UNIVERSITY OF CALIFORNIA

Los Angeles

Predictive Modeling of Stormwater Runoff Quantity and Quality for a Large Urban Watershed

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

by

Simon Joonho Ha

2006

The dissertation of Simon Joonho Ha is approved.

Keith D. Stolzenbach

Eric M.V. Hoek

Richard F. Ambrose

Michael K. Stenstrom, Committee Chair

University of California, Los Angeles

2006

To my Father

To my Mother

To Terry and Tammy

To Vivianna

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

VITA

June 13, 1959	Born, Gwangju, Republic of Korea
1983	B.A., Chemical Engineering Chonnam National University Gwangju, Republic of Korea
1985	M.S., Chemical Engineering Chonnam National University Gwangju, Republic of Korea
2001	M.S., Computer Science California State University, Northridge Los Angeles, California
2002	M.S., Civil Engineering University of California, Los Angeles Los Angeles, California
2002-06	Graduate Student Researcher Department of Civil Engineering University of California, Los Angeles
2001	Teaching Assistant Department of Physics and Astronomy University of California, Los Angeles

PUBLICATIONS AND PRESENTATIONS

M. Kayhanian, S. Ha, M. K. Stenstrom (2005), Constituents' Annual Load Estimation from Highways, *International Conference on Urban Drainage*, August, Copenhagen/Denmark.

J. Ma, S. Khan, Y. Li, L. Kim, S. Ha, S. Lau, M. Kayhanian and M. K. Stenstrom (2002), First Flush Phenomena for Highways: How it can be meaningfully defined, *Proceedings of 9thInternational Conference on Urban Drainage,* Portland, Oregon.

R. J. Lorentz and S. Ha (1999), Data Structures for $2 \times N$ Go, *Game Programming Workshop in Japan,* October, Hakone, Kanagawa, Japan

ABSTRACT OF THE DISSERTATION

Predictive Modeling of Stormwater Runoff Quantity and Quality
for a Large Urban Watershed

By

Simon Joonho Ha
Doctor of Philosophy in Civil Engineering
University of California, Los Angeles, 2006
Professor Michael K. Stenstrom, Chair

Urban and agricultural areas have been recognized by the US EPA as major national problems due to their highly polluted stormwater runoff. Remediation of this runoff has not occurred in part because assessment of nonpoint source pollution is inherently complex and expensive. One approach to deal with this complex nonpoint source pollution problem is to improve understanding through modeling and information management using Geographic Information Systems (GIS).

In this research a predictive model for stormwater runoff volume was implemented in an ArcGIS platform based on the Rational Method and Browne's empirical relation for soil

characteristics. The heterogeneity of the watershed was quantified by dividing the watershed into many small sub-areas and applying lumped parameters for each of them. Characterization of pollutant load contributions of landuse types to total loads of the upper Ballona Creek watershed was achieved through zeroth-order regularization and L-BFGS-B optimization techniques. Relative form was used in the objective function to compensate for strong contributions of high magnitude variables. Model predictions showed reasonable agreement with total Zn, TKN, and TSS loadings measured at the mass emission site for the upper Ballona Creek watershed. Two additional categories, highways and local roads, which have not been routinely used as landuse categories, were separately studied.

Best Management Practices (BMP) strategies were evaluated a typical storm event, which exceeded total zinc TMDL by over 70%. The model was used to compare optimized BMP applications to the simplest application, which would treat all areas equally. Approximately 44 % removal efficiency with treatment of the entire runoff would be needed to meet the TMDL.

To show how the model can be used to improve BMP application, two subwatersheds with high leverages (4.3) were identified and Austin sand filters were simulated. By assuming typical total Zn removal efficiency of 45 %, total Zn removal at the mouth of the watershed was 12.4 % with the treatment of 5 % of the upper Ballona Creek watershed area. If the same leverage and runoff per area are assumed for additional areas in the upper Ballona Creek watershed, about 20 % of the watershed area would need to be treated to meet the TMDL. This would require about $240 million if one assumes Caltrans' predicted construction cost per unit volume of runoff.

Modified infiltration trenches were simulated for local roads and the TMDL could be met with treatment of 68 % of local roads area, which is equivalent to 7 % of the upper Ballona Creek watershed area. The construction cost was about $58 million, which is less than one-fourth of the cost needed for the subwatershed approach.

1. INTRODUCTION

Over the past ten years, stormwater management has become the major priority for water pollution control. This has occurred in part because of the completion of wastewater treatment plants and in part because of the recognition of the pollutant contributions from stormwater. Stormwater management is intrinsically more challenging than wastewater treatment because of the distributed nature of stormwater and because of the "ownership" of stormwater problems is not clear. Various approaches are being taken to mitigate stormwater impacts and these include individual actions, such as the industrial stormwater management permits as well as agency or city-wide actions, such as Los Angeles' recent passage of Proposition O, which will provide $500,000,000 for stormwater management.

In 1987, Congress amended Clean Water Act (CWA) to require implementation of a comprehensive national program for addressing stormwater discharge. The program required National Pollutant Discharge Elimination System (NPDES) permits for stormwater discharge. More recently, Total Maximum Daily Loads (TMDLs) are being developed. The concept of the TMDL is to obtain the maximum allowable or permissible discharge rate by optimally minimizing all sources. In this way, a more economical design can be obtained by reducing the easily controlled sources and minimizing controls for small sources or sources that cannot be effectively treated. One approach to deal with this complex Non-Point Source (NPS) pollution problem is through modeling.

Among the many urban stormwater pollution modeling program, the United States Environmental Protection Agency (US EPA) Storm Water Management Model (SWMM) continues to be widely used throughout the world since its first development in 1971. This model tries to simulate NPS pollutions problem as detailed as possible. However, this approach requires that many parameters be measured. These parameters are influenced by many temporal and aerial factors that are related to the heterogeneous characteristics of the subcatchments. The measurements required by this type of model are costly and time consuming, especially for a large watershed. If there is insufficient measured data available, inaccurate predictions can be expected (Donigian and Huber, 1991).

The other approach to this inherently complex problem is to adopt a Rational Method, where all runoff is assumed to have a constant concentration for a given pollutant. The changes in pollutant concentration during a storm event have small impact on most receiving waters due to their relative insensitivity to rapid changes. The insensitivity may be due in part to the large dilution that can occur in a receiving water or to the dispersed load that stormwater runoff creates. This insensitivity makes the assumption of time invariant concentrations a viable option for receiving water quality analysis. This method has been used to estimate loads from the Santa Monica watershed (Wong et al. 1997), larger areas in Southern California (Ackerman and Schiff, 2003; McPherson et al., 2005), and many other areas (Wu et al., 1998; Shinya et al., 2000; Choe et al., 2002).

In order to better understand stormwater management and to create a tool for TMDL development, a Geographic Information System (GIS)-based model was developed. A GIS was used extensively to incorporate inherent heterogeneity of urban watershed surfaces and climatic conditions. The model contains several layers which characterize landuse, rainfall, slope and elevation, and hydrologic soil group. An empirical runoff model, based largely on the rational method, is used to estimate pollutant loads. The model will be used to estimate pollutant loads and compare contributions from different landuse types. In this way, costly and time consuming field surveys can be avoided.

The model was optimized with existing monitoring data using regression analysis and optimization techniques. After optimization, the model was used to evaluate the effectiveness of potential Best Management Practices (BMP), as well as explore the feasibility of proposed TMDLs.

The objectives of this research are (1) to develop a model to predict stormwater runoff quantities taking into account the spatial variability of watersheds, (2) to develop a methodology for predicting stormwater runoff qualities, and (3) to show that there exist Best Management Practices (BMP) strategies using the developed models.

2. LITERATURE REVIEW

2.1 MODELING TECHNIQUES

In the 1970's, point source pollution was considered the most significant problem resulting in water quality degradation in receiving water bodies. The 1972 CWA established extensive programs to address point source pollution, and the majority of federal investments went to control point source pollution. By the late 1980s, programs focusing on treatment facilities resulted in better controls of point source pollution. However, water quality was still impacted because of the growing impacts of nonpoint source pollution.

Urban and agricultural areas have been recognized by the US EPA as major national problems due to their highly polluted runoff (Browne, 1990), which results in impaired receiving waters. The inventory of the 1996 Report to Congress indicated that approximately 40 percent of the Nation's assessed rivers, lakes, and estuaries were impaired (US EPA, 1998). Urban runoff or storm sewers were found to be a source of pollution in 13 percent of impaired rivers, 21 percent of impaired lakes, ponds, and reservoirs, and 45 percent of impaired estuaries. In 1987, Congress amended the CWA to require implementation of a comprehensive national program for addressing storm water discharges. Phase I of the program required NPDES permits for stormwater discharge from large municipal separate storm sewer systems generally serving populations of 100,000 or more. The Phase II Final Rule covered all small municipal separate storm sewer systems located within an urbanized area (US EPA, 2000b). An urbanized area is land area that has a residential population of at least 50,000 and an overall population density of at least 1,000 people per square mile.

The large number of NPDES permits required by the CWA amendments creates challenges for regulators. Assessment of nonpoint source pollution is inherently complex and expensive and the existence of many permittees makes the problem even more difficult. One approach to deal with this complex nonpoint source pollution problem is through modeling and information management using Geographic Information Systems (GIS).

2.1.1 Rational Method

The Ration Method is perhaps the oldest method for estimating rainfall runoff. All runoff is assumed to have a constant concentration for a given pollutant when using the Rational Method, which is also known as a unit load. The changes in pollutant concentration during a storm event have small impact on most receiving waters due to their relative insensitivity to rapid changes, as illustrated in Table 2.1. The insensitivity may be due in part to the large dilution that can occur in a receiving water or to the dispersed load that stormwater runoff provides. This insensitivity makes the assumption of time invariant concentrations a viable option for receiving water quality analysis.

Table 2.1 Required Temporal Variations for Receiving Water Analysis (Donigian and Huber, 1991)

Type of Receiving Water	Key Pollutants	Response Time
Lakes, Bays	Nutrients	Weeks - Years
Estuaries	Nutrients, OD[*], Bacteria	Days - Weeks
Large Rivers	OD, Nitrogen	Days
Streams	OD, Nitrogen	Hours - Days
	Bacteria	Hours
Ponds	OD, Nutrients	Hours - Weeks
Beaches	Bacteria	Hours

[*] OD = Oxygen demand, e.g., BOD, that affects dissolved oxygen.

Unit loading rates are represented by mass per area per time. For example, the Universal Soil Loss Equation (Wischmeier and Smith, 1978) was developed to estimate sediment loss in tons per acre per year from land surfaces. More recently, this method has been used to estimate loads from the Santa Monica watershed (Stenstrom et al., 1998), larger areas in Southern California (Ackerman and Schiff, 2003; McPherson et al., 2005), and many other areas (Wu et al., 1998; Shinya et al., 2000; Choe et al., 2002).

Unit loading rates are extremely variable and difficult to apply over shorter periods of time, and the approach is best suited to estimation of long term loads, such as seasonal or annual loads. Simple prediction methods generally perform better over a long averaging time and poorly at the level of a single storm event.

2.1.2 US EPA Statistical Method

The US EPA Statistical Method assumes that event mean concentrations (EMCs) are not constant but tend to exhibit a lognormal frequency distribution at a site and across a selection of sites. The concentrations are characterized by their median value and coefficient of variation (CV). CV is defined as a standard deviation divided by a mean value. This method is useful only if the measured dataset follows a lognormal distribution. For example, the US EPA introduced the delta-lognormal distribution to incorporate non-detected measurements since the lognormal distribution is not defined at zero. As shown in Figure 2.1, non-detected measurements are represented separately at zero. The delta in the name refers to the percentage of non-detects.

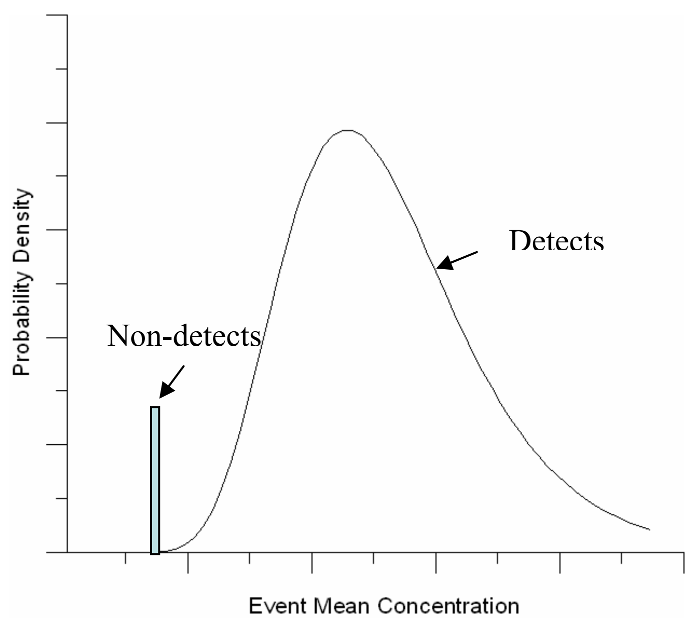


Figure 2.1 Delta-Lognormal Distributions.

Instead of using zero value for the non-detected measurements, the detection limit value is used in the adapted delta-lognormal distribution. Later, the US EPA introduced the modified adapted delta-lognormal distribution as shown in Figure 2.2 to account for non-detected measurements of the iron and steel industry. This distribution modeled the non-detected measurements as a discrete distribution made up of multiple concentrations (US EPA, 2000a). The pollutant EMC frequency distribution is useful for the analysis of the exceedence frequency of water quality standards.

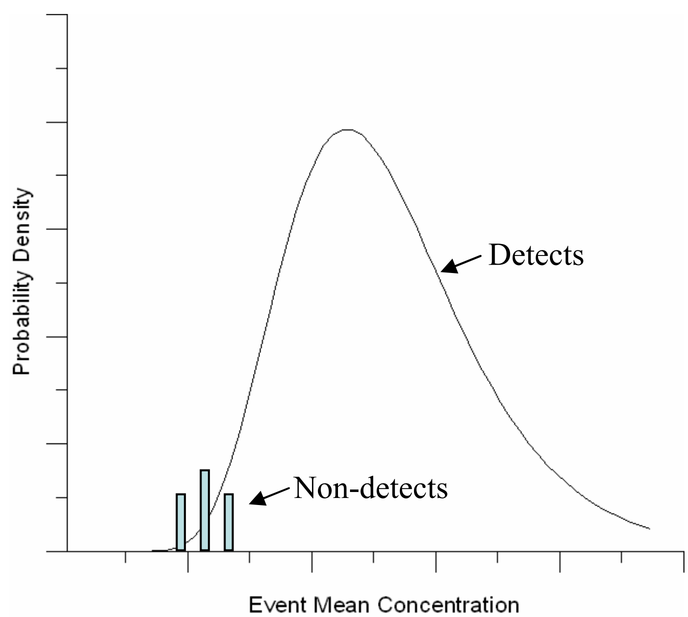


Figure 2.2 Modified Adapted Delta-Lognormal Distribution.

The US EPA statistical method based on lognormal distribution was used as the primary screening tool in the National Urban Runoff Program (NURP) studies (US EPA, 1983). EMCs were measured from 81 sites in 28 urban areas in the United States. They concluded that EMCs are essentially not correlated with stormwater runoff volumes and mass loadings are strongly influenced by the runoff volumes, which imply that constant concentration assumption is adequate for urban stormwater runoff studies. They recommended that the most reliable basis for characterizing seasonal and annual mass loads is on the basis of the EMC and site specific rainfall and runoff characteristics.

2.1.3 Regression Method

Regression, also called a rating curve approach, has often been used to relate total pollutant load and runoff volume of a storm event. The United States Geological Survey (USGS) assembled rainfall, runoff, water quality, and landuse characteristics databases for 98 urban stations in 30 metropolitan areas for multiple regression analysis (Driver and Tasker, 1988). Thirty four linear regression models for stormwater runoff volumes and loads, thirty one models for stormwater runoff mean concentrations, and ten models for mean seasonal or annual loads were developed by analyzing long term rainfall records. They concluded that total rainfall and total contributing drainage area were the most significant explanatory variables. Other significant variables in the models were impervious area, landuse, and mean annual climatic characteristics. Runoff volume models were more accurate than the load models. The average variance of prediction for estimating mean seasonal or annual loads ranged from -63 to 171 percent. Although these models incorporated vast amounts of data, the authors recommended that model limitations be considered when applying them.

Recently, Brezonik and Stadelmann (2002) assembled a database for 68 watersheds in the Twin Cities metropolitan area, which encompassed seven counties surrounding Minneapolis and Saint Paul, MN. EMCs were obtained for 562 events at 65 sites. Event loads were obtained for 360 events at 43 sites. They found only weak correlations between explanatory variables and stormwater loads and EMCs. The most important variables in their regression models to predict event loads were rainfall amount, intensity, and drainage area but uncertainty was high for both EMC and load models. Many researchers are critical of regression models and quotes such as "Regression approaches are notoriously difficult to apply beyond the original data set from which the relationships were derived" (Donigian and Huber, 1991) are common.

2.1.4 Buildup and Washoff Method

The buildup mechanism refers to the complex phenomenon that accumulates pollutants on land surfaces between storm events. This dry weather process includes deposition, suspension, and redeposition of pollutant particles. During a storm event, these accumulated pollutants are subject to washoff mechanisms. Models that include buildup and washoff mechanisms have to incorporate conceptual representations of these mechanisms because the mechanisms are not well known. Thus, a buildup and washoff model requires calibration with measured pollutographs.

For certain investigations, such as first flush phenomenon (Kim et al., 2003; Kim et al., 2005), a buildup and washoff model is preferable because variations of pollutant concentrations or loads during a storm event can be simulated. However, when a tributary watershed of receiving waters is large, pollutant concentrations of receiving waters are relatively insensitive to the variations of upstream pollutant concentrations during a storm event (Lee and Bang, 2000). In this case, the rational method is adequate and preferred for its simplicity.

This review so far has discussed only how to deal with pollutant concentrations for storm events. For the pollutant load calculations, all of the above methods needed to be coupled with stormwater runoff volume, which requires analysis of rainfall and contributing watershed area. Furthermore, these two variables are the most influential for the calculation of pollutant loads as mentioned in Section 2.1.3.

The US EPA SWMM incorporates effects of rainfall and contributing area as parameters. It was first developed in 1971 and continues to be widely used throughout the world. SWMM is a dynamic, non-linear, and lumped model for a single event or a long-term simulation for primarily urban areas. Major control parameters in the Runoff Block of SWMM are presented in Table 2.2 (Choi and Ball, 2002). Many temporal and aerial factors that are related to the heterogeneous characteristics of the subcatchments influence these control parameters. The authors also presented suggestions for influencing factors for these control parameters, which are shown in Table 2.3.

Table 2.2 SWMM Runoff Control Parameters (Choi and Ball, 2002)

Measured parameters	Inferred parameters
Subcatchment area	Impervious area factor
Length of channel/pipe	Subcatchment length/width ratio
Shape and bed slope of channel/pipe	Subcatchment slope
Characteristic dimension of conduit	Maximum and minimum infiltration
Manhole type	Impervious area Manning's roughness coeff.
Catchment soil type	Pervious area Manning's roughness coefficient
Catchment land-use type	Impervious area detention storage
Rainfall depth within last record period	Pervious area detention storage
	Percentage of imperviousness of subcatchment
	Conduit roughness coefficient
	Decay rate of infiltration curve

Table 2.3 Factors Influencing Control Parameter Values (Choi and Ball, 2002)

Control parameters	Influential factor	
Subcatchment width	Subcatchment area Topographic characteristics Over land flow length	
Subcatchment impervious fraction	Land use Imperviousness	
Depression storage	Impervious area	Interception Antecedent conditions
	Pervious area	Soil type Surface vegetation Antecedent conditions Interception Infiltration
Manning's roughness coefficient	Impervious area	Roof material type Road material type Other impervious surface types
	Pervious area	Vegetation Soil type Ground cover type
	Drainage system	Channel/pipe type
Infiltration (Horton's equation) [a]	f_c : Soil type f_0 : Soil type, initial moisture content k : Initial moisture content (Surface wetness)	
Slope	Elevation difference Overland flow length Slope model	

[a] f_0 : Maximum initial infiltration rate; f_c : Min. infiltration rate; k : Infiltration decay rate

Obviously, some type of averaging is needed to obtain the parameters and this averaging process requires the model to be calibrated against measured data. If there is insufficient measured data available, inaccurate predictions can be expected (Donigian and Huber, 1991).

To simplify the problem, the Rational Method was adopted in our research. As mentioned earlier, this approach is preferred when dealing with larger areas and over larger time periods. To incorporate inherent heterogeneity of urban watershed surfaces and climatic conditions, a Geographic Information System (GIS), which is discussed in the following section, was used extensively.

2.2 GEOGRAPHIC INFORMATION SYSTEM

A GIS is a database and a map that is made up of layers, which is a collection of geographic objects that have similar features. For example, a GIS world map might be made up of cities, rivers, countries, and oceans layers. One of the most interesting features of geographic information systems is the ability to overlay layers. As long as required layers for a problem are constructed, correlations between different features from different layers can be examined easily through overlaid layers. Without GIS layers, extracting hydrologic parameters from a database involves complex empirical or physics-based relation analysis (DeVantier and Feldman, 1993). However, compiling features into a layer can be labor intensive even with digitizing hardware and software. Geographic features can be represented in a GIS layer using vector, raster, or Triangular Irregular Network (TIN).

2.2.1 GIS Data Models

In a vector data model, each feature is represented as a set of points, lines, and polygons. Feature shapes are defined by (x, y) coordinates in space. Features can be discrete events at a location, lines, areas, or others. For example, EMCs of storm events at a sampling site are event features located at a location feature. A vector layer typically has an attribute table for topology, which contains attribute information of the features. It stores information as point, polyline, or polygon features with associated attribute tables. A vector data model is useful for representing and storing discrete features such as buildings, pipes or parcel boundaries. Statistical and relational database operators are typically well developed in vector systems, and attribute query is a common means of accomplishing analysis and report generation.

A raster model, also known as a raster dataset, is a storage intensive data system in which the topography of the data is represented as a matrix of cells in a continuous space. The raster is based on grid structure and each cell has a value, which represents the topography of a raster dataset, such as elevation. Digital Elevation Models (DEMs) are commonly represented with raster. Raster DEMs have been widely used to analyze hydrologic problems, such as delineation of subwatersheds and flow networks. However, there is an inherent problem working with a raster DEM when the dataset has sinks resulting from noisy data sources. A sink is defined as a cell that has lowest value compared to those of surrounding cells. If a raster DEM contains a sink, the resulting stream flow terminates at the sink. This erroneous termination can be remedied by creating a depressionless DEM. The most widely adopted DEMs in the United States are the 30×30 m Level 1 DEMs produced by the U.S. Geological Survey (USGS), which uses photogrammetric techniques. More accurate USGS Level 2 DEMs are available for selected geographic areas. Recently, Hodgson et al. (2003) demonstrated that a Light Detecting and Ranging (LIDAR) DEM model, which uses laser measurements, showed highest overall absolute elevation accuracy.

Topography is represented in a different way in a Triangular irregular network (TIN) model. Features are represented as a collection of irregularly spaced points and connecting lines that create triangles as shown in Figure 2.3. Surfaces can be modeled more accurately with less computer memory. Stream paths are identified through the slopes of triangles or edges between two triangles. Thus, more continuous stream paths can be identified and erroneously terminated stream flows at depression points in a raster DEM can be avoided with a TIN model.

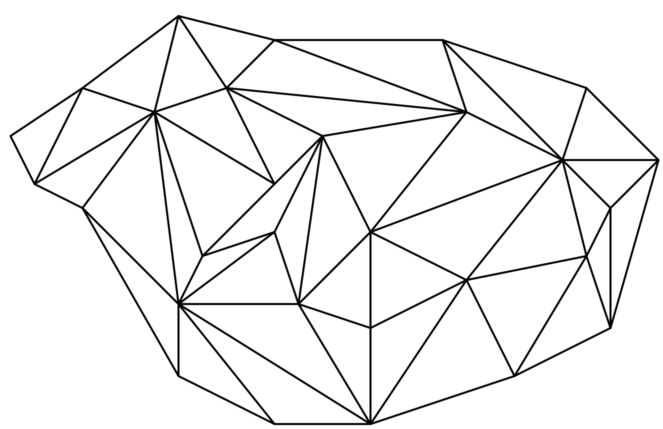


Figure 2.3 A Triangular Irregular Network.

2.2.2 Visual Basic for Application

Visual Basic for Application (VBA) is one of many object oriented programming languages, which is included with ArcGIS Desktop (Redlands, CA). The main difference between VBA and other object oriented programming languages is that VBA was designed to be embedded within applications facilitating the use of ArcObjects. ArcObjects are a set of program objects designed for programming with ArcGIS applications. ArcObjects include layers, tables, points, polylines, polygons, and other objects. The ArcObjects library, which consists of over 1,000 classes and 2,000 interfaces, is designed to provide researchers with the functionality to extend ArcGIS for their specific applications. The extent of the library is overwhelming and difficult to know for beginners. Fortunately, Exploring ArcObjects, published by Environmental Systems Research Institute (Zeiler, 2001), is a helpful reference for beginners as well as advanced users.

2.2.3 Stormwater Runoff Modeling using GIS

Stormwater runoff modeling uses many types of spatially distributed data, such as heterogeneity of watershed surfaces and climatic conditions, and recent researchers have used GIS to incorporate these data. Some researchers have used GIS for hydrologic analysis. Barco et al. (2006) coupled GIS and SWMM for a large urban watershed. They concluded that the time required for data management is dramatically reduced by incorporating GIS procedures. Djokic and Maidment (1993) constructed GIS network routines for steady state flow computations. This network describes channels and artificial links that have been added to represent overland flow paths. They concluded that the GIS network routines are useful for steady-state flow computations. Shea et al. (1993) implemented a Surface Water Management Plan in a GIS because of its ability to implement hydrologic and hydraulic modeling, manipulate spatial data, perform spatial operations, and support a relational database. They were successful in reviewing drainage plans for new developments, revising floodplains, supporting a county's landuse plan, and performing hydrologic analysis for bridge and road construction. Brun and Band (2000) coupled the Hydrological Simulation Program – FORTRAN (HSPF) and a GIS to assess the effects of landuse change on watershed behavior. They used the relationships among baseflow, percent impervious cover, and percent soil saturation. They found that the increase in impervious cover in the upper Gwynns Falls watershed in Baltimore, MA over the 17-year period did not significantly affect the runoff ratio.

