Characterization of Urban Runoff Quality: A Toronto Case Study

Pradeep Kumar Behera, James Y. Li and Barry J. Adams

This chapter presents an overview of the characterization of urban runoff quality constituents. Characterization includes descriptive statistics, correlation analysis, frequency analysis and regression analysis of event mean concentrations of various quality constituents from separated and combined sewer catchments. Emphasis is placed on the procedures required to determine not only summary statistics but also complete descriptions in the form of probability density functions. Metropolitan Toronto runoff quality databases are used to illustrate these procedures. The event mean concentrations of fifteen quality constituents representing chemical and bacteriological pollutants, nutrients and heavy metals were studied. Three probability distributions (exponential, gamma and lognormal) were fitted to the data and goodness-of-fit was assessed using the Kolmogorov-Smirnov test. In many studies, the lognormal probability distribution has been assumed to describe the runoff quality constituents. However, in this study, in addition to the lognormal distribution, it is observed that gamma and exponential probability distributions can also adequately describe runoff quality constituents.

Behera, P., J. Li and B.J. Adams. 2000. "Characterization of Urban Runoff Quality: A Toronto Case Study." *Journal of Water Management Modeling* R206-14. doi: 10.14796/JWMM.R206-14.

© CHI 2000 www.chijournal.org ISSN: 2292-6062 (Formerly in *Applied Modeling of Urban Water Systems*. ISBN: 0-9683681-3-1)

Portions of this chapter are reprinted from **URBAN STORMWATER MANAGEMENT PLANNING WITH ANALYTICAL PROBABILISTIC MODELS** by B.J. Adams, © 2000 John Wiley and Sons Inc., with permission of John Wiley and Sons Inc.

14.1 Introduction

Numerous studies on runoff quality characterization conducted in different parts of the world have concluded that urban runoff carries relatively high concentrations of a variety of pollutants. These pollutants originate from diverse sources, both natural and anthropogenic, categorized by various boundary inputs, pollution processes and human activities that occur on the urban catchment. The impacts of stormwater pollution have been recognized by many agencies and municipalities as limiting the full and beneficial uses of receiving waters in urban environments (Ellis, 1986). Characterization of stormwater pollution is necessary (Heaney et al., 1998) to understand the impacts of stormwater on receiving waters, to develop stormwater management alternatives, to operate and maintain sustainable drainage systems and to predict the ultimate fate of pollutants in the environment.

Studies conducted to address the control of runoff quantity and quality problems in urban areas include: estimation of runoff quantity; characterization of quality constituents, sources, transport and fate of contaminants; and their corresponding impacts on receiving waters. Runoff quantity modeling issues have been effectively addressed by simulation models (e.g. SWMM, Huber and Dickinson, 1988) and analytical models (e.g. SUDS, Adams and Bontje, 1984; and ASWMM, Guo and Adams, 1998) using available meteorological and hydrological data. The collection of such data has become a relatively routine task of government agencies, and the procedures for data collection are fairly well established. In contrast, achievements in addressing runoff quality issues have been more limited for two reasons. The first is the availability of long-term runoff quality data is limited by the expense of data collection along with the absence of widely accepted procedures/protocols for data collection and analysis. In addition, most runoff quality data collection efforts have been confined to individual urban catchments. Second, early approaches to runoff quality modeling such as pollutant generation and transport were physically based; the results from such urban runoff quality models have often been regarded with skepticism. Consequently, the complex nature of the processes involved and the difficulties in the representation of those processes in physically-based models have led to statistical approaches to modeling runoff quality and its impact on receiving waters. To facilitate these approaches, databases have been established with the objective of characterizing stormwater pollution generation, transport and fate. These databases have furthered the need for more reliable stormwater quality data analysis methods to derive the maximum utility from the available data.

Examples of major investigations into urban runoff quality are the Nationwide Urban Runoff Program (NURP) undertaken by the US EPA (1983) and the Urban Stormwater-Quality Investigations of the U.S. Geological Survey (Jennings and Miller, 1986). The NURP study investigated runoff quality problems by

collecting data from 81 sites in 22 cities geographically distributed across the United States, with a total of 2300 separate storm events. Similar studies were also conducted in Canada including the Pollution from Land Use Reference Group (PLUARG) of the International Joint Commission (Novotny and Olem, 1994), which conducted a series of studies to identify pollution in the Great Lakes basin from surface runoff, atmospheric and wastewater sources. In addition, a number of individual studies have been conducted by different levels of government and by consultants to support stormwater management plans in urban watersheds. For example, the Toronto Area Watershed Management Strategy Study (MOE, 1984) was conducted by the Ontario Ministry of the Environment to develop a comprehensive water quality management plan for Toronto area watersheds.

Water pollution caused by discharges from urban drainage systems, both combined and separate, and their consequent impacts are highly variable as they occur from time to time with different magnitudes. Hence, urban drainage system performance from a runoff quality control perspective is related to the magnitude and frequency of pollutant mass discharges from combined sewer overflows (CSOs) and stormwater discharges to receiving waters (Adams and Zukovs, 1986). Runoff pollution constituents are numerous and include: floatable material, suspended solids, oxygen-demanding materials, nutrients, pathogenic microorganisms, toxicants such as heavy metals and pesticides, petroleum hydrocarbons and other hazardous contaminants. Although it is not a constituent per se, the thermal enrichment of runoff is also considered to be a form of water pollution. These pollutants have different types of effects on receiving waters and over different time scales. Typically, water quality concerns related to stormwater discharges are classified as short-term acute effects and long-term cumulative effects (Harremoes, 1988). Acute effects are characterized by high pollutant concentrations resulting from individual events measured in hours, such as those caused by certain toxicants. Cumulative effects are characterized by a gradual build-up of pollutant mass and concentration over a long period of time leading to detrimental environmental effects after certain threshold levels have been exceeded. Examples of such effects include the accumulation of plant nutrients or persistent toxicants (Harremoes, 1988). The above observations suggest that the evaluation of stormwater pollution discharges and the performance of urban drainage systems should be based on both extreme and average statistics of pollutant concentrations and loads.

