figure 3.3. Histogram for BOD, COD, Al and Cu(Unit of the parameters are mg/L).

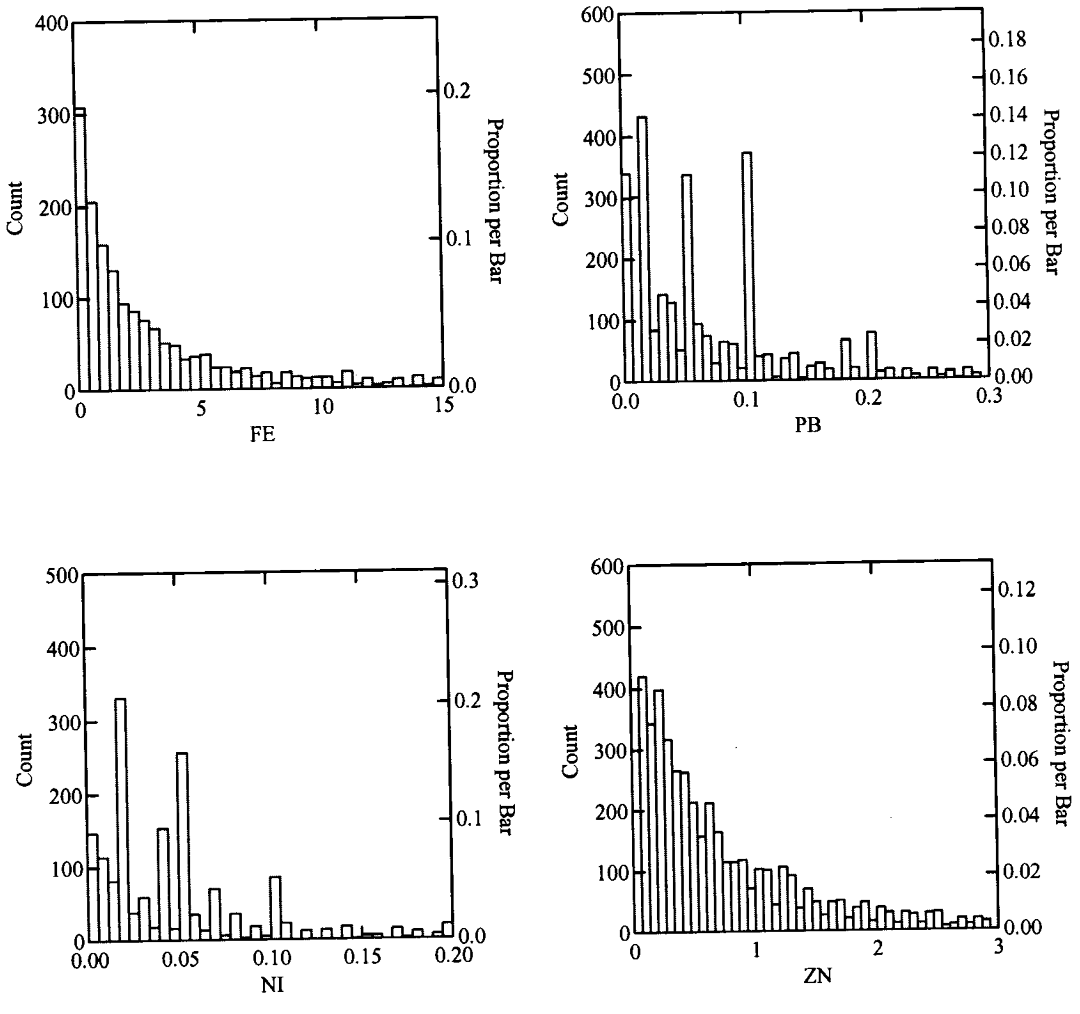


Figure 3.4. Histogram for Fe, Pb, Ni, and Zn (Unit of the parameters are mg/L)

severely skewed toward the low values. Oil and Grease, copper, lead, and nickel readings show a highest percentage of values at 5mg/l, 0.05 mg/l, 0.014 mg/l, and 0.016 mg/l respectively. The large number of cases in these lower intervals probably represents observations that fall at or below detection limits. The detection limits are not know and are probably different for each case, since a variety of contract labs were probably used by the permittees.

To see the relationship between parameters, the Pearson correlation value is obtained using Systat 10. The procedure compares each parameter to every other parameter, which means that a correlation coefficient is obtained for every possible combination of two parameters. Table 3.3 shows the correlation value and the number of observations. There is a close relationship between aluminum and iron, copper and lead, copper and nickel, and lead and nickel. The rest of them are poorly related.

3.2 Reevaluation of existing data

Eight industrial sectors were selected based on their prevalence in the data set, which means some SIC codes are more numerous than others. The selected eight industrial sectors are shown in Table 3.4. The distributions of the conventional water quality parameters and metals as a function of their industrial sectors for nine years are displayed as Figures 3.5 through 3.7. The number of cases for all parameters varies from as many as 2887 for TSS to only 9 for aluminum. The dashed line in the figures indicates benchmark levels (BLs) that have been suggested by US EPA for storm water. All the observations for TOC and nickel were below the BL at all major sector industries. The BL for TOC, 110 mg/L, is above the highest graphed value and is not shown. Oil and Grease remained below BL at most major facilities. Median concentration of zinc exceeds BL at all major facilities. The median concentration of Oil and Grease, Specific Conductivity, and TOC were highest for wholesale trade-durable good (WT) facilities. The median concentration of aluminum and iron were highest for food and kindred products (FKP) but their observation number is fewer than the other parameters (see Table 3.5). In general, distribution of conventional water quality parameters and metals

Table 3.3. Pearson correlation value (lower part) and the observation number (upper part) for nine years (1992 -2001)

	pH	TSS	SC	O & G	TOC	Al	Cu	Fe	Pb	Ni	Zn
pH		23850	23358	18275	9549	1595	3290	1816	3443	1802	5060
TSS	0.008		22803	17956	9409	1581	3236	1801	3417	1788	4991
SC	-0.001	0.027		17437	9250	1568	3163	1774	3320	1690	4851
O & G	-0.001	0.015	0.001		5219	1305	2695	1411	2829	1450	3933
TOC	-0.052	0.067	0.029	0.162		331	988	557	1078	677	1519
Al	0.019	0.126	0.013	0.042	0.128		668	1113	535	187	1250
Cu	0.014	0.003	-0.003	0.001	0.089	0.14		677	2397	1519	3043
Fe	-0.056	0.293	0.197	0.006	0.007	■	0.21		650	275	1168
Pb	0.009	0.002	-0.008	-0.001	0.011	0.4	■	0.02		1544	2783
Ni	0.001	0.005	-0.009	-0.001	-0.009	0.024	■	0.33	■		1513
Zn	-0.085	0.01	0.061	0.003	0.197	0.121	0.321	0.671	0.143	0.165	

1) Pearson correlation value is highlighted.
2) Dark shell indicates a close relationship between two parameters.

Table 3.4. Description of the selected eight major sectors

Description	SIC code
Food and kindred products (FKP)	20
Chemical and allied products (CAP)	28
Primary metal industries (PMI)	33
Fabricated metal products, except machinery and transportation equipment (FMP)	34
Transportation equipment (TE)	37
Motor freight transportation and warehousing (MFTW)	42
Electric, gas, and sanitary services (EGSS)	49
Wholesale trade-durable goods (WT)	50

Table 3.5. Observation number according to the parameter after clipping outliers for 1992-2001

	Upper inner fence	FKP	CAP	PMI	FMP	TE	MFTW	EGSS	WT
TSS	294.5	1451	2293	1586	2887	1498	1674	1173	1235
SC	641.0	1497	2180	1585	2751	1474	1685	1254	1245
O&G	20.8	905	1629	1257	2335	1093	1241	1173	898
TOC	84.5	774	1062	448	965	545	708	677	342
Al	6.8	9	69	287	505	71	29	21	302
Cu	0.44	16	53	478	425	279	159	322	570
Fe	15.1	9	82	153	480	72	30	262	288
Pb	0.3	19	152	211	415	392	136	362	696
Ni	0.2	15	38	111	361	185	92	264	116
Zn	2.9	42	446	536	974	469	207	339	614

1) The description of the selected sectors was described in Table 3.4.
2) Unit of the parameters is described in table 2.

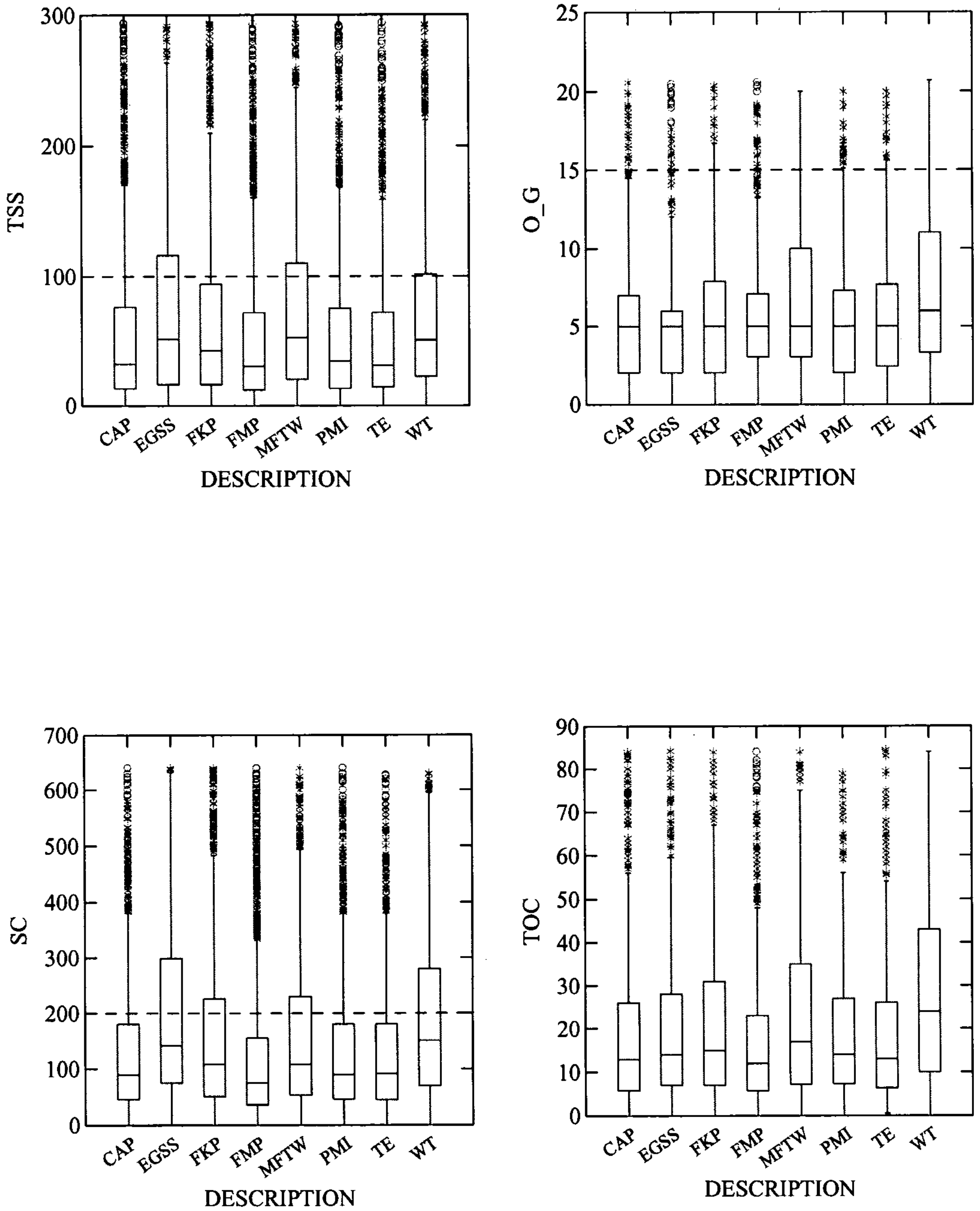


Figure 3.5. Distribution of TSS, O_G (Oil and Grease), SC (Specific Conductance) and TOC for nine years