For water quality analysis, a GIS has been proven to be very efficient for data preparation, model parameter extraction, and model results visualization (Reinelt et al., 1991; Seiker and Klein, 1998; Wang, 2005; Martin et al., 2005). A GIS is most commonly used to generate input data needed for a model, such as SWMM, HSPF, Better Assessment Science Integrating point and Nonpoint Sources (BASINS), and other models. As described in Section 2.1.4, these models require many parameters that represent heterogeneous watershed surfaces and climate conditions. Meyer et al. (1993) attempted to link a GIS and the Runoff block of SWMM. They concluded that GIS provided an improved assessment of the reliability of estimated parameters compared to traditional methods. Wong et al. (1997) developed landuse runoff model using a GIS coupled with an empirical runoff model. They stated that the integration of a nonpoint source model and a GIS offered a powerful tool to assist watershed managers in developing control strategies to improve water quality within local drainage systems.

HSPF is a comprehensive model of watershed hydrology and water quality that allows the integrated simulation of land and soil contaminant runoff processes. It is a lumped parameter model, but the spatial variability of a watershed is considered by dividing the watershed into subwatersheds and applying the lumped model to them. This lumped model is based on the Stanford watershed model. Rahman and Salbe (1995) applied HSPF to model Hawkesbury River catchments in Sydney and water quality in the associated stream reaches because it incorporates the wide range of significant processes involved. HSPF model was successfully calibrated and performed frequency duration analysis for output time series to determine the proportion of time that output values exceed any specified level.

BASINS (US EPA, 2001) was developed to facilitate examination of environmental information in a GIS framework. It is a multipurpose environmental analysis system that includes an interface to HSPF. A new interface to the HSPF is called WinHSPF (US EPA, 2001) in BASINS Version 3.0. This interface allows users to partition a watershed into land segments and river reaches, parameterize hydrologic model, and display results. Tong and Chen (2002) adopted BASINS to model the plausible effects of landuse on water quality in a local watershed in the East Fork Little Miami River Basin in southwestern Ohio. They concluded that the impacts of landuse on water quality could be modeled effectively using BASINS in smaller watersheds.

Shen et al. (2004) adopted a relational database to handle a large number of input data sets representing the environmental condition of a watershed and parameters to quantify underlying physical and biogeochemical processes. This relational database was used to link both GIS tools and a Loading Simulation Program in C^{++} (LSPC). LSPC is a modified version of the Mining Data Analysis System (MDAS) originally developed by the US EPA Region 3 and has been applied to mining applications and TMDL studies. The relational database allowed users to access and customize the system, which includes modifying the data set, extracting model parameters, dealing with relationships among data sets, and analyzing model results. The relational database approach facilitated data management and model simulations.

A stormwater runoff model needs to incorporate a lumped approach at some detailed subwatershed level. It cannot be modeled with distributed approach alone because of the heterogeneity of a watershed and the stochastic nature of nonpoint source problems. Most water quality models described incorporated heterogeneity of a watershed by dividing it into subwatersheds and applying lumped models to each subwatershed. However, this approach has limitations with regard to handling large scale watersheds with numerous subwatersheds because a large number of model parameters are required for these subwatersheds. Our approach to this problem is to model stormwater runoff of a large watershed within a GIS platform using VBA, thus, eliminating a large number of parameter generation problems.

2.3 ARTIFICIAL NEURAL NETWORK

An ANN is a mathematical model that simulates the operation of the human brain. ANNs consist of many simple arithmetic computing elements corresponding to neurons, and the network as a whole corresponds to a collection of interconnected neurons. The connections have associated numeric weights. Weights are the primary means of long term storage in neural networks, and learning usually takes place by updating the weights. Function approximation by neural networks begins in a random state and learns using repeated processing of a training set. The training set is a set of inputs and target outputs. Learning occurs because the error between an ANN output and a target output is calculated and used to adjust the weights. This continues until errors are sufficiently small or until no more improvement is possible. The trained ANN can be used with new inputs for estimation and prediction.

2.3.1 ANN Popularity

ANNs have recently gained significant popularity because it has been proven that an ANN can create non-linear mappings between input and output variables under certain conditions. More specifically, Scarselli and Tsoi (1998) proved that continuous functions can be approximated up to any degree of precision with a finite number of hidden nodes in a three-layered feedforward neural network. They also mentioned that four or more layer feed forward neural networks are rarely used in practice but they are universal approximators. Reich and Barai (2000) overviewed recent studies and stated that continuous functions on a bounded range can be modeled arbitrarily closely using a three-layered ANN and arbitrary functions can be modeled with four-layered ANN. However, theoretical results do not provide guidelines about ANN configuration or the number of data required to achieve a desired approximation (Scarselli and Tsoi, 1998). Consequently, choosing an appropriate evaluation method is critical in the success of an ANN modeling.

2.3.2 Evaluation of ANN models

Tools for fitting parametric models to data such as non-linear regression require a selection of a parametric model prior to function approximation. This selection can introduce bias that might have significant impact on the success of the function approximation. An ANN, which is a non-parametric model, avoids this problem at the potential cost of additional computational resources and increased variance. However, varying the number of hidden units effectively creates models that vary from parametric to nonparametric for feed-forward neural networks (Reich and Barai, 2000). Smoothing can be performed using a lesser number of hidden units, while overfitting can result from using too many hidden units. Overfitting is characterized by a continuous decrease in a training error while a testing error increases when both data sets are representative of a data set to be modeled. There is no direct way to estimate confidence intervals for real data analysis. However, sensitivity analysis can address the variability of model results.

Error estimation methods influence tradeoffs between bias and variance. Brief descriptions about these methods are as follows:

a. *Resubstitution test*: A single data set is used for training and testing.
b. *Holdout test*: A data set is randomly divided into disjoint training and testing sets. It is common to select 2/3 of the set for training and the remaining 1/3 for testing.
c. *k-fold cross validation*: A data set is divided into k subsets of roughly equal size. A network is trained k times, each time leaving out one of the subsets from training and using it for testing.
d. *Leave-one-out cross validation*: Similar to *k-fold cross validation*, except that k is equal to the number of elements in the data set. Therefore, the training data set contains k-1 elements and the testing data set contains only one element. Training and testing is repeated k times.

The detailed explanations of these methods are described by Reich and Barai (1999). Tradeoffs between bias and variability of these methods are summarized in Table 2.4. The number of internal iterations in the table denotes the number of executions performed over the entire data set. Resubstitution estimate is very optimistic, meaning it is biased very high. Holdout test estimate is pessimistic, which means low-biased.

Table 2.4 Properties of Error Evaluation Methods (Reich and Barai, 1999)

Estimation method	Size of training set	Size of testing set	Number of internal iterations	Method variability	Method bias
Resubstitution	n	n	1	Very high	Very optimistic
Holdout test	$(0.6–0.8) \cdot n$	$(0.2–0.4) \cdot n$	1	High	Pessimistic
k-fold cross-validation	$n(k - 1)/k$	n/k	k	Moderate-high	Nearly unbiased
Leave-one-out	$n - 1$	1	n	Moderate-high	Nearly unbiased

Generally, *10-fold cross validation* is used for sample size greater than a hundred and *leave-one-out cross validation* is used for smaller data sets. Some error evaluation methods are better than the others for specific problems, but no method is always superior (Reich and Barai, 1999). The authors also suggested the following procedure for choosing the best ANN for a data set and its operational parameters:

1. Divide the data set into training and testing subsets.
2. Select the best ANN and its operational parameters using the following method:
    a. Test each combination of ANN and/or operational parameters with a k-fold cross validation test.
    b. Select the combination that leads to the best cross validation performance.
3. Assessment of the best ANN:
    a. Create a model from all training data using the best ANN and its operational parameters.
    b. Test the model on the testing data set.

It has been common to compare developed ANN models with another ANN or regression model. Cannon and Whitfield (2002) developed empirical downscaling models for stream flow at 21 stations in British Columbia. They showed that ensemble neural network models either outperformed or yielded the same performance as stepwise linear regression models at 19 out of 21 stations. Raid et al. (2004) used a multilayer perceptron neural network with a backpropagation algorithm to predict the drainage basin runoff. They demonstrated that the ANN is more suitable to predict runoff than a classical regression model even in arid and semiarid regions with a very irregular rainfall and runoff characteristics. The ability of ANNs to account for non-linear patterns and irregular seasonal variation in a data set makes them well suited for use in hydrological modeling applications (Maier and Dandy, 2000).

2.4 STORMWATER RUNOFF STUDIES AT UCLA

Several previous Ph.D. students have worked on non-point source pollution problem. A brief summary is provided in order to show the different approaches and goals adopted in their projects.

Wong (1999) developed an integrated GIS and empirical urban runoff model. This empirical model is a modified version of the SWMM model. The integrated model uses the Runoff and Transport Blocks of SWMM model and incorporates GIS as a preprocessor and a postprocessor. EMCs are used for the stormwater runoff quality analysis. Lee (2002) investigated relationships between stormwater quality and landuse types using Neural Networks. The developed neural model effectively classified landuse types with water quality data. Park (2004) studied empirical methods to estimate stormwater pollution from estimated landuse using Bayesian networks. The results, which include runoff coefficients, EMCs, and pollutant loadings, were visualized in a GIS. Lourdes (2005) developed an expert system that classifies landuse categories from satellite imagery.

In this research, landuse contributions to the total pollutant loadings in upper Ballona Creek watershed are discussed based on predictive stormwater runoff quantity and quality models. Predictive model for stormwater runoff quantity was developed in an ArcGIS platform. Unlike other models shown in Table 2.5, input data requirements for this model are simplified. More distributed approach was adopted to incorporate heterogeneity of a watershed and stochastic nature of nonpoint source problems. Predictive model for stormwater runoff quality was developed using multiple linear regressions and optimization techniques for estimation and prediction. Using these models for stormwater runoff quantities and qualities, effective BMP strategies are discussed.

Table 2.5 Models for Storm Water Runoff Quantities and Qualities

Model	Developer	Main characteristics	Advantages	Disadvantages	Applications
HEC-1	US Army Corps. of Engineers (1968-1990)	▪ Rainfall-runoff simulations ▪ DOS-based ▪ Unit hydrograph technique	▪ Variety of rainfall-runoff methods ▪ Several common channel and storage routing techniques ▪ Dam safety options ▪ Flood damage analysis ▪ Hydrologic features	▪ Limited hydraulic capabilities ▪ No water quality features ▪ No Tail-water effects ▪ Simulates only a limited number of urban hydraulic structures ▪ Not user-friendly ▪ Editing and graphical capabilities are limited	▪ Accepted by FEMA ▪ Flood insurance studies ▪ Watershed master planning ▪ Multiplan-Multiflood analysis
TR-55	NRCS (1975-1986)	▪ Runoff hydrographs and peak discharges ▪ DOS-based ▪ Used in small urban catchments ▪ Unit hydrograph technique	▪ Storage effects analysis ▪ Calculations of area weighted, time of concentration, and travel time ▪ Detention pond analysis ▪ Preliminary pond sizing ▪ Easy-to-use	▪ SCS unit hydrograph only ▪ Editing and graphical capabilities are not available ▪ No hydraulic features ▪ No water quality features	▪ Peak rate and the total volume studies in small watersheds
HEC-RAS	US Army Corps. of Engineers (1995)	▪ One-dimensional hydraulic calculations ▪ Windows-based	▪ User-friendly ▪ Floodplain elevations and floodway encroachments ▪ Graphical interface, tabular, and report formats ▪ Subcritical, critical, and supercritical flow regimes ▪ Effects of in-stream structures	▪ Not GIS-based data processing ▪ No hydrologic features ▪ No water quality features	▪ Flood insurance studies ▪ Watershed master planning ▪ Multiplan-Multiflood analysis
MOUSE	Danish Hydraulic Institute (1995)	▪ Analysis of storm water quantity and quality ▪ Windows- based ▪ Developed in 1985	▪ User-friendly ▪ Graphical interface, tabular, and report formats ▪ Single-events and continuous events ▪ Manholes surcharging and pressure	▪ Input data requirements for the model are extensive	▪ Return periods of overloading sewer system ▪ Determine causes of overloading ▪ Impacts of replacing sewer

		and subsequently improved with several additional capabilities	flow ▪ Open channel simulations ▪ Modules for transport of pollutants and sediments ▪ Off-line and on-line modeling of real time control		▪ Environmental impact of changing operation strategy ▪ Water quality protection studies ▪ Sedimentation deposition ▪ Pollutant concentration after storms
Hydro Works	HR Wallingford Ltd (1997)	▪ Analysis of storm water quantity and quality ▪ Wastewater network simulation	▪ User-friendly ▪ Graphical interface	▪ Input data requirements for the model are extensive	▪ Design of sewer system ▪ Impacts of replacing sewer ▪ Water quality protection studies
WSPRO	USGS (1988)	▪ One-dimensional, gradually varied, steady flow ▪ DOS-based	▪ Easy-to-use ▪ Analysis and design of bridge opening, embankment, open channels, and culverts ▪ Profiles in sub-critical, critical, and supercritical flow regimes ▪ Flow through bridges and culverts and pressure flow under bridges can be simulated	▪ No hydrologic features ▪ Editing and graphical capabilities are not available	▪ Flood insurance studies ▪ Water surfaces profiles analysis for highway design ▪ Establishing stage-discharge relationships
US EPA SWMM	US EPA (1969-1971)	▪ Analysis of storm water quantity and quality ▪ DOS and Windows-based	▪ **Public domain** ▪ **Open source code** ▪ Storm water, sanitary, and combined sewer system ▪ Single-events and continuous events ▪ Manholes surcharging and pressure flow ▪ Open channel simulations ▪ Buildup, washoff, and pollutants routing ▪ Propriety extension with graphical interfaces, such PCSWMM, XPSWMM, and MIKE-SWMM	▪ The graphical capabilities of original version are limited ▪ Input data requirements for the model are extensive	▪ FEMA-approved model for NFPI studies ▪ Flood control studies ▪ Water quality protection studies ▪ Impact of inflow and infiltration on sanitary sewer overflows ▪ Effectiveness of BMPs ▪ Designing control strategies for minimizing combined sewer overflows

Culvert Master	Haestad Methods (1999)	▪ Culvert simulation and design program ▪ Calculate flow peak using the rational method ▪ Windows-based	▪ Easy-to-use ▪ Pressure or free surface flow condition in sub-critical, critical, and supercritical flow conditions ▪ A Variety of culvert shapes and sections types are available ▪ Tail-water effects are considered	▪ The user must enter all rainfall and runoff information ▪ No hydrologic features ▪ No water quality features	▪ Culvert analyzer studies ▪ Culvert designer
Flow Master	Haestad Methods (2000)	▪ Hydraulic pipe and channel design program ▪ Windows-based	▪ Calculates hydraulics of weirs, orifices, gutter, and ditch flow	▪ No hydrologic features ▪ No water quality features	▪ Design and analysis of pipes, ditches, open channels, weirs, orifices, and inlets
QUAL2E	US EPA (1991-1996)	▪ Steady state or dynamic conditions ▪ DOS-based	▪ **Public domain** ▪ Uncertainly analysis ▪ Impact analysis of waste loads on water quality ▪ Up to 15 water quality constituents can be modeled ▪ Analysis of diurnal variation in water quality	▪ Complex model and data requirements for a simulation ▪ Editing and graphical capabilities are not available ▪ No hydrologic features	▪ Water quality planning

- NRCS: Natural Resources Conservation Service.
- FEMA: Federal Emergency Management Agency.
- Most of the models presented in the table were improved. For example, US EPA developed the new version SWMM5, which is user-friendly and can be integrated with GIS.
- Software such as PCSWMM, XP-SWMM, MIKE-SWMM, MIKE-URBAN, and WMS are propriety extensions with graphical interfaces for US EPA SWMM, EPA SWMM5, HEC-1, TR-55, HEC-RAS, and WSPRO.

3. METHODOLOGY

Nonpoint source pollution problems are inherently complex. One approach is to model its complexity as closely as possible; however, complex models require many quantity and quality prediction parameters, such as soil characteristics, chemical and biological properties, partition coefficients, reaction rates, and buildup and washoff parameters. The propagated or accumulated error from these parameters may result in poor predictions, especially when compared to simple models that might have fewer and more easily measured empirical parameters. In this study, a simple methodology for modeling pollutant loadings of storm events that accounts for the heterogeneity of a watershed is developed. For the simulation of the model, Los Angeles County Department of Public Works (LADPW) and California Department of Transportation (Caltrans) data were used.

Each county in the United States including Los Angeles County is required by the NPDES program to monitor pollutant loadings to its receiving water body. The current monitoring program of Los Angeles County consists of the Santa Monica Bay receiving water impacts study, mass emission monitoring, landuse runoff monitoring, and critical industry monitoring. For example, the LADPW monitored four drainage areas near their ocean discharges to comply with their 1990 and 1996 NPDES Permits. These four sites, called mass emission sites are the Los Angeles River, San Gabriel River, Ballona Creek, and Malibu Creek. The Ballona Creek site drains a highly urbanized watershed, which is the ideal subject for this research project. It has been the subject of previous studies conducted by University of California, Los Angeles (UCLA) investigators (Wong et al., 1997; Lau and Stenstrom, 1999; Stenstrom and Lee, 2005), and was chosen for this study. Most of the previously collected data and the County's monitoring program used a mass emission monitoring station located on Ballona Creek between Sawtelle and Sepulveda Boulevards in the City of Los Angeles. This station was chosen to avoid tidal influences and has a United States Geological Survey (USGS) gauge. The upstream tributaries of Ballona Creek, which is called upper Ballona Creek watershed, drain 23,211 hectare (89.6 squire miles). This research project will utilize these accumulated data to estimate total pollutant loadings in a new, automated fashion.

3.1 DATA REQUIREMENTS

To estimate total loads of a storm event, accumulated monitoring data need to be incorporated into an ArcGIS layer, which is a tool to store and access large data sets that have geographical references or standalone databases. The following data were incorporated into the model as ArcGIS layers:

1. Digital Elevation Model (DEM)
2. Event Mean Concentration (EMC) data for each landuse categories
3. Hydrologic soil group
4. Dams locations
5. Mass emission monitoring station location

6. Rain gauge data for each rainfall event
7. Rain gauge locations
8. Southern California Association of Government (SCAG) Landuse
9. Watershed area boundaries

Available GIS layers were downloaded through the web. The sources are as follows:

1. USGS (http://seamless.usgs.gov/) – DEM was downloaded from the Seamless Data Distribution System. This system is operated by USGS and EROS Data Center.
2. UCLA (http://www.ats.ucla.edu/) – SCAG 6 County landuse 2000, watershed 20, and watershed 22 were downloaded from UCLA ArcSDE/Oracle database. This database is supported by the UCLA Academic Technology Services (ATS).
3. California Spatial Information Library (CaSIL, http://gis.ca.gov) – The local roads layer has been made available through the CaSIL. It is derived from the US Census Bureau Tiger 2K (June 7, 2002 Version) information.
4. NCDC (http://www.ncdc.noaa.gov/) – Rain gauge locations and its hourly rainfall records were downloaded from National Climatic Data Center (NCDC).
5. LADPW (http://ladpw.org/) – Locations of rain gauge stations operated by LADPW along with rainfall data and EMCs for mass emission and landuse monitoring stations.

The spatial metadata of the GIS layers used in this study are shown in Table 3.1.1. Users will not ordinarily need to know the metadata, but they are included for users or reviewers who have advanced knowledge of GIS or users who might wish to modify the program.

Table 3.1.1 Metadata of the GIS Layers Used

Datum	NAD 27
Projection	Albers
Units	Meters
1st Std. Parallel	34 00 00 (34.0 degrees N)
2nd Std. Parallel	40 30 00 (40.5 degrees N)
Longitude of Origin	- 120 00 00 (120.0 degrees W)
Latitude of Origin	00 00 00 (0.0 degrees)
False Easting (X shift)	0
False Northing (Y shift)	- 4,000,000
Source	USGS digital line graph (DLG) digital series
Source Media	Mylar maps
Source Projection	Universal Transverse Mercator Zones 10 & 11
Data Structure	Vector or Raster

3.2 SOFTWARE AND HARDWARE REQUIREMENTS

The model was implemented in VBA, which is included with ArcGIS Desktop Version 9.0 (Redlands, CA). The program should run with later versions of ArcGIS Desktop, although this has not been verified. The developer recommends installing ArcGIS Desktop Developer Kit, which provides essential tools for developers. The most useful tools are Object Browser and Library Locator. It is also necessary to have Microsoft Access 2002 or later. For satisfactory execution, the computer should be a Pentium with 2GHz or higher, 1GB memory, 2GB swap space, and about 4GB free disk space (850MB for the ArcGIS Desktop installation, 350MB for Developer Kit installation, and 3GB for the model). Since most ArcGIS operations are memory intensive, dual memory access architecture is also recommended.

3.3 RATIONAL METHOD

This study is based on conventional rational method and EMCs for each landuse categories. However, to account for the heterogeneity of a watershed, a distributed approach was taken. A watershed is divided into many small sub-areas to account for the heterogeneity of a watershed and lumped parameters were applied on each of them. Thus, the conventional rational method can be rewritten as follows:

$$L = \sum_{j} EMC_{j} \times \left(\sum_{i} A_{i} \times P_{i} \times RC_{i} \right)_{j} \qquad (3.1)$$

where L = Pollutant loading

j = Landuse type

EMC_j = EMC of landuse type j

A_i = Homogeneous sub area

P_i = Precipitation in the sub area A_i

RC_i = Runoff coefficient of the sub area A_i

The division of a watershed depends on the availability of data that can be applied to each of the divided sub area. Each sub area A_i is homogeneous in terms of the available data if the watershed information is sufficiently detailed so that each sub area has unique parameters. In this way, heterogeneity of a watershed is accounted for up to available datasets. Available datasets and its division into smaller sub area will be discussed in the following sections.

3.4 STORMWATER RUNOFF QUANTITY ANALYSES

The developed GIS model is used to calculate runoff quantities from each landuse type. Total runoff quantity is then calculated by adding these quantities for a storm event. In the following section, procedures of calculating watershed area, rainfall, and runoff coefficients are described. Data needed for these calculations are DEM, hydrologic soil groups, landuse types, rainfall data of rain gauge stations around the watershed, and locations of the gauge stations. The calculated stormwater runoff volumes were compared to the measured volumes at the emission site of the upper Ballona Creek watershed.

3.4.1 Area

The top eight landuse types based on total area were selected by LADPW through the field survey conducted in 1996 (Woodward, 1996):

1. Educational Facilities
2. High Density Single Family Residential
3. Light Industrial
4. Multifamily Residential
5. Mixed Residential
6. Retail/Commercial
7. Transportation
8. Vacant

To apply EMCs monitored by LADPW (see Equation 3.1), the Southern California Association of Governments (SCAG) landuse types were reclassified to the eight landuse types. SCAG completed the 2000 landuse data update using 2000 digital aerial imagery, which is shown in Figure 3.4.1 for the study area. The reclassified landuse types are shown in Figure 3.4.2. The relationship between LADPW and SCAG landuse types are listed in the Appendix B. The distribution of landuse types calculated from the reclassified landuse types layer is shown in Table 3.4.1.

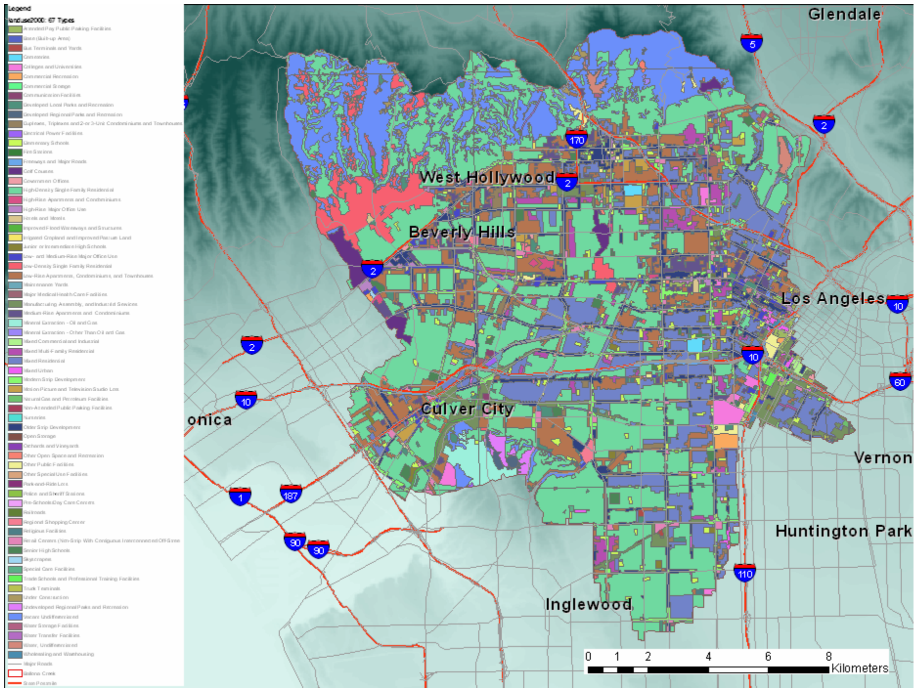


Figure 3.4.1 SCAG Landuse Types for Upper Ballona Creek.