This chapter presents a statistical framework for the characterization of urban runoff quality constituents. Characterization includes descriptive statistics, correlation analysis, frequency analysis and regression analysis of event mean concentration (EMC) of various quality constituents from separated and combined sewer catchments. Emphasis is placed on the procedures required to determine not only summary statistics such as means and standard deviations, but also complete descriptions in the form of probability distributions. A key

aspect of this characterization is to establish the appropriate probability distributions of various pollutants in urban runoff for use in the long-term management of stormwater pollution. An extensive Metropolitan Toronto runoff quality database is used to illustrate these procedures.

14.2 Runoff Quality Concentration Estimation

The primary quantitative measure of a runoff quality constituent is its concentration. It is defined as the mass of a constituent per unit volume of water and is mathematically represented as:

$$C = \frac{M}{V} = \frac{\text{Quantity (mass) of a constituent}}{\text{Volume of water}} \quad (14.1)$$

The concentration of the constituent is generally measured in mg/L. For certain other constituents, concentration is not expressed in terms of mass units; for example, bacterial concentration is measured as number counts (most probable number or MPN) per unit volume.

The event mean concentration of a pollutant is defined as the total pollutant mass M (constituent load) discharged during a runoff event divided by the total runoff volume during a given storm event. Typically, the EMC is obtained from flow-weighted composite samples taken during a storm, mathematically represented as:

$$EMC \equiv \bar{C} = \frac{M}{V} = \frac{\int C(t)Q(t)dt}{\int Q(t)dt} \quad (14.2)$$

where:

$C(t)$	=	time variable concentration of the pollutant
$Q(t)$	=	time variable volumetric flow rate during the event
M	=	pollutant mass as defined in Equation 14.1
V	=	runoff volume as defined in Equation 14.1

It is emphasized that the representation of mean concentration as an EMC is a flow-weighted average concentration, not a time-averaged concentration during the event. The variability of concentration during a storm event is related to rainfall driven runoff which generally follows the temporal and spatial characteristics of rainfall. It is noted that the instantaneous pollutant concentration during a storm can be higher or lower than the EMC, but the use of the EMC as an event

characterization substitute the actual time variation of concentration in a storm event with a pulse of constant concentration having the same mass and duration as the actual event (Huber, 1993). Concentration-based pollutant load estimation models employ the EMC of the pollutant. When the EMC is multiplied by event runoff volume, an estimate of event pollutant load is obtained. Such models are also useful in the estimation of annual loads.

14.3 Study Area and Runoff Quality Data

The degradation of Great Lakes water quality is a major concern to both Canada and the United States. Both countries have identified several polluted areas or Areas of Concern (AOCs) located in the Great Lakes basin. Most of these AOCs are located in the proximity of large urban centres. Stormwater discharges and CSOs from these areas are considered as one of the major sources of toxic contaminants in the Lakes. Potential problems from these sources were rated as medium or high in 11 of the 17 Canadian AOCs and as very high for Hamilton Harbour and Metro Toronto Waterfront (Weatherbe and Sherbin, 1994). The delisting of many of these AOCs will depend on successful abatement of stormwater and CSOs pollution (Marsalek and Kok, 1997). In order to establish the degree to which stormwater discharges and CSOs contribute to impairment problems and to recommend appropriate cost effective abatement strategies in AOCs, a number of pollution control studies have been conducted over last two decades. The Ontario Ministry of the Environment (MOE) commissioned two such studies to characterize and to quantify contaminants from urban runoff discharged into the Metropolitan Toronto Waterfront (MTW). Paul Theil Associates and Beak Consultants (1992) and Aquafor Engineering (1995) conducted these studies in two phases. The study area comprises waterfront catchments in the Cities of Toronto, Scarborough and Etobicoke. The MTW, approximately 50 kilometers in length, receives direct flow from: over 100 waterfront sewer outfalls, of which 31 are combined sewer outfalls; effluents from three water pollution control plants; backwash water from three water treatment plants; and flow from six streams (Humber, Don and Rouge Rivers, and Etobicoke, Mimico and Highland Creeks). At present, both separated systems and/or combined sewer systems service the MTW catchments. Due to the large number of outfalls along the waterfront and limited resources, sixteen representative outfalls were selected to characterize the contaminants from various land uses. Of these sixteen outfalls, seven were predominantly combined sewer catchments, and nine were separated storm sewer catchments. The catchments were selected based on the drainage area, land use, type of sewershed (e.g. combined or separate), geographic region and complexity of the sewer system. The catchment areas varied from 7 to 181 hectares. The field program

was conducted during the fall of 1989 and 1990 by monitoring flow rates and collecting flow-weighted samples of runoff at the outfalls. However, the monitoring program was continued at two of these outfalls, representing combined and separated catchments, for a complete calendar year.

Montedoro Whitney Q-Logger flow-monitoring equipment was used to continuously record the flow depth and velocity. The flow monitor was interfaced with an automatic wastewater sampler (ISCO Model 2700) for the collection of flow dependent sample aliquots. The sampler-triggering device on the Q-Logger was set to sample more frequently during the periods of higher flow rates in the sewer. Twenty-litre flow-proportional water quality samples were collected during each wet weather event. The sampling methods followed the MOE sample collection, preservation and submission protocol (MOE, 1988). Composite samples were analyzed by the MOE Laboratory Services Branch. The flow-weighted runoff quality samples were analyzed for a comprehensive list of parameters, including general water chemistry, bacteriology, nutrients, heavy metals and trace organic compounds, and their EMC values were reported. From the database, 15 quality constituents of general interest are presented herein. The constituents chosen are commonly found in runoff, they represent most pollutant types and they have reasonable detection limits. Furthermore, these constituents are consistently present in the data sets.