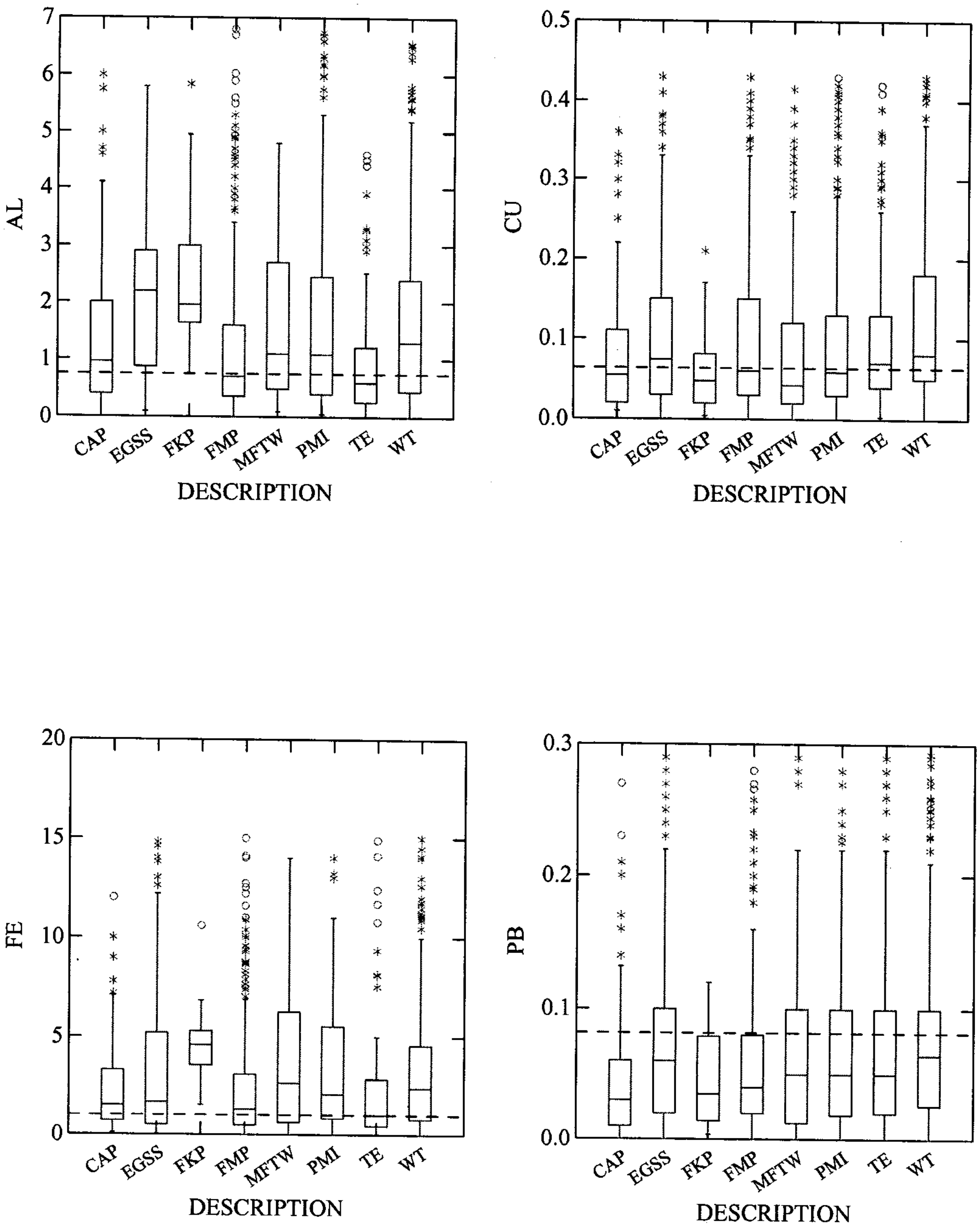


Figure 3.6. Distribution of Al, Cu, Fe, and Pb for nine years

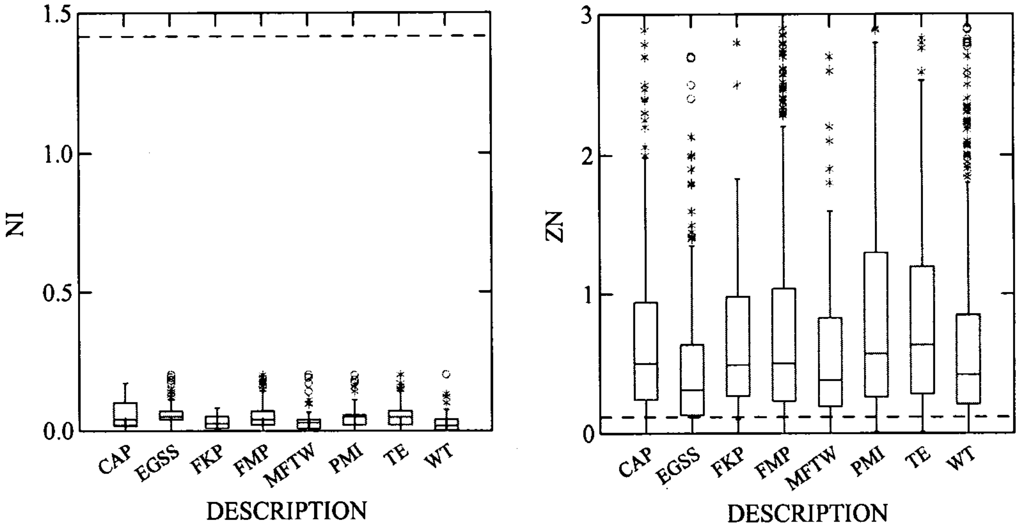


Figure 3.7. Distribution of Ni and Zn nine years

are similar between various industrial activities. However, observation of metals is limited, because only certain facilities must analyze for metal.

Figures 3.8 and 3.9 show the number and percentage of outliers by sector for the conventional water quality parameters and metals. Over 20 percent of Electric, Gas and Sanitary services (EGSS) observations are outliers for SC and TSS. Over 20 percent of Primary Metal Industries (PMI) observations are outliers for lead and zinc. Over 20 percent of Transportation Equipment (TE) samples are outliers for copper.

Discriminant analysis (Systat 10) was used as a preliminary classification approach to discriminate among the eight sectors. The data cases that contain both the mandatory conventional water quality parameters (pH, TSS, SC, TOC, Oil and Grease) and metals are very limited. For this reason, discriminant analysis was performed only for conventional water quality parameters. The analysis indicates which parameter is more valuable to differentiate among the different class. Figure 3.10 shows the result of discriminant analysis. The more useful parameters for this purpose are pH, SC, and Oil & Grease. However, only 19 percent of the data were correctly classified. The distribution of the eight sectors using a linear combination of the three parameters totally overlaps.

Figure 3.11 shows an example of successful discriminant analysis. The example uses well-known Iris flower data. The patterns allow one to recognize three types of flowers from four parameters.

To further seek a relationship between water quality data and various activities of industries, a supervised neural network, multi-layer perceptron model was used. Neural Connection 2.1 (SPSS Inc. and Recognition System Inc, Chicago, IL) was used to build the neural model. The neural model was extensively trained with various architectures. The performance was only a little better than using the discriminant analysis (see Table 3.6). The performance was still very poor and only 15 to 31 percent of the cases were correctly classified for each sampling season.

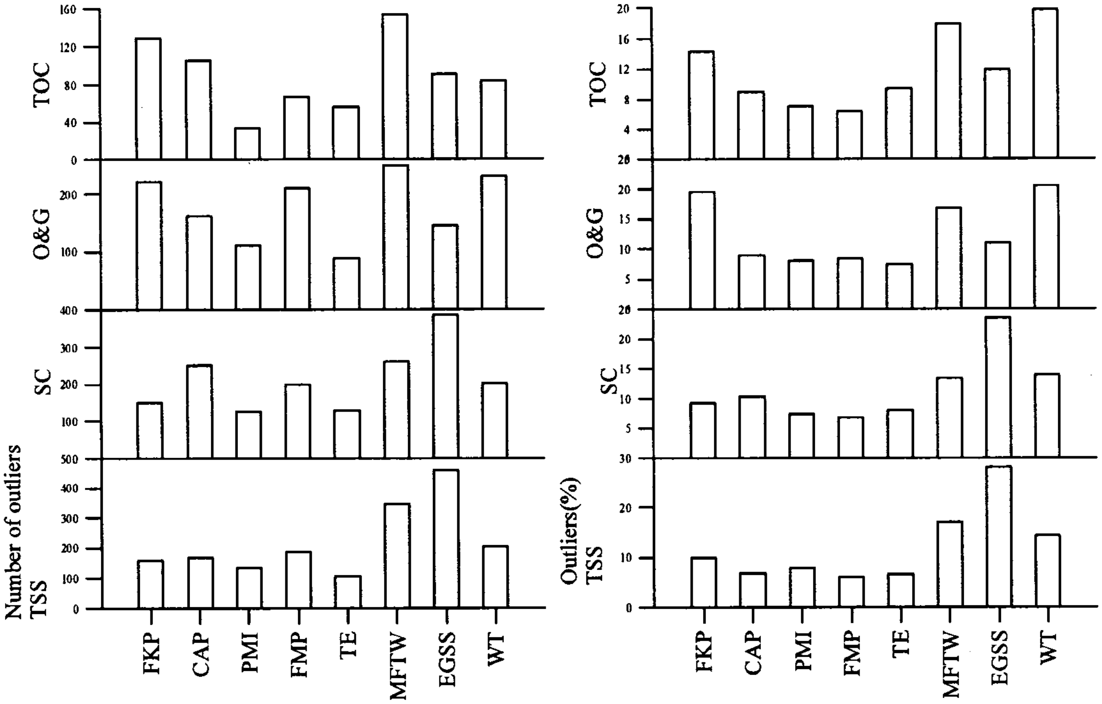


Figure 3.8. Distribution of number and percentage of outliers according to its sector for TSS, SC (Specific Conductance), O&G (Oil and Grease) and TOC

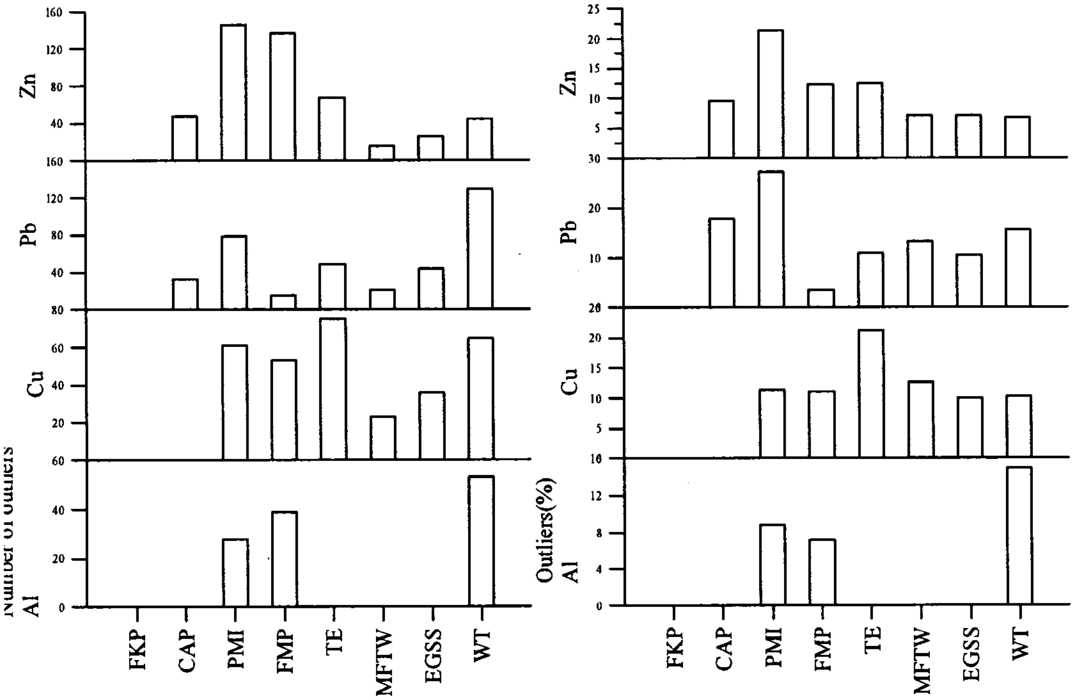


Figure 3.9. Distribution of number and percentage of outliers according to its sector for Al, Cu, Pb, and Zn

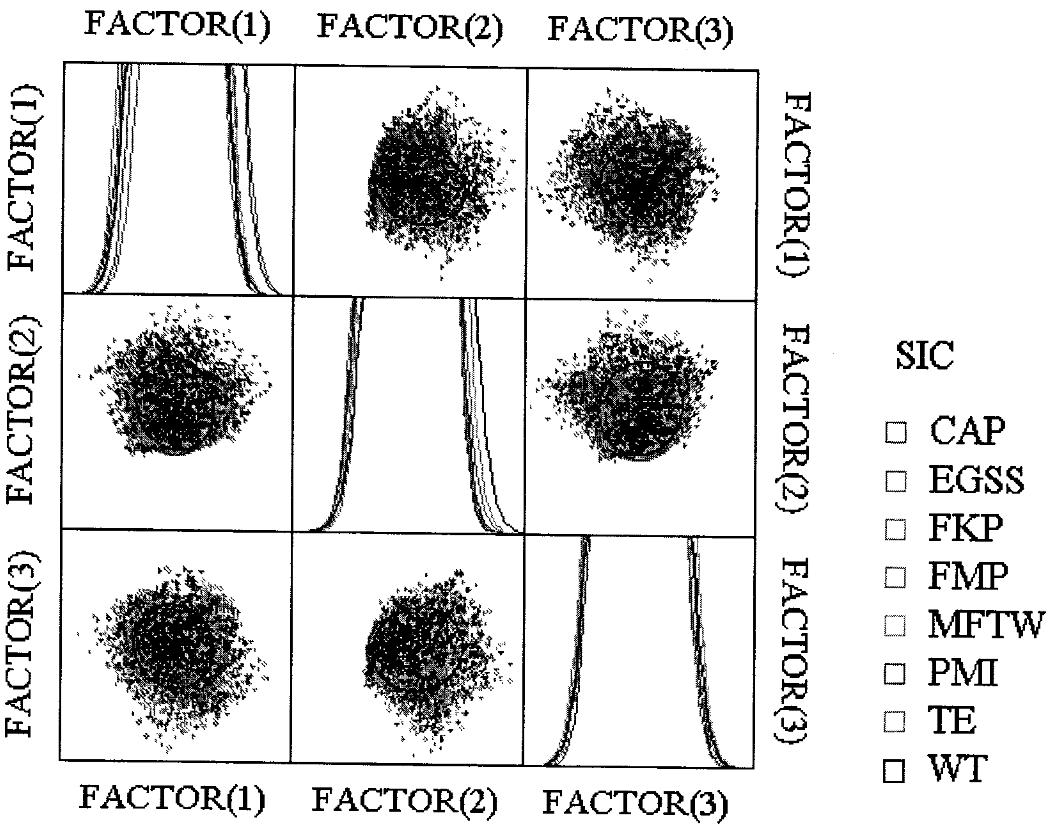


Figure 3.10 The result of the discriminant analysis

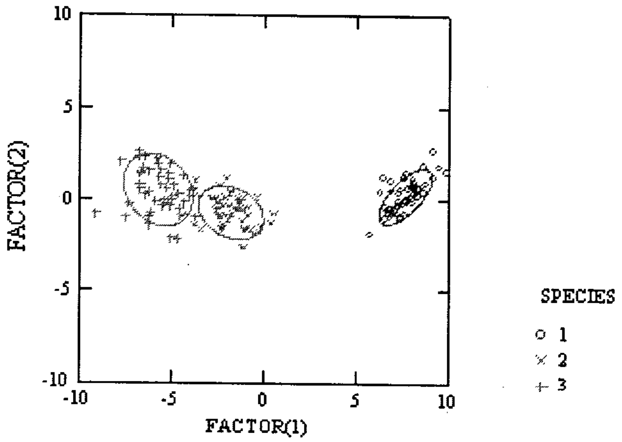


Figure 3.11 An Example of discriminant analysis using the Iris data

Table 3.6. Results of a neural model, multi-layer preceptron.