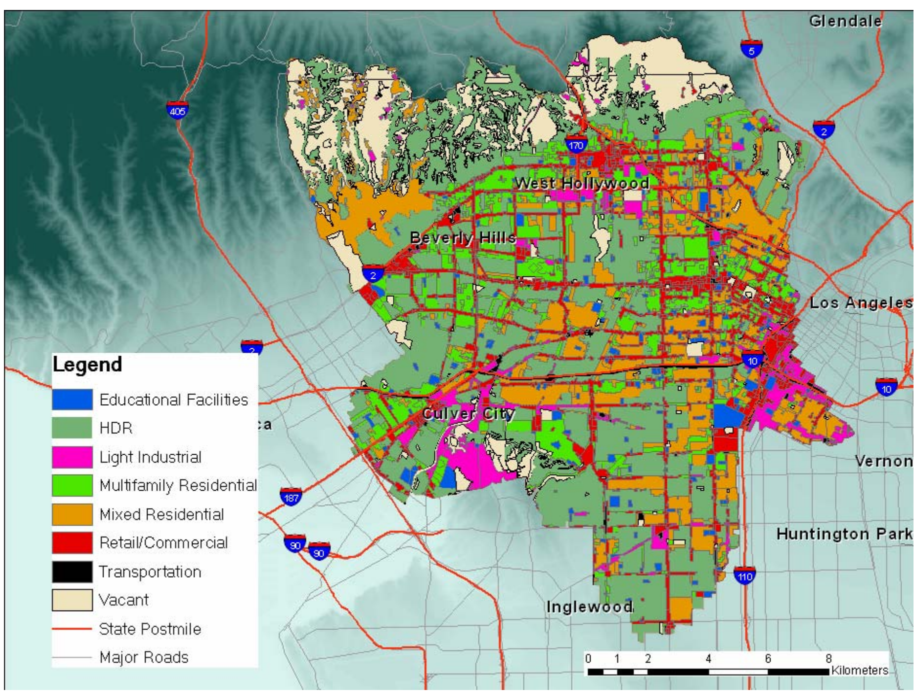


Figure 3.4.2 Reclassified Landuse Types.

Table 3.4.1 Distribution of Landuse Types

Landuse	Area [Hectare]	% Area
Vacant	3,094	13.3
Educational Facilities	667	2.9
High Density Residential	7,838	33.8
Light Industrial	1,217	5.2
Multi-Family Residential	2,898	12.5
Mixed Residential	3,868	16.7
Retail/Commercial	3,215	13.8
Transportation	415	1.8
Total	23,211	100.0

3.4.2 Runoff Coefficients

Spatial variability of runoff coefficients in a watershed is incorporated using Browne's relation (1990) shown in Appendix C, which relates runoff coefficient with hydrologic soil group, slope, and landuse type:

Runoff Coefficient = f (Hydrologic soil group, Percent slope, Landuse)

Soils are classified by the Natural Resource Conservation Service into four hydrologic soil groups based on the soil's runoff potential, which are A, B, C and D. Group A has generally the smallest runoff potential and group D has the greatest. As shown in Figure

3.4.3, the upper Ballona Creek is made up of entirely Group D soils, which have very low infiltration rates and consist chiefly of clay soils.

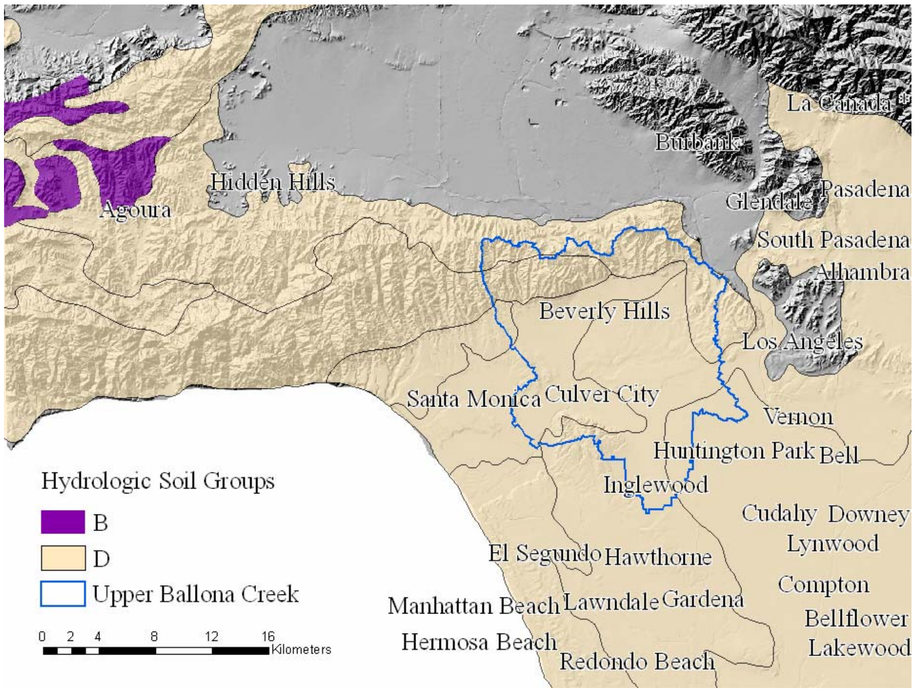
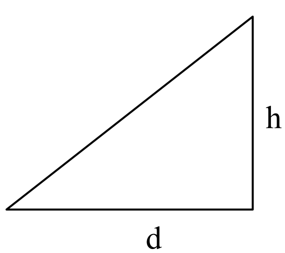


Figure 3.4.3 Hydrologic Soil Group around the Upper Ballona Creek Watershed.

Browne relation uses three categories of percent slopes, which are 0-2%, 2-6%, and 6+%. The ArcGIS Spatial Analyst Toolbox provides a tool to calculate percent slopes, which are defined as in Eq. (3.2).



$$\text{Percent Slope} = \frac{h}{d} \times 100 \tag{3.2}$$

A DEM downloaded from USGS seamless data distribution system (http://seamless.usgs.gov/) is shown in Figure 3.4.4. This DEM raster was used to calculate a percent slope raster with the slope tool. The slope tool calculates average percent slope from each cell to its neighbors.

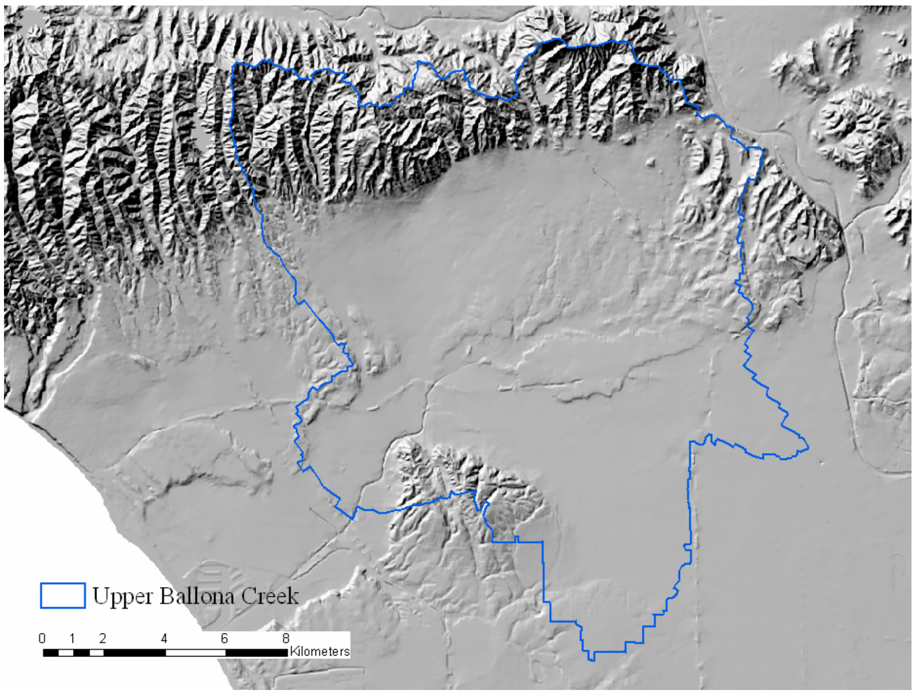


Figure 3.4.4 A DEM downloaded from the USGS Seamless Data Distribution System.

The percent slope raster was then reclassified into three categories defined in the Browne relation using reclassify tool in the Spatial Analyst Toolbox. Both the percent slope raster and the reclassified percent slope raster are shown in the Figure 3.4.5 and Figure 3.4.6, respectively.

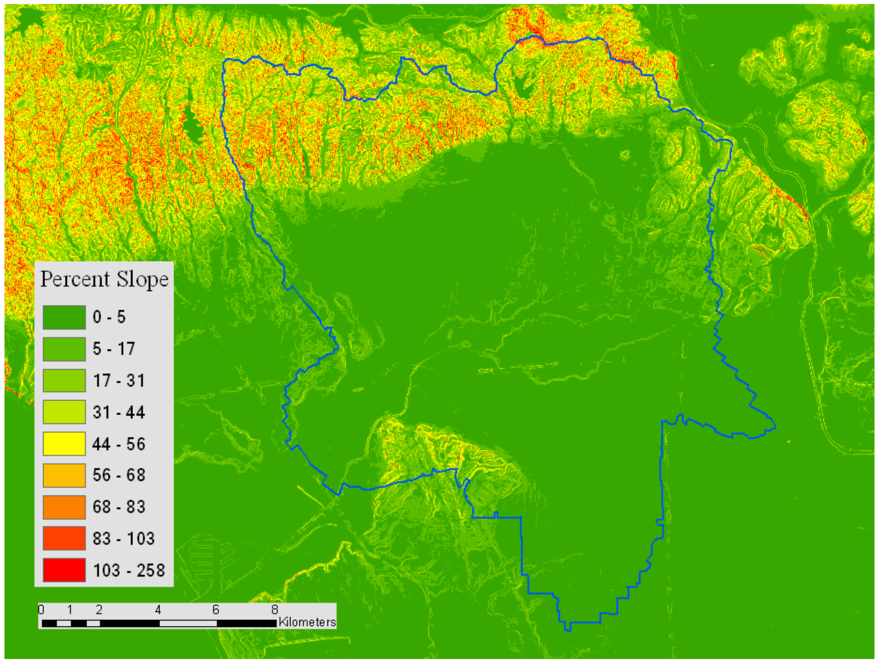


Figure 3.4.5 Percent Slope calculated from the DEM.

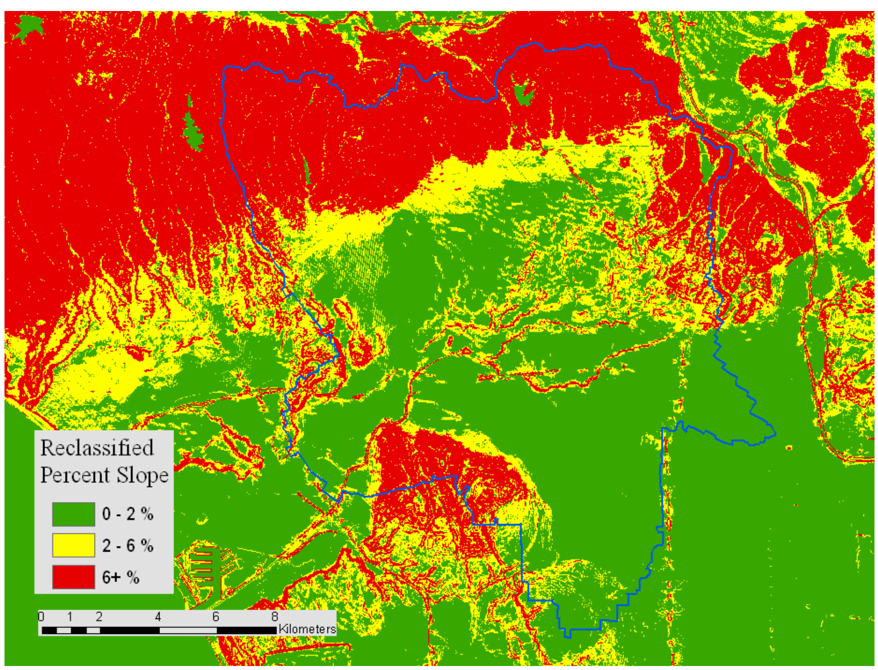


Figure 3.4.6 Reclassified Percent Slope.

The landuse category polygons shown in Figure 3.4.2 were overlaid with a hydrologic soil group layer and a percent slope layer created from the DEM, which are shown in Figure 3.4.5 and Figure 3.4.6, respectively. This overlay created approximately 33,000 polygons for the upper Ballona Creek watershed as shown in Figure 3.4.7.

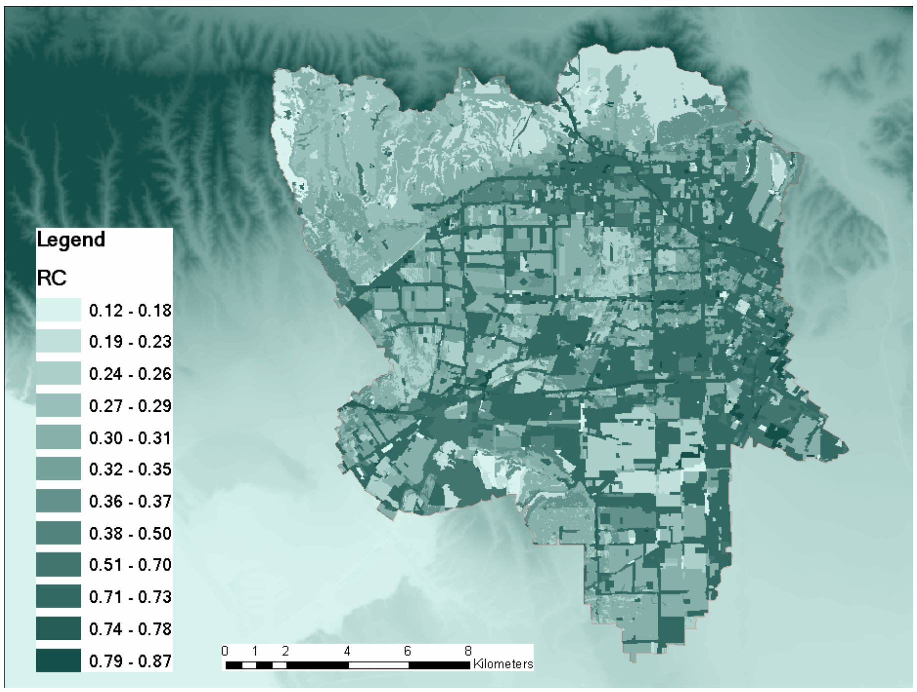


Figure 3.4.7 Runoff Coefficients Distributions in the Upper Ballona Creek.

The resulting polygons are internally homogeneous in terms of runoff coefficients and EMCs. However, the layer does not yet contain runoff coefficient values. To create a runoff coefficient field in the overlaid layer, runoff coefficient calculation module was implemented in the VBA. This module identifies three fields in the overlaid layer, which are reclassified landuse type, hydrologic soil group, and reclassified percent slope fields. The module then creates a new runoff coefficient field using the Browne relation.

3.4.3 Precipitation

Hourly precipitation data from the seven rain gauge stations, which are shown in Table 3.4.2, located around the upper Ballona Creek are used. Gauge station locations are shown in Figure 3.4.8 along with the extent of the upper Ballona Creek watershed. The precipitation data were gathered from NCDC and LADPW.

Table 3.4.2 Rain Gauge Stations

Station Name	Latitude	Longitude	Elevation [m]
Burbank Valley Pumping Plant [1]	34° 11′ 11″	118° 20′ 54″	199.6
Los Angeles Airport [1]	33° 56′ 25″	118° 23′ 44″	30.5
Los Angeles Civic Center [1]	34° 03′ 09″	118° 14′ 13″	82.3
Sepulveda Dam [1]	34° 10′ 06″	118° 28′ 11″	204.2
Ballona Creek Mass Emission Site [2]	34° 00′ --″	118° 24′ --″	-
Malibu Creek Mass Emission Site [2]	34° 05′ --″	118° 43′ --″	-
Santa Monica Pier [2]	34° 00′ 43″	118° 29′ 27″	28.7

[1]: NCDC gauge station; [2]: LADPW gauge stations

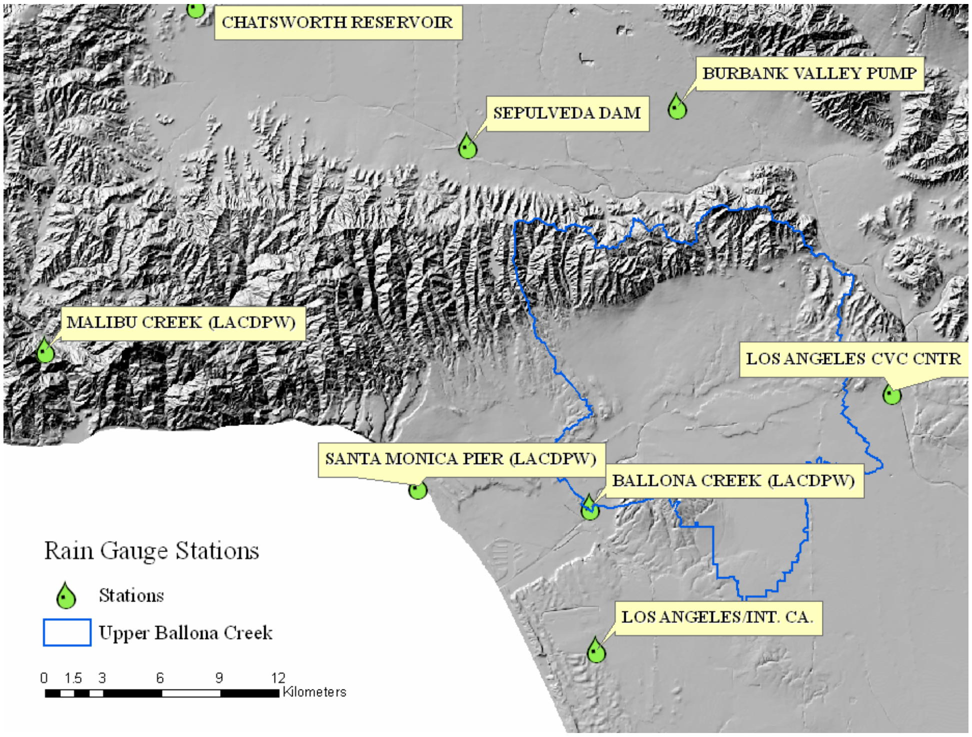


Figure 3.4.8 Rain Gauge Station Locations

3.4.3.1 Data Screening

LADPW measured total runoff volumes and their EMCs for mass emission sites starting from 1998, which are available from their web site (http://ladpw.org/WMD/npdes/report_directory.cfm). The measured total runoff volume and its corresponding precipitation data gathered from the seven rain gauge stations are shown in Table 3.4.3. Among the available 49 storm events from 1998 to 2003 storm seasons, four storm events without flow measurement data was excluded.

Table 3.4.3 Measured Rainfall for Seven Gauges and Total Runoff Volumes (1998-2003)

Event [mm]	Malibu Creek	Sepulveda Dam	Burbank Valley Pump	LA Civic Center	LA /INT. CA.	Santa Monica	Ballona Creek	Measured [$10^3 m^3$]
12/1/1998	11	3	5	7	4	1	4	378
12/6/1998	6	5	5	5	10	12	11	528
1/20/1999	2	5	3	3	6	28	4	144
1/25/1999	22	10	20	9	10	6	25	2526
1/31/1999	13	17	15	14	5	10	15	1306
2/4/1999	12	0	0	2	5	1	5	280
2/9/1999	23	6	8	9	4	8	5	606
3/9/1999	1	0	0	1	4	2	2	50
3/15/1999	21	19	20	11	17	1	18	1479
3/20/1999	12	9	15	6	8	6	8	615
4/6/1999	15	15	15	19	11	2	36	2577
4/8/1999	2	0	0	2	2	2	2	99
4/11/1999	13	29	25	32	34	30	35	1888
1/25/2000	23	23	0	11	15	15	12	1771
1/30/2000	5	3	0	3	5	7	7	925
2/10/2000	25	9	20	12	8	12	14	1983
2/12/2000	31	23	23	16	13	17	15	3678
2/16/2000	11	11	23	17	14	17	16	2488
2/20/2000	39	56	33	22	17	28	20	1749
2/23/2000	59	42	46	28	19	26	23	4046
3/5/2000	34	37	48	44	29	25	34	5775
3/8/2000	23	18	15	18	22	15	10	2139
[1] 10/12/2000	0	0	0	0	0	0	2	798
10/29/2000	23	0	15	13	15	11	10	433
1/8/2001	5	4	5	7	6	5	4	556
1/10/2001	143	104	66	55	53	84	99	8907

Table 3.4.3 – Continued –

Event [mm]	Malibu Creek	Sepulveda Dam	Burbank Valley Pump	LA Civic Center	LA /INT. CA.	Santa Monica	Ballona Creek	Measured $[10^3 m^3]$
1/24/2001	13	6	13	8	7	8	7	839
1/26/2001	2	12	8	18	17	19	1	1477
[2] 2/10/2001	119	13	10	9	8	16	105	990
2/19/2001	20	9	3	14	6	21	10	1413
[2] 2/24/2001	67	12	8	5	7	8	58	8069
[2] 3/4/2001	122	11	8	1	2	7	33	3412
11/12/2001	14	11	0	9	9	0	7	761
11/24/2001	55	30	0	20	15	0	19	2950
11/29/2001	16	6	0	6	8	0	8	443
12/2/2001	7	2	0	3	3	0	8	495
12/20/2001	15	10	0	10	2	12	18	1544
[2] 1/27/2002	39	0	23	0	0	21	12	2231
11/8/2002	58	29	20	46	24	0	30	5218
[1] 12/16/2002	70	36	33	50	26	35	27	133
[2] 2/11/2003	130	26	18	22	15	0	7	11600
[3] 3/15/2003	128	60	66	104	38	81	84	13065
10/31/2003	13	8	14	8	0	20	14	469
[1] 12/25/2003	36	21	24	32	0	26	51	13815
1/1/2004	0	12	11	5	0	9	16	1146

[1] Unexpectedly high or low total runoff volume compared to rainfall records.
[2] Rainfall patterns are extremely irregular across the rain gauge stations.
[3] Large storm event compared to the other events.

For the rest 45 storm events, nine storm events were filtered, which were indicated with superscripts in the table, according to the following filtering criteria:

1. Unexpectedly high or low total runoff volume compared to rainfall records, suggesting an error in the data.
2. Rainfall patterns are extremely irregular across the rain gauge stations.
3. Large, atypical storm event compared to the other events.

The main reason behind these filtering criteria is to exclude the possibility of perturbation errors, such as measurement errors, approximation errors, or equipment malfunction errors. For example, the approximation error occurs when measuring total runoff volume of a storm event. LADPW uses previously established rating table to calculate flow rate by measuring water elevation in a storm drain or calculated with an equation such as Manning's. But their stormwater flow measurement efforts indicated that all stations required multiple storm events to gather the data necessary for calibration of the measurement devices. Thus, filtering relatively large storm events is reasonable, since those cases are rare. Simple correlation tests were conducted with the rainfall records of the seven gauge stations shown in Table 3.4.3. LA Civic Center rainfall records showed best correlation to the total runoff volume of upper Ballona Creek watershed. A histogram of the Civic Center rainfall records is shown in Figure 3.4.9. The filtered large event is the one event around 100 mm of rainfall. If the rainfall patterns are extremely irregular across the rain gauge stations, calculations based on these rainfall records may cause large approximation error.

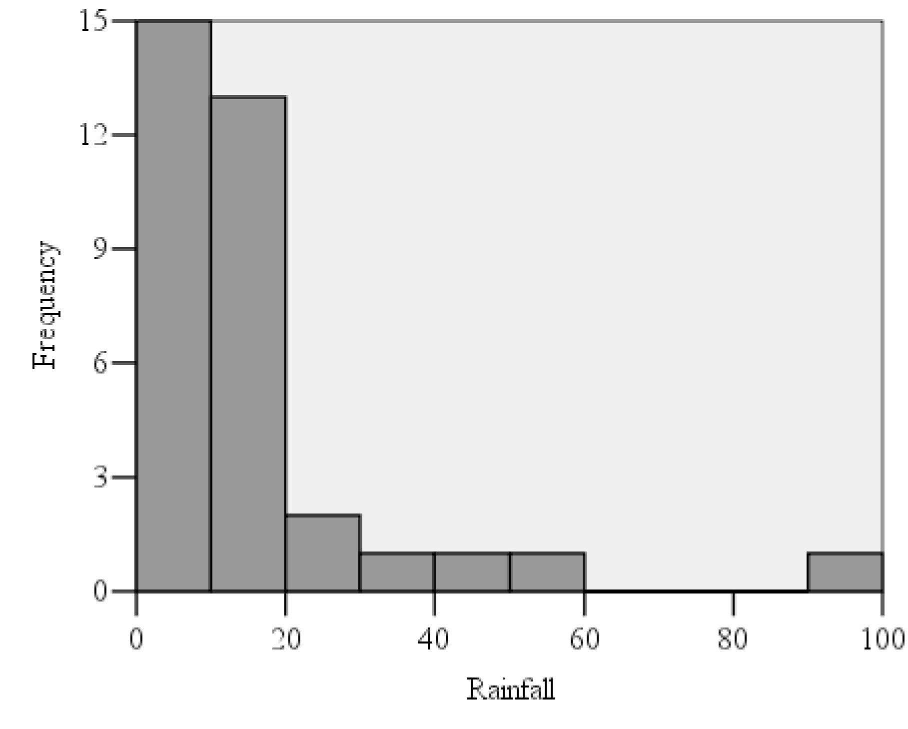


Figure 3.4.9 Histogram of Civic Center Rainfall Records (Unit: mm)

3.4.3.2 Interpolation of Precipitation Data

Interpolation of the precipitation data gathered from the gauge stations is important because precipitation is one of the most influential parameters in the calculation of total pollutant loads. Among many interpolation techniques, Kriging was chosen for the first attempt. Unlike the other interpolation techniques, Kriging assumes that the spatial variation is statistically homogeneous throughout the surface. For example, Inverse Distance Weighted (IDW) interpolation is based on the weighted distance of each data points from a point being estimated. The disadvantage of the IDW interpolation is that it treats all points the same way regardless of their geographic orientations. Kriging uses different weighting function depending on the distance to each data point and spatial relationships among these points. This is done by creating empirical semivariogram which is computed from an input dataset.

Semivariance for a separation distance of h is the average squared difference in Z value between pairs of input points separated by h. Semivariances are calculated from input data with the following equation (McCoy et al., 2004):

$$\gamma(h) = \frac{1}{2n}\sum_{i=1}^{n}\left\{Z(x_i) - Z(x_i + h)\right\}^2 \tag{3.3}$$

where n is the number of pairs of data points separated by a distance h. An empirical semivariogram is constructed by fitting the semivariances with a model, which is similar to regression analysis. This fitting with a model is also known as structural analysis or variography. The empirical semivariogram is used to calculate weights for the Kriging interpolation. Since Kriging is based on a statistical model, it can provide some measure of the certainty of the predictions, which is the variance raster created by Kriging. This raster contains Kriging variances at each output raster cells.