14.4 General Statistics

The EMC data sets for the 15 water quality parameters from nine separated sewer catchments and seven combined sewer catchments were analyzed. The initial step in the analysis was to check the data sets for outliers. The procedure was to: (i) group the data sets according to combined and separated sewer catchments; (ii) calculate the pooled mean and standard deviation of each quality parameter; (iii) set the outlier limit (the value above which the data points were deleted) as the mean plus three standard deviations for each pollutant; and (iv) delete from data sets data points greater than the outlier limit. The outliers constitute less than 4% of the total data points. After the outliers were removed, the catchments having approximately similar land use characteristics were grouped for each of the combined and separate sewer systems. A general statistical analysis consisting of estimation of means, standard deviations and coefficients of variation, was conducted on the quality parameters for both types of catchments.

A summary of the calculated runoff quality statistics from combined sewer and storm sewer catchments is given in Table 14.1. The EMC data sets used in the study were frequently detected and contained mostly non-censored data. Probability distribution estimation techniques were used to estimate the mean and variance for the data sets containing censored data (Paul Theil Associates and

Table 14.1 EMC statistics for various pollutants from combined sewer and storm sewer catchments.

Pollutant	Combined Sewer Catchments				Storm Sewer Catchments			
	Range	Mean	Standard Deviation	Co. Var.	Range	Mean	Standard Deviation	Co. Var.
Suspended Solids(mg/L)	3 -1,250	256	248	0.97	10 - 482	133	116	0.87
Total Solids (mg/L)	184 -2,870	822	538	0.65	64 - 3, 420	796	683	0.86
E. Coli (counts/100ml)	170 - 4.5×10^6	8.1×10^5	1.08×10^6	1.33	200 - 2.2×10^6	1.22×10^5	4×10^5	3.27
Fecal Coliforms (counts/100ml)	170 - 7×10^6	1.1×10^6	1.6×10^6	1.45	200 - 8.7×10^5	60, 000	1.3×10^5	2.23
Fecal Streptococci (counts/100ml)	110 - 9.5×10^5	1.6×10^5	2×10^5	1.24	370 - 7.7×10^5	55, 000	1.45×10^5	2.64
Nitrates (mg/L)	0.06 - 8.35	1.57	1.64	1.04	0.05 -37.7	2.55	5.21	2.04
Total Kjeldahl Nitrogen (mg/L)	0.3 -16	4.84	3.34	0.69	0.4 -13.5	2.84	2.24	0.79
Total Phosphorus (mg/L)	0.04 -3.5	1.12	0.81	0.72	0.1 - 2.15	0.62	0.45	0.73
Chemical Oxygen Demand (mg/L)	16 -3,842	782	846	1.08	4 - 4,818	809	1, 077	1.33
Phenol (mg/L)	1.2 - 34	8.5	6.16	0.72	1-50	12.0	10.30	0.86
Aluminum (mg/L)	0.04 - 9.7	2.7	2.20	0.81	0.07 - 8.4	1.37	1.46	1.07
Barium (mg/L)	0.02 - 0.22	0.1	0.05	0.50	0.006 - 0.13	0.05	0.03	0.60
Copper (mg/L)	0.004 - 0.3	0.08	0.06	0.75	0.0021- 0.11	0.04	0.02	0.50
Iron (mg/L)	0.19 - 22	6.75	5.58	0.83	0.16 - 15	3.0	2.65	0.88
Zinc (mg/L)	0.011 - 0.77	0.27	0.16	0.59	0.011 - 0.41	0.13	0.08	0.62

Beak Consultants, 1992 and Aquafor Engineering, 1995). The metal concentrations represent the total water concentration. The descriptive statistics shows that pollutant concentrations vary considerably between events, between catchments of similar type of drainage system and between the types of catchments. The coefficients of variation of runoff quality concentration range from about 0.5 to 1.0 for many of the pollutants, which is consistent with NURP (1983) results. For bacterial organisms, typically higher variability is expected, although a portion of this may be due to sampling error and measurement. The EMC data presented in Table 14.1 are valuable in assessing the magnitudes of the specific problems related to runoff quality. From these results, it can be seen that the average EMCs of various pollutants in combined sewer catchments are generally higher than those in separated storm sewer catchments. However, the variability of EMCs is generally higher in separated sewer catchments than combined sewer catchments. Suspended solids in the combined sewer catchments are twice as high as in the storm sewer catchments; however, total solids are approximately the same for both types of catchments. Generally separated sewer catchments contribute levels of bacteriological pollutants within an order of magnitude of those from combined sewer catchments. The magnitudes of nutrients and heavy metals are higher in the combined sewer catchments than in the separated sewer catchments.

14.5 Correlation Analysis

In order to determine the degree of dependence between pairs of pollutants, correlation analyses were conducted for the one-year EMC data sets on the separated and combined sewer catchments. The correlation between pollutants is used to identify possible relationships between individual pollutants as well as between different categories of pollutants. This could be useful for studying areas where limited quality parameters are measured. For instance, when two pollutants have a high correlation, a high value of EMC of one pollutant may indicate that another unmeasured pollutant would also have a respectively high value. The product moment correlation coefficients were estimated for pairs of the various pollutants. Although correlation coefficients were estimated for all of the pollutants, only the more significant relationships are presented in Table 14.2. From these results, it is seen that suspended solids are positively correlated with nutrients and heavy metals, and that nutrients are positively correlated with heavy metals. The pollutants within the bacteriological and heavy metal categories are highly correlated among themselves. These relations are consistent in both types of sewersheds.