Season	Correct classification (%)	Case number of training file	Parameters = input variables	Categories based on SIC = output variables
92-93	30.64	235	pH, TSS, SC	20,28,33,34,37,42,49
93-94	24.57	814	pH, TSS, SC	20,28,33,34,37,42,49
94-95	18.76	1125	pH, TSS, SC	20,28,33,34,36,37,42,49
	28.45	840	pH, TSS, SC, Oil&Grease	20,28,33,34,36,37,42,49
95-96	17.08	890	pH, TSS, SC, Oil&Grease	20,28,33,34,37,42,49
96-97	28.40	960	pH, TSS, SC	20,28,33,34,37,42,49
	31.31	674	pH, TSS, SC, Oil&Grease	20,28,33,34,37,42,49
97-98	14.65	314	pH, TSS, SC, Oil&Grease	20,28,33,34,37,42,50

SIC; Standard Industrial Classification

The data set was also examined to determine if a seasonal first flush could be identified. California has a Mediterranean climate, typified by winter and spring precipitation and summer drought. Most of western parts of California including Los Angeles are dry from May through August. Figure 3.12 shows the monthly average rainfall for six locations in California during 1971- 2000. Records of the monthly average rainfall were obtained online (**http://www.ncdc.noaa.gov/oa/climate/online/ccd/nrmlprcp.html**). This rainfall pattern creates a long period for pollutant build-up and the first storm of the season usually has higher pollutant concentrations, which is called a seasonal first flush.

To determine a seasonal first flush phenomenon for GISP storm water discharges, the recent three years of storm water monitoring data were examined. TSS, specific conductance, TOC, and Zn for the three years are shown in Figures 3.13 through 3.15. Since flow monitoring is not required, no flow data are shown. The daily precipitation record from Los Angles Civic Center station was used for the rainfall and antecedent dry days instead of real site rainfall data, which is not monitored. Records of the daily precipitation data were obtained online (**http://www.nwsla.noaa.gov/climate/climate.html**). Although, rainfall will vary by site, the record from Civic Center was used as representative of the general pattern of rainfall for Los Angeles area.

The GISP requires the first storm and one later storm to be sample. In some cases the first sample does not represent the first rainfall event. Several events for each sampling year were sampled by the GISP permittees, which means some storm events have many cases than other storm events. The permittees had no direct guidance for selecting storm events, other than an early event in the rainfall year.

The guidance to select the first storm of the year virtually insures that high data will be observed. Median concentrations of all parameters were highest in the initial event of storm water and decreased except during the 1998-1999 season. The four parameters show similar trends for all events. The rainfall pattern in California is such that most portions of the state are likely to experience a seasonal first flush. Therefore, data collected using sampling strategies based upon first storm data can be biased high.

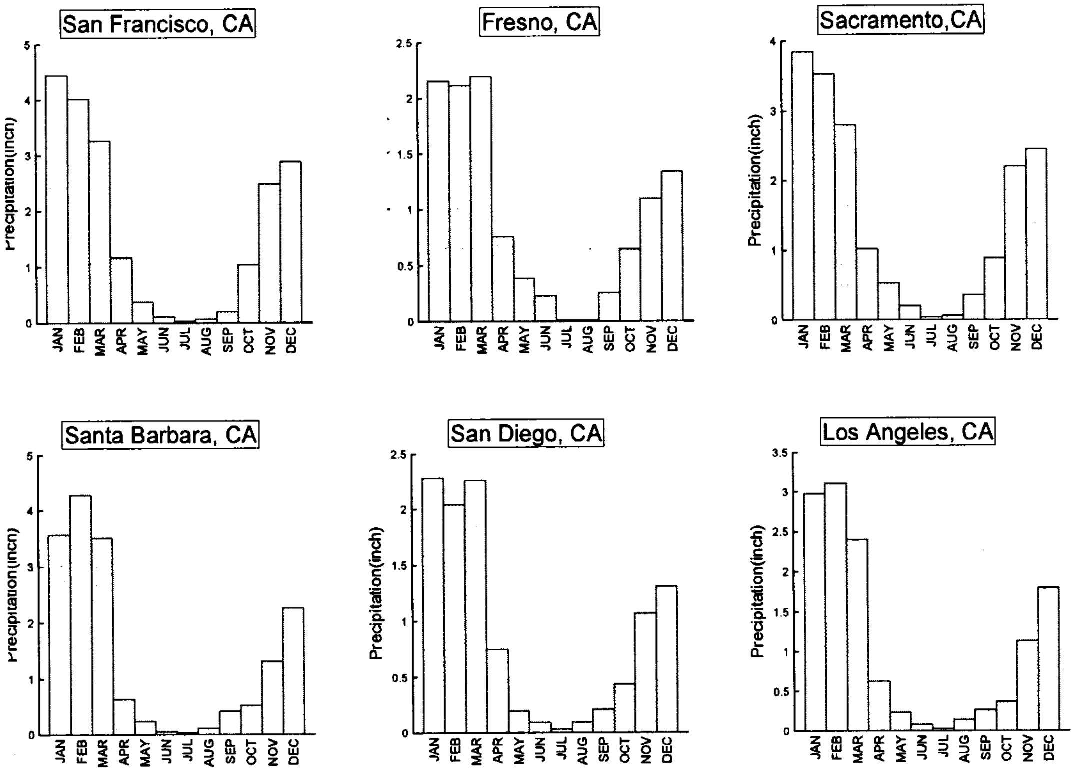


Figure 3.12. Monthly average rainfall in California, 1970-2000

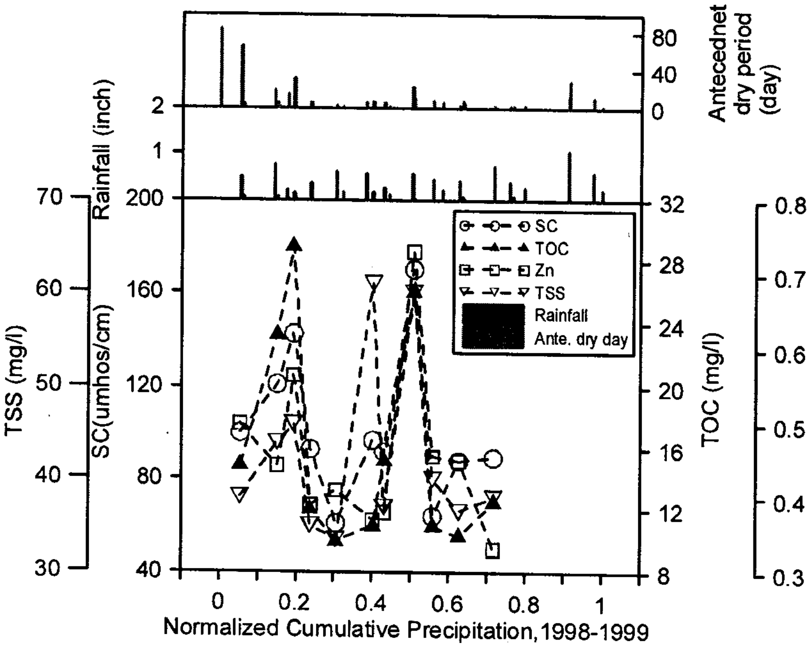


Figure 3.13. Distribution of TSS, SC (Specific Conductance), TOC and Zn for 1998-1999

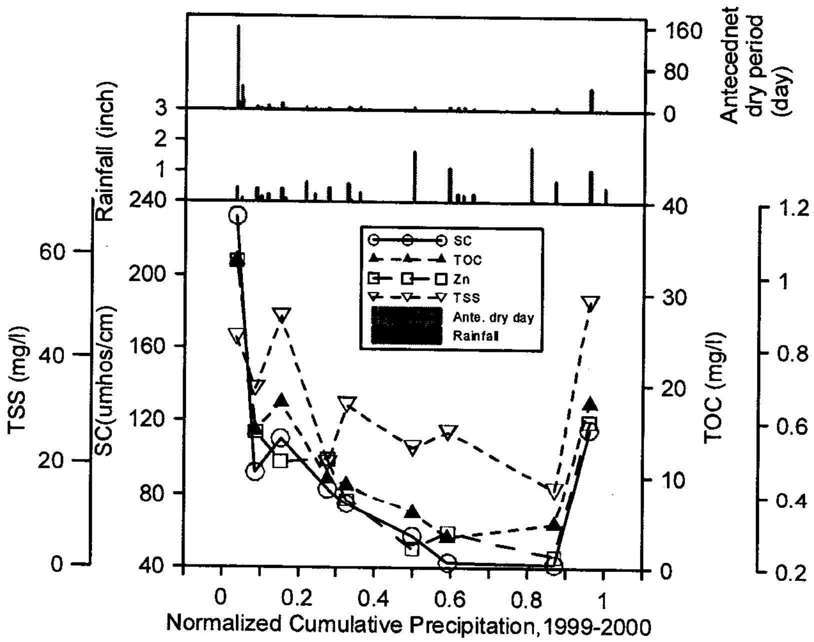


Figure 3.14. Distribution of TSS, SC (Specific Conductance), TOC and Zn for 1999-2000

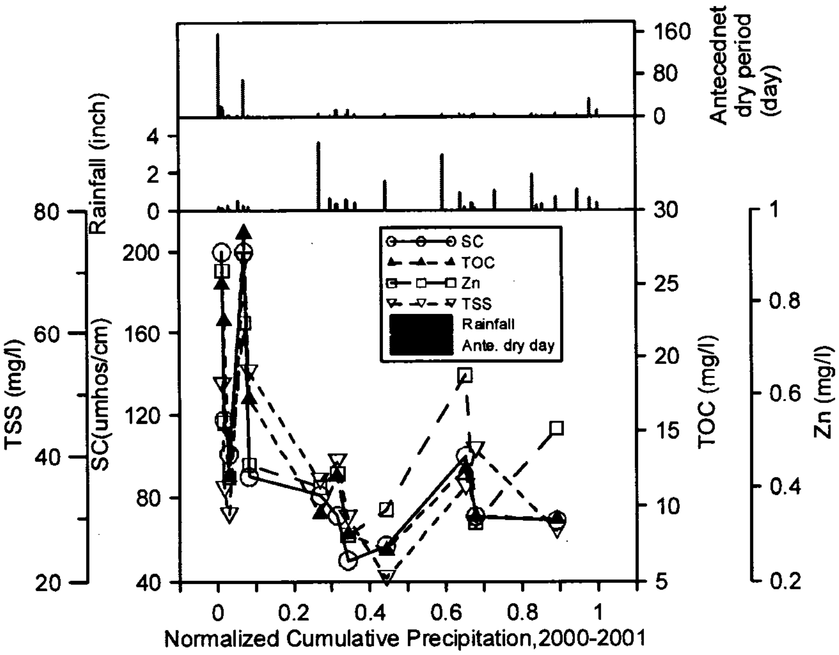


Figure 3.15. Distribution of TSS, SC (Specific Conductance), TOC and Zn for 2000-2001

3.3 Summary of Existing Data

The existing data show very limited utility. The variability among reported results is so great that differences among different industries and landuses are not possible to identify. Although quantitative evidence for differences in storm water runoff from different SIC codes has not yet been documented in the published literature, it is reasonable to assume that in most cases there should be differences. Differences in storm water quality from different landuses have been well documented for more than 20 years (Stenstrom, et al., 1984, Fam et al., 1987) and have been become an important tool for prioritizing BMPs. Therefore the failure to identify quantitative differences in pollution based on SIC codes is considered a monitoring program failure, rather than the inference that there are no differences in storm water pollutant emissions. The power of the database provided by the monitoring program is simply too weak to show the differences that are of concern for regulators and storm water program managers.

To illustrate this effect Table 3.7 shows the basic statistics from three GISP programs (Los Angeles County, Sacramento County and Connecticut State). The data are all lumped together for all SIC codes. The table illustrates differences among programs (e.g., facilities in the County of Los Angeles have a greater number of observations, by far), but the purpose of the table is to show the coefficients of variation (CV). The coefficient of variation is the standard deviation of the data divided by the mean of the data. Since both statistics are in the same units, the ratio is dimensionless, which allows the variability of data with different magnitudes to be directly compared. The CVs in Table 3.7 range from 0.2 to 15. This shows that in most cases, the variability in the data is greater than the mean value of the data.

This high variability may be surprising to professionals who have been monitoring water and wastewater treatment plant influents and effluents. Figure 3.16 shows CVs for a typical, large west coast wastewater treatment plant influent, a large typical west coast water treatment plant, and the two storm water programs being evaluated. The greater variability in storm water quality is dramatic.

Table 3.7. Basic statistics for three General Industrial Storm Water permit programs.