Zimmerman et al. (1999) conducted an investigation to compare the spatial interpolation accuracy of Kriging and IDW. Their result showed that Kriging interpolation consistently and substantially outperformed the IDW interpolation. The relative superiority was greatest when the surface was most regular, noise was low, and spatial correlation was high. However, since Kriging is based on the assumption that the spatial variation in the phenomenon is statistically homogeneous throughout the surface, data sets that have rapidly changing values are not appropriate for Kriging interpolation. In this case, IDW interpolation, which does not correlate spatial arrangement among the data points, was used.

Isohyets and corresponding variances generated with Kriging interpolation for a storm event occurred in Dec. 1, 1998 is shown in Figure 3.4.10. The maximum variance for the upper Ballona Creek watershed was about 0.01 mm of rainfall at most.

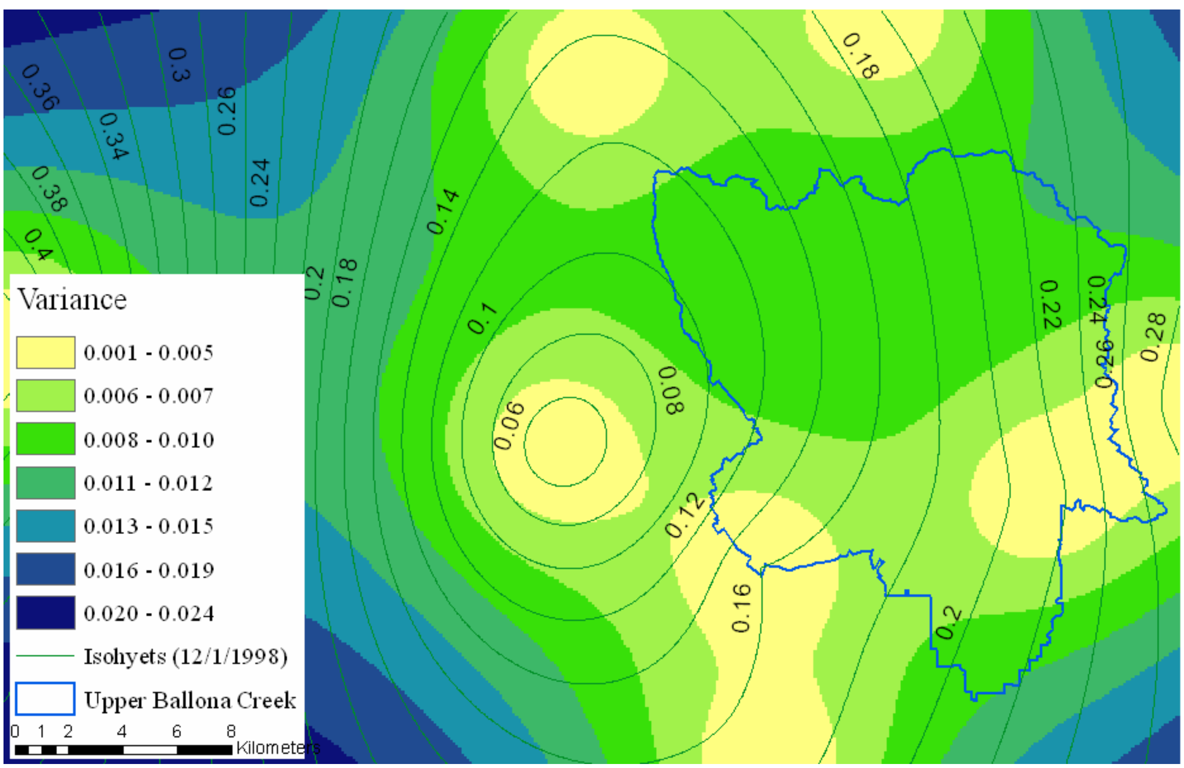


Figure 3.4.10 Isohyets and Variances for a Storm Event (12/1/1998; Unit: in).

3.4.4 Implementation of a GIS Model

A GIS model to predict the total stormwater runoff volume of storm events is implemented in the VBA. The VBA is one of the object oriented programming languages. Although the ArcObjects library consists of over 1,000 classes and 2,000 interfaces to provide developers with the ability to extend ArcGIS functionality, the extent of the library is overwhelming and difficult to know where to begin. Fortunately, there are tools available. The most useful tools are Object Browser and Library Locator as mentioned in Section 3.2 along with Exploring ArcObjects (Zeiler, 2001). The main objective of this model is to account for the heterogeneity of a watershed up to available datasets. This approach also avoids time consuming and costly field surveys. Besides the LADPW landuse types, two additional categories, which are highways and local roads, were separately studied. Thus, results from these two categories can be used as comparison purposes. Methodology adopted in the analyses of three different categories will be discussed in the following sections.

3.4.4.1 Runoff Volumes of the Eight Landuse Types

Reclassified landuse layer (Figure 3.4.2) and the runoff coefficient layer (Figure 3.4.7) were overlaid again to get internally homogeneous polygons in terms of landuse type, runoff coefficient, and EMCs. The isohyets layer (Figure 3.4.10) was then spatially referenced from each polygons of the overlaid layer. Figure 3.4.11 illustrates how the isohyets are referenced in the program. The figure is for illustration purpose only. Actual isohyets layer used in the program have more dense isohyets to reduce errors.

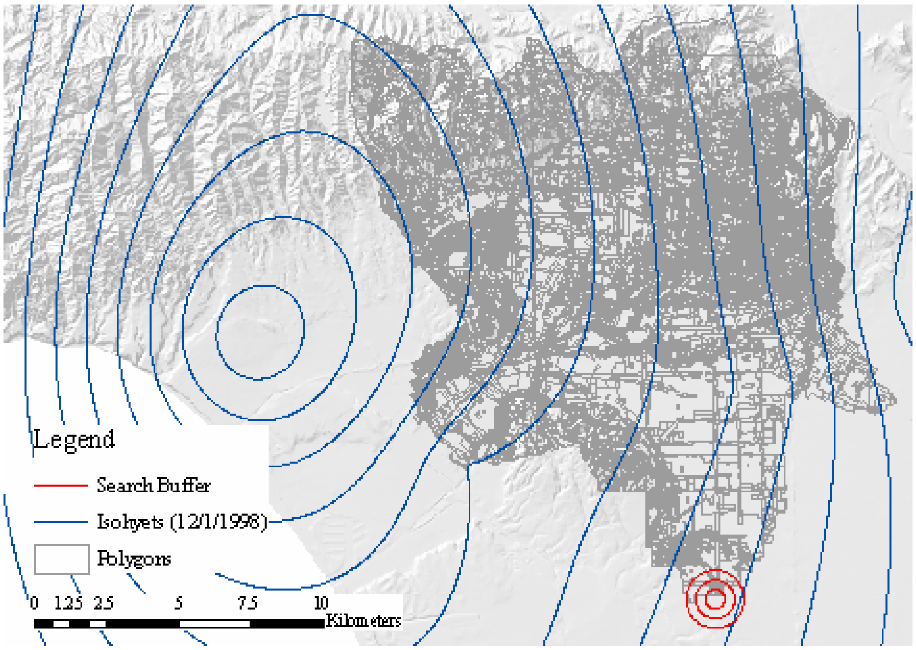


Figure 3.4.11 Spatial Reference of Isohyets

Each polygon in the overlaid layer is scanned to see if any isohyets cross the polygon. If more than one crosses the polygon, the average of the crossed isohyets is used as a rainfall. If no isohyets cross the polygon, the polygon buffer is increased as shown in Figure 3.4.11 until at least one isohyets cross the buffered polygon. If more than one cross the buffered polygon, average value is used again. Since each polygon is one of the eight landuse types, runoff volumes from each landuse type can be calculated.

The better way to get the isohyets is to convert isohyets raster to polygon features using Raster to Polygon tool in the Spatial Analyst Toolbox and then overlay the isohyets polygons with the overlaid polygons mentioned above. However, the Raster to Polygon tool currently has a limit on the number of grids in a raster that can be converted to polygon features and the number of grids in the isohyets raster exceeds the limit. The limit is not documented anywhere and the GIS function usually do not issue reasonable error message when this type of limits are violated. This is one of the frustrating aspects when programming in the VBA.

3.4.4.2 Runoff Volumes for Highways

Caltrans performed a statewide stormwater runoff characterization study from 1997 to 2003 for 171 highway sites throughout California. To analyze contributions from the highway areas to total loading rates, a module that calculates highway runoff volume was implemented. The program was developed for any given watershed or combination of watersheds in California. Total highway area in a watershed including shoulders was calculated from the Caltrans state postmile layer that contains numbers of lanes information for each segment of highways. For the width of highway areas, the following assumptions were made:

1. Typical width of lane = 3.66 m (12 ft)
2. Typical width of shoulder (One shoulder in each side of highways)

   = 2.44 m × 2 (8 ft × 2)
3. Highway runoff coefficient = 0.9

An isohyets layer was spatially referenced from each segment of highways as shown in the Figure 3.4.11.

3.4.4.3 Runoff Volumes of Local Roads

Schematic diagram of typical urban road is shown in Figure 3.4.12. Although there are other runoffs besides runoffs from local road surface and side walk, those contributions are minimal in urban area. If runoffs from other sources, such as lawn or hillside, are negligible, we can calculate runoffs from local roads separately with the following assumptions:

1. Local road width = 9.14 m (30 ft)
2. Local road runoff coefficient = 0.9

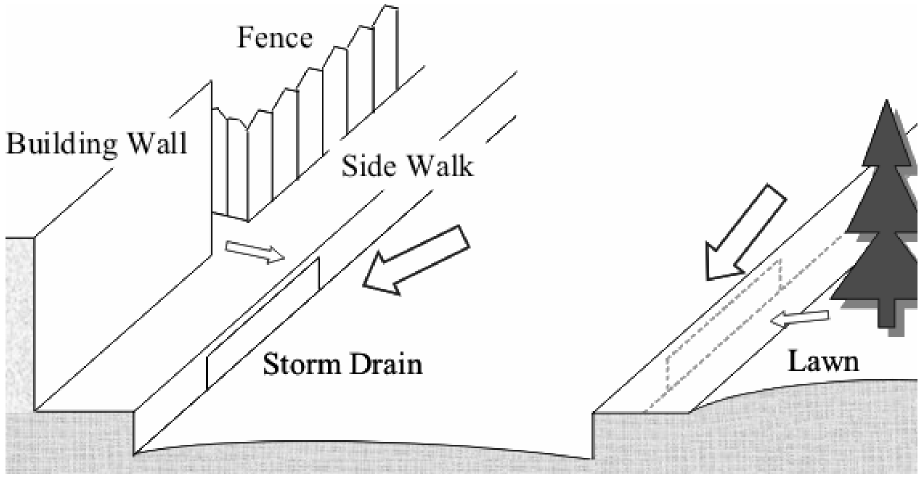


Figure 3.4.12 Schematic Diagram of Typical Urban Road

The local roads layer used in this study has been made available through the California Spatial Information library (http://gis.ca.gov). It is derived from the US Census Bureau Tiger 2K (June 7, 2002 Version) information.

The local roads layer was clipped using Clip Tool in the Analysis Toolbox programmatically. For each segment of the local roads, isohyets were identified as in highway runoff calculation and multiplied by the assumed road width, length of the segment, and the assumed runoff coefficient to get a runoff from the segment. Finally, all the runoffs from each segment were added to obtain the total runoff from the clipped local roads.

3.4.5 Stormwater Runoff Quantity Prediction using the Implemented GIS Model

Total storm events available from the LADPW after the data screening were 36 as shown in Table 3.4.3 from 1998 to 2003 storm seasons. The GIS model, which does not require costly and time consuming field surveys, mentioned in Section 3.4.4.1 was used to predict stormwater runoff quantities for the 36 storm events. The results were compared with the measured runoff quantities at the upper Ballona Creek emission site, which is shown in Figure 3.4.13. For the heterogeneity of such a large watershed (23,211 hectare) and the variability of rainfall, the predicted runoff volumes were in good agreement with the measured data and had R^2 of 0.86, and generally were within +184/-54 percent of the measured runoff.

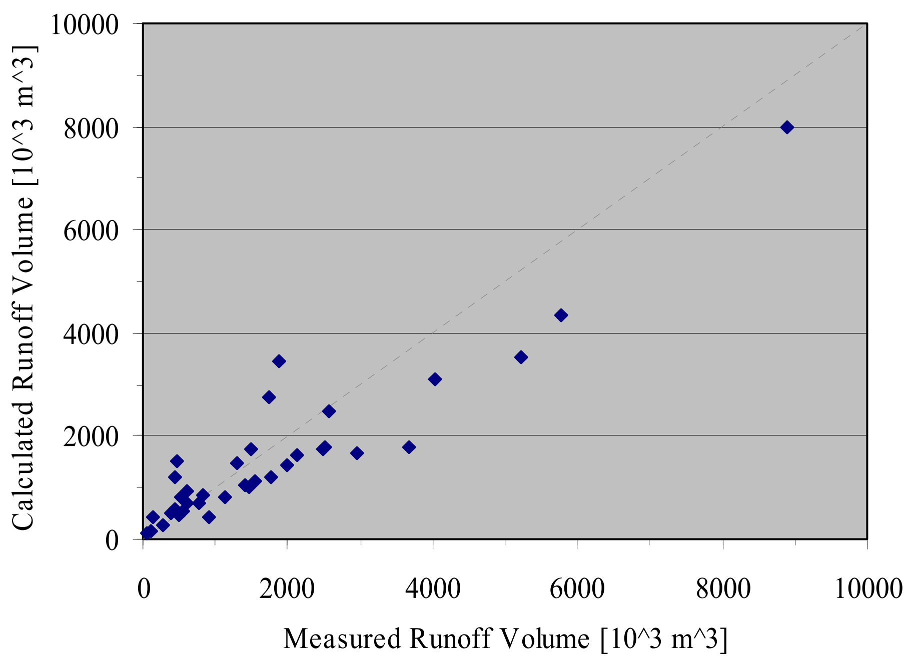


Figure 3.4.13 Calculated versus Measured Stormwater Runoff Quantities (1998-2004). The Diagonal shows the One-to-One Correspondence.

The implemented model showed improved runoff prediction (- 8 % of the measured runoff) than the model defined by landuse definitions (+ 54 %; Wong et. al, 1997) for an example analysis of Dec. 2, 2001 storm event. The model over predicted runoff from small storms and under predicted runoff from large storms, which is a well-known characteristic of the Rational Method. The model has the advantage of reduced data needs, when compared to a runoff model such as SWMM. For example, Barco et al. (2006) required approximately 30,000 input data points for the characteristics of drain network and about 2,500 subcatchments to model the upper Ballona Creek Watershed. The new model required only four GIS layers, which are DEM, hydrologic soil group, landuse type, and isohyets layers. More importantly, only the fourth layer needed to be developed for the modeling activity.

3.5 STORMWATER RUNOFF QUALITY ANALYSES

LADPW was required under the 1990 NPDES Permit to sample representative areas that were predominantly of a single landuse type. The landuse sites chosen by the LADPW are scattered around LA County. Locations of the sampling sites are shown in Appendix Figure A.8 for reference. In the following sections, sampling results of the landuse sites and the calculated runoff quantities of the GIS model are used to predict the total pollutant loads at the mass emission site of the upper Ballona Creek Watershed.

3.5.1 LADPW Landuse Sampling Sites

Eight landuse sampling sites were selected for each landuse type describe in Section 3.4.1 by LADPW through the field survey conducted in 1996 (Woodward, 1996). The largest area of landuse monitoring site was 1341.6 hectare, which is shown in Figure 3.5.1. Seven other sites were selected and they are shown in the Appendix A.

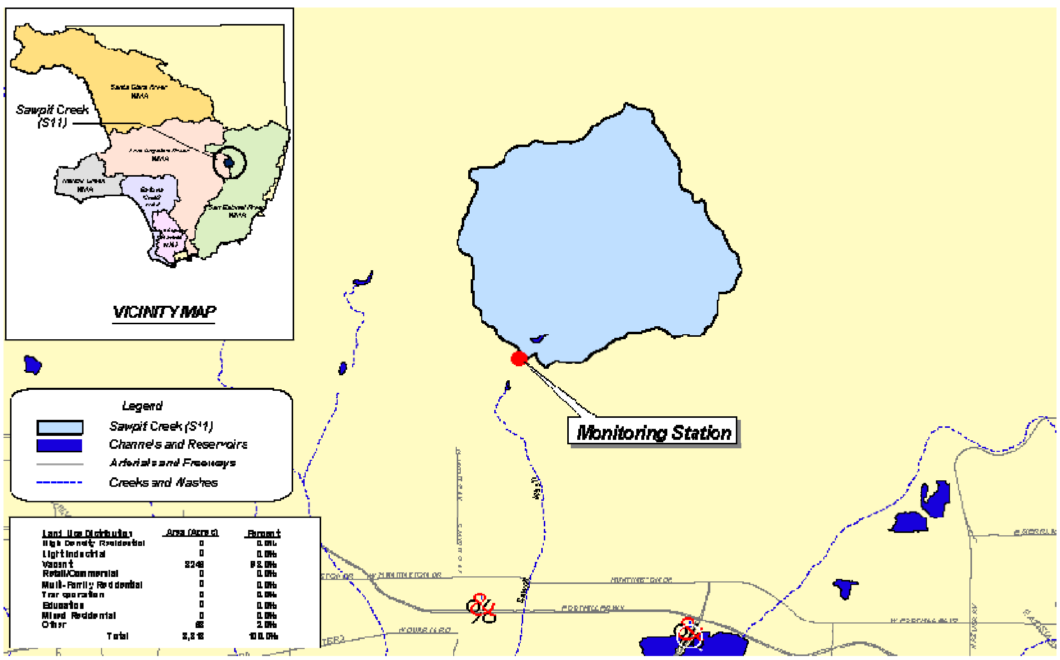


Figure 3.5.1 Vacant Landuse Monitoring Site at Monrovia Creek, Monrovia (LADPW)

The percent landuse distributions of the monitoring catchments were shown in Table 3.5.1. The percent of each representative landuse types for the monitoring catchments ranged from 54% (Retail/Commercial) to 100% (High Density Single Family Residential).

Table 3.5.1 Percent Landuse Distributions of Monitored Catchments (LADPW)

Station Name	Station No.	Drainage Area (Hectare)	V	LI	MR	T	RC	EF	HDR	MFR	Other
Santa Monica Pier	S08	33.7				4	<u>54</u>			5	37
Sawpit Creek	S11	1341.6	<u>98</u>								2
Project 620	S18	67.3							<u>100</u>		
Dominguez Channel	S23	349.7		18		<u>73</u>		1	1		7
Project 1202	S24	277.1	3	<u>74</u>		3			1		19
Project 474	S25	106.2						<u>90</u>	8	2	
Project 404	S26	88.1	1			2	17		6	<u>74</u>	
Project 156	S27	51.8			<u>77</u>		5	5	4	4	5

Underlined numbers indicate landuse types for each station.
Empty cells are land use percentages less than 0.5%
V = Vacant
HDR = High Density single-family Residential
MFR = Multi-Family Residential
MR = Mixed Residential
EF = Educational Facilities
T = Transportation
RC = Retail/Commercial
LI = Light Industrial

One site for each of the land uses was retained for continued sampling. The sampling protocol required that EMCs be measured, which are obtained from flow-weighted composite samples. Automatic samplers with flow meters were used over the duration of each storm event. The permit required continued sampling until one of two provisions were met: achieving an EMC at an error rate of 25% or less or detecting a constituent less than 25% of the time for 10 consecutive samples. Sampling can be discontinued for constituents that meet either criterion. Otherwise, 200 station samples are required. Cumulative EMCs for each landuse type over 1994-2000 storm seasons are shown in Appendix D.

3.5.2 Storm Characteristics of Landuse Catchments

As mentioned earlier, sampling sites for eight landuse types are scattered around the LA County. To apply analysis results of landuse types to the upper Ballona Creek watershed as a function of site specific conditions, four types of storm characteristics were used as independent variables: Antecedent Dry Days (ADD), total rainfall, rainfall duration, and maximum hourly rainfall intensity. Site specific characteristics that may influence pollutant loads such as subcatchment impervious fraction, depression storage, infiltration, and roughness coefficient were not used in the analyses to simplify input to the model and to avoid expensive and time consuming field surveys. The storm characteristics are calculated from hourly rainfall records downloaded from NCDC. For the storm characteristics calculation, a procedure was written in VBA and it is shown in Appendix E. Two parameters affect the storm characteristic calculations, which are:

1. Minimum rainfall that can be considered as a storm event.
2. Minimum dry period between two consecutive rainfall events to make these two events as separate storm events.

Effects of minimum rainfall parameter on total rainfall and ADD calculations are shown in Figure 3.5.2 and 3.5.3, respectively. For some events, total rainfall and ADD change a lot when the parameter becomes more than 1 mm. Also, runoff usually starts with 1 mm of rainfall from our field experience. Thus, 1 mm (0.04 in) of rainfall was chosen for the first parameter. For the second parameter, 6 hours was chosen. This parameter does not significantly affect the calculated results.

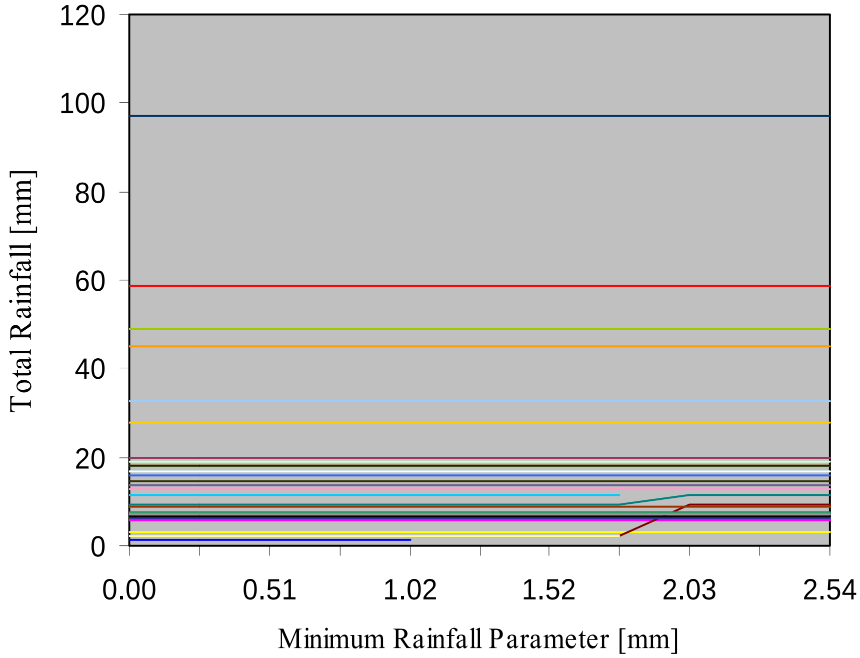


Figure 3.5.2 Effect of Minimum Rainfall Parameter to Total Rainfall Calculations

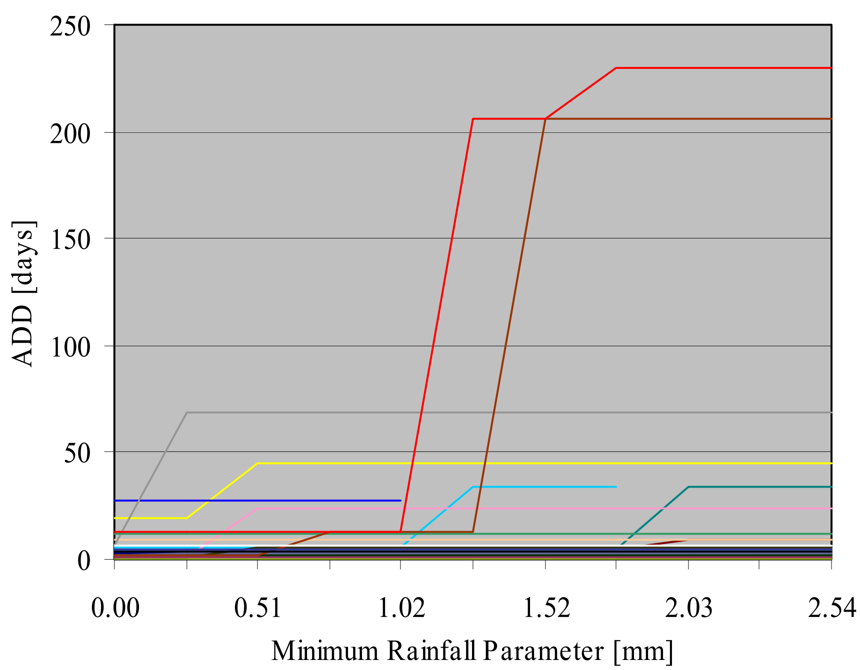


Figure 3.5.3 Effect of Minimum Rainfall Parameter to ADD Calculations

Table 3.5.2 summarizes storm characteristics of the sampled storm events from 1998 to 2001 for eight landuse type monitoring catchments, which are available from the LADPW web site. LADPW reported average storm intensities and NCDC rain gauge is not available around the Retail/Commercial monitoring catchments. Thus, maximum intensity data for this landuse type is not available. For the other sites, nearby NCDC rainfall data was used to calculate storm characteristics including hourly maximum intensities.

Table 3.5.2 Ranges of Storm Characteristics of the Sampled Storm Events for Landuse Type Monitoring Catchments

Landuse Types	ADD [days]	Rainfall [mm]	Duration [hrs]	Intensity [mm/hr] (1)	N (2)
Vacant	0.3 – 158.1	2.5 – 74.7	1 - 35	0.5 – 12.2	22
Light Industrial	0.3 – 122.8	1.5 – 50.3	2 - 26	1.0 – 18.3	31
Mixed Residential	0.4 – 157.8	2.5 – 66.0	1 - 37	2.5 – 17.8	25
Transportation	0.3 – 175.8	2.0 – 46.7	3 - 33	1.0 – 17.3	33
Retail/ Commercial	1.2 – 179.0	5.1 – 35.6	1 – 45	(3)	14
Educational Facilities	0.3 – 157.0	3.3 – 132.3	1 - 77	1.0 – 12.2	21
High Density Single Family Residential	0.4 – 157.8	2.5 – 66.0	1 - 37	2.5 – 12.7	13
Multi-Family Residential	0.3 – 158.1	1.3 – 74.7	2 - 35	0.5 – 11.7	25

(1) Peak one-hour intensity of a storm event in inches/hour calculated from the NCDC hourly rainfall data.