Table 14.2 Correlation among pollutants in separate and combined sewer catchments.

Correlation among pollutant EMCs in separated sewer catchment									
	TSS	E-Coli	F-Coli	F-Strep.	TKN	TP	Al	Cu	Fe
E-Coli	0.15								
F-Coli	0.19	0.98							
F-Strep.	0.18	0.86	0.93						
TKN	0.77	-0.05	0.09	0.28					
TP	0.74	-0.11	0.05	0.29	0.86				
Al	0.67	-0.10	-0.12	0.13	0.73	0.63			
Cu	0.60	-0.04	-0.15	0.33	0.62	0.65	0.75		
Fe	0.70	-0.07	-0.11	0.13	0.73	0.62	0.91	0.61	
Zn	0.58	-0.02	-0.19	0.28	0.64	0.61	0.77	0.84	0.64
Correlation among pollutant EMCs in combined sewer catchment									
	TSS	E-Coli	F-Coli	F-Strep.	TKN	TP	Al	Cu	Fe
E-Coli	0.36								
F-Coli	0.23	0.92							
F-Strep.	0.50	0.65	0.67						
TKN	0.69	0.63	0.58	0.79					
TP	0.70	0.40	0.36	0.77	0.96				
Al	0.70	0.30	0.28	0.62	0.66	0.60			
Cu	0.74	0.60	0.50	0.81	0.82	0.76	0.76		
Fe	0.84	0.39	0.46	0.63	0.72	0.67	0.83	0.80	
Zn	0.82	0.54	0.55	0.74	0.74	0.67	0.77	0.91	0.85

The presence of correlation between pollutant levels and rainfall characteristics could be useful in identifying rainfall characteristics contributing to high pollutant levels and possibly designing control measures. Correlation analyses were therefore conducted between EMCs and rainfall characteristics of the Toronto data; however, no significant relations were observed. This may be due to the unavailability of on-site rainfall data as the analyses were performed using the closest meteorological station records. Thus, it is important to record the on-site rainfall with stormwater quality data.

14.6 Frequency Analysis

The purpose of EMC frequency analysis is to select an appropriate probability distribution function (PDF) to describe the pollutant, which may be achieved by fitting theoretical PDFs to the observed data. In order to perform such analyses, a long-term EMC record, at least a calendar year should be used. The reason for this is to incorporate any seasonal variability into EMC values because the concentrations of some pollutants, such as nutrients, may vary from season to season. Furthermore, if one year of EMC data is used, the probability of exceeding some concentration level per runoff event on an annual basis can be estimated. Such information is useful to engineers and planners in assessing the magnitude of runoff pollution in the catchment on an extreme event basis.

Four important uses arise from the application of PDFs of EMCs in stormwater management analysis: (i) computation of pollutant loads; (ii) determination of the probability of compliance with water quality criteria; (iii) assessment of runoff controls; and (iv) processing of censored stormwater quality data (Van Buran et al., 1997).

First, the PDF of the pollutant load may be derived from the PDFs of runoff volume and EMC. Using derived probability distribution theory (Benjamin and Cornell, 1970) the distribution of pollutant load is obtained by integrating the joint distribution of runoff volume and concentration. Alternately, the PDF of the pollutant load may be derived from the PDFs of rainfall characteristics and the PDF of EMC. A constant concentration approach is often adopted by replacing the PDF of EMC distribution with a constant mean value in the estimation of the load distribution (Li, 1991).

Second, water quality standards are defined in terms of the magnitude (maximum concentration), the allowable frequency of exceedence, and the allowable time interval during which the standard can be exceeded (Novotny, 1997). Examples include permissible frequency and duration of toxicity criteria adopted by the U.S. EPA. The degree of compliance is typically related to the nature of such requirements, which generally range from mandatory standards

to recommended guidelines. This problem may be addressed by fitting a probability distribution to the observed data. For example, probability distributions of observed microorganism densities were used to assess the levels of fecal bacteria pollution from urban runoff in the St. Clair River, Sarnia, Ontario to make inferences about the compliance with recreational water quality guidelines (Marsalek et al., 1994).

Third, the assessment of the performance of runoff control measures, including best management practices (BMPs), by field measurements may be accomplished by comparing the distributions of concentrations of upstream versus downstream locations, inflow versus outflow, and wet weather versus dry weather. For example, the load distributions of inflow and outflow concentrations may be developed to assess the performance of detention basins and wet lands.

Finally, most of the priority pollutants in runoff occur with highly variable concentrations, some of which may fall below detection limits of analytical methods. In order to deal with such censored data, it is advantageous to approximate the EMCs by a probability distribution which can be utilized to estimate the sample statistics and to treat the data below detection limit.

14.6.1 Methodology for EMC Frequency Analysis

This section describes the procedure for fitting and selecting an appropriate probability distribution for a runoff quality constituent concentration. Among the probability distributions used for stormwater quality parameters, the two-parameter lognormal distribution is very commonly assumed (e.g. Driscoll 1986; US EPA, 1983). There are two reasons for this assumption. First, the realization of a random variable, X (i.e. EMC of a quality constituent) is a result of the joint action of many causative meteorological and geographical factors, which may be expressed as a product of a number of independent variables. For example, Chow (1957) used the Central Limit Theorem to show that the logarithm of X is normally distributed when the number of causative factors is infinitely large. Secondly, the probability distribution of environmental concentrations are often skewed to the right, because low values are common and high values are rare, as represented by a lognormal distribution (Parkhurst, 1998).