Parameters [a]	Los Angeles County			Sacramento County			Connecticut State		
	Sample No.	Mean.	CV	Sample No.	Mean.	CV	Sample No.	Mean.	CV
pH (pH unit)	24851	7.01	0.95	857	7.16	0.17	9617	6.32	0.20
TSS	24144	376.15	11.85	769	185.49	2.86	9617	124.02	6.59
SC (μmhos/cm)	23585	561.68	8.13	846	204.20	2.27			
Oil & grease	18637	16.57	14.25	286	11.26	1.61	9561	5.66	14.57
TOC	9714	50.13	5.23	399	31.44	2.12			
COD	1834	271.29	2.77	50	154.40	1.58	9606	80.67	3.44
Aluminum	1618	10.12	12.21	46	3.44	1.66			
Copper	3354	1.01	16.50	83	0.18	2.31	9596	0.13	7.56
Iron	1844	25.48	6.39	82	7.49	2.03			
Lead	3525	2.96	14.12	78	4.48	3.82	9563	0.06	8.37
Zinc	5163	4.96	13.85	141	2.23	7.59	9614	0.51	7.79
Phosphorous							9606	0.45	4.30
TKN							9608	2.50	3.11
NO_3-N							9613	1.19	2.73
24hr LC_{50} (%)							9628	82	0.38
48hr LC_{50} (%)							9628	75	0.45

[a] Unit is mg/l unless otherwise noted

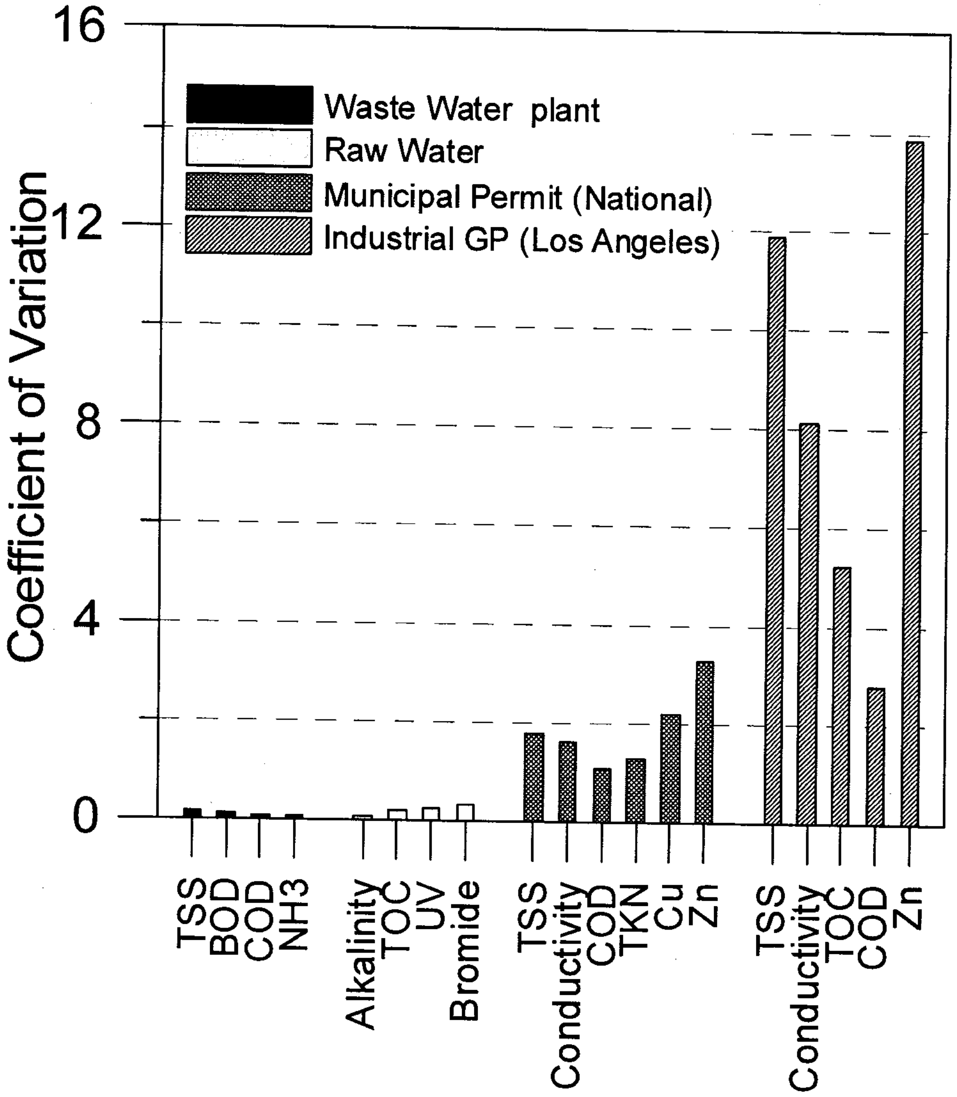


Figure 3.16 Typical variability of water and wastewater treatment plant influents as compared to variability in storm water quality.

To illustrate the impact of the greater variability on hypothesis testing, Figure 3.17 is provided. Each line on Figure 3.17 shows the required number of cases or observations to conclude with 95% confidence that there is a significant difference in two means. The assumptions are for two-tailed normal distributions. Each line represents a different CV. The horizontal axis shows the desired difference in the means. For example, if one wanted to detect a 50% difference in the mean pollutant discharge from two sources (e.g., categories or facilities, etc.), only 12 observations are required if the CV is 0.4, which according to Figure 3.16 is typical for water and wastewaters. For storm waters however, the CVs are often 6 or more. For this high a CV, 2,270 observations are required. This graph shows very dramatically that it is impractical to interpret and perform hypothesis testing with such data. Table 3.5 showed that few categories had this many observations, even if all nine years of the monitoring program are used. If the storm water monitoring program is to be successful in detecting differences among permittees and categories, a different type of monitoring with lower variability must be developed.

The use of composite samplers and professionally trained monitoring personnel is one candidate solution to the monitoring problem. Many monitoring programs use composite or flow-weighted composite samplers. Figure 3.18 shows data collected in such a program. Four metals copper, zinc, iron and nickel are show. Figure 3.19 shows data from the industrial monitoring program using grab samples and untrained sampling personnel. The difference in variabilities are dramatic and log scales are required to show the ranges.

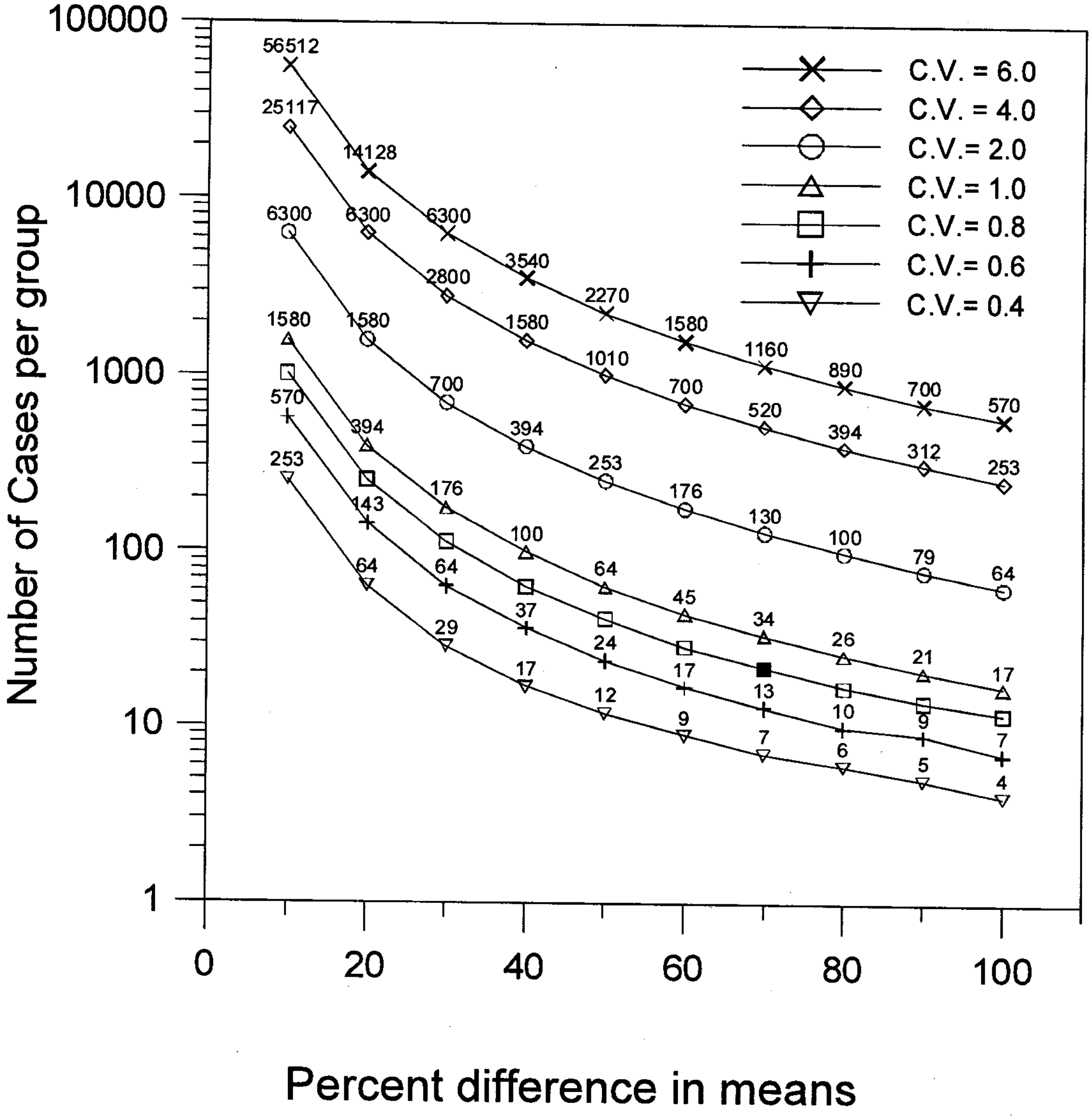


Figure 3.17. Number of observations (cases required to detect differences in means with different variability.

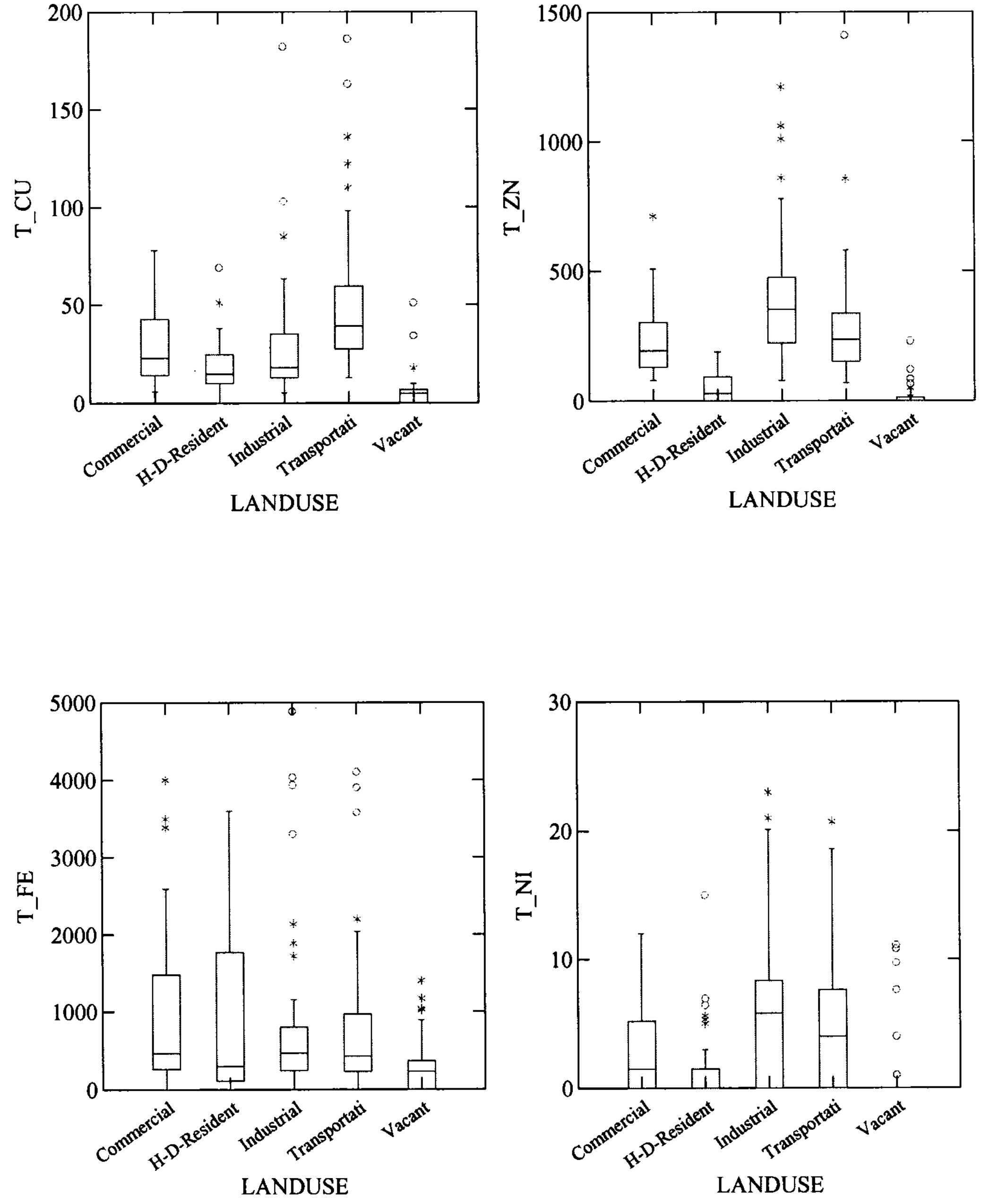
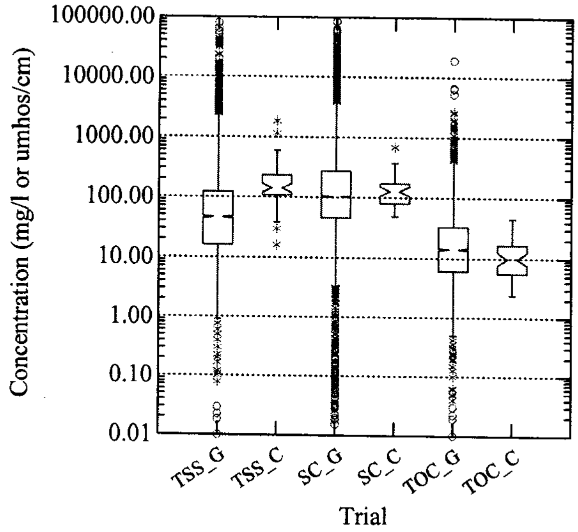
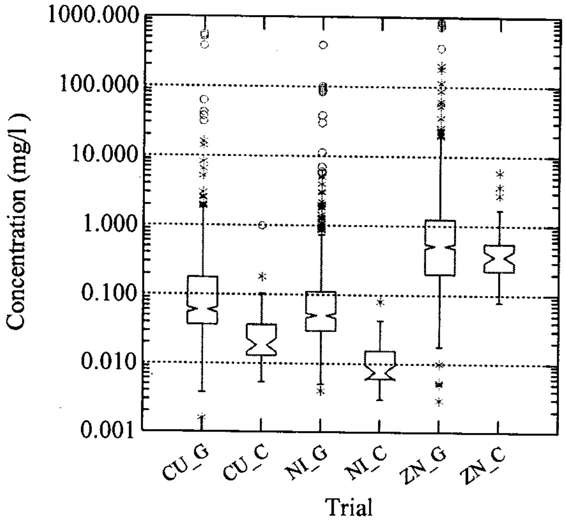


Figure 3.18. Distribution of Cu, Zn, Fe, and Ni in various landuse. Data are from the landuse monitoring by LACDPW. Unit of the parameters are ug/l.