(2) Number of storm events used (LADPW stormwater runoff monitoring data from1998 to 2001).

(3) Data not available.

Although ANNs are more suitable than regression models for non-linear patterns and irregular seasonal variation in a data set as mentioned in Section 2.3.2, the number of data needed for proper evaluation of an ANN should be over 50 (Reich et al., 1999). The numbers of monitored storm events for each landuse types that are available for water quality analysis are limited as shown in Table 3.5.2. Therefore, regression analyses were performed to characterize landuse contributions to total loads of upper Ballona Creek watershed in the following section.

3.5.3 Multiple Linear Regression Models of the Monitored Landuse EMCs

Multiple linear regression analyses were performed to derive correlations between Total Zn landuse EMCs and the storm characteristics discussed in the previous section. The results are shown in Table 3.5.3. The EMC unit calculated with the regression models is $\mu g / L$.

Table 3.5.3 Multiple Linear Regression Models of Total Zn EMCs [μg / L] for LADPW Landuse Monitoring Catchments

Landuse Types	Coefficients										R^2	N (3)	Models (4)
	Intercept	P (1)	ADD [days]	P	Rainfall [cm]	P	Duration [hrs]	P	Intensity [cm/hr] (2)	P			
Vacant (5)	25.00											22	Constant
Light Industrial	202.66	0.08	24.60	0.00	-1.44	0.99	-1.25	0.89	-0.39	0.99	0.84	31	Linear
Mixed Residential	123.29	0.00	3.51	0.00	-5.75	0.78	1.30	0.69	-36.73	0.48	0.85	25	Linear
Transportation	241.51	0.00	7.34	0.00	-77.95	0.26	-3.61	0.53	189.25	0.22	0.88	33	Linear
Retail/ Commercial	(6) 218.45	0.00			16.84	0.10	-309.64	0.07			0.27	14	Inverse
Educational Facilities	64.81	0.00			2.22	0.17					0.10	21	Inverse
High Density Single Family Residential	24.53	0.05			-1.98	0.03			2.10	0.05	0.39	13	Inverse
Multi-Family Residential	56.96	0.00							0.79	0.00	0.37	25	Inverse

(1) Significance.

(2) Peak one-hour intensity of a storm event in centimeter per hour calculated from the NCDC hourly rainfall data.

(2) Number of storm events used (LADPW stormwater runoff monitoring data from1998 to 2000 storm seasons).

(3) Best fit regression models for each landuse types.

(4) 21 out of 22 events were below detection limit, 50 μg / L .

(5) Underlined coefficients were used as optimization parameters.

Total Zn EMCs of the Light Industrial, Mixed Residential, and Transportation landuse types showed high correlations with storm characteristics ($R^2 > 0.8$), while Retail/Commercial, Educational Facilities, High Density Single Family Residential, and Multi-Family Residential landuse types showed low correlations ($R^2 < 0.4$). Regression coefficients of the landuse types that showed low correlations were used as part of optimization parameters since EMCs calculated with these coefficients may not represent the EMCs of the corresponding landuse types. Total Zn EMCs for the Vacant landuse were mostly below detection limit. Thus, half of the detection limit (25 $\mu g/L$) was used as a constant for the total load prediction of a storm event for the upper Ballona Creek watershed.

3.5.4 Optimization of the Upper Ballona Creek Total Loadings

As mentioned in Section 3.5.1, the percent of each representative landuse types for the monitoring catchments ranged from 54% (Retail/Commercial) to 100% (High Density Single Family Residential). As a result, the regression models developed from these eight landuse types does not represent single landuse type. In contrast, total runoff volumes calculated for each landuse type by the GIS model discussed in Section 3.4 represents runoff volumes from single landuse types. Thus, to apply the regression models to the total loading calculations of the upper Ballona Creek watershed, optimization techniques are needed. Optimization techniques used for this purpose are described in the following sections.

3.5.4.1 Data Screening

Outlier detection is important for effective modeling. If all the available data are included in the model fitting, the fitted model may be poor. Outliers can be attributable to equipment errors and human errors. For example, rain gauge, automatic sampler, or stage monitoring equipment, may malfunction and this can be especially important for stormwater monitoring, since stormwater may contain debris and other material that can foul or plug sensors. Also, a measurement might be correct but represents a rare event. There may also be unusual human or animal activities which cause abnormal quality peaks. For example, wildfires in Southern California may cause episodic high loads. To determine the potential impacts of outliers caused by possible error sources, Cook's distance versus leverage plot was used (Cook, 1977; 1982).

Cook's distance is a measure of the influence of a case. It measures the effect of deleting a given observation. Cook's distance for the *i*-th observation is based on the differences between the predicted responses from the model constructed from all of the data and the predicted responses from the model constructed by deleting the *i*-th observation. Leverage values are the measures of the influence of a point on the fit of regression. High leverage gives an extra weight in the computation of regression line. High Cook's distance indicates that the case affects the slope of the regression line. Thus, cases with high leverage and high Cook's distance were identified as outliers in a preliminary test. Figure 3.5.4 shows an example plot of Cook's distance versus leverage for TSS. The identified outlier in this example is a storm event that occurred on 1/10/2001. Figure

3.5.5 shows the residual frequency distribution constructed from the dataset without the identified outlier. If the residual frequency distribution does not follow normal curve, Cook's distance versus leverage plot needs to be reconstructed from the dataset without identified outliers. In this case, the frequency distribution is a little skewed to the right. Thus, Cook's distance versus leverage plot is needed again with the remaining dataset to check existence of outliers. The procedure should continue until all outliers are removed.

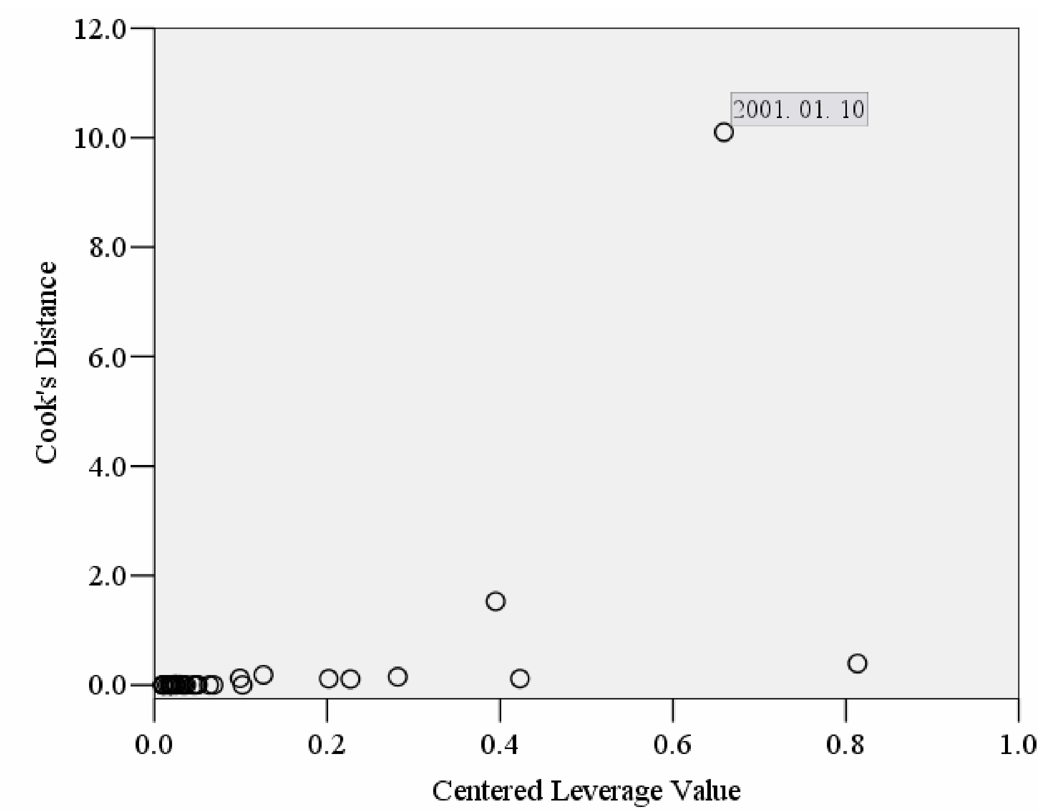


Figure 3.5.4 Cook's Distance versus Leverage.

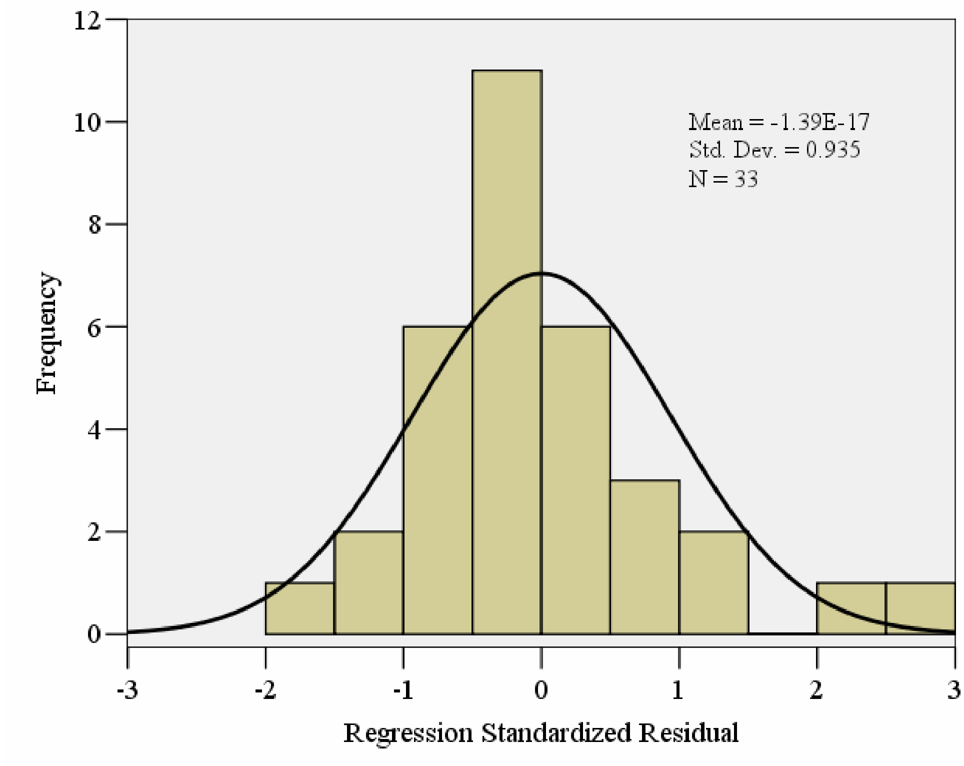


Figure 3.5.5 Regression Standardized Residual Frequency Distribution.

3.5.4.2 Limited memory Broyden-Fletcher-Goldfarb-Shanno Bound constrained (L-BFGS-B) Nonlinear Optimization Technique

L-BFGS-B was developed at the Optimization Technology Center, a joint venture of Argonne National Laboratory and Northwestern University (Zhu et. al., 1994). This algorithm is a quasi-Newton algorithm good for a limited memory large scale bound constrained optimization problem. The user is required to calculate the objective function value and its gradient.

Outline of the algorithm:

1. A quadratic function is defined using the objective and the gradient functions.
2. Gradient projection method is used to identify a set of active variables.
3. Quadratic model is approximately minimized with respect to the free variables.
4. Search direction is defined to be the vector leading from the current iterate to this approximate minimizer.
5. Line search is performed along the search direction.

Advantages of L-BFGS-B:

1. The structure of the objective function is not needed.
2. The storage requirements are modest and can be controlled by the user.
3. The cost of the iteration is low.

Disadvantages of L-BFGS-B:

1. A large number of function evaluations may be required for problems that converge slowly.
2. On highly ill-conditioned problems, it may fail to converge.
3. It cannot use knowledge about the structure of the problem to accelerate convergence.

Detailed algorithm is described by Byrd (1994). The ending criteria is based on the change of the objective function F or projected gradient, which is the projection of the gradient vector onto the space tangent to the active bounds. The program terminates when changes of objective function F or the norm of the projected gradient becomes sufficiently small. The projected gradient is zero at a local minimum of the bound constrained problem.

3.5.4.3 Zeroth-Order Regularization Method

Solutions of ill-posed inverse problems inevitably end up with active bounds. This active bounds problem can be avoided with the zeroth-order regularization method (Press et. al., 1986). The objective function shown in Eq. 3.4 is to minimize both the residual norm (first term) and the solution norm (second term). The regularization parameter λ in the equation acts as a weight between these two norms and is often acquired empirically. Increasing λ pulls the solution away from minimizing residual norm in favor of minimizing solution norm, which means initial values of optimization parameters are weighted high. Also, relative least square form is used in the objective function to compensate strong contributions of high magnitude state variables (Saez et. al., 1992).

Objective Function:

$$F = \sum_i \left(\frac{P_i - M_i}{P_i} \right)^2 + \lambda \times \sum_n \left(\frac{x_n - x_n^*}{x_n^*} \right)^2 \qquad (3.4)$$

$$P_i = \sum_j x_j (EMC_j \times Q_j) \qquad (3.5)$$

$$EMC_j = C_{1,j} + C_{2,j} \times ADD^k + C_{3,j} \times R^k + C_{4,j} \times D^k + C_{5,j} \times I^k \qquad (3.6)$$

where, I = Subscript, which denotes storm events

P_i = Predicted total loadings using regression equations for landuse types [Kg]

M_i = Measured total loadings [Kg]

λ = Regularization parameter

n = Subscript, which denotes optimization parameters

x_n, x_j = Optimization parameters

x_n^* = Initial values of optimization parameters

j = Subscript, which denotes landuse types

EMC_j = Event mean concentration for each landuse types

Q_j = Total runoff volume from each landuse types

$C_{1..5,\ j}$ = Constants from regression equations. These constants also serve as optimization parameters for the Retail/Commercial, Educational Facilities, High Density Single Family Residential, Multi-Family Residential landuse types. These constants become initial values when used as optimization parameters.

K = Superscript, which denotes

0 for Vacant landuse type,

1 for Light Industrial, Mixed Residential, and Transportation landuse types,

-1 for the other landuse types

ADD = Antecedent dry days [day]

R = Rainfall [in]

D = Storm duration [hours]

I = Maximum intensity [in/hr]

The optimization parameters x_j in Eq. (3.5) behave as weighting factors. Contributions of each landuse types to the total loadings of upper Ballona Creek watershed will be adjusted by these weighting factors x_j. This is to account for the variability of the percent of each representative landuse types for the monitoring catchments, which were ranged from 54% to 100%, as mentioned in the beginning of Section 3.5.4.

3.5.4.4 Initial Values and Bound Constraints of Optimization Parameters

Initial values of the optimization parameters x_n were given with the underlined coefficients of the regression equations shown in Table 3.5.3. Standard deviations of regressions were used as bound constraints for these parameters. Example of these standard deviations for the Retail/Commercial landuse is shown in Table 3.5.4. Also, the initial values of the optimization parameters x_j were given all ones. The upper and lower bounds were 1.7 and 0.3, respectively.

Table 3.5.4 Example of Total Zn Regression Analysis for the Retail/Commercial Landuse

Model	Unstandardized Coefficients		Standardized Coefficients	Sig.
	B	Std. Error	Beta	
Constant	218.45	56.92		.00
Rainfall	108.66	60.72	.92	.10
Duration	-309.64	155.22	-1.02	.07

3.5.4.5 The L-Curve

A graphical tool for the analysis of ill-posed problems is called L-curve, which is shown in Figure 3.5.6. This figure was constructed by trying many different λ values in the L-BFGS-B optimization program and plotting calculated solution and residual norms in terms of λ values.

Solutions of real world problems are always contaminated by various types of perturbation errors, such as measurement errors and approximation errors. The L-curve shows the compromise between minimization of perturbation error and the error caused by regularization. For the horizontal part of the curve, solution changes a little with the regularization parameter, which means that the error caused by regularization dominates. In contrast, the vertical part of the curve corresponds to solutions that are dominated by the perturbation error, where the solution norm changes much with the regularization parameter and residual norm changes little.

Hansen (1992) showed that the optimal regularization parameter is not far from the regularization parameter that corresponds to the L-curve's corner. Thus, by selecting λ=0.5, we can compute regularized solution with optimal balance between perturbation errors and regularization errors.

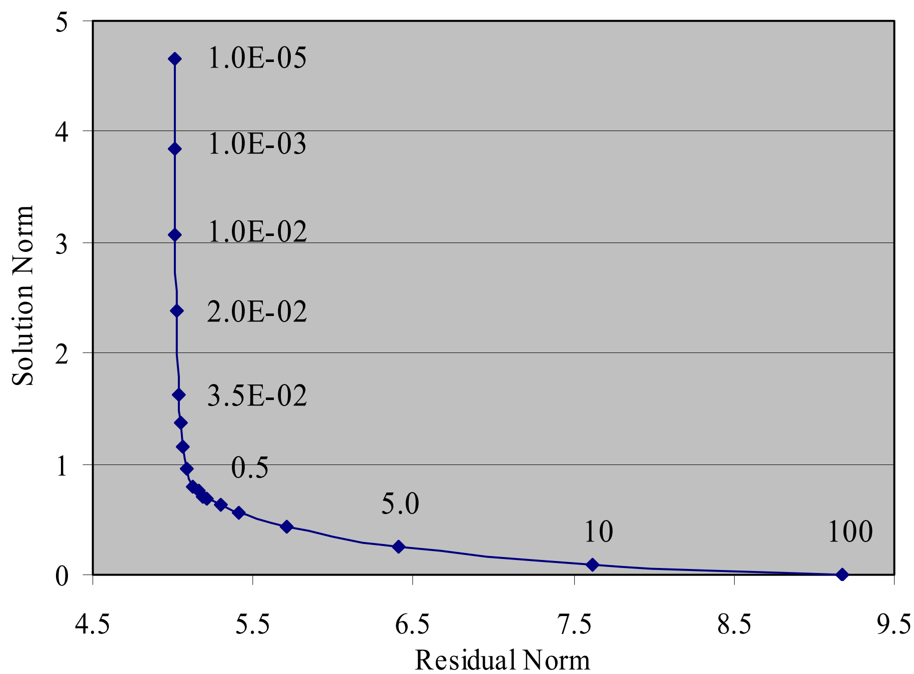


Figure 3.5.6 The L-Curve with λ Values as Data Labels.

3.5.5 Stormwater Runoff Quality Prediction using the Optimization Techniques

Preliminary regression analysis was performed with the available 36 storm events (See Section 3.4.5) and storm characteristics calculated with LA Civic Center rainfall records

(See Section 3.4.3.1) for the total Zn measured at the emission site of upper Ballona Creek watershed. Two outliers (1/24/2001, 10/31/2003) were dropped by this preliminary test (See Section 3.5.4.1). For two storm events (4/6/1999, 2/20/2000), no rainfall records were available from the LA Civic Center rainfall gauge station. Among the remaining 32 storm events, 25 events from 1998 to 2001 were used for optimization and 7 events from 2001 to 2004 were used for validation.

The regularization parameter λ used was 0.5, which corresponds to the corner of the L-curve as mentioned in the previous section. Table 3.5.5 shows optimization program parameters after a successful execution. There were no active bounds. The program stopped execution because the norm of projected gradient was sufficiently small.

Table 3.5.5 Test Results of L-BFGS-B for 18 Parameters

Parameters	Iterations	Segments Explored	Active Bounds	Norm of Projected Gradient	F
18	209	216	0	6.767×10^{-6}	5.529

The optimized 18 parameters are shown in Table 3.5.6. The percent of representative landuse for each landuse types, which was shown in Table 3.5.1, are shown again for convenience. In general, weighting factors were closer to unity when the representative landuses for corresponding landuse types were closer to 100%.

Table 3.5.6 Optimized Parameters and Regression Coefficients

Landuse Types	Coefficients					x_j	Percent of Representative Landuse
	Intercept	ADD [days]	Rainfall [cm]	Duration [hrs]	Intensity [cm/hr]		
Vacant	25.00					0.96	98
Light Industrial	202.66	24.60	-1.44	-1.25	-0.39	0.63	74
Mixed Residential	123.29	3.51	-5.75	1.30	-36.73	0.68	77
Transportation	241.51	7.34	-77.95	-3.61	189.25	0.94	73
Retail/ Commercial	167.42		16.24	-309.40		0.32	54
Educational Facilities	62.13		2.22			0.96	90
High Density Single Family Residential	22.21		-2.01		2.10	0.92	100
Multi-Family Residential	52.13				0.79	0.92	74

* Underlined numbers indicate optimized parameters.

Total Zn loadings of the upper Ballona Creek watershed for the 32 storm events calculated using the optimized regression coefficients and weighting factors along with the storm characteristics calculated from the LA Civic Center rainfall records were plotted against the measured loadings in Figure 3.5.7. The diamonds indicate 25 storm events used in optimization and the triangles indicate 7 storm events used to validate the optimized quality prediction model. Predicted and measured loadings showed a reasonably good agreement without any site specific information. Furthermore, this model can be applied to LA County since the landuse monitoring sites are scattered over the LA County.

To illustrate effectiveness of the regularization method, 10 regression coefficients of the landuse types for R-square values less than 0.4 were used in optimization without regularization. In this analysis, all of the 32 storm events were used in the optimization process. Nevertheless, the predictions of total Zn loadings generally over estimate the corresponding measured ones, which are shown in Figure 3.5.8. The RMS error of the optimized model without regularization for total Zn was 24 Kgs and only 9 Kgs with regularization, which is low compared to the mean discharge of 77 Kgs per storm event.

To check the validity of the optimized regression coefficients, EMCs of Retail/Commercial landuse monitoring catchments were calculated using the coefficients optimized with regularization and the results are shown in Figure 3.5.9. For the same landuse type, EMCs were calculated using the coefficients optimized without regularization and the results are shown in Figure 3.5.10. Again, calculated EMCs using optimization with regularization show better predictions overall than the calculated EMCs using optimization alone.

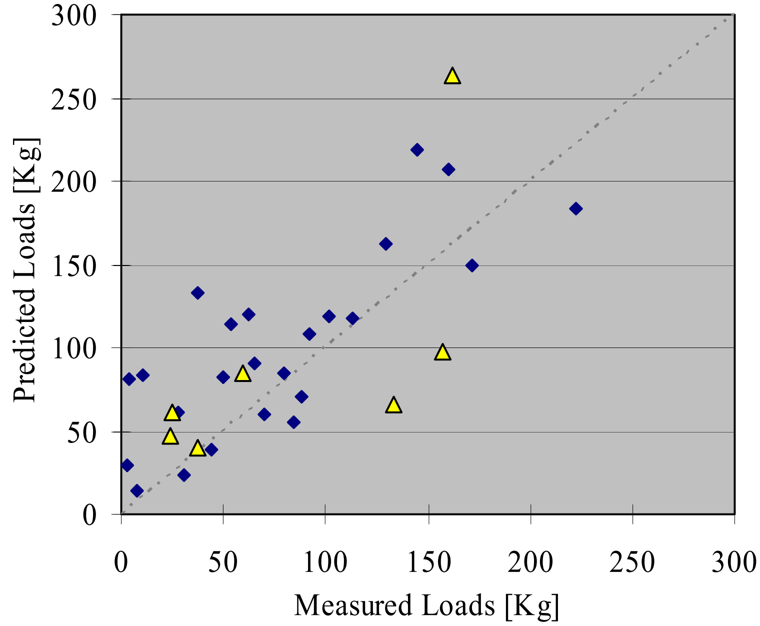


Figure 3.5.7 Comparisons of Predicted and Measured Loads of Total Zn for the Upper Ballona Creek Watershed using Regularization and 18 Optimization Parameters.

◆ : 25 storm events used for optimization (1998 – 2000 storm seasons).

△ : 7 storm events used for validation (2001 – 2003 storm seasons).

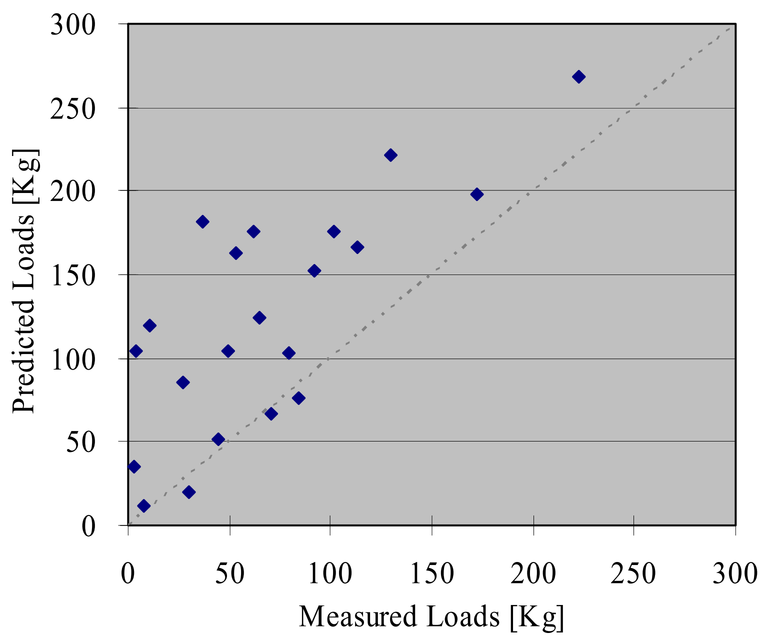


Figure 3.5.8 Comparisons of Predicted and Measured Loads of Total Zn for the Upper Ballona Creek Watershed using 10 Optimization Parameters.