Hall et al. (1990) reported that the sediment associated pollutants from Danish and French catchments are represented by mixtures of two normal distributions. More recently, Van Buran et al. (1997) suggested that the lognormal distribution may not be appropriate for all stormwater constituent concentrations and in-line storage (e.g. detention ponds) may modify the form of the distribution. Although past studies have frequently suggested the lognormal distribution for stormwater quality parameters, the Toronto EMC data sets are utilized to explore the suitability of other alternative distributions.

The concentration of pollutants in storm runoff and overflows varies within a storm and between storms at a site. The intra-storm variation of concentrations is replaced with the EMC of the quality parameter, and the frequency analysis is performed on the resulting time series of EMCs from a number of storms as illustrated in Figure 14.1.

Several conditions must be met before performing a frequency analysis on the EMC data series of various pollutants (Wilson, 1988). The general conditions are:

1. The EMCs are randomly drawn from some underlying population;
2. The EMCs are assumed to be homogeneous (i.e. the EMCs are drawn from the same population);
3. The EMCs are independent of each other;
4. The involved processes which generate the pollutant mass are stationary;
5. The size of sample is large enough to warrant the estimation of parameters of the underlying probability distribution with the required reliability;
6. The estimates of population parameters are unbiased.

Although it was not possible to statistically verify all of the above conditions with the data set analyzed, it was assumed that they were reasonably met.

In the frequency analyses of EMC data, the following steps were used to obtain the appropriate probability distributions:

1. Compute the empirical cumulative distribution function (CDF) of the observed sample.
2. Hypothesize the underlying PDF of the random variable.
3. Select a goodness-of-fit test and the significance level of test statistic.
4. Estimate the sample moments and construct the hypothesized CDF.
5. Assess the goodness-of-fit between the empirical and hypothesized CDFs by the selected goodness-of-fit test.
6. Repeat steps 2-5 for all assumed hypothesized distributions.
7. Select the hypothesized distribution which has the best fit to the empirical distribution. A best-fit criterion may be selected on the basis of the minimum deviation between the empirical and hypothesized distributions in accordance with the test statistic of the goodness-of-fit test.

The above procedure was applied to EMC data obtained from one year of sampling at two representative combined and separated catchments located in the Metropolitan Toronto region. Three probability density functions (exponential, gamma and lognormal) were fitted to the observed EMCs of 15 runoff quality parameters.

Figure 14.1 EMC time series.

Statistical goodness-of-fit tests are used to test the hypothesis that a sample is drawn from a specified population distribution. These tests include the Chi-Square, Kolmogorov-Smirnov (K-S) and Cramer-von Mises tests. In these tests, the test statistic is an index of agreement between an observed sample distribution and a hypothesized population distribution. The K-S test is often preferred because it is strictly valid for continuous probability distributions and also advantageous when the model hypothesized is wholly independent of the data (Benjamin and Cornell, 1970). In addition, the K-S test compares the data in an unaltered form and evaluates the deviations between the empirical and the hypothesized CDFs.

14.6.2 Results from the EMC Frequency Analysis

The above procedure was implemented in a spreadsheet program and applied to the Toronto data. K-S test statistics for all hypothesized distributions were computed and, at the significance level of 5%, the accepted PDFs for each pollutant were identified. Tables 14.3 and 14.4 present the results of the computations for the separated and combined sewer catchments, respectively. It is seen from Table 14.3 that for suspended solids, all three distributions are statistically accepted since the K-S test statistic is larger than the maximum deviation between empirical and fitted CDF for all distributions. In the case of COD, only the gamma and lognormal are statistically acceptable distributions because the maximum deviation between the empirical and fitted exponential is

Table 14.3 Frequency analyses of pollutant EMCs for a separated storm sewer catchment.

Pollutants	Number of Events	K-S Test Statistic C_α ($\alpha=5\%$)	Maximum Deviation, D, between Empirical Distribution and Fitted Distribution		
			Exponential Distribution	Gamma Distribution	Lognormal Distribution
Suspended Solids(mg/L)	40	0.210	<u>0.093</u>	0.108	0.128
Total Solids (mg/L)	42	0.210	0.222	<u>0.175</u>	0.480
E. Coli (counts/100ml)	42	0.210	0.564	0.464	<u>0.193</u>
Fecal Coliforms (counts/100ml)	30	0.240	0.522	0.315	<u>0.193</u>
Fecal Streptococci (counts/100ml)	32	0.234	0.462	0.442	<u>0.155</u>
Nitrates (mg/L)	26	0.256	<u>0.289</u>	0.294	0.292
Total Kjeldahl Nitrogen (mg/L)	42	0.210	0.128	<u>0.083</u>	0.138
Total Phosphorus (mg/L)	33	0.231	0.190	<u>0.099</u>	0.119
Chemical Oxygen Demand (mg/L)	39	0.210	0.271	<u>0.102</u>	0.112
Phenol (mg/L)	27	0.252	0.146	0.137	<u>0.112</u>
Aluminum (mg/L)	43	0.207	<u>0.094</u>	0.133	0.112
Barium (mg/L)	44	0.205	0.332	<u>0.143</u>	0.151
Copper (mg/L)	43	0.207	0.220	<u>0.130</u>	0.210
Iron (mg/L)	45	0.203	0.105	0.098	<u>0.082</u>
Zinc (mg/L)	43	0.207	0.164	0.153	<u>0.150</u>

larger than the K-S test statistics. The lognormal distribution seems to best represent the bacteriological quality parameters. From these analyses, it is observed that the gamma and lognormal distributions are the most accepted distributions for most of the pollutants. Similar results are obtained for the combined sewer catchments (Table 14.4).