Comparison of grab sample and composite sample



Comparison of grab sample and composite sample

Figure3.19 Comparison of grab samples from the General Industrial Storm Water Permit data for 1992-1998 to flow-weighted composite samples from the industrial landuse monitoring data by LACDPW for 1996-2001. G indicates a grab sample and C indicates composite sample.

4. REVIEW COMMITTEE AND RECOMMENDATIONS

Midway through the project, a review committee was assembled. The objective of the review was to was to provide expert guidance as well as providing stakeholder insight. The following individuals attended a meeting on the UCLA campus on November 17, 2003. Their agency or employer is shown.

Name	**Agency or Employer**
Michael Drennan	Brown and Caldwell (consulting firm)
Mark Gold	Heal-the Bay (NGO)
Gerry Green	City of Downey
Kosta Kaporis	City of Los Angeles, Bureau of Sanitation
Ken Schiff	Southern California Coastal Water Research Project (Joint Powers Agency monitoring coastal waters)
Eric Strecker	Geosyntec, Inc (consulting firm)
Xavier Swamikannu	California Regional Water Quality Control Board, Los Angeles Region
Tracy Wilcox	MWH, Inc (consulting Firm) Surfrider Foundation (NGO)

Most of the material contained in section 3 was presented at the meeting and the following suggestions were made as desirable or necessary inclusions in a new permit.

1. The new permit should clearly state its objectives. Three objectives are recommended: being able to identify high emitters (polluters), encouraging pollution prevention and creating data for use by planners and particularly by developers of TMDLs.
2. The monitoring program should have more specifications. Among those recommended are requirements for the use of certified laboratories, except for field measurements, trained sampling personnel, quality assurance plans, improved sampling protocols, and specification of analytical techniques, including minimum detection limits.
3. Expand the list of monitored pollutants to included industry-specific pollutants (e.g., metals analysis for those industries likely to be discharging metals). Include

parameters in the Clean Water Act Section 303(d) listing for industries in impacted watersheds. Allow requirements for monitoring specific pollutants to expire if they are routinely not detected.

4. The committee was reluctant to require flow weighted composite sampling of all permittees. Only the larger permittees should be required to perform such sampling. For smaller permittees, a method of estimating flow should be developed, and included as part of the permit. An example is estimating the amount of impervious area and other factors that affect runoff. In this way, a person reviewing the data collected by the permittees will be able to estimate mass emissions from flow rate and concentrations.
5. Make monitoring results available in real-time to regulators. A web-based reporting method was proposed. Such a procedure should be able to eliminate transcription errors by flagging implausible values (e.g., the specific conductivity data shown earlier ranges from the best distilled water to sea water, neither of which is likely for storm water runoff). The monitoring program should also be able to include the data into a database, accessible by authorized personnel, so that routine reports can be generated. Also the program should be able to identify problematic data, for early investigation.
6. Answer the question, "How much more data is required to make decisions." This was done and was included as Figures 3.17 in the previous chapter. The group was interested to know if simply collecting more samples would solve or partially solve the problem, or if better protocols are also required.
7. The new permit will likely increase cost and training requirements. While the new monitoring program requirements are being implemented, the permittees should be exempt from monitoring (anticipate to be 12 to 24 months, maximum). The State Board should develop a new guidance and training document, which should include training material on the new requirements. Method to select a good sampling site and estimating flow rate are two new topics that will benefit from guidance and training.

The group felt that of all the recommendations made, the two most important were that that the new permit should be able to identify high emitters (polluters) and be web based to speed up reporting and eliminate implausible values.

5. PROPOSED REQUIREMENTS OF THE NEW MONITORING PROGRAM.

New provisions are proposed for the new monitoring program. These are based in large part on the recommendations of the review committee and also in part upon the experience of the research team and the findings of the project. They are divided into three general sections: sampling methodology, parameter selection, and reporting method. There are also recommendations for future research and development that will improve monitoring techniques.

5.1 Sampling Methodology

The sampling methodology is probably the most lacking part of the existing permit. The new permit should require a minimum level of training for sampling personnel for all but the smallest of facilities. Training is required to insure that representative sites are chosen. The skills to do this are beyond most industrial employees, who have no reason to be trained in this area before. Alternatively, a certified laboratory or professional engineering can be employed.

Certified laboratories or consultants may offer the best opportunity to implement the new training program. Certified laboratories and consultants are already familiar with many sampling issues, since they may be sampling or monitoring results in other environmental sampling programs. Typical issues such as sampling preservation and holding time should be well known to them.

A model is envisioned for a certified laboratory to perform the sampling which will involve the permittee as well as laboratory personnel. Laboratory personnel can complete a training program provided by the State Board, where they receive instruction on the monitoring program requirements as well as general knowledge of representative sampling techniques. The permittee contracts with a laboratory for sample analysis and in so doing is contracting with the laboratory to set up a representative sampling program. In this way the certified laboratory provides a "package" – sampling plan, analytical services and reporting.

A model exists for permittees to use consultants for monitoring. The California Department of Transportation (Caltrans) was required to monitor freeway runoff in many areas of the state. They successfully contracted with consultants to perform the monitoring, quality assurance and reporting. The State Board may wish to adopt a similar approach, utilizing monitoring funds paid by industrial permittees to perform targeted sampling. Through this process, a subset of representative sites could be selected which will result in financial savings to the permittees.

The frequency of sampling - two storms per season maybe adequate, but more guidance is needed. For example, collecting the first runoff of an early storm creates a bias. The sampling plan should provide some why of compositing samples. Even if only several grab samples are composited, this will be an improvement.

5.2 Parameter Selection

The rationale for choosing the existing parameters is not known to the research team, but it appears that some were selected for convenience, low cost and ability to measure in the field. Their environmental significance is sometimes limited. For example, the specific conductivity is of little importance in predicting the environmental impact of storm water on receiving waters. Only the most pristine of receiving waters will have conductivities lower than most storm water. The value of the conductivity measurements is counter-intuitive. It has shown that an easily measured parameter, using reliable but inexpensive instruments, can not be measured reliably and accurately. The following list of parameters is proposed. Some of the measurements are included because of environmental significance but others are included because they are either simple, inexpensive and/or are useful in determining the reliability of the monitoring program. The specific conductivity analysis is retained because it can serve as a validity check and because is it very inexpensive to perform. Oil and grease has been omitted for two reasons: it is not a trivial test and requires care in sampling, since oil adsorption to glass ware and sampling equipment can be significant; the environmental significance of oil and grease and efficiency of the measuring technique depends upon the source (Stenstrom et al., 1986, Fam, et al., 1987). Also, for oil and grease from highway runoff, COD or TOC are adequate surrogates (Khan, et al., 2005).

The following minimum list of parameters is proposed:

Field measurements
    Specific Conductivity
    pH
    Turbidity
Laboratory measurements
    Total suspended solids
    Chemical oxygen demand or total organic carbon
    Total metals – cadmium, chrome, copper, lead, nickel, zinc

Optional measurements

The Regional Board or another regulator needs to be able to add specific parameters when they may be needed. This would include parameters on the Clean Water Act Section 303(d) impacted waters list. Nutrients, such as nitrogen and phosphorous are the likely candidates. Additionally, pesticides from industries that are likely to be using large amounts of pesticides in their business, or industries whose runoff discharges into sensitive surface waters, are also good candidates. These measurements will be made in the field or in a laboratory as required.

Toxicity is also a good candidate for additional measurements. While toxicity is quite expensive, it has the ability to detect small differences, and as shown earlier, was one of the more useful parameters.

5.3 Reporting Method

One of the most significant improvements that can be made to the existing monitoring program is to require web based reporting. This will require the State Board to develop a program and maintain servers, but this effort should be less costly than managing a paper based reporting system.

The web-based reporting has several important advantages. First, it provides data to the regulatory agencies in real-time. The server can be programmed to provide periodic summary reports. Secondly, permittees will most likely find it easier to use web based reporting. If certified laboratories are used, they can do the web based reporting.

Web based reporting has another advantage that may not be obvious. The parameters being reported all have plausible values. The range of specific conductivity for storm water is reasonably known, and ranges outside this value are probably errors. Many are obviously reporting or transcription errors – the wrong units are used or a person's handwriting is hard to read. A web based reporting system can incorporate a simple expert system that can query the user.

For example, if an obviously implausible or unlikely value of pH is reported (e.g., 11 or higher), the person entering the data can be asked to confirm the entry. If it is still implausible, the person can be asked a second question, such as confirming the units, or the person can suspend the reporting and come back after checking the data. Finally, if no error is found, the program accepts the implausible data and marks it for attention by the appropriate person at the regulatory agency. The web based system can also be programmed to create summary or annual reports.

5.4 Co-Sampling Programs

There are many Municipal Separate Storm Sewer System (MS4) permittees in the State and they are required to develop and implement municipal storm water monitoring programs that include receiving water monitoring requirements. The MS4 permittees are required to operate or contract with state certified laboratories. The MS4 permittees are trained and equipped to perform the industrial monitoring required by the existing permit and the new proposed permit. The MS4 permittees are also charged with meeting TMDLs.

The TMDL requirement provides an incentive for MS4 permittees to locate discharges that are causing water quality violations. Therefore they have an incentive to locate and correct high emitters.

An alternative that should be available is for industrial permittees to fund the MS4 permittees to do the required monitoring and reporting. The economies of scale will insure that the cost of funding MS4 programs will be less than the industrial permittees would have to pay to do the work themselves. The MS4 permittees would have the responsibility for selecting the number of industrial permittees to be sampled and through this process, attempt to locate and correct high emitters. This could be a desirable option for small facilities, and might help MS4 permittees more easily meet TMDLs as well as providing funding to support technicians or laboratories.

6. CONCLUSIONS

Data collected from the California statewide General Industrial Storm Water Permit (GISP) was examined over the nine-year period from 1992 to 2001. The data were evaluated to determine if it could be used to identify permittees with high emissions and if the storm water loads from various classes of industries could be characterized in order to create rankings and typical emission rates. The data were also compared to data collected in other monitoring programs. The following conclusions are made:

1. The data collected by the permittees are highly variable, with coefficients of variation as high as 15. This compares to coefficients of variation, generally less than 0.5 for other environmental monitoring programs, such as water and wastewater treatment plant influents.
2. There are several sources for the variability and the use of grab samples, untrained sampling personnel, and a limited selection of monitored parameters are among the largest sources.
3. The data generally do not allow for hypothesis testing and generally could not be used to identify high dischargers, using statistical tests with confidences of 0.05 or greater. The data also could not be used to identify differences in discharges from different types of industries.
4. The variability in the collected data is so great that the collection of additional data points, up from two to ten or more storms per year, will still not provide the needed precision. Improving the precision of sampling, by using composite samples for example, is a more promising approach.
5. The data collected in Los Angeles has greater means, medians and variability than data collected by two similar programs (Sacramento, CA and the state of Connecticut). The frequency of exceeding US EPA target levels is also greater for the Los Angeles data.
6. A review committee composed of experts from consulting, government and NGOs suggested improvements for a new permit. Among the most important of those were requirements for trained sampling personnel, certified laboratories, and web-based reporting. The most important goal for the new permit is to be able to identify high dischargers.

The requirements of a new monitoring program are proposed. They include a broadened suite of parameters, use of composite samples and certified laboratories, joint sampling programs and a web-based reporting system. It is also proposed that the current monitoring program be suspended while the State Board develops the new program.

7. REFERENCES

Bay, S., Jones, B.H., Schiff, K. and Washburn, L. (2003) Water quality impacts of stormwater discharges to Santa Monica Bay. *Marine Environmental Research*, **56**, 205-223.

Davis, A.P., Shokouhian M. and Ni, S. (2001) Loading estimates of lead, copper, cadmium, and zinc in urban runoff from specific sources. *Chemosphere*, **44**, 997-1009.

Duke, L.D., Buffleben, M. and Bauersachs, L.A. (1998) Pollutants in storm water runoff from metal plating facilities, Los Angeles, California. *Waste Management*, **18**, 25-38.

Fam, S,. Stenstrom, M.K. and G. Silverman, G "Hydrocarbons in Urban Runoff," *Journal of the Environmental Engineering Division*, ASCE, Vol. 113, pp 1032-1046, 1987.

Khan, S., Lau, S-L, Kayhanian, M. and Stenstrom, M.K. (2004) Oil and Grease Measurement in Highway Runoff – Sampling Time and Event Mean Concentrations. submitted to the *Journal of Environmental Engineering,* ASCE.

Lee, H. and Stenstrom, M.K. (2005) Utility of stormwater monitoring. *Water Environmental Research,* Vol. 77, No. 1.

Lee, H., Lau, S-L., Kayhanian, M. and Stenstrom, M.K. (2004) Seasonal first flush phenomenon of urban stormwater discharges. *Water Research*, **38**, 4153-4163.

Lee, J.H., Bang, K.W., Ketchum, L.H., Choe, J.S., and Yu, M.J. (2002) First flush analysis of urban storm runoff. *The Science of the Total Environment*, **293**,163-175.

Leecaster, M.K., Schiff, K. and Tiefenthaler, L.L. (2002) Assessment of efficient sampling designs for urban stormwater monitoring. *Water Research*, **36**, 1556-1564.