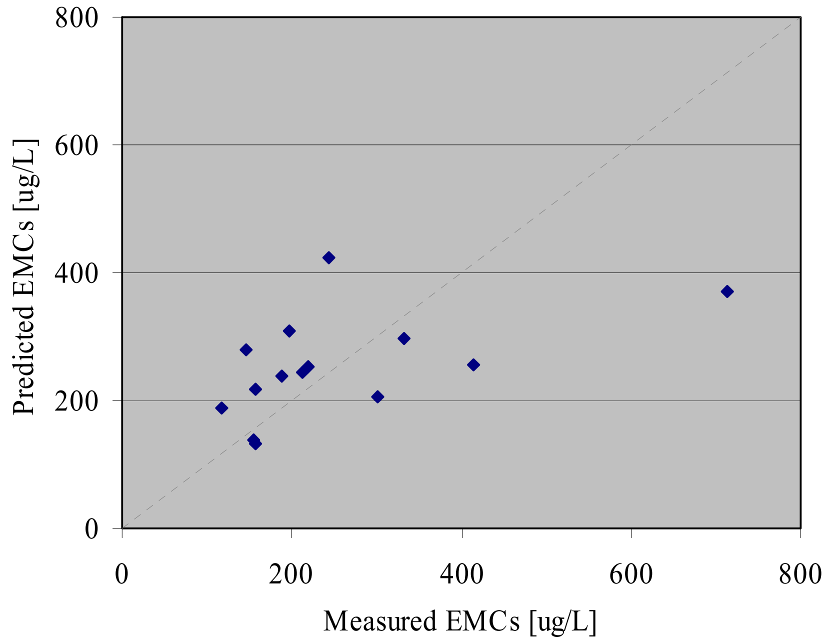


Figure 3.5.9 Comparisons of Predicted and Measured EMCs of Total Zn for Retail/Commercial Landuse using Regularization and 18 Optimization Parameters.

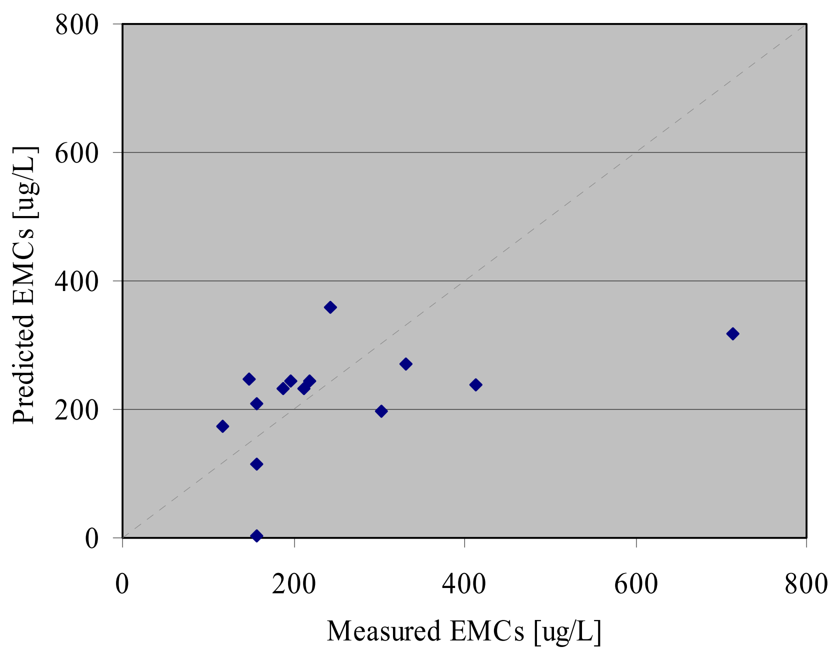


Figure 3.5.10 Comparisons of Predicted and Measured EMCs of Total Zn for Retail/Commercial Landuse using 10 Optimization Parameters.

4. DISCUSSION

The approach used in this dissertation can be used for a wide range of pollutants. In order to produce a manageable problem, a subset of the common pollutants was chosen. Among the conventional pollutants, metals and nutrients, Total Zn, Total Kjeldahl Nitrogen (TKN), and Total Suspended Solids (TSS) were chosen for study.

Trace metals have toxic effects on aquatic plants and animals. They accumulate in the sediments of rivers, lakes, and seabed as well as in fish tissue. Zinc is a byproduct of vehicle tire wear and motor oils. It is also produced by the corrosion of galvanized iron. A TMDL was set for total zinc for the Ballona Creek watershed by Los Angeles Regional Water Quality Control Board and US EPA Region 9 in 2005. Metals are usually more toxic to aquatic organisms in their dissolved form. Total zinc was used for the TDML as a compromise since it is less costly to measure the total form than measuring both forms. Recent study showed that there is a strong relationship between dissolved and total zinc. Often the fraction of the total concentration that is dissolved is a predictable fraction and Han et al. (2006a) showed that this was true for zinc in highway runoff. The Los Angeles Regional Water Quality Control Board, by regressing LADPW’s composite stormwater samples from December 1994 to January 2002, showed that the dissolved zinc averaged 79% of the total zinc.

TKN is contained in animal and human waste, decaying organic matter, and living organic material like algae cells. Nitrogen promotes algae blooms, which harm aquatic life by depleting oxygen in the water and decreasing light penetration for desirable photosynthetic organisms. TSS are solids in water that can be trapped by a 0.45 μ filter. High concentrations of TSS decrease light penetration and increase water temperature because the suspended particles absorb heat from sunlight. Also, high TSS in a water bodies is often associated with higher concentrations of bacteria, nutrients, and metals. Model results for these three pollutants are discussed in this section.

4.1 REGRESSION MODELS OF EMCs FOR LANDUSE MONITORING CATCHMENTS

To characterize pollutant contributions of landuse types to total loads of the upper Ballona Creek, multiple linear regression analyses were performed using EMCs obtained from the landuse monitoring catchments. Four types of storm characteristics, which are antecedent dry days, total rainfall, rainfall duration, and maximum hourly rainfall intensity, were used as independent variables. Site specific characteristics were not used in the analyses to avoid expensive and time consuming field surveys. Developed regression models for TKN and TSS EMCs are shown in Tables 4.1.1 to 4.1.4 along with standard errors and significances of regression coefficients. Regression models for Total Zn EMCs are shown in Table 3.5.3 (See Section 3.5.3).

Table 4.1.1 Multiple Linear Regression Models of TKN EMCs [mg / L] for LADPW Landuse Monitoring Catchments

Landuse Types	Coefficients										R^2 (3)	N (4)	Models (5)
	Intercept	P (1)	ADD [days]	P	Rainfall [cm]	P	Duration [hrs]	P	Intensity [cm/hr] (2)	P			
Vacant	1.00	0.00	0.075	0.52	0.093	0.06	-3.32	0.09	-0.03	0.11	<u>0.25</u>	22	Inverse
Light Industrial	2.45	0.00	0.062	0.00	-0.869	0.12	0.01	0.81	1.31	0.42	0.62	32	Linear
Mixed Residential	3.19	0.00	0.033	0.00	0.419	0.26	-0.07	0.22	-2.30	0.02	0.71	24	Linear
Transportation	2.01	0.00	0.035	0.00	-1.091	0.05	0.05	0.29	1.19	0.33	0.75	35	Linear
Retail/ Commercial	8.65	0.00	0.003	0.91	-2.120	0.08	0.01	0.92	(6)		<u>0.42</u>	14	Linear
Educational Facilities	1.49	0.00	0.052	0.00	0.117	0.55	-0.02	0.32	-0.19	0.79	0.85	21	Linear
High Density Single Family Residential	3.59	0.02	0.064	0.00	0.140	0.89	-0.04	0.79	-2.46	0.38	0.80	13	Linear
Multi-Family Residential	1.89	0.00	-0.306	0.20	0.128	0.04	-2.79	0.41	-0.03	0.18	<u>0.48</u>	22	Inverse

(1) Significance of independent variables.

(2) Peak one-hour intensity of a storm event in centimeter per hour calculated from the NCDC hourly rainfall data.

(3) Regression coefficients of underlined R-square values ($R^2 < 0.6$) were used as optimization parameters.

(4) Number of storm events used (LADPW stormwater runoff monitoring data from1998 to 2001).

(5) Best fit regression models for each landuse types.

(6) Data not available (LADPW does not measure maximum intensity and NCDC rain gauge station is not available near this site).

Table 4.1.2 Standard Error of TKN Regression Coefficients

Landuse Types	Station No.	Intercept	ADD	Rainfall	Duration	Intensity
Vacant	S11	0.24	0.115	0.30	1.85	0.11
Light Industrial	S24	0.56	0.011	3.53	0.05	10.35
Mixed Residential	S27	0.45	0.006	2.33	0.06	5.85
Transportation	S23	0.37	0.004	3.41	0.04	7.73
Retail/Commercial	S08	1.56	0.023	6.88	0.10	n/a
Educational Facilities	S25	0.42	0.006	1.24	0.02	4.54
High Density Single Family Residential	S18	1.22	0.013	6.59	0.15	17.12
Multi-Family Residential	S26	0.45	0.230	0.37	3.31	0.13

* n/a: Data not available

Table 4.1.3 Multiple Linear Regression Models of TSS EMCs [mg / L] for LADPW Landuse Monitoring Catchments

Landuse Types	Coefficients										R^2 (3)	N (4)	Models (5)
	Intercept	P (1)	ADD [days]	P	Rainfall [cm]	P	Duration [hrs]	P	Intensity [cm/hr] (2)	P			
Vacant	232.70	0.03	23.33	0.63	32.19	0.12	-1299.81	0.07	-8.23	0.25	0.26	24	Inverse
Light Industrial	156.18	0.00	5.68	0.00	-50.14	0.17	-1.34	0.65	100.78	0.34	0.50	30	Linear
Mixed Residential	73.05	0.03	2.22	0.01	3.17	0.90	-2.23	0.61	-15.57	0.79	0.44	22	Linear
Transportation	59.87	0.08	1.57	0.00	-18.60	0.69	1.51	0.70	28.28	0.79	0.38	35	Linear
Retail/ Commercial	90.71	0.01	-150.35	0.27	7.48	0.05	-153.81	0.03	(6)		0.44	13	Inverse
Educational Facilities	127.60	0.00	-0.35	0.47	38.14	0.20	-4.33	0.19	-34.13	0.61	0.19	20	Linear
High Density Single Family Residential	76.36	0.17	-10.55	0.59	-67.64	0.03	1684.44	0.02	2.58	0.51	0.58	12	Inverse
Multi-Family Residential	-20.18	0.60	12.43	0.41	3.26	0.34	514.02	0.10	-1.05	0.39	0.45	21	Inverse

(1) Significance of independent variables.

(2) Peak one-hour intensity of a storm event in centimeter per hour calculated from the NCDC hourly rainfall data.

(3) All regression coefficients were used as optimization parameters since R-square values are low for all landuse types ($R^2 < 0.6$).

(4) Number of storm events used (LADPW stormwater runoff monitoring data from1998 to 2001).

(5) Best fit regression models for each landuse types.

(6) Data not available (LADPW does not measure maximum intensity and NCDC rain gauge station is not available near this site).

Table 4.1.4 Standard Error of TSS Regression Coefficients

Landuse Types	Station No.	Intercept	ADD	Rainfall	Duration	Intensity
Vacant	S11	101.56	48.08	128.22	668.04	45.09
Light Industrial	S24	37.53	1.33	226.66	2.94	665.43
Mixed Residential	S27	30.46	0.74	161.71	4.27	377.68
Transportation	S23	32.82	0.38	300.07	3.87	680.40
Retail/Commercial	S08	29.61	126.89	20.84	57.46	n/a
Educational Facilities	S25	35.19	0.47	184.85	3.16	419.21
High Density Single Family Residential	S18	49.68	18.59	153.90	574.27	24.18
Multi-Family Residential	S26	37.95	14.74	21.29	291.78	7.55

* n/a: Data not available.

The strength of the relationship between regression model and dependent variables are explained with R-square values. R-square is called the multiple correlation or coefficient of multiple determinations and is the percent of the variance in the dependent variable explained by the independent variables. Table 3.5.3 shows that the regression explained more than 80% of the variability in total Zn EMCs for Light Industrial, Mixed Residential, and Transportation landuses. The TKN EMC correlations better explained ($R^2 > 0.8$) the variability in other landuse types: Educational Facilities and High Density Single Family Residential landuses. R-square values for the Retail/Commercial and Multi-Family Residential landuses were below 0.5 for both total Zn and TKN. For TSS EMC correlations, all R-square values were low ($R^2 < 0.6$).

Another measure of the regression model strength is the F test, which is used to test the significance of independent variables. If a significance, which is the probability of F, is less than 0.05, the variable is considered significantly better explanatory than would be expected by chance. In general, ADD was the most significant explanatory variable for the three pollutant EMCs analyzed. Total rainfall was the next significant variable and the maximum hourly rainfall intensity was the least significant for all three pollutant EMCs.

The number of storm events used in these correlations ranged from 12 to 35, which were measured for the storm season 1998 to 2001 by the LADPW. For most of the landuse types, the ranges of storm characteristics for the sampled storm events are limited as shown in Table 3.5.2 (See Section 3.5.2). For example, total rainfall for the

Retail/Commercial landuse ranged from 5.1 mm to 35.6 mm. Therefore, the developed regression models are applicable to storm events whose storm characteristics do not deviate too much from the ranges shown in Table 3.5.2. For example, the storm event occurred on 3/15/2003, whose total rainfall is 104 mm as shown in Table 3.4.3, does not fit with the developed regression models.

4.2 POLLUTANT LOAD ANALYSIS USING STORMWATER RUNOFF QUALITY MODELS

Characterization of pollutant load contributions of landuse types to total loads of the upper Ballona Creek was achieved through optimization techniques described in Section 3.5.4. Preliminary correlation analysis, L-curves, optimization parameters, and total load predictions for the three pollutants are described in this section.

4.2.1 Preliminary Correlations of Total Loads and Storm Characteristics

Multiple linear regression analyses were performed to determine if the total pollutant loadings can be described with storm characteristics. If total loadings of upper Ballona Creek watershed can be described with storm characteristics calculated from the LA Civic Center rainfall records, landuse contributions to the total loadings might be characterized. Also, data screening can be performed with these analyses, as mentioned in Section 3.5.4.1.

Results of regression analyses are shown in Table 4.2.1. The R-square value of total Zn loading was the highest (0.75) among the three pollutants but less than one-half of the variance in loadings was explained for TSS by the storm characteristics. Total rainfall was a significant variable for total Zn and TKN loadings, while maximum hourly rainfall intensity was significant for TSS loadings.

Table 4.2.1 Multiple Linear Regression Models of Total Pollutant Loads for Upper Ballona Creek Watershed

Pollutant Loads	Coefficients										R^2	N (3)
	Intercept	P (1)	ADD [days]	P	Rainfall [cm]	P	Duration [hrs]	P	Intensity [cm/hr] (2)	P		
Total Zn [Kg]	25.61	0.22	0.35	0.60	45.21	0.00	-0.76	0.46	-9.58	0.83	0.75	32
TKN [Ton]	-0.17	0.91	-0.02	0.40	1.71	0.03	-0.07	0.39	6.18	0.09	0.63	34
TSS [Ton]	-92.26	0.26	0.99	0.67	18.94	0.75	7.71	0.08	448.61	0.05	0.44	30

(1) Significance of independent variables.

(2) Peak one-hour intensity of a storm event calculated from the NCDC hourly rainfall data.

(3) Number of storm events used (LADPW stormwater runoff monitoring data from1998 to 2003 storm seasons).

(4) Identified outliers:

Total Zn: 1/24/2001 and 10/31/2003

TSS: 1/10/2001, 11/24/2001, 11/8/2002, and 10/31/2003.

4.2.2 L-Curves

The behavior of the L-curve is more easily observed using a log-log scale and the criterion for choosing the regularization parameter is to choose a point on this curve that is at the corner of the vertical piece as shown in Figure 4.2.1 to Figure 4.2.3 (Hansen and O'Leary, 1993). The figures show steep horizontal part, which is due to noisy data according to Hansen. For noiseless data, the L-curve on a log-log scale is flat. The λ values for total Zn, TKN, and TSS were 0.5, 7.6, and 9.7, respectively.

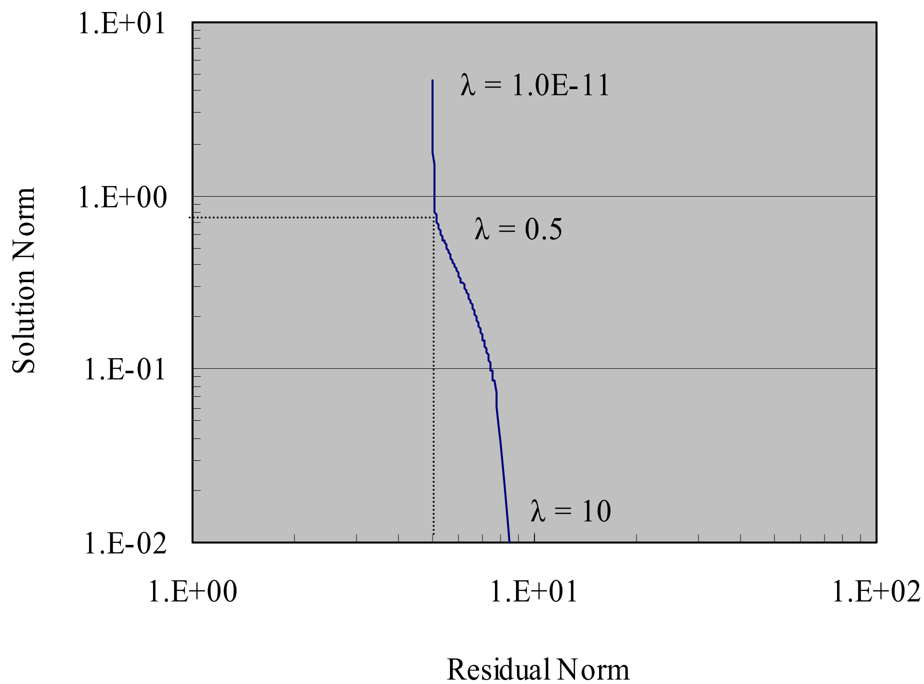


Figure 4.2.1 L-Curve for Total Zn

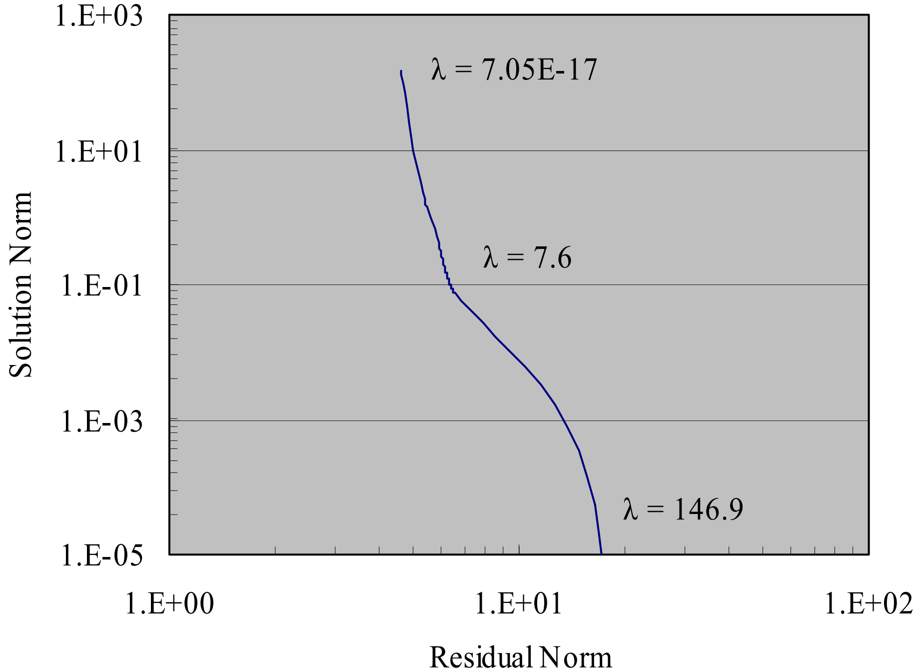


Figure 4.2.2 L-Curve for TKN

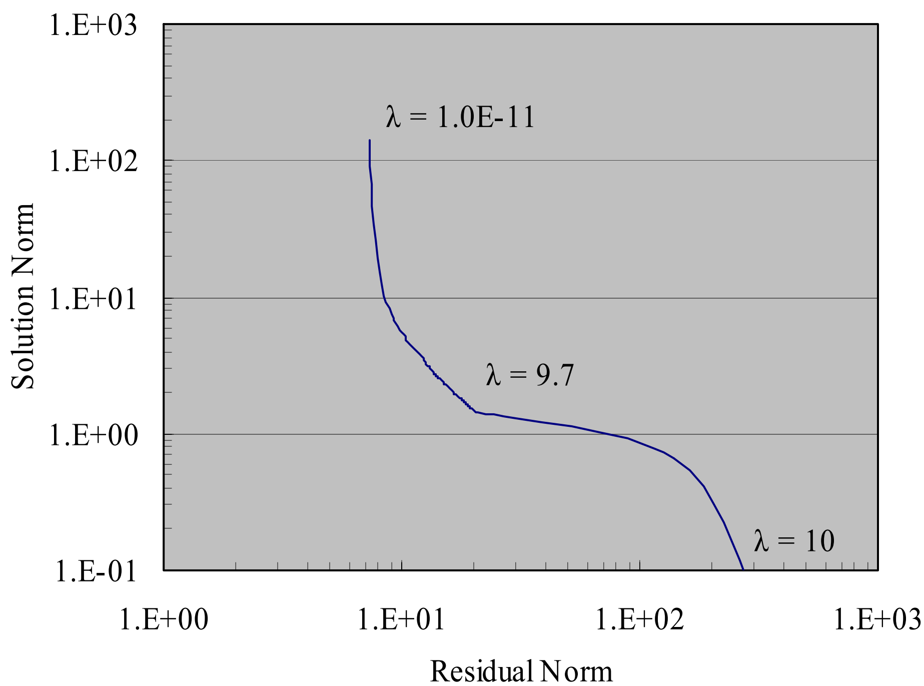


Figure 4.2.3 L-Curve for TSS

4.2.3 Optimized Parameters

Using the λ values obtained in the previous section, optimization parameters were calculated for TKN and TSS, and are shown in Table 4.2.2 and 4.2.3. Optimized parameters for total Zn are shown in Table 3.5.6. Regression coefficients of landuses with R-square values lower than 0.6 were used as optimization parameters for both TKN and TSS. That is, regression coefficients of Vacant, Retail/Commercial, Multi-Family Residential landuses for TKN and all regression coefficients for TSS were used as optimization parameters. The calculated residual norms for total Zn, TKN, and TSS were 5.1, 6.5, and 13.6, respectively. Thus, TSS correlation is expected to be worse than for the other pollutants. Landuse-type weighting factors for TKN showed the same tendency as total Zn. The weighting factor of the Retail/Commercial landuse, which has the lowest percent representative landuse, had the greatest deviation. However, weighting factors for TSS did not show any tendency.

Table 4.2.2 Optimized Parameters and Regression Coefficients for TKN EMCs

Landuse Types	Coefficients					x_j	Percent of Representative Landuse
	Intercept	ADD [days]	Rainfall [cm]	Duration [hrs]	Intensity [cm/hr]		
Vacant	1.01	0.075	0.09	-3.33	-0.03	0.997	98
Light Industrial	2.45	0.062	-2.21	0.01	3.31	1.012	74
Mixed Residential	3.19	0.033	1.06	-0.07	-5.85	0.994	77
Transportation	2.01	0.035	-2.77	0.05	3.03	1.005	73
Retail/ Commercial	10.21	0.003	-1.74	0.01	n/a	1.135	54
Educational Facilities	1.49	0.052	0.30	-0.02	-0.48	1.004	90
High Density Single Family Residential	3.59	0.064	0.36	-0.04	-6.25	1.005	100
Multi-Family Residential	1.93	-0.306	0.13	-2.81	-0.03	1.002	74

- Underlined numbers indicate optimized parameters.
- n/a: Data not available.

Table 4.2.3 Optimized Parameters for TSS EMCs

Landuse Types	Coefficients					x_j	Percent of Representative Landuse
	Intercept	ADD [days]	Rainfall [cm]	Duration [hrs]	Intensity [cm/hr]		
Vacant	285.91	23.63	34.86	-1205.52	-6.28	1.121	98
Light Industrial	176.91	5.77	-46.65	-1.31	103.82	1.099	74
Mixed Residential	87.50	2.26	3.21	-1.96	-15.34	1.133	77
Transportation	61.07	1.57	-18.41	1.53	28.38	1.025	73
Retail/ Commercial	111.68	-118.51	7.66	-150.96	n/a	1.115	54
Educational Facilities	135.54	-0.35	39.31	-4.15	-33.92	1.048	90
High Density Single Family Residential	97.13	-10.34	-44.83	2152.11	3.27	1.524	100
Multi-Family Residential	-19.80	12.51	3.27	526.07	-1.01	0.989	74

- n/a: Data not available.

4.2.4 Stormwater Runoff Quality Predictions

TKN and TSS loadings calculated for the upper Ballona Creek watershed using the optimized regression coefficients and weighting factors along with the storm characteristics calculated from the LA Civic Center rainfall records were plotted against the measured loadings in Figure 4.2.4 and 4.2.5, respectively. The diamonds indicate storm events used in optimization and the triangles indicate storm events used to validate the optimized prediction model. Corresponding total Zn loading comparisons were shown in Figure 3.5.7. Predicted and measured loadings for TKN showed reasonable agreement without using site specific information. The RMS error for TKN was 0.5 Kgs, which is low compared to the mean discharge of 2.9 Kgs per storm event.

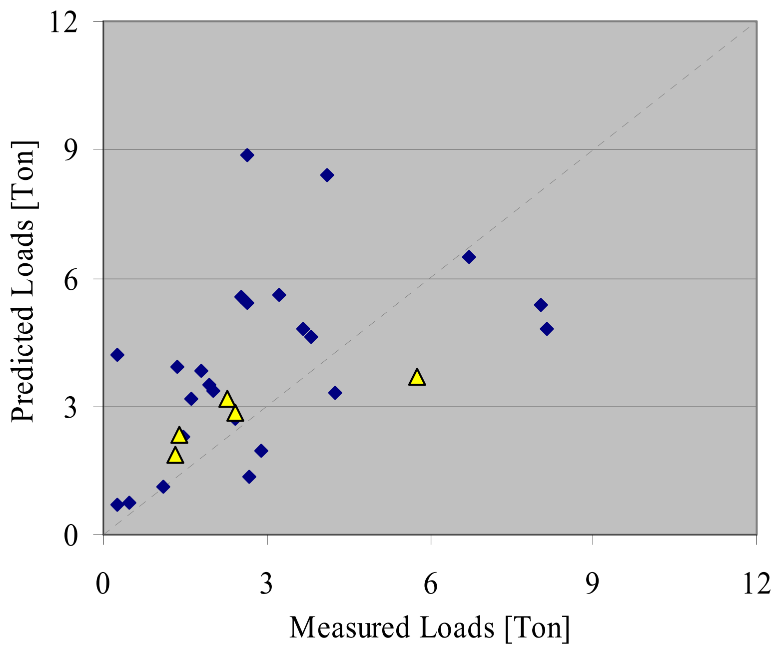


Figure 4.2.4 Comparisons of Predicted and Measured Loads of TKN for the Upper Ballona Creek Watershed using Regularization and 22 Optimization Parameters.