Figure 14.2 illustrates the distributions fitted to some of the pollutants in separated and combined sewer catchments. For example, suspended solids in the separated system can be acceptably modeled by all three probability density functions as illustrated in Figure 14.2(a). The empirical CDF of COD in the combined sewer system is closely fitted with all three hypothesized CDFs [Figure 14.2(c)] while total phosphorus in the separated system is statistically fitted with lognormal and gamma distributions [Figure 14.2(d)].

From the analysis of the Toronto data, it was observed that several distributions might provide statistically acceptable fits for many of the pollutants in both types of catchments. Although it is difficult to identify the “true” or “best”

Table 14.4 Frequency analyses of pollutant EMCs for a combined sewer catchment.

Pollutants	Number of Events	K-S Test Statistic C_α ($\alpha=5\%$)	Maximum Deviation, D, between Empirical Distribution and Fitted Distribution		
			Exponential Distribution	Gamma Distribution	Lognormal Distribution
Suspended Solids (mg/L)	45	0.203	<u>0.191</u>	0.194	0.203
Total Solids (mg/L)	45	0.203	0.284	<u>0.10</u>	0.112
E. Coli (counts/100ml)	34	0.228	0.276	<u>0.215</u>	0.222
Fecal Coliforms (counts/100ml)	36	0.222	0.284	<u>0.194</u>	0.222
Fecal Streptococci (counts/100ml)	33	0.231	0.223	<u>0.152</u>	0.171
Nitrates (mg/L)	18	0.310	0.246	<u>0.239</u>	0.240
Total Kjeldahl Nitrogen (mg/L)	41	0.212	0.179	<u>0.172</u>	0.398
Total Phosphorus (mg/L)	34	0.228	0.233	<u>0.075</u>	0.124
Chemical Oxygen Demand (mg/L)	43	0.207	0.180	<u>0.099</u>	0.154
Phenol (mg/L)	33	0.230	0.215	<u>0.070</u>	0.095
Aluminum (mg/L)	43	0.207	<u>0.127</u>	0.157	0.220
Barium (mg/L)	44	0.205	0.422	<u>0.153</u>	0.156
Copper (mg/L)	42	0.210	0.128	<u>0.103</u>	0.163
Iron (mg/L)	45	0.203	<u>0.163</u>	0.177	0.168
Zinc (mg/L)	42	0.210	0.171	<u>0.161</u>	0.205

distribution in the case of a specific water quality parameter, the theoretically “best” distributions on the basis of the K-S test statistic are underlined in Table 4 and Table 5 for separated and combined catchments respectively. It is found that none of the pollutant from the combined sewer catchment is best fitted by lognormal distribution. The results of the Toronto EMC data analyses suggest that the gamma and exponential probability distributions may be used for runoff quality parameters in addition to lognormal distribution. The EMCs of suspended solids, as an indicator pollutant, for both types catchments are best fitted by the exponential distribution.

Frequency Analysis of Regional EMC Data

Frequency analysis of EMC data is often problematic in urban runoff quality analysis because sufficient information is seldom available at a site to adequately determine the frequency of extreme events. For a large urban watershed, runoff

Figure 14.2(a) Fitted CDFs of suspended solids in a separated sewer system.

Figure 14.2(b) Fitted CDFs of suspended solids in a combined sewer system.

quality data is typically collected at few representative sites and the information is extrapolated to unmeasured sites. Given that sufficient data will seldom be available at the catchment of interest, it is becomes necessary to use the data from nearby similar catchments. Akin to regional flood frequency analyses, a regional EMC frequency analyses may be performed for a large urban watershed. Such analyses were performed for the Metropolitan Toronto region by aggregating the data from approximately similar land use catchments for each type of sewer system. The PDFs were identified for the 15-runoff quality parameters applying

Figure 14.2(c) Fitted CDFs of COD in a combined sewer system.

Figure 14.2(d) Fitted CDFs of total phosphorus in a separated sewer system.

the procedures described in the previous section. The results of regional frequency analyses of separated and combined sewer catchments are presented in Table 14.5 and Table 14.6. The best distribution based on K-S test statistic is underlined for each of the pollutant for both types of catchments. It is found again with these data sets that the gamma and exponential distributions are appropriate, in addition to the lognormal distribution, in representing runoff quality constituents. However, the exponential probability distribution is the best fitting distribution for

Table 14.5 Frequency analyses of all separated sewer catchments of similar land uses.

Pollutants	Number of Events	K-S Test Statistic C_α ($\alpha=5\%$)	Maximum Deviation between Empirical Distribution and Fitted Distribution		
			Exponential Distribution	Gamma Distribution	Log-Normal Distribution
Suspended Solids(mg/L)	83	0.149	<u>0.075</u>	0.083	<u>0.075</u>
Total Solids (mg/L)	86	0.147	0.109	<u>0.081</u>	0.093
E. Coli (counts/100mL)	55	0.183	0.502	0.468	<u>0.138</u>
Fecal Coliforms (counts/100mL)	55	0.183	0.434	0.271	<u>0.137</u>
Fecal Streptococcus (counts/100mL)	54	0.185	0.402	0.416	<u>0.134</u>
Nitrates (mg/L)	52	0.189	0.190	0.407	<u>0.119</u>
Total Kjeldahl Nitrogen (mg/L)	89	0.144	0.184	0.295	<u>0.108</u>
Total Phosphorus (mg/L)	78	0.154	0.137	0.293	<u>0.081</u>
Chemical Oxygen Demand (mg/L)	82	0.150	0.250	<u>0.058</u>	<u>0.058</u>
Phenol (mg/L)	72	0.160	0.096	<u>0.078</u>	0.109
Aluminum (mg/L)	94	0.140	0.200	0.294	<u>0.096</u>
Barium (mg/L)	95	0.140	0.206	0.110	<u>0.086</u>
Copper (mg/L)	92	0.142	0.152	<u>0.092</u>	0.131
Iron (mg/L)	93	0.141	0.091	0.075	<u>0.057</u>
Zinc (mg/L)	94	0.140	0.202	<u>0.109</u>	0.122

suspended solids for separated and combined sewer catchments in the case of the Metropolitan Toronto data. The empirical and hypothesized CDFs for suspended solids are presented for lumped separated and combined sewer catchments in Figure 14.3(a) and 14.3(b) respectively.