Ma, J-S., Khan, S., Li, Y-X., Kim, L-H., Ha, H., Lau, S-L., Kayhanian, M., and Stenstrom, M.K. (2002) Implication of oil and Grease measurement in stormwater management systems. Presented at the 9th ICUD, Portland, Oregon, September 2002.

Pitt, R., Maestre, A. and Morquecho, R. (2003) Compilation and review of nationwide MS4 stormwater quality data. *Proceedings of 76th Water Environment Federation Technical Exposition and Conference*, Los Angeles, CA, 2003.

Pitt, R (2004) http://unix.eng.ua.edu/~rpitt/Research/ms4/Final%20Table%20NSQD%20v1_1%20020704.xls .

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

Stenstrom, M. K., Silverman, G.S. and Bursztynsky, T.A., Oil and Grease in Urban Stormwaters, *Journal of the Environmental Engineering Division*, ASCE, Vol. 110, No. 1, pp 58-72, February 1984.

Stenstrom, M.K., Fam S., and Silverman, G.S.Analytical Methods for Quantitative and Qualitative Determination of Hydrocarbons and Oil and Grease in Wastewater, *Environmental Technology Letters*, Vol. 7, pp 625-636, 1986.

Stenstrom, M.K. and Strecker, E. (1993) Assessment of Storm Drain Sources of Contaminants to Santa Monica Bay. Vol. I, Annual Pollutants Loadings to Santa Monica Bay from Stormwater Runoff, UCLA-ENG-93-62, May 1993, Vol. I, pp 1-248.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2000) First Flush Stormwater runoff from Highways. *Research report for California Department of Transportation*, September, 2000.

Stenstrom, M.K., Lau, S-L., Ma, J-S., Ha, H., Lim, L-H., Lee, S., Khan, S. and Kayhanian, M. (2001) First Flush Stormwater runoff from Highways. *Research report for Department of California Transportation*, September, 2001.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2002) First Flush Stormwater runoff from Highways (year3). *Research report for Department of California Transportation*, July, 2002.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2003) First Flush Stormwater runoff from Highways (year4). *Research report for Department of California Transportation*, July, 2003.

APPENDIX

Design of storm water monitoring programs for Industrial sites

Haejin Lee and Michael K. Stenstrom

A manuscript submitted to Water Research

Design of storm water monitoring programs for Industrial sites

Haejin Lee[a], Seung-jai Kim[b] and Michael K. Stenstrom[c*]

[a] *United States Department of Agriculture, Natural Resources Conservation Service,*
El Centro Office, 177 N. Imperial Ave., El Centro, California 92243, USA
[b]*Department of Environmental Engineering, Chonnam National University,*
300 Yongbong-dong, Buk-gu, Gwangju 500-757, KOREA
[c] *Department of Civil and Environmental Engineering, 5714 Boelter Hall, University of*
California, Los Angeles, Los Angeles, California 90095, USA

Abstract

Storm water runoff is now the leading source of water pollution in the United States, and storm water monitoring programs have only recently been developed. This paper evaluates more than 20 storm water monitoring programs and datasets to determine their usefulness in characterizing discharges and achieving their ultimate goal of reducing storm water pollution. The monitoring results are highly variable, with coefficients of variation that are 2 to 60 times higher than those observed in water or wastewater monitoring programs. Industrial landuse is an important source of metals, although the monitoring programs could not differentiate metals contribution from different types of industries. Data from California, which has distinct dry periods showed a seasonal first flush; data from Connecticut did not show a seasonal first flush. Recommendations for improving monitoring programs include using composite samplers, selecting alternate or additional water quality parameters, alternate timing of sample collection, and strategies that sample a subset of the total permittees using more sophisticated methods.

Keywords: First flush; Industrial General Permit; Monitoring; Municipal Permit; Storm water

1. Introduction

The completion of wastewater treatment plants mandated by the Clean Water Act has reduced pollution from point sources to the waters of the United States. As a result, non-point sources pollution such as storm water runoff are now the major contributor to pollution of receiving waters. The problem of storm water pollution is growing worse because of continuing development, which results in increased impervious surface area. In order to reduce storm water pollution, regulatory agencies are requiring storm water monitoring programs. The programs are implemented through the National Pollutant Discharge Elimination System (NPDES) Storm water Permit.

The overall goal of the storm water monitoring includes the identification of high risk dischargers, but is also for the development of a better understanding of the mechanisms and sources of storm water pollution, with the long-term goal of reducing pollutants to less harmful levels. Most storm water monitoring programs are relatively new and evaluations of their usefulness for satisfying these goals are only now possible (Duke et al., 1998; Lee and Stenstrom, 2003; Pitt et al., 2003). Lee and Stenstrom (2003) recently evaluated the General Industrial Storm Water and Municipal Storm Water Permits for Los Angeles County for three wet seasons (1998 - 2001) and five wet seasons (1996 - 2001), respectively. The results of our previous study suggest that parts of the current industrial storm water monitoring programs will not be helpful to identify high dischargers, nor will they be useful in developing Total Maximum Daily Loads (TMDLs). The design and requirements of the monitoring program do not produce data with sufficient precision for decision-making.

In this paper, we evaluated a number of storm water monitoring programs to determine their usefulness in achieving their dual goals, as well as making recommendations for improvement. Our analysis included more than 20 monitoring programs and datasets, which covered three General Industrial Storm water Permit programs, US EPA's Multi-Sector Storm Water General permit (MSGP) program, Municipal Storm water Permit programs from 17 states, and Caltrans' first flush highway runoff characterization study. The programs are summarized in Table 1. We believe our results will be helpful to planners and regulators to interpret existing datasets and programs as well as providing recommendations for improving the future programs.

2. Background

2.1. General Industrial Storm Water Permit Programs

The General Storm water Industrial Permit (GISP) Programs require facilities that discharge storm water associated with the industrial activities directly or indirectly to apply and obtain coverage under the GSIP. Industries are categorized by Standard Industrial Classification (SIC) codes. Industrial monitoring programs generally vary by state. Currently, there are approximately 3,000 permittees within Los Angeles County. Under the monitoring requirements, permittees must collect water quality samples from two storms per year and analyze for four conventional parameters: pH, specific conductance (SC), total suspended soils (TSS), and oil and grease (O&G). Total organic carbon (TOC) can be substituted for oil and grease, and certain facilities must analyze for specific additional pollutants such as metals. Permittees are requested to collect storm water samples during the first hour of discharge from the first storm event of the wet season, and at least one other storm event later in the wet season. In our previous study (Lee and Stenstrom, 2003), we analyzed data from three recent wet seasons (1998-2001) of data. In this study we extended the evaluations to the available data, which covered a total of nine wet seasons (1992-2001), and we also analyzed eight wet seasons (1993-2001) of data from a similar GSIP in Sacramento County.

Nine years (1995-2003) of storm water data were available from the State of Connecticut. Their program requirements are different from California's requirements. Permittees in Connecticut must analyze for three metals (total copper, lead, and zinc), nutrients (total phosphorous, total Kjeldahl nitrogen and nitrate as nitrogen), aquatic toxicity (LC_{50}) and conventional parameters (pH, COD, O&G and TSS). They must collect a grab sample within the first 30 minutes of runoff from at least one storm per year.

The U.S. Environmental Protection Agency (USEPA) issued the Multi-Sector Storm Water General permit (MSGP) for storm water discharges associated with most industrial activities in September 1995 (revised in 2000). The permit covers industrial activities in states and territories that have not been authorized to run the NPDES general permitting program. The types of monitoring and required parameters vary among industry sectors and sub-sectors. Grab samples may be used except at airports which must collect a flow-weighted composite, in addition to a grab sample. Grab samples are to be collected within the first 30 minutes of discharge. Permittees are required to sample every other year. Six years (1998-2003) of data were available.

2.2. Municipal Storm water Permit Programs

The Storm Water Program for Municipal Separate Storm Sewer Systems (MS4) is designed to monitor and reduce sediment and pollution that enters surface and ground waters from storm sewer systems to the maximum extent practicable. Storm water discharges associated with MS4s are regulated using NPDES permits. Pitt et al (2003) under US EPA sponsorship compiled and evaluated storm water data from a representative number of NPDES MS4 storm water permitees. Over ten years of data from more than 200 municipalities throughout the United States were assembled. The areas were primarily located in the southern, Atlantic, central and western parts of the United States. Only data from well-described storm water outfall locations were used in the database. Most pollutants were characterized using flow-weighted composite samples except for pollutants having restrictive holding time requirements, such as bacterial indicators, which were sampled using grab samples. Numerous pollutants were analyzed including typical conventional pollutants, heavy metals, and organic toxicants. Los Angeles County MS4 data were not included in the database.

The Los Angeles County Department of Public Works (LACDPW) has had its own municipal monitoring program since the early 1970s. In 1994 they began an improved program, which was designed to determine total pollutant emissions to coastal waters as well as landuse specific discharges (Stenstrom and Strecker, 1993). Total emissions are estimated from flow-weighted composite samples that are collected at five sampling stations (four stations are required under the 1996 NPDES Municipal Permit and one station remains from an earlier permit.). The stations are equipped with flow monitoring equipment and operate unattended in secure facilities. Many water quality parameters are measured, including indicator organisms, general minerals, nutrients, metals, semi-volatile organic compounds and pesticides.

2.3. Caltrans First Flush Monitoring Study

The Department of Civil and Environmental Engineering at UCLA has characterized storm water runoff from three highway sites for California Department of Transportation (Caltrans) since 1999 (Stenstrom et al., 2000, 2001, 2002, and 2003). The study was conducted to assess runoff water quality and quantity from California freeways with particular emphasis on characterizing runoff during the early stage of storm events. Grab samples were collected every 15 minutes during the first hour of the storm. After the first hour, additional grab samples were collected each hour for up to 8 hours. An automatic composite sampler was also used after the first year to collect flow-weighted composite samples. A large suite of constituents including indicator bacteria, general minerals, nutrients, oil and grease, organic and metals were monitored.

3. Utility of the programs

3.1. Variation

Our early results showed that storm water data sets are fundamentally different that data sets derived from water and wastewater monitoring programs. The variability of storm water data and especially industrial storm water data is much greater than commonly found in potable water or wastewater datasets. This results in part because of the time-varying nature of storms but is also due to the use of grab samples and less experienced monitoring personnel. Most of the storm water monitoring programs allow for self-monitoring, which is similar to water and wastewater programs. The important difference is that many storm water permittees usually have no experience with water quality monitoring, whereas the water and wastewater permittees are usually formally trained and certified in the operation of a water or wastewater plant.

Figure 1 illustrates this difference in monitoring programs and shows the coefficient of variation (CV) for various routinely monitored water quality parameters. The leftmost bars show the CVs for influent wastewater quality parameters, which are typical for large west coast wastewater treatment plants. The CVs range from 0.09 to 0.2. The next set of bars show the CVs for raw water quality parameters for a large Los Angeles area potable water treatment plant. The CVs range from 0.07 to 0.3. The next group of bars shows the CVs from the Municipal Storm water Permit which range from 1.1 to 3.3 (Pitt 2004). The final group of bars shows the CVs from the GISP data for Los Angeles County. The CVs range from 2.8 to 14.

This graph dramatically illustrates the difficulty in using storm water monitoring data. The variability in the data is several times the mean value. With such high variability, it is virtually impossible to make statistical inferences, even for such simple matters as identifying high dischargers. The sampling method is a major source of much of the variation. For GISP monitoring, grab samples are allowed while flow-weighted composite samples were collected in the Municipal Storm Water Permit program. Water and wastewater treatment plant data are generally collected with composite samplers. Reducing the variability in data collection is the first problem that needs to be solved to improve the current monitoring program. In the recommendation section, we discuss the problem in detail and suggest ways to reduce the variability for improving future programs.

3.2. Regional Differences

Figure 2 shows the percentage of the data collected under the GISP and Multi-sector General Permit (MSGP) that falls outside the US EPA's benchmarks. Only the data from Massachusetts are shown for the MSGP , which had the largest number of observations; most other states had too few data to analyze. Data from Los Angeles County have the highest percentage outside the benchmark for all parameters except pH. Zinc exceeds the benchmark level in over 50 percent of samples for all the monitoring results, and in over 90 percent of the observations from Los Angeles County. The data from Connecticut has the lowest percentage outside benchmark, except for pH. Table 2 shows the basic statistics for the three major industrial monitoring. In general, samples from Los Angeles have the highest mean concentration and the largest range of CV for all parameters among the three GISP programs.

Climatic conditions such as the existence of long dry periods may greatly impact pollutant emissions from urban storm water discharges. Figure 3 shows the monthly average rainfall for Los Angels and Hartford, Connecticut during 1971-2000 (http://www.ncdc.noaa.gov/oa/climate/online/ccd/nrmlprcp.html).
Most of western parts of California including Los Angeles and Sacramento are dry from May through September. This rainfall pattern creates a long period for pollutant build-up, and therefore the initial storm of the wet season may have higher pollutant concentrations than in later events (Lee et al., 2004). This phenomenon is called a seasonal first flush, and is illustrated in Figure 4. Figure 4 shows the median concentrations of Specific Conductance, TOC, TSS, and Zn, as a function of the normalized cumulative precipitation during 2000-2001 wet season for GISP data from Los Angeles County. The first event of the wet season clearly has greater pollutant concentrations, and if only the first storm is monitored, it will overestimate pollutant concentrations in later storms. Areas with more uniform rainfall, such as Connecticut, did not shown a seasonal first flush. This presence of a seasonal first flush is a dilemma for monitoring programs, which attempt to identify high dischargers as well as estimate central tendencies, such as yearly emissions.