◆ : 24 storm events used for optimization (1998 – 2000 storm seasons).

△ : 5 storm events used for validation (2001 – 2003 storm seasons).

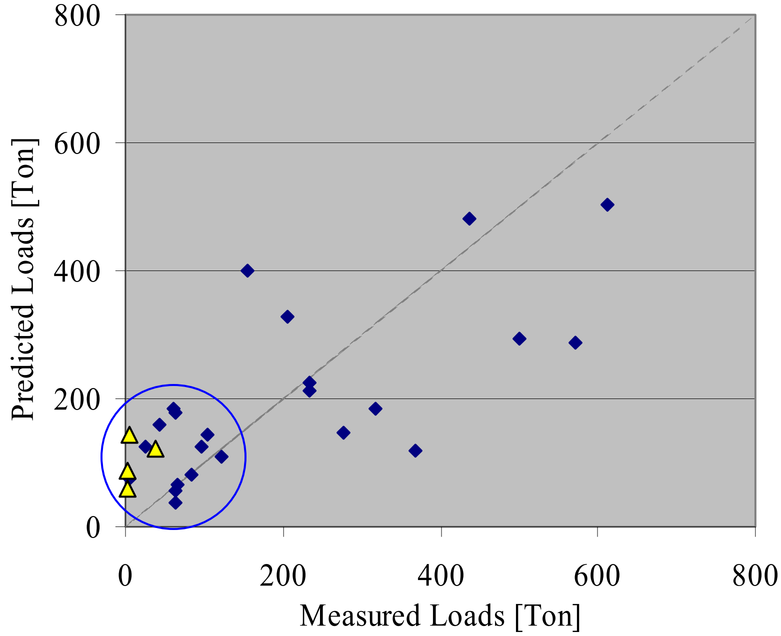
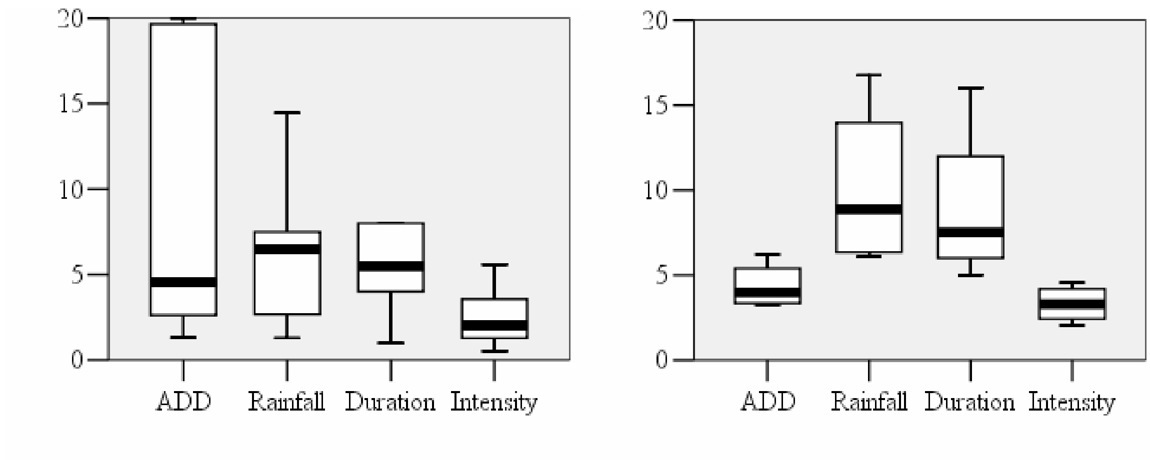


Figure 4.2.5 Comparisons of Predicted and Measured Loads of TSS for the Upper Ballona Creek Watershed using Regularization and 47 Optimization Parameters.
◆ : 24 storm events used for optimization (1998 – 2000 storm seasons).
△ : 4 storm events used for validation (2001 – 2003 storm seasons).

The RMS error for TSS was 25 Kgs for the mean discharge of 220 Kgs. However, measured TSS loads for the validation dataset, which are shown with triangles in Figure 4.2.5, were much lower than the predicted loads. To investigate possible causes to this discrepancy, box plots of the storm characteristics are shown in Figure 4.2.6 for optimization and validation datasets within the circle in the figure. If we look at the maximum hourly rainfall intensity, which is the most significant explanatory variable for TSS loads, median of validation dataset is higher than the median of optimization dataset. This means that the measured loads for the validation dataset should be higher than the measured loads for the optimization dataset. This discrepancy might be caused by measurement errors. Also, it might be an indication of other activities in the watershed, which would require further investigation to understand.



(a) Optimization Dataset (b) Validation Dataset

Figure 4.2.6 Box Plots of Storm Characteristics for Optimization and Validation Datasets. Lower and upper box boundaries: 25 and 75 percentiles, respectively; Top and bottom horizontal lines: Maximum and minimum of the observed values, respectively; Heavy black line: Median.

4.2.5 Landuse Contributions to Total Loads

Runoff volumes from eight LADPW landuse types were calculated using the model implemented in ArcGIS platform for Dec. 2, 2001 storm event as an example. These runoff volumes and the optimized model for total Zn described in previous section were used to calculate loadings of each landuse types. Percent contributions from eight

LADPW landuse types to the upper Ballona Creek watershed are shown in Figure 4.2.7. Light Industrial, Retail/Commercial, Mixed Residential, and High Density Single Family Residential landuses were the major pollutant contributors.

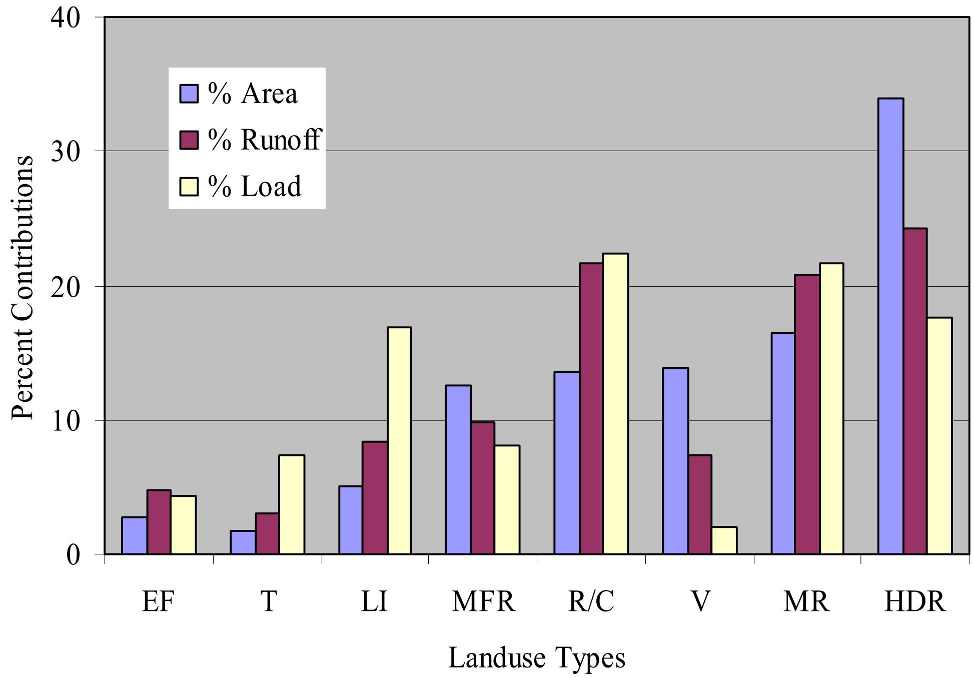


Figure 4.2.7 Percent Contributions of Landuses to Total Zn Load of Upper Ballona Creek

To better compare the pollution contributions as a function of landuse and area, the concept of "leverage" is used. Leverage is defined as the percent load divided by percent area and is shown in Figure 4.2.8 for various landuse types. The leverage is a good indicator of pollution level. If the watershed is uniformly polluted, leverage for all landuse types is equal to one. If the cost of applying BMPs is proportional to area, landuses with higher leverages, such as Transportation and Light Industrial landuses in Figure 4.2.8 are better candidates for further BMP considerations.

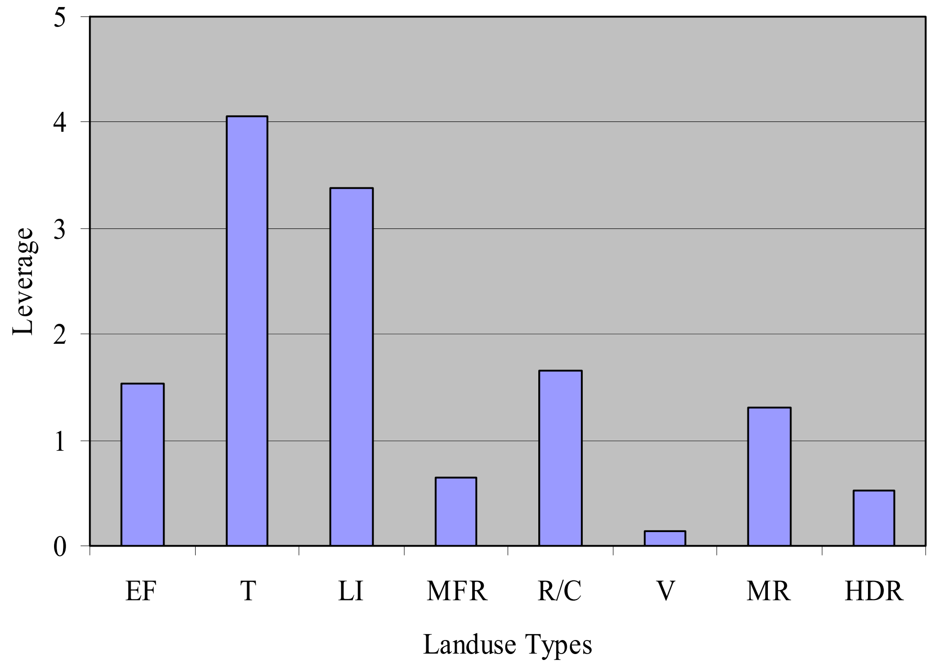


Figure 4.2.8 Total Zn Leverage for Eight Landuse Types

If landuses with higher leverages are clustered together, those areas would be the first target of BMP application. To identify whether there are clustered areas of higher leverage landuses, three landuse types with higher leverages, which are Transportation, Light Industrial, and Retail/Commercial landuses were chosen and the distributions of the three landuses in the watershed are shown in Figure 4.2.9. Among the three landuse types shown in Figure 4.2.9, Light Industrial landuse shows some clusters. Two candidate subwatersheds were identified for BMP considerations and showed in the figure with its drain locations. The subwatersheds in the figure were named as Industrial and Downtown, which are labeled in the figure. These subwatershed names will be used in the subsequent discussions.

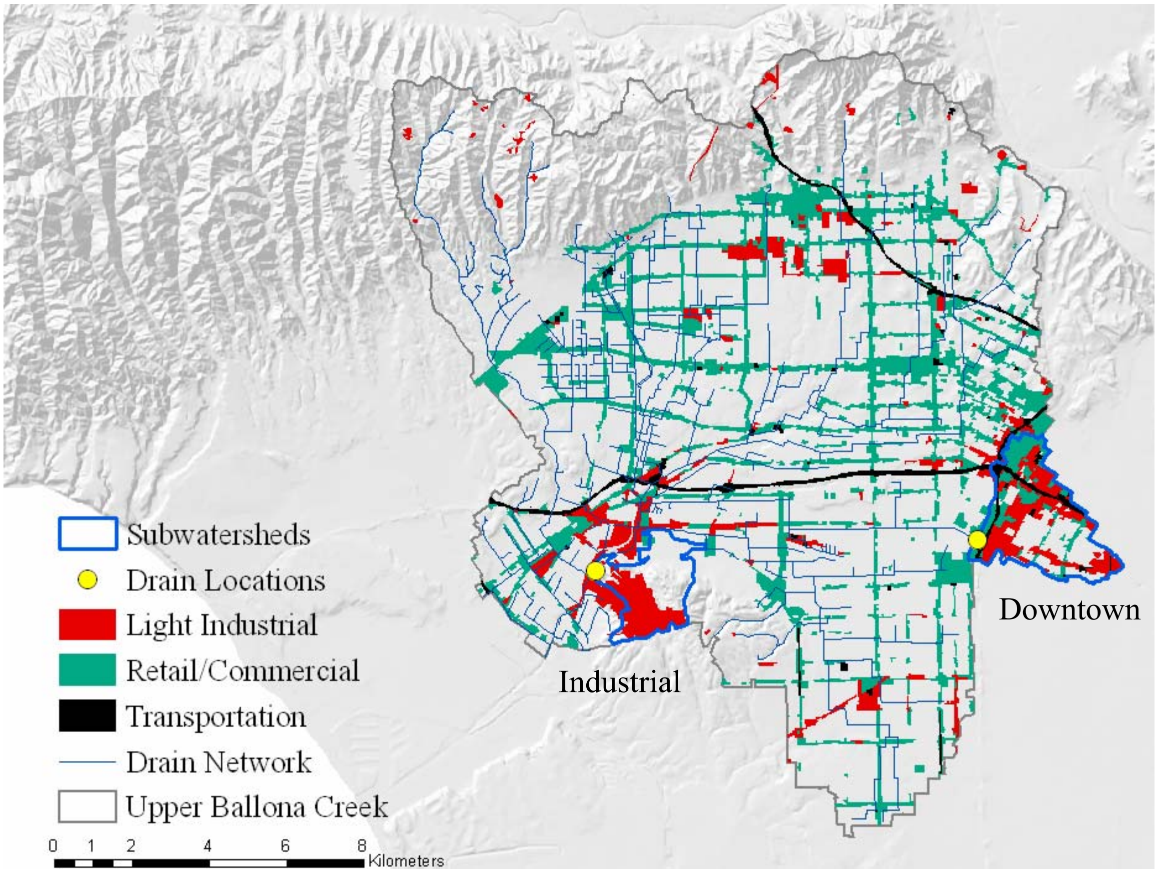


Figure 4.2.9 Distribution of High Leverage Landuses (Light Industrial, Retail/Commercial, and Transportation Landuses)

The two subwatershed areas for each landuse type and their runoff volumes are shown in Table 4.2.4, for the Dec. 2, 2001 storm event, as an example analysis. Load contributions from the two subwatersheds were calculated from the runoff volumes and compared with load contributions from highways and local roads in the subsequent discussions.

Table 4.2.4 Area of Landuse Types and its Runoff for Dec. 2, 2001 Storm Event

Landuse Types	Downtown		Industry	
	Area [Hectare]	Runoff [m^3]	Area [Hectare]	Runoff [m^3]
Vacant	2.0	10	134.5	2,130
Educational Facilities	30.5	770	5.1	240
High Density Single Family Residential	1.4	10	88.3	1,730
Light Industrial	274.9	6,500	183.3	8,410
Multi-Family Residential	9.1	100	36.2	710
Mixed Residential	205.3	4,970	1.2	20
Retail/Commercial	183.3	4,720	4.2	190
Transportation	75.6	1,990	0	0
Total	782.1	19,070	452.8	13,430

Besides the LADPW landuse types, two additional categories, which are highways and local roads, were separately studied as mentioned in Section 3.4.4. The EMCs in Table 4.2.5 are averages of six Caltrans sampling sites around the upper Ballona Creek watershed. The six sampling sites used are shown in Figure 4.2.10. These EMCs were used for load calculations for highway runoffs within the watershed. For the load calculation for local roads, EMCs of LADPW Transportation landuse are used.

Table 4.2.5 Average EMCs of Six Caltrans Highway Sites around the Upper Ballona Creek Watershed

Dissolved P	Total P	Nitrate-N	TKN	TSS	Dissolved Cu	Total Cu	Total Pb	Dissolved Zn	Total Zn
0.14	0.51	1.06	3.81	120.73	29.04	54.77	46.26	185.46	282.02

- Metal EMC units are $\mu g / L$. All other units are mg / L.

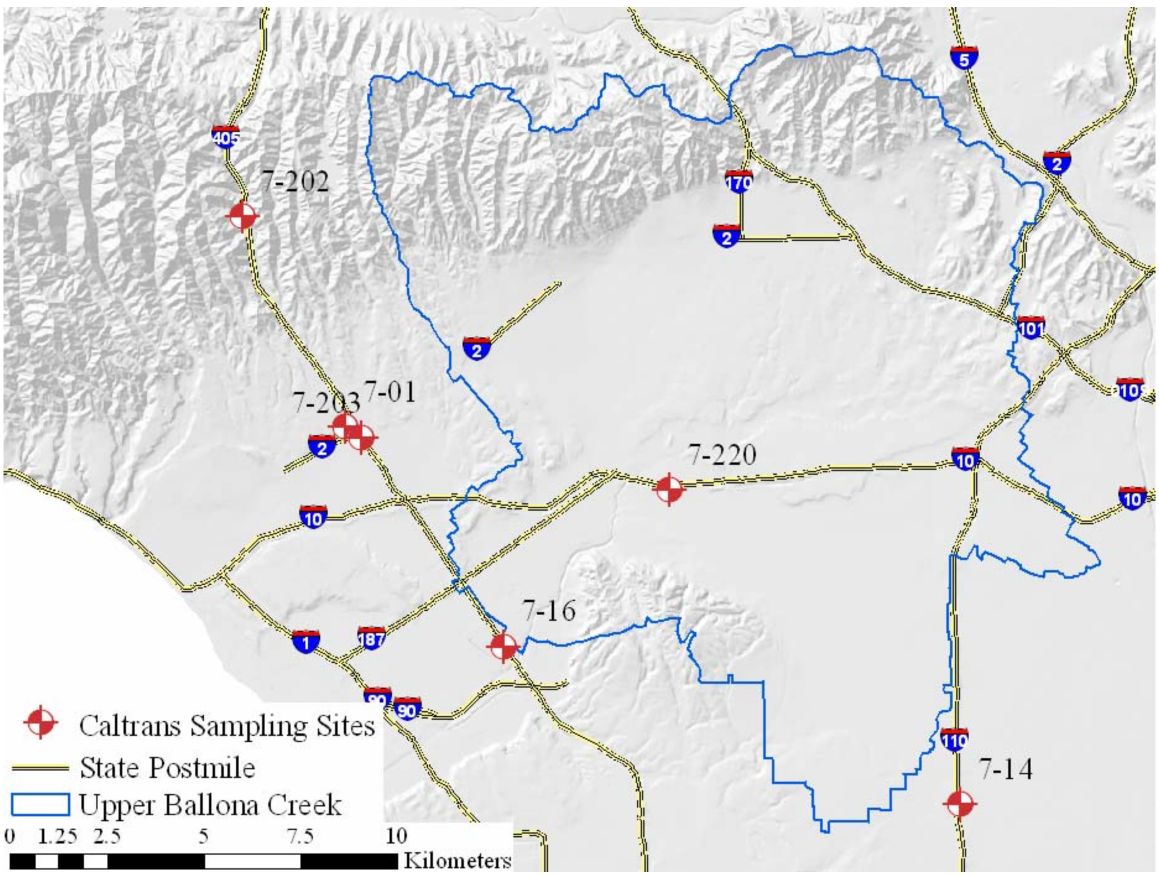


Figure 4.2.10 Caltrans Sampling Sites

Calculated results for the subwatersheds, highways, local roads, and upper Ballona Creek watershed are shown in Table 4.2.6. For comparison purpose, percent contributions and leverages for the three categories are shown in Figure 4.2.11 and 4.2.12, respectively.

Table 4.2.6 Calculated Results of Subwatersheds, Highways, and Local Roads

	Downtown	Industry	Highways	Local Roads	Upper Ballona Creek
EMC [$\mu g / L$]	–	–	282.0	212.1	–
Length [Km]	–	–	47	2,754	–
Area [Hectare]	782	453	152	2,518	23,212
Runoff [m^3]	19,100	13,400	6,077	98,000	454,400
Event Load [Kg]	2.5	1.7	1.7	24.7	39.7

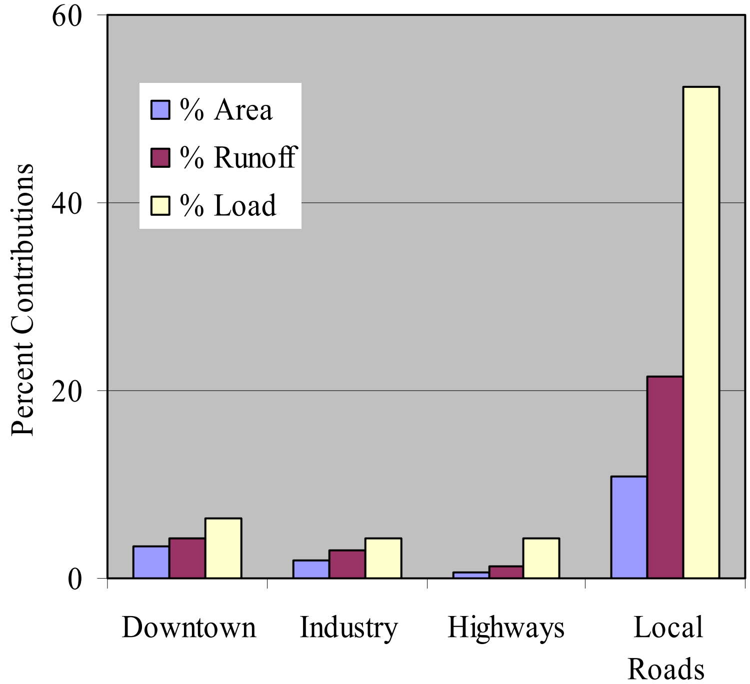


Figure 4.2.11 Percent Contributions of Total Zn for Subwatersheds, Highways, and Local Roads

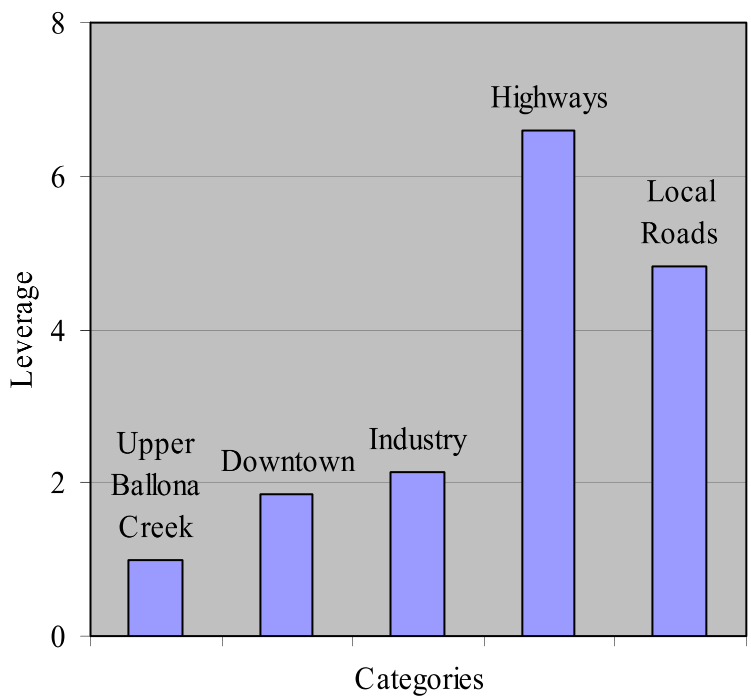


Figure 4.2.12 Total Zn Leverage for Subwatersheds, Highways, and Local Roads

Leverages of Downtown and Industry subwatersheds are both about two, which means doubled load reduction could be achieved by treating these two subwatersheds as compared to the average. However, percent load contributions of Downtown and Industry subwatersheds are only 6.3 % and 4.3 %, respectively. Thus, reduction in total load would be small. For the Highway category, leverage is greater (6.6) than the others (the next highest was 4.1 for Transportation among LADPW landuse types) but percent contribution to the total watershed load is also small, which is only 4.3 %. This is attributable to the runoff contribution of highways, which is only 1.3 % of the total watershed runoff. Local road area is only 10.9 % and its runoff volume is 21.6 % of the total and the load contribution from local roads to the total watershed load is surprisingly high. Local roads contributed 52.4 % of the total load as shown in Figure 4.2.11. This makes local roads the best candidate for further BMP considerations.

4.3 BEST MANAGEMENT PRACTICES (BMP)

The purpose of this section is to show that there exist better and worse BMP strategies. To approach the problem of BMP evaluation, currently defined TMDLs and load trends of upper Ballona Creek watershed were studied first.

4.3.1 Total Maximum Daily Load (TMDL)

For certain toxic pollutants, the US EPA has established numeric criteria that serve as water quality standards for California's inland surface waters. The US EPA established the numeric criteria in the California Toxics Rule (CTR) to protect aquatic life in all of California's inland surface waters. If a pollutant is present in surface water at a level higher than a CTR criterion, then the surface water is toxic. Acute freshwater quality objective established in the CTR for dissolved Zn is $94\,\mu g/L$ for Ballona Creek watershed. Chronic criteria are typically based on exposures, which occur over a four day time interval and acute criteria are typically based on a shorter time interval. Numeric targets for the TMDL was calculated based on this numeric objective excluding direct atmospheric deposition effect (0.6 %). Freshwater wet weather numeric target for total Zn, which is equivalent to EMC based on flow weighted composite, is $119\,\mu g/L$. The total Zn TMDL was prepared by Los Angeles Regional Water Quality Control Board and US EPA Region 9 (http://www.waterboards.ca.gov/losangeles) and is defined by the following equation.

$$\text{Total Zn TMDL [g/day]} = (1.18 \times 10^{-4}) \times \text{Daily storm volume [L]} \qquad (4.1)$$

4.3.2 Comparisons of Total Zn Loads and TMDLs

TMDLs were calculated using Eq. 4.1 and compared with measured total Zn loads for the 1998 to 2003 storm seasons. The results are shown in Figure 4.3.1.

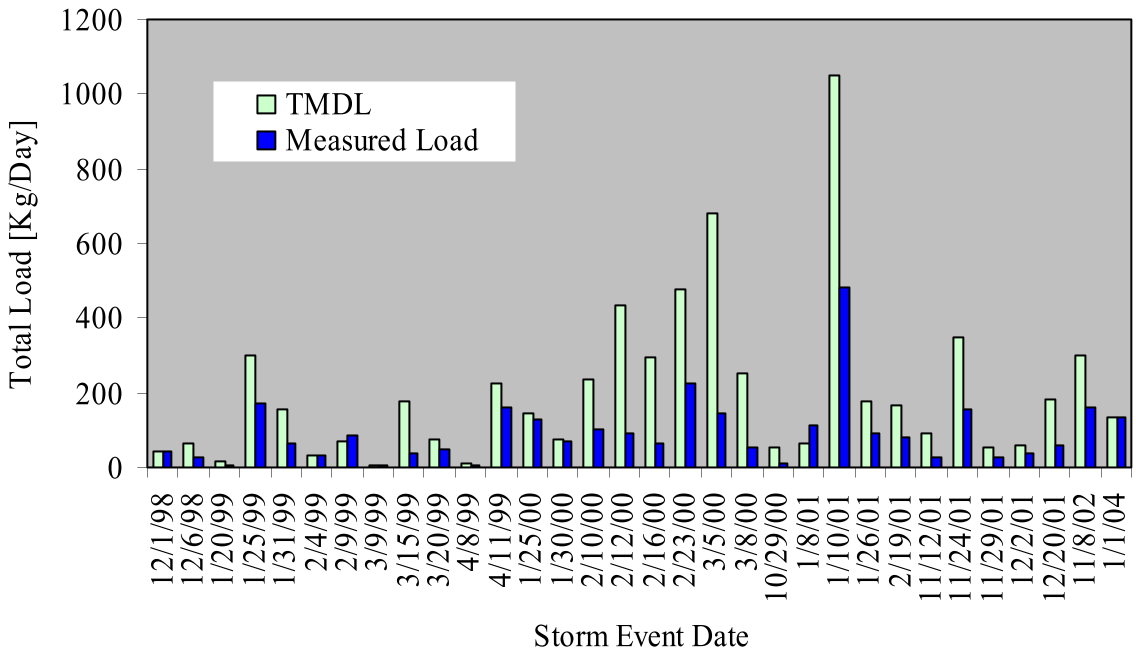


Figure 4.3.1 Comparisons of Total Zn Loads and TMDLs for Ballona Creek Watershed

Since TMDLs are defined in terms of EMCs, they are likely to be violated by small storm events. Among the 32 storm events, two events exceeded TMDLs. Total Zn load of the Feb. 9, 1999 storm event exceeded the TMDL by 18% and the Jan. 8, 2001 storm event exceeded the TMDL by over 70%. Total rainfall for these two storm events were 9.4 and 6.9 mm, respectively. Therefore, BMP strategies focusing the treatment of small storm events or first flush of large storm events are preferred. Structural BMP applications will be discussed for these two storm events in the subsequent sections.

4.3.3 Structural BMPs

If the total zinc concentration of the Jan. 8, 2001 storm event is uniform over the upper Ballona Creek watershed, and over the duration of the storm event, BMPs with pollutant removal efficiencies higher than 44 % will be needed to treat the entire runoff to meet the required TMDL. Fortunately, recent studies have identified strong first flush phenomena for the highways, which show that the most polluted stormwater runoff occurs during the first part of the runoff (Han et. al, 2006b). Also, the leverages shown in Figure 4.2.12 were higher for highways and local roads. Therefore, among the BMPs shown in Table 4.3.1, a sand filter was selected and applied to the two subwatersheds mentioned in Section 4.2.5. It has high pollutant removal and is widely applicable for small sites. Later in this dissertation, a strategy utilizing leverage and first flush phenomena is evaluated.

Table 4.3.1 Comparison of Issues for BMP Selection (Weld et. al., 1997)

BMP	Pollutant Removal Reliability	Longevity	Maintenance Requirements	Applicability to Sites	Environmental Concerns	Comparative Cost	Special Considerations
(Extended) Detention Basin	Moderate	20+ years	Low	Widely applicable, larger drainage areas (10+ acres)	Possible downstream warning; low bacteria removal	Low to moderate	Available land area, design considerations, sediment forebay
Wet (Retention) Pond	Moderate to high	20+ years	Low to moderate	Widely applicable, larger drainage areas (7+ acres)	Possible downstream warning; low bacteria removal	Moderate to high	Available land area, design considerations, sediment forebay
Constructed Stormwater Wetland	Moderate to high	20+ years	Low to moderate	Widely applicable, larger drainage areas (7+ acres)	Possible downstream warning; wildlife benefits	Marginally higher than wet ponds	Available land area, design considerations, sediment forebay
Water Quality Swale	Moderate	20+ years	Low to moderate	Widely applicable	Restricted use for hotspots	Low to moderate	Pretreatment, check dams, careful design
Infiltration Trench	Moderate to high	High rates of failure within first 5 years	High	Highly restricted; small sites, proper soils, depth to water table and bedrock, slopes	Potential for ground water contamination; restricted use for hotspots	High; rehabilitation costs can be considerable	Recommended with careful site evaluation and pretreatment
Infiltration Basin	Moderate	High rates of failure within first 5 years	High	Highly restricted; small sites, proper soils, depth to water table and bedrock, slopes	Potential for ground water contamination; restricted use for hotspots	Moderate, rehabilitation costs can be high	Not widely recommended until longevity is improved

Table 4.3.1 – Continued –

BMP	Pollutant Removal Reliability	Longevity	Maintenance Requirements	Applicability to Sites	Environmental Concerns	Comparative Cost	Special Considerations
Organic Filters	Moderate to high	20+ years	High	Widely applicable for small sites	Minor	High; frequent maintenance	Recommended with careful design; pretreatment
Sand Filters	Moderate to high	20+ years	High	Widely applicable for small sites	Minor	High; frequent maintenance	Recommended with careful design; pretreatment
Water Quality Inlets	Low	20+ years	Moderate to high	Small, highly impervious areas (< acres)	Resuspension of PAH loads, Disposal of residuals	Moderate to high	Pretreatment technology, off-line
Sediment Trap (Forebay)	Low	20+ years	Moderate	Widely applicable as pretreatment	Resuspension of accumulated sediment if not maintained	Low to moderate	Pretreatment technology
Drainage Channel	Low	20+ years	Low to moderate	Low density development and roads	Erosion, resuspension	Low	Pretreatment technology, with check dams
Deep Sump (Modified) Catch Basin	Low	20+ years	Moderate	Small, highly impervious areas (<2 acres)	Resuspension of accumulated sediment if not maintained	Low to moderate	Pretreatment technology, design modified with sump

4.3.3.1 BMP Construction Costs

Average installation costs for BMPs are shown in Table 4.3.2 (Caltrans, 2004). The average costs in the table were calculated from one to five examples of each BMP type. The costs per cubic meter of stormwater runoff to be treated are highly variable between jurisdictions as shown in the table and are relatively higher for smaller units. Among these construction costs, Caltrans costs were used in the BMP cost analyses because they were calculated from smaller BMP unit constructions. Smaller BMPs may be better for highly urbanized cities, where large areas for sitting BMPs may not exist. In the following sections, strategies using small BMP units are discussed.

Table 4.3.2 Comparison of Mean Unit Costs and Runoff Volumes

BMP	Caltrans		Nationwide Survey [a]		MD SHA [b, c]	
	Cost [$/m^3]	Runoff Volume [m^3]	Cost [$/m^3]	Runoff Volume [m^3]	Cost [$/m^3]	Runoff Volume [m^3]
Austin sand filter	1,447	168	82	12,123	[d] 33	[d] 1,140
Delaware sand filter	1,912	120	200	1,836		
Extended-detention basin	590	293	5	99,537	18	32,279
Infiltration trench	733	199	46	2,485	11	4,304
Bio-filtration swale	752	748	[d] 9	[d] 2,066		
Wet pond	1,731	259	8	44,833	9	20,391

Notes:
Table adapted from Caltrans (2004). Values are in 1999 dollars.
[a] Means for all entries in the Third Party Cost nationwide survey where stormwater runoff volume is available.
[b] Means for all Maryland State Highway Administration (MD SHA) BMPs where stormwater runoff volume is available.
[c] MD SHA had a retrofit policy that capped retrofit costs at $30,000 per hectare.
[d] Based on a single installation.

4.3.3.2 Austin Sand Filter

The Austin sand filter was chosen among the BMPs shown in Table 4.3.1 because its pollutant removal reliability is moderate to high and widely applicable for small sites, which are important criteria in a highly urbanized city. A typical sand filter system consists of sedimentation and filtration processes. Runoff is temporarily stored in a sedimentation chamber, and then directed to a sand filter, where additional pollutants are filtered through a sand bed. Sand filter systems are preferred over infiltration systems where ground waters are nearby. A drawback of this system is that the filter media may need to be replaced in 3 to 5 years. In general, sand filters are good options in highly urbanized areas because they require small space. Among the currently available sand filters, the Austin sand filter was chosen for this study because it is suited to a larger range of drainage areas, up to 20 hectares, compared to the other sand filters.

Table 4.3.3 shows the calculated results for Downtown and Industry subwatersheds before and after the application of Austin sand filters. Typical total Zn removal efficiency, 45 %, was used to calculate total Zn removed by the sand filters (US EPA, 1999b). As shown in the Case A of the table, total Zn removal at the mouth of the watershed was 12.4 % with the treatment of 5 % of the upper Ballona Creek watershed area (or equivalently 9 % of the runoff volume), which includes Downtown and Industry subwatershed areas. Efficiency can be enhanced by avoiding low leverage areas within the two subwatersheds (See Figure 4.2.9). For example, 23 Austin sand filter units are needed to treat about 453 hectares of Industry subwatershed if we assume one filter covers 20 hectares. Leverages of Vacant and High Density Single Family Residential landuses in the Industry subwatershed are less than one. Total area of these two landuses is about 223 hectares. Thus, if only 12 filters are used for the high leverage areas, pollutant reduction rate is decreased to 10.2 % while area treated is reduced to 4 %.

Table 4.3.3 Simulation Results of Austin Sand Filters for Industry and Downtown Subwatersheds (January 8, 2001 Storm Event)

Case A	Downtown		Industry		Total*
Area [Hectare]	782	3 %	453	2 %	23,212
Runoff [m^3]	35,500	7 %	9,000	2 %	549,000
Load without sand filters [Kg]	19.9	17 %	6.8	6 %	117.9
Load with sand filters [Kg]	10.9	11 %	3.7	4 %	103.3
Construction Cost [$ Million]	51.4		13.0		64.4
Treatment Effectiveness [Kg Zn/$ Million]	150		200		160
Case B	Downtown		Industry without Low Leverage Areas		Total*
Area [Hectare]	782	3 %	230	1 %	23,212
Runoff [m^3]	35,500	7 %	6,400	1 %	549,000
Load without sand filters [Kg]	19.9	17 %	6.6	6 %	117.9
Load with sand filters [Kg]	10.9	10 %	3.6	3 %	105.9
Construction Cost [$ Million]	51.4		9.3		60.7
Treatment Effectiveness [Kg Zn/$ Million]	150		270		170

* Totals are based on model predictions and percentages are based on these totals.
- TMDL = 66 Kg
- Measured total Zn load = 113 Kg
- Measured total runoff volume = 556,000 m^3
- LA Civic Center rainfall record = 6.9 mm

To meet the TMDL requirement, more than 30 % load reduction needs to be achieved besides the treatment of the two subwatersheds. If the same leverage and runoff per area are assumed for additional area in the upper Ballona Creek watershed, this would require about four times the construction costs than the costs estimated for the two subwatersheds and about 20 % of the watershed area (or equivalently about 36 % of the runoff volume) would need to be treated.

For the treatment effectiveness calculation, average annual rainfall was calculated from the LA Civic Center rainfall records for the 1998 – 2003 storm seasons. The calculated average annual rainfall (290 mm/year) was then applied to the assumed 20 years of sand filter lifetime. Also, the load reduction rate shown in Table 4.3.3 was assumed in the calculation of Total Zn removed by the sand filters.

4.3.3.3 BMPs for Local Roads

As mentioned in previous sections, if we do not utilize high leverage areas and first flush phenomena, all the stormwater runoff of the Jan. 8, 2001 storm event had to be treated with BMPs that have 44 % or better pollutant reduction rates. Since over 50 % of total Zn comes from local roads for the Dec. 2, 2001 storm event as shown in Figure 4.2.11, BMP applications for local roads are considered in this section.

The landscape approach was taken by the City of Portland Bureau of Environmental Services (2004) for the impervious area of local roads. Figure 4.3.2 shows one of their facilities where vegetated areas were used as BMPs. The figure shows a shallow vegetated curb extension installed in the street parking lane. Impervious drainage area for the extension was 465 square meters and the curb extension area was 25 square meters. A flow test was performed with 25-year design storm (48 mm rainfall in 6 hours) for the curb extension. Pervious areas were not included because their hydrologic modeling results for the design storm using XPSWMM showed negligible runoff contribution from those areas. Their simulated results with the design storm showed that the extension captured 85 % of the runoff volume during the flow test. The vegetated curb extension's high runoff volume reduction was due to high saturated infiltration rate of the soil, which was 51 mm/hour.

Figure 4.3.2 Stormwater Curb Extension on Siskiyou Street, Portland (Portland Bureau of Environmental Services, 2004)

Saturated infiltration rates for hydrologic soil groups are shown in Table 4.3.4 to check whether the curb extension technique can be applied to upper Ballona Creek watershed.

The effective water capacity of a soil in the table is the fraction of the void spaces available for water storage. The data presented in the table are based on the analysis of over 5,000 soil samples by the United States Department of Agriculture (USDA). Since the hydrologic soil group of the entire upper Ballona Creek watershed is type D as mentioned in Section 3.4.2, the watershed's saturated infiltration rate is up to 2.3 mm per hour as shown in the table. The curb extension technique can not be applied directly to the watershed because of the watershed's low saturated infiltration rate.

Table 4.3.4 Hydrologic Soil Properties (Rawls, 1982)

Hydrologic Soil Group	Soil Texture	Saturated infiltration rate [mm/hr]	Effective Water Capacity [-]
A	Sand	210.1	0.35
	Loamy Sand	61.2	0.31
	Sandy Loam	25.9	0.25
B	Loam	13.2	0.19
	Silt Loam	6.9	0.17
C	Sandy Clay Loam	4.3	0.14
D	Clay Loam	2.3	0.14
	Silty Clay Loam	1.5	0.11
	Sandy Clay	1.3	0.09
	Silty Clay	1.0	0.09
	Clay	0.5	0.08

4.3.3.4 Infiltration Trench

Infiltration trenches can be generally described as trenches filled with media designed to maximize infiltration of runoff to the ground. Infiltration trenches are especially attractive because they provide the highest level of surface water quality protection (Caltrans, 2004). They not only reduce the runoff volume but also restore natural water balance of the site. Design criteria for infiltration trenches are as follows (US EPA, 1999a; Clar, 2004):

1. Soils should have a low silt and clay content and have infiltration rates greater than 13 mm per hour.
2. A minimum of 1.2 meters to the seasonally high water table to prevent potential ground water contamination.
3. Trenches should be located at least 30.5 meters away from water supply wells.
4. Trenches should be constructed with pretreatment to prevent site failure due to clogging.
5. Recommended maximum depth for an infiltration trench is 2.5 m (Schueler, 1987).
6. If the drainage area exceeds 2 hectares, other BMPs should be carefully considered.

Besides the infiltration rates of soils, hotspot landuses are not recommended without pretreatment because of possible ground water contamination and failure due to clogging of trenches. Even though traditional infiltration trenches may not appropriate for the highly urbanized Ballona Creek watershed, a modified infiltration trench, such as the one shown in Figure 4.3.3, may be feasible.

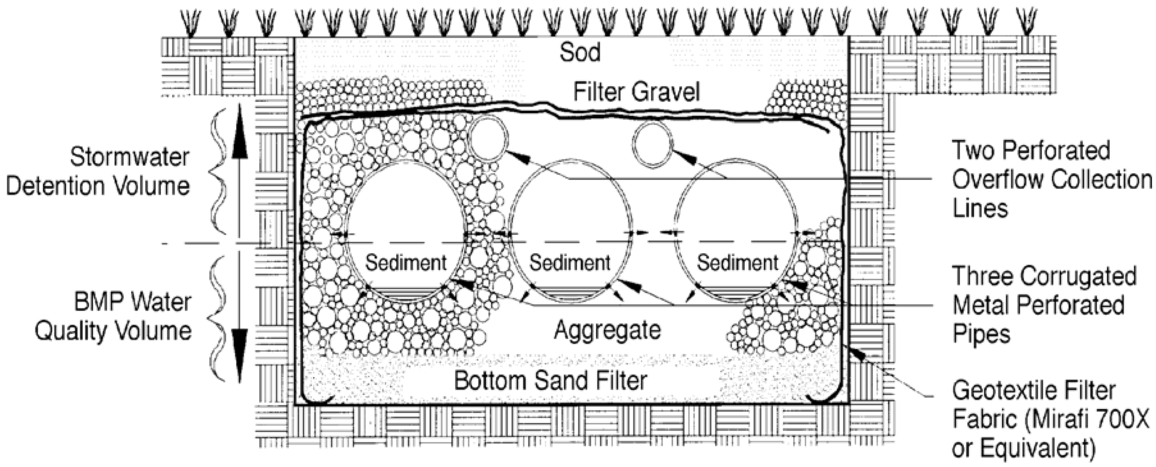


Figure 4.3.3 Modified Infiltration Trench (US EPA, 1999a)

The modified infiltration trench uses a layer of peat or loam to the trench subsoil, which enhances the removal of metals and nutrients through adsorption. The stone aggregate used in the trench normally provides a void space of 40% (Harrington, 1989; Schueler, 1987). This layer is then covered with a geotextile filter fabric. Pea gravel can be used in the top soil to improve sediment filtering. The depth of top layer is about 30 cm. If the trench becomes clogged, only the top layer is replaced.

Infiltration trenches are commonly designed to drain prior to the next storm event. This practice is not possible in Ballona Creek watershed because of its low infiltration rate. It is better to design trenches to capture first flush of a storm event, whose partial EMCs at 30 and 60 minutes into the storm may be 1.9 to 7.4 times higher than the EMCs (Han, 2006b). For example, the antecedent dry days of the two storm events occurred on Feb. 9, 1999 and Jan. 8, 2001 mentioned in Section 4.3.2 was about 5 and 69 days, respectively. The total zinc loads of the two storm events exceeded TMDLs by 18 % and 70 %, respectively. Since the ADD of Jan. 8, 2001 storm event was long enough to drain the trenches, the Feb. 9, 1999 storm event is considered here. The LA Civic Center rainfall record for this storm event was 9.4 mm. If we assume that the saturated infiltration rate of upper Ballona Creek watershed is 1.5 mm/hr, trench water level drops about 18 cm prior to the Feb. 9, 1999 storm event. If we further assume that the drainage area is 500 square meters, runoff coefficient of local roads is 0.9, and the first flush of a storm event is covered with 5 mm of rainfall, infiltration trench area can be calculated as the following:

Runoff volume of the first flush

$= 500\ m^2$ (Drainage area) $\times$ 0.005 m (Rainfall) $\times$ 0.9 (Runoff coefficient)

= Trench Area $\times$ 0.18 m (Dropped water level for 5 days) $\times$ 0.4 (Void space)

The calculated infiltration trench area for this example is 31 m^2 for the assumed 500 m^2 local road drainage area. Thus, by choosing appropriate trench area, first flush of the storm event can be captured. In turn, this will provide enough coverage to avoid the 18 % exceedence with an ADD of 5 days. Other design criteria listed above are also assumed to be satisfied by other means in the following analysis.

Feasibility of infiltration trenches as a BMP for local roads was investigated for the Jan. 8, 2001 storm event. The EMC was calculated using the optimized parameters and the weighting factor for Transportation landuse (See Table 3.5.6):

Transportation landuse EMC

= Calculated Transportation EMC using optimized parameters $\times$ Weighting Factor

= 661 $\mu g / L$

The load contribution of local roads was then calculated by multiplying the EMC and the runoff from local roads, which was 77 Kg as shown in Table 4.3.5. The total Zn load that needs to be removed by infiltration trenches is 52 Kg (Total Zn load of the watershed – TMDL), which is 68 % of the local road contribution ($\frac{52}{77} \times 100\ \%$). Local road runoff and area that needs to be treated by trenches are also assumed to be 68 %. The rainfall record can not be used in these calculations because the hydrology model described in Section 3.4 uses local rainfall from isohyets. As shown in the table, the TMDL was met with the treatment of 7 % of the upper Ballona Creek watershed area (or equivalently about 14 % of the runoff volume).

Table 4.3.5 Simulation Results of Infiltration Trench Installations for Local Roads (January 8, 2001 Storm Event)

	Local Roads Area Covered by Infiltration Trenches		Total Local Roads		Upper Ballona Creek Watershed
Area [Hectare]	1,712	7 %	2,518	11 %	23,212
Runoff [m^3]	79,356	14 %	116,700	21 %	549,000
Load without Trenches [Kg]	52	44 %	77	65 %	118
Load with Trenches [Kg]	0	0 %	25	38 %	(-44 %) 66
Construction Cost [$ Million]	58				58
Treatment Effectiveness [Kg Zn/$ Million]	320				320

- All calculations are based on model predictions.
- TMDL = 66 Kg
- Measured total Zn load = 113 Kg
- Measured total runoff volume = 556,000 m^3
- LA Civic Center rainfall record = 6.9 mm

Caltrans' cost per unit runoff volume was used (See Table 4.3.2) for the construction cost calculation. For the treatment effectiveness calculation, average annual rainfall frequency was calculated from the LA Civic Center rainfall records from 1998 to 2003 storm seasons. The calculated average annual rainfall frequency (18 storm events/year) was then applied to the assumed 20 years of trench lifetime. If the trench is filled up with stormwater runoff, no further load reduction occurs. Thus, Total Zn removed by the trench for each storm event was assumed to be equal to the load reduction shown in Table

4.3.5, which was 52 Kg per storm event. This assumption is conservative because first flush EMCs may be 1.9 to 7.4 times higher than event EMCs (Han et. al, 2006). By the application of infiltration trenches for stormwater runoff from local roads, treatment effectiveness is reduced to half of the effectiveness resulted in by the subwatershed approach, which is discussed in Section 4.3.3.2, while construction cost is reduced to one-fourth of the cost shown in the Case A of Table 4.3.3.

As mentioned earlier, infiltration rate of upper Ballona Creek watershed is low. However, first flush of storm events can be captured with appropriate design of infiltration trenches. According to Han et al. (2006a), total metals, COD, and DOC were generally well correlated with mass fist flush for stormwater runoff from highways. Thus, high concentration of those pollutants can be removed by capturing early portion of runoff from local roads with infiltration trenches.

5. CONCLUSIONS

Stormwater quantity and quality models were developed to estimate runoff volumes and pollutant loads for large urban watersheds. The models were based on the rational method and landuse types but used improvements to account for build up between storm events, ground slope, hydrologic soil group, and storm characteristics. The models were integrated with a GIS (ArcGIS) to manage subcatchments (~33,000; Figure 3.4.7) and optimization techniques were used for calibration. The models successfully predicted stormwater volumes and pollutant loads for total suspended solids (TSS), total Kjeldahl nitrogen (TKN) and total zinc. The water quality model was calibrated using 24 storm events from the 1998 to 2000 storm seasons and validated with 7 additional storm events from the 2001 to 2003 storm seasons for the upper Ballona Creek watershed (23,211 ha). Calibration and validation were performed by comparing model predictions to measurements from the Los Angeles Department of Public Works' (LADPW) mass emission site at intersection of Sawtelle Blvd and Ballona Creek (34° 00′, 118° 24′). The following specific conclusions are made:

1. The Rational Method and Browne's empirical relation for calculating runoff from DEM, landuse type, hydrologic soil group, and isohyets layers were integrated into the ArcGIS platform using Visual Basic. The implemented model showed improved runoff prediction (- 8 % of the measured runoff) than the model defined by landuse definitions (+ 54 %; Wong et. al, 1997) for an example analysis of Dec. 2, 2001 storm event.

2. The predicted runoff volumes using the new model were in good agreement with the measured data (Figure 3.4.13) and had R^2 of 0.86, and generally were within +184/-54 percent of the measured runoff. The model over predicted runoff from small storms and under predicted runoff from large storms, which is a well-known characteristic of the Rational Method.

3. The new model has the advantage of reduced data needs, when compared to a runoff model such as SWMM. For example, Barco et al. (2006) required approximately 30,000 input data points for the characteristics of drain network and about 2,500 subcatchments to model the upper Ballona Creek Watershed. The new model required only four GIS layers, which are DEM, hydrologic soil group, landuse type, and isohyets layers. More importantly, only the fourth layer needed to be developed for the modeling activity. The first three layers were available from other sources.

4. To calibrate pollutant mass emissions as a function of site specific conditions, four types of storm characteristics were used as independent variables: antecedent dry days, total rainfall, rainfall duration, and maximum hourly rainfall intensity. Site specific characteristics such as subcatchment impervious fraction, depression storage, infiltration, and roughness coefficient were not used in the analyses to simplify input to the model and to avoid expensive and time consuming field surveys. Multiple linear regression analyses were performed to derive correlations between the pollutants' landuse EMCs and the storm characteristics. The number of storm events

used in these correlations ranged from 12 to 35, which were measured at eight landuse monitoring sites for the storm season 1998 to 2001 by the LADPW.

The regression explained more than 80% of the variability in total Zn EMCs for Light Industrial, Mixed Residential, and Transportation landuses. The regressions poorly explained ($R^2 < 0.4$) the variability in total Zn for Retail/Commercial, Educational Facilities, High Density Single Family Residential, and Multi-Family Residential landuses. Whereas, the TKN EMC correlations fit Educational Facilities and High Density Single Family Residential landuses well ($R^2 > 0.8$), but R-square values for the Vacant, Retail/Commercial and Multi-Family Residential landuses were below 0.5. For TSS EMC correlations, all R-square values were low ($R^2 < 0.6$).

5. In general, Antecedent Dry Days (ADD) were the most significant explanatory variable for the three pollutant EMCs analyzed. Total rainfall was the next significant variable and the maximum hourly rainfall intensity was the least significant for all three pollutant EMCs.

6. To improve the model predictions of pollutant