14.7 Regression Analysis for Concentration Estimation

As an alternative to site-specific concentration data, statistically based regression equations have been developed from runoff quality databases to estimate both EMCs and mean pollutant loads, (Hodge and Armstrong, 1993). Such regression equations may take either linear or non-linear forms. In these equations, the

Table 14.6 Frequency analyses of all combined sewer catchments of similar land uses.

Pollutants	Number of Events	K-S Test Statistic C_α ($\alpha=5\%$)	Maximum Deviation between Empirical Distribution and Fitted Distribution		
			Exponential Distribution	Gamma Distribution	Log-Normal Distribution
Suspended Solids(mg/L)	105	0.133	<u>0.078</u>	0.082	0.126
Total Solids (mg/L)	105	0.133	0.240	0.082	<u>0.050</u>
E. Coli (counts/100mL)	76	0.156	0.223	<u>0.108</u>	0.134
Fecal Coliforms (counts/100mL)	81	0.151	0.239	<u>0.075</u>	0.132
Fecal Streptococcus (counts/100mL)	81	0.151	0.164	<u>0.075</u>	0.151
Nitrates (mg/L)	42	0.210	0.125	0.145	<u>0.102</u>
Total Kjeldahl Nitrogen (mg/L)	103	0.134	0.161	<u>0.061</u>	0.131
Total Phosphorus (mg/L)	96	0.139	0.133	<u>0.057</u>	0.091
Chemical Oxygen Demand (mg/L)	106	0.132	<u>0.072</u>	0.101	0.148
Phenol (mg/L)	88	0.145	0.206	0.083	<u>0.064</u>
Aluminum (mg/L)	110	0.130	0.134	<u>0.075</u>	0.146
Barium (mg/L)	111	0.129	0.267	<u>0.056</u>	0.070
Copper (mg/L)	108	0.131	0.130	<u>0.062</u>	0.084
Iron (mg/L)	112	0.129	0.091	<u>0.084</u>	0.124
Zinc (mg/L)	110	0.130	0.181	<u>0.057</u>	0.115

dependent variable, either load or concentration of a pollutant is functionally related to one or more independent variables. The independent variables may include storm event characteristics, hydrologic characteristics, land use, percent impervious area, street cleaning practices, population density, particulate fallout rate, drainage system characteristics (number of catch basins / hectare). Regression analyses were also performed on the Metropolitan Toronto runoff quality data hypothesizing that the quality of stormwater runoff from a catchment varies with the type of land uses. The average EMC of a pollutant was considered as the dependent variable and four land use categories (%Residential, %ICI, and %Road and %Open area) were used as independent variables. The low, medium and high density residential areas of a catchment were aggregated to represent the %residential area. Similarly, industrial, commercial and institutional (ICI) areas were aggregated to represent %ICI area. Approximately similar land use catchments were entered into the regression analyses and a step-wise linear

Figure 14.3(a) CDFs of suspended solids for the regional separated sewer system (lumped data).

Figure 14.3(b) CDFs of suspended solids for the regional combined sewer system (lumped data).

regression procedure was employed. Relationships derived from the regression analyses could not be accepted with a high reliability. The reliability of such analyses may be improved by increasing the number of catchments in the analyses; however, these results are perhaps not surprising because of the larger number of other confounding factors such as age and condition of roads and sewer systems, atmospheric fallout rates of pollutants, wind effects and so on.

14.8 Conclusions

Long-term characterizations of stormwater quality parameters are outlined in a statistical framework. These analyses are necessary to support planning and management decisions related to the remediation of stormwater quality problems in urban areas. Metropolitan Toronto stormwater quality databases were employed to characterize the runoff quality from separated and combined sewer catchments. The EMCs of fifteen quality constituents including chemical, bacteriological, nutrients and heavy metals are analyzed. It was observed that the average EMCs of various pollutants in combined sewer catchments were generally higher than those in separated sewer systems. However, the variability of EMCs was generally higher in separated sewer systems than combined systems. The correlation analyses between the quality constituents indicated that suspended solids were positively correlated with nutrients and heavy metals and that nutrients were positively correlated with heavy metals. Furthermore, correlation existed between the pollutants of bacteriological and heavy metal categories. The frequency analyses of EMCs of both combined and separated systems were conducted. The choice of the appropriate probability distribution should be made from the information provided by the observed data in performing the frequency analysis and from appropriate goodness-of-fits tests. From the analyses of Metropolitan Toronto stormwater quality data it was concluded that gamma and exponential probability distributions, in addition to the lognormal distribution, can describe runoff quality constituents. It was observed that the suspended solids concentrations for both separated and combined sewer catchments were best represented by the exponential probability distribution.

Acknowledgments

The authors gratefully acknowledge the financial support of this work provided by the Canadian Commonwealth Scholarship Program, International Council of Canadian Studies, Canada and the Natural Sciences and Engineering Research Council of Canada.

References

- Adams, B.J., and Bontje, J. B. 1984. "Microcomputer Application of Analytical Models for Urban Stormwater Management," *Emerging Computer Techniques in Stormwater and Flood Management*, W. James, (Ed.), ASCE, New York, 1984. pp 138-162.
- Adams, B.J., and Zukovs, G. 1987. "Probabilistic Models for Combined Sewer Rehabilitation Analysis," in *System Analyses in Water Quality Management*, M.B.