3.3. Facility landuse differences

Our previous study (Lee and Stenstrom, 2005) used several statistical procedures and several types of neural networks to identify relationships between water quality data and various landuses for three wet seasons (1998-2001). In this study we extended the analysis to the previous six wet seasons (1992-1998), and we also analyzed eight wet seasons (1993-2001) of Sacramento County's data and the nine years (1995-2003) of Connecticut's data. Eight major industries were selected; food and kindred products, chemical and allied products, primary metal industries, fabricated metal products, transportation equipment, motor freight transportation and warehousing, electric, gas, and sanitary services, and whole trade-durable goods. Outliers in the data set were first eliminated, and were defined as greater than 1.5 times the interquartile range plus the 75% data, or less than 25% data minus 1.5 times the interquartile range. The lower value was generally zero. The neural models were extensively trained using various architectures; however, the performance of all models was very poor, as previously observed, confirming that no relationship exists between the industrial categories based on SIC code and the storm water runoff quality.

To further evaluate relationships between water quality and various types of industries, a t-test for two industrial categories was used. The food and kindred products and primary metal industry were selected for demonstration because they are both major industries with many observations, and representative industry for a light and heavy industry, respectively. Also, the nature of the businesses suggests that there should be differences in storm water runoff quality. Data from two GISP programs, Los Angels County and Connecticut were analyzed.

Table 3 shows the case number and basic statistics for the two categories from the two regions after removal of outliers. The mean concentrations for all parameters are dramatically reduced as compared to the original data (Table 2). The table shows the t-value and the associated probability of obtaining a false result (Type I error) from a two-group t-test. The t-value is positive if the mean concentration of the food and kindred group is larger than the mean concentration of the primary metal group, and negative if it is smaller. The exception is the aquatic toxicity test, LC_{50} which has the opposite meaning of normal concentrations, with higher LC_{50} indicating less toxicity.

The significant t-values (alpha = 0.05) suggest that the conventional water quality parameters (TSS, SC, P, BOD and COD) from food and kindred products industries are higher than from primary metal industry for both Los Angeles County and Connecticut programs (the only exception is for pH in Los Angeles County). For several metals, the t-value from primary metal industry is higher than from food and kindred products industry. Unfortunately, the evaluation is unbalanced, with many more observations for primary metals industries. The combination of removal of outliers (approximately 5 to 10% of the observations for food and kindred products and 20 to 25% for primary metal industry) and the unbalanced and limited number of observations, reduces the confidence of any statistical comparisons for metals. The data do not support the commonly held view that metals-related industries discharge more metals that food producing industries. Four of six comparisons implicate metals industries in Los Angeles County and two of three implicate metals industries in Connecticut. Overall there is no basis for comparison due to problems with the number of observations and outliers in the datasets.

The aquatic toxicity test, LC_{50} is a more useful comparison parameter. The LC_{50} from food and kindred products industries are higher (i.e., less toxic) than from primary metal industry. The results support the commonly held view that storm waters from primary metal industries are likely to be more toxic than storm waters from food and kindred products. The other metal related industry in the Connecticut database, fabricated metal products industry, also has lower LC_{50} than the other major industries (data not shown). Unlike the metals concentration data, the toxicity data provides a stronger basis for comparison.

In general, the level of confidence associated with comparisons of large numbers of parameters needs to be higher than 0.95, which is an often-used value to indicate significance. For the 24 comparisons shown in Table 3, there is likely to be at least one Type I error using alpha = 0.05. In general, the t-values and levels of confidence shown in Table 3 are not sufficient to differentiate the two industries as categories. For example, the range of t- value for the four parameters from the Iris data set, which is the best known database in the pattern recognition literature (Systat 10.2, SPSS Inc., Chicago, IL), ranges from approximately 9 to 40 (alpha = 0.01). With the large value of the t-value, the iris data set can be successfully classified in three different groups using a typical supervised neural network model or other classification tool.

To illustrate the effect of the large CVs of the storm water data, and the associated difficulty of hypothesis testing, the required number of observations to differentiate parameters was calculated. Figure 5 shows the required sample number per group as a function of mean difference in percentage over a range of CVs. For example, the required sample number for the pooled CV of 1.0 and the mean difference of 50 % is 64 for each group or category. If the CV increases by a factor of two, the number of required samples increases approximately four times. The data generated from nine years of sampling that is shown in Table 3 is insufficient to detect differences of less than 50% in the mean for most parameters.

4. Recommendations

In this section, we discuss and offer recommendations to improve the overall utility of monitoring programs. The sampling method, sampling time, and sampling frequency are discussed and ways to improve accuracy and reduce variability are recommended. Reducing sampling variability is a key object in order to make monitoring programs useful for developing and enforcing TMDLs. The parameters to be monitored are also discussed.

4.1. Sampling method

A large part of the variability associated with the industrial monitoring programs is the reliance on grab samples. The CVs for municipal program were only 15% to 40% of the CVs for the same parameters monitored in the industrial monitoring program (Figure 1). It is believed that the major reason for this reduction is the use of flow-weighted composite samples in the Municipal Storm Water Permit program. Grab samples are discrete samples taken within a short period of time, usually less than 15 minutes. They capture both the variability within the storm (i.e., the first flush), as well as random fluctuations in runoff quality.

It is universally recognized that collecting flow weighted composite samples is better for storm water monitoring. The result is often called an event mean concentration (EMC), which is not only more representative than a grab sample, but can also be used to estimate pollutant loading, since the product of EMC and total runoff volume is the pollutant load (Sansalone and Buchberger, 1997; Lee et al., 2002). However, collecting flow-weighted composite samples is more difficult and expensive. An automatic flow weighted sampler, costing thousands of dollars, must be used. The location and installation of the sampler generally requires engineering and construction before the storm. Suitable locations must be found that allow flow measurement as well as security for the equipment. It will burden all industrial permittees to construct composite sampling facilities. Additionally, several water quality parameters such as oil and grease, toxicity and indicator bacteria are not easily measured by an automated composite sampler.

Using automated composite samplers may have the additional benefit of employing more skillful sampling crews. At present, most industrial permittees use ordinary employees, who may have no formal or practical training, to collect samples. Many aspects of storm water sampling require a skilled specialist. The training required to program a flow-weighted composite sampler could include other basic sampling skills, such as using appropriate containers (e.g., glass or plastic), handling and transportation and sample preservation.

4.2. Sampling frequency

The sampling time among storm events is also problematic. In areas like southern California, with extended dry periods, a seasonal first flush exists. If only the first storm is sampled, a biased result is obtained. The early storms have higher pollutant concentration than in later storms (Figure 4). Sampling only one or two storms each year is not likely to be representative. Leecaster et al., (2002) recommended sampling seven storms per year to obtain small confidence intervals. Increasing the number of samples will create a burden to all industrial permittees, and permittees will understandably question the benefits.

An alternative approach might be to select a subset of representative industrial facilities from each major category. A larger number of events could be sampled using improved sampling technology such as flow-weighted composite samplers. The cost of such an approach could be shared, and the overall cost might even be lower than current costs. Even if only 10% of the permittees were sampled using composite samplers, the CVs should be reduced. The remainder of the permittees could continue using grab samples or use some other program, as mandated. A regulatory agency or "broker" will have to develop appropriate methods for selecting sampling sites and distributing costs.

4.3. Sampling Time

Sampling time during a storm event is important to avoid the bias of the first flush and properly characterize the event. The industrial sites are generally small watersheds, and will experience an event first flush. Figure 6 shows the impact of sampling time for TSS and total Zn from highway sites. The ratio of observed concentration to EMC is shown for more than 30 events during 2000-2001 and 2001-2002 wet seasons. The line at 1.0 is the ideal ratio, if the grab samples were the same as EMC. In general the ratio of the observed concentration to EMC is much higher than 1.0 in the beginning of storm and declines as the storm progresses. It is obvious that collecting a sample in the early part of the storm overestimates the EMC and the total load. Khan et al. (2004) has examined this effect for sampling oil and grease, which must be collected as a grab sample. They concluded that collecting a grab sample 2 to 3 hours into a typical storm more closely approximates the EMC than sampling earlier or later in the storm. If grab samples are to be used, the most appropriate time to sample should be investigated.

4.4. Monitored Parameters

The choice of monitored parameters will depend on the expected loads, which can be related to landuse. Industrial and transportation landuses are generally known as greater sources of heavy metals (generally six metals: cadmium, copper, chromium, lead, nickel and zinc) than other landuses such as residential. Figure 7 shows this relationship by plotting metal concentrations as a function of landuse (data adapted from Pitt et al., 2003). Industrial landuse has the highest median concentrations for all metals except copper, which is highest from highways. Similar results were found from Los Angeles County's Municipal Storm Water Permit program, which were not included in the survey. Industrial landuse had the highest concentrations of aluminum, nickel, and zinc. Copper was highest in transportation landuse (not shown in this paper). All the metals can usually be analyzed from a single sample using an instrument called an ICP/MS.

Including metals in storm water monitoring programs is also important to assess environmental impact. Industrial permittees exceeded the US EPA's storm water benchmark concentrations for metals more frequently than for basic water quality parameters (Figure 2). For example, zinc concentrations exceeded the benchmark concentration in over 50 percent of samples for all the GISP monitoring results. In addition, Bay et al. (2003) reported that zinc was the primary cause of toxicity in both the storm water in Ballona Creek and the near shore waters where it discharges. Ballona Creek is the largest stormdrain to Santa Monica Bay.

Including metals in industrial storm water monitoring programs should be a high priority but varies among states and industrial categories. It is mandatory in Connecticut for all industrial categories, whereas only certain industrial facilities are required to analyze for metals in California. There may be logic for including or excluding metals in specific industrial permits, but the current monitoring results showed no remarkable differences among industrial categories. Additionally, the source of metals from industrial landuse is not only from industrial activity but also may be from building materials such as roof material and siding material of the industrial facilities (Davis et al., 2001). In anticipation of controlling metals for TMDLs and other purposes, any future permit should require monitoring for metals for the six metals previously noted for all industrial landuses.

5. Conclusions

This paper has examined several General Industrial Storm Water Permit monitoring programs, the Multi-Sector Storm Water General permit, Municipal Permit programs from 17 states and Los Angeles County, and Caltrans' first flush highway runoff characterization study. We evaluated the storm water monitoring programs to determine their usefulness in characterizing discharges and achieving their ultimate goal of reducing pollutants in storm water discharges. The following conclusions and recommendation are made for improving future programs:

1. The variability of storm water data and especially the General Industrial Storm water Permit monitoring data collected with grab samples is much greater than commonly found in potable water or wastewater monitoring programs. Regulators and others need to understand that decision-making using the storm water monitoring results is limited due to its high variability. Reducing the variability in data collection is the first problem that needs to be solved to improve current monitoring programs.
2. Data from Los Angeles County have the highest percentage of observations outside the US EPA benchmarks among the four Industrial storm water monitoring programs. The only exception was pH. Zinc exceeded the benchmark concentration in over 50 percent of samples for all the Industrial monitoring results, and in over 90 percent of the observations from Los Angeles County.
3. Data from the Los Angeles County, which has distinct dry periods from May through September, had higher pollutant concentrations in early storm events than

in later events, documenting the existence of a seasonal first flush. Data from Connecticut, with more uniform rainfall throughout the year, did not show a seasonal first flush.

4. Several techniques were used to differentiate industrial categories based upon monitoring results. Neural models were extensively trained using various architectures to identify relationships between water quality data and various landuses, but were unable to detect relationships. Variability in monitoring data obscured any relationship that might exist between the industrial categories based on SIC code and the runoff water quality. To further evaluate relationships between water quality and various types of industries, a t-test for two distinct industrial categories was used. In general, the t-values and levels of confidence were not sufficient to differentiate the two industries.
5. A power analysis was performed to determine how many additional observations are needed to differentiate landuse or industry types. The number of required additional observations was several times the current number of observations, confirming that measures to reduce variability are preferred to additional observations.
6. Several recommendations for improved monitoring programs were made. Selecting a subset of each monitored category using more advanced sampling techniques such as flow-weighted composite samplers is a reasonable approach and may result in lower overall cost with improved accuracy and variability.
7. The timing of grab sample collection will affect results. Samples collected early in the storm will have higher concentrations than the EMC, and will be lower than

the EMC if collected late in the storm. Areas with long dry periods are likely to show a seasonal first flush. Grab samples need to be collected at the appropriate time.

8. Industrial landuse have the highest concentrations for all metals except copper. In addition, concentrations of metals exceeded the storm water benchmark values suggested by US EPA more frequently than the basic water quality parameters. In anticipation of controlling metals for TMDLs and other purposes, future permits should require monitoring for metals for all industrial landuses.

We believe our results will be helpful to planners and regulators to interpret existing datasets and programs as well as providing recommendations for improving future programs.