- Beck, (Ed.), Pergamon Press, Oxford, England. pp. 97-108.
- Aquafor Engineering Ltd. 1995. "Metropolitan Toronto Waterfront Wet Weather Outfall Study- Phase II," Report to Ontario Ministry of the Environment, Toronto, Ontario.
- Benjamin, J.R., and Cornell, C.A. 1970. "Probability, Statistics and Decision for Civil Engineers," McGraw-Hill, New York, NY.
- Chow, V. T. 1954. "The Log-probability Law and its Engineering Applications," Proceedings, ASCE, Vol. 80. pp. 1-25.
- Driscoll, E.D. 1986. "Lognormality of Point and Non-Point Source Pollutant Concentrations", Proceedings, Urban Runoff Quality - Impact and Quality Enhancement Technology, B. Urbonas and L.A. Roesner (Eds.), ASCE, New York, NY. pp. 438-458.
- Ellis, J.B. 1986. "Pollutional Aspects of Urban Runoff,": Proceedings, Urban Runoff Pollution, H. C. Torno et al. (Eds.), Springer-Verlag Berlin Heidelberg, New York. pp.1-38.
- Guo, Y. and B.J. Adams. 1998. "The Analytical Stormwater Management Model (ASWMM)." *Journal of Water Management Modeling* R204-19. doi: 10.14796/JWMM.R204-19.
- Hall, M.J., Ellis, J.B., Brizio, M.G. 1990. "On the Statistical Distribution of Event Mean Concentrations of Pollutants in Stormwater Runoff," Proceedings, 5th International Conference of Urban Storm Drainage, Y. Iwasa and T. Sueishe (Eds.), Osaka, Japan, Vol. I. pp.317-324.
- Harremoes, P. 1988. "Stochastic Models for Estimation of Extreme Pollution from Urban Runoff," *Water Research*, Vol. 22, No. 8. pp. 1017-1026.
- Heaney, J.P., Wright, L., and Sample, D. 1998. "Wet Weather Flow Management- A Research Needs Survey for Urban Areas", Project 96-IRM-1, Water Environmental Research Foundation, Alexandria, VA.
- Hodge, T.A. and L.J. Armstrong. 1993. "Use of a Multiple Linear Regression Model to Estimate Stormwater Pollutant Loading." *Journal of Water Management Modeling* R175-09. doi: 10.14796/JWMM.R175-09.
- Huber, W.C. 1993. "Contaminant Transport in Surface Water," in *Handbook of Hydrology*, D. R. Maidment (Ed.), McGraw-Hill, New York, NY.
- Huber, W.C., and Dickinson, R.E.: 1988. "Stormwater Management Model, Version 4: User's Manual", Environmental Research Laboratory, Office of Research and Development, U.S. EPA, Athens, Georgia, August.
- Jennings, M.E., and Miller, T.L. 1986. "Urban Stormwater Quality Investigations by the USGS", Proceedings, Urban Runoff Quality - Impact and Quality Enhancement Technology, B. Urbonas and L.A. Roesner (Eds.), ASCE, New York, NY. pp. 29-33.
- Li, J.Y. 1991. "Comprehensive Urban Runoff Control Planning," Ph.D. Thesis, Department of Civil Engineering, University of Toronto, Toronto, Ontario.
- Masalek, J., and Kok, S. 1997. "Stormwater Management and Abatement of Combined Sewer Overflow Pollution", *Water Quality Research Journal Canada*, Vol. 32, No. 1, pp. 1-5.
- Masalek, J., Dutka, B.J., and Tsanis, I.K.: 1994. "Urban Impacts on Microbiological Pollution of the St. Clair River in Sarnia, Ontario," *Water Science and Technology*, Vol. 30, No.1. pp. 177-184.
- MOE, Ontario Ministry of Environment, 1984. "Humber and Toronto Area Water

- Quality," Technical Report No. 4, Toronto, Ontario.
- MOE, Ontario Ministry of Environment, 1988. "MOE Analytical Methods", Toronto, Ontario.
- Novotny, V., and Olem, H. 1994 "Water Quality Prevention, Identification, and Management of Diffuse Pollution," Van Nostrand Reinhold, New York, NY.
- Novotny, V. 1997. "Integrated Watershed Management and Diffuse Pollution Control," Proceedings, Workshop on Research Needs for Coastal Pollution in Urban Areas, E. R. Christensen and C. R.O, Melia (Eds.), Milwaukee, WI. pp. 130-146.
- Paul Theil Associates Limited and Beak Consultants Limited, 1992. "Metropolitan Toronto Waterfront Wet Weather Outfall Study- Phase I," Report prepared for the Ontario Ministry of the Environment, Toronto, Ontario.
- Parkhurst, D.F. 1998. "Arithmetic versus Geometric Means for Environmental Concentration Data," *Environmental Science and Technology*, Vol. 32, No. 3, February. pp. 92-98.
- U.S. Environmental Protection Agency (EPA), 1983. "Results of the Nationwide Urban Runoff Program, Vol. I, Final Report, Water Planning Division, US EPA, NTIS No. PB84-185552, Washington, D.C.
- Van Buren, M.A., Watt, W.E., and Marsalek, J. 1997. "Application of the Log-Normal and Normal Distributions to Stormwater Quality Parameters", *Water Research, IAWQ*, Vol. 31, No.1. pp. 95-104.
- Weatherbe, D. G. and Sherbin, I.G. 1994. "Urban Drainage Control Demonstration Program of Canada's Great Lakes Cleanup Fund", *Water Science Technology*, Vol. 29, No. 1-2, pp. 455-462.
- Wilson, B.G. 1988. "A Critical Review of the Method of Moments in Hydrology," M.A.Sc. Thesis, Department of Civil Engineering, University of Toronto, Ontario, Canada.