Table 1 Summary of the monitoring data sets for this study

	Name of monitoring program	Monitoring area	Primary landuse	Observation year in this study
Industrial Storm water Monitoring	General Industrial Permit	County of Los Angeles, CA	Industrial	1992 - 2001
	General Industrial Permit	County of Sacramento, CA	Industrial	1993 - 2001
	General Industrial Permit	Connecticut	Industrial	1995 - 2003
	Multi-Sector General Permit	States and territories that have not been authorizes to run NPDES general permit program	Industrial	1998 - 2003
Municipal Storm water Monitoring	Municipal Permit	17 states in U.S.	Various landuse	1991 - 2002
	Municipal Permit	County of Los Angeles, CA	Various landuse	1996 - 2001
Other Storm water Monitoring	First Flush Highway Runoff Characterization	405 and 101 freeway near UCLA, CA	Transportation (highway)	1999 - 2003

Table 2 Basic statistics for three General Industrial Storm Water Permit programs

Parameters [a]	Los Angels County			Sacramento County			Connecticut State		
	Sample No.	Mean.	CV	Sample No.	Mean.	CV	Sample No.	Mean.	CV
pH (pH unit)	24851	7.01	0.95	857	7.16	0.17	9617	6.32	0.20
TSS	24144	376.15	11.85	769	185.49	2.86	9617	124.02	6.59
SC (μmhos/cm)	23585	561.68	8.13	846	204.20	2.27			
Oil & grease	18637	16.57	14.25	286	11.26	1.61	9561	5.66	14.57
TOC	9714	50.13	5.23	399	31.44	2.12			
COD	1834	271.29	2.77	50	154.40	1.58	9606	80.67	3.44
Aluminum	1618	10.12	12.21	46	3.44	1.66			
Copper	3354	1.01	16.50	83	0.18	2.31	9596	0.13	7.56
Iron	1844	25.48	6.39	82	7.49	2.03			
Lead	3525	2.96	14.12	78	4.48	3.82	9563	0.06	8.37
Zinc	5163	4.96	13.85	141	2.23	7.59	9614	0.51	7.79
Phosphorous							9606	0.45	4.30
TKN							9608	2.50	3.11
NO$_3$-N							9613	1.19	2.73
24hr LC$_{50}$ (%)							9628	82	0.38
48hr LC$_{50}$ (%)							9628	75	0.45

[a] Unit is mg/l unless otherwise noted

Table 3 Basic statistics and t-value between food and kindred products and primary metal industry

Program	Parameters	Food and kindred products			Primary metal industry			t-value	Probability
		Sample No.	Mean.	S.D.	Sample No.	Mean.	S.D.		
Los Angeles County	pH	1594	6.740	0.63	1649	6.898	0.64	-7.08	0.000
	TSS	1451	66.402	66.59	1586	56.035	60.22	4.49	0.000
	SC	1497	159.222	146.65	1585	134.780	127.62	4.92	0.000
	O&G	905	5.803	4.59	1257	5.525	4.24	1.44	0.151
	TOC	774	21.509	19.01	448	19.862	17.19	1.55	0.121
	BOD	62	38.942	36.98	25	17.652	18.20	3.58	0.001
	COD	65	179.986	167.65	57	86.544	93.65	3.86	0.000
	Al	9	2.624	1.69	287	1.664	1.63	1.68	0.129
	Cu	16	0.063	0.06	478	0.095	0.10	-2.09	0.052
	Fe	9	4.839	2.70	153	3.484	3.40	1.44	0.182
	Pb	19	0.048	0.04	211	0.068	0.07	-2.08	0.046
	Ni	15	0.032	0.03	111	0.052	0.04	-2.57	0.017
	Zn	42	0.713	0.63	536	0.851	0.74	-1.35	0.183
Connecticut State	pH	168	6.421	1.07	699	5.847	1.24	6.07	0.000
	TSS	153	38.608	37.87	657	25.129	31.76	4.08	0.000
	COD	146	42.795	34.80	668	36.473	31.29	2.02	0.044
	P	153	0.192	0.15	659	0.144	0.13	3.65	0.000
	TKN	145	1.666	1.12	670	1.405	0.96	2.61	0.010
	NO_3	150	0.829	0.72	664	0.768	0.60	0.97	0.333
	Cu	151	0.035	0.03	526	0.040	0.03	-2.14	0.033
	Pb	163	0.019	0.02	646	0.021	0.02	-1.15	0.250
	Zn	139	0.236	0.18	589	0.199	0.16	2.20	0.029
	24hr LC_{50}	156	0.926	0.17	518	0.845	0.22	4.78	0.000
	48hr LC_{50}	168	0.786	0.32	700	0.570	0.38	7.59	0.000

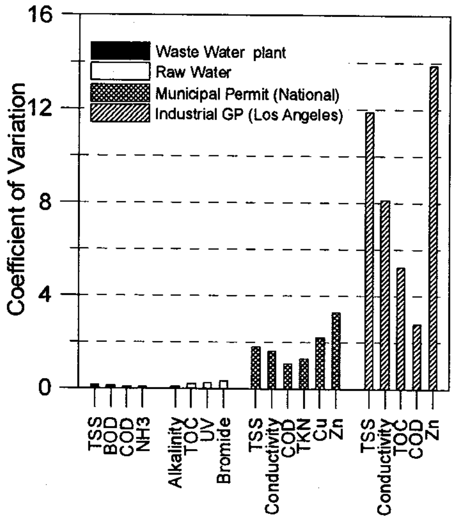


Fig. 1. Coefficient of variation in various water sampling programs.

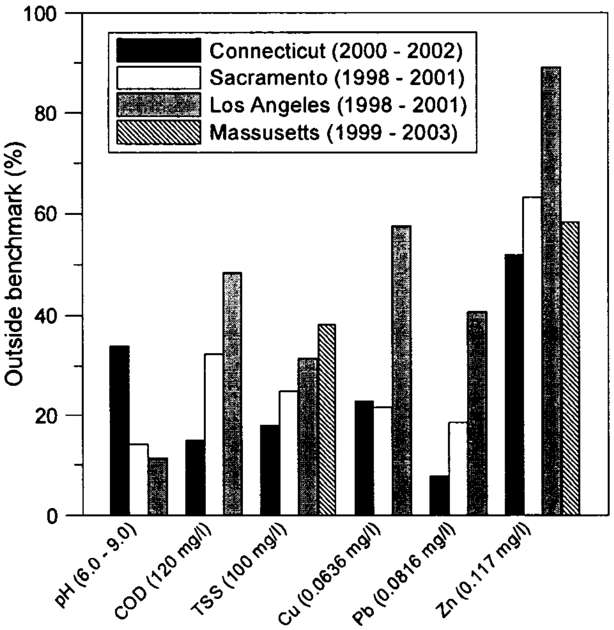


Fig. 2. Percentage of observations outside the US EPA benchmark for Industrial General Permit programs. Concentration in parenthesis on the x-axis indicates the storm water benchmark concentration.

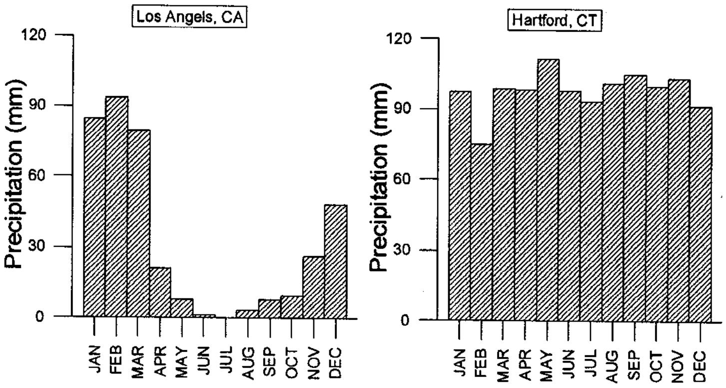


Fig. 3. Monthly average rainfall in Los Angeles and Hartford during 1971-2000.

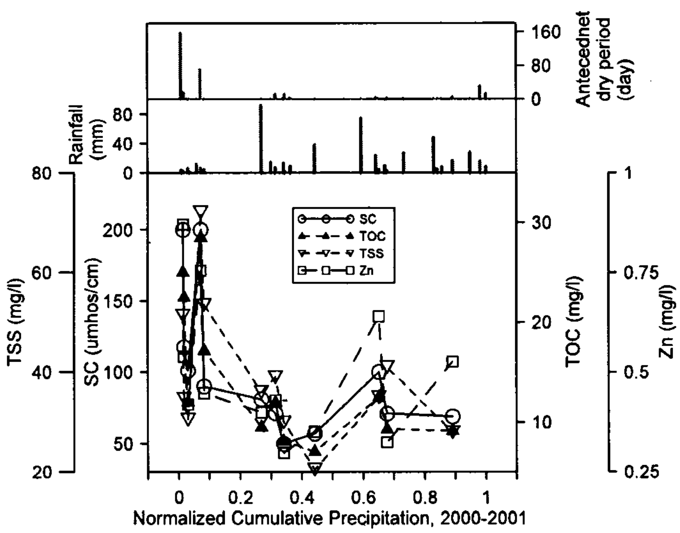


Fig. 4. Concentrations versus normalized cumulative rainfall during 2000-2001wet season for Los Angeles General Industrial Storm Water Permit monitoring.

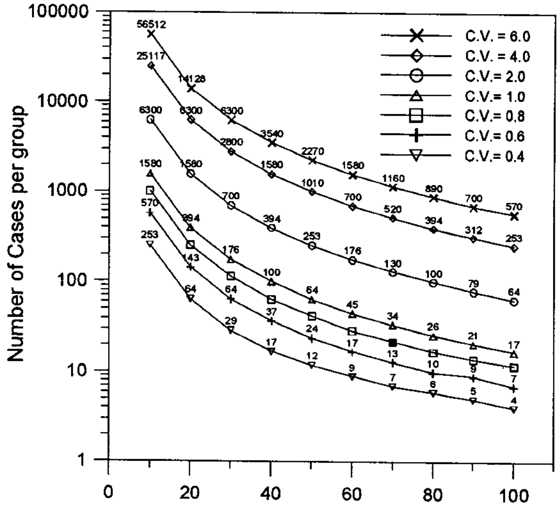


Fig. 5. The required sample number per group as a function of mean differences in percentage at alpha = 0.05 and power = 0.8.

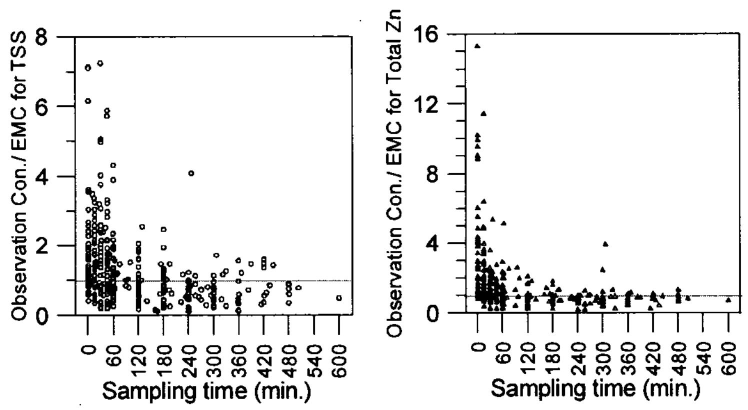


Fig. 6. The ratio of observed concentration to EMC for TSS (left) and Total Zn (Right) from UCLA first flush Study during 2000-2001 and 2001-2002 wet seasons.

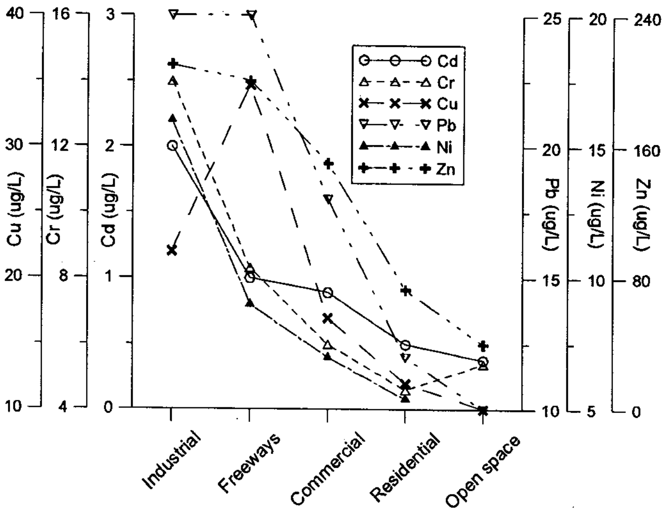


Fig. 7. Median concentrations of metals in various landuse from the Municipal Stormwater Permit program.

REFERENCES

Bay, S., Jones, B.H., Schiff, K. and Washburn, L. (2003) Water quality impacts of stormwater discharges to Santa Monica Bay. *Marine Environmental Research*, **56**, 205-223.

Davis, A.P., Shokouhian M. and Ni, S. (2001) Loading estimates of lead, copper, cadmium, and zinc in urban runoff from specific sources. *Chemosphere*, **44**, 997-1009.

Duke, L.D., Buffleben, M. and Bauersachs, L.A. (1998) Pollutants in storm water runoff from metal plating facilities, Los Angeles, California. *Waste Management*, **18**, 25-38.

Khan, S. and Stenstrom, M.K. (2004) Oil and Grease Measurement in Highway Runoff – Sampling Time and Event Mean Concentrations. submitted to the *Journal of Environmental Engineering,* ASCE.

Lee, H. and Stenstrom, M.K. (2005) Utility of stormwater monitoring. *Water Environmental Research,* Vol. 77, No. 1.

Lee, H., Lau, S-L., Kayhanian, M. and Stenstrom, M.K. (2004) Seasonal first flush phenomenon of urban stormwater discharges. *Water Research*, **38**, 4153-4163.

Lee, J.H., Bang, K.W., Ketchum, L.H., Choe, J.S., and Yu, M.J. (2002) First flush analysis of urban storm runoff. *The Science of the Total Environment*, **293**,163-175.

Leecaster, M.K., Schiff, K. and Tiefenthaler, L.L. (2002) Assessment of efficient sampling designs for urban stormwater monitoring. *Water Research*, **36**, 1556-1564.

Pitt, R., Maestre, A. and Morquecho, R. (2003) Compilation and review of nationwide MS4 stormwater quality data. *Proceedings of 76th Water Environment Federation Technical Exposition and Conference*, Los Angeles, CA, 2003.

Pitt, R (2004) http://unix.eng.ua.edu/~rpitt/Research/ms4/Final%20Table%20NSQD%20v1_1%20020704.xls .

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

Stenstrom, M.K. and E. Strecker (1993) Assessment of Storm Drain Sources of Contaminants to Santa Monica Bay. Vol. I, Annual Pollutants Loadings to Santa Monica Bay from Stormwater Runoff, UCLA-ENG-93-62, May 1993, Vol. I, pp 1-248.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2000) First Flush Stormwater runoff from Highways. *Research report for California Department of Transportation*, September, 2000.

Stenstrom, M.K., Lau, S-L., Ma, J-S., Ha, H., Lim, L-H., Lee, S., Khan, S. and Kayhanian, M. (2001) First Flush Stormwater runoff from Highways. *Research report for Department of California Transportation*, September, 2001.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2002) First Flush Stormwater runoff from Highways (year3). *Research report for Department of California Transportation*, July, 2002.

Stenstrom, M.K., Lau, S-L., Lee, H-H., Ma, J-S., Ha, H., Kim, L-H., Khan, S. and Kayhanian, M. (2003) First Flush Stormwater runoff from Highways (year4). *Research report for Department of California Transportation*, July, 2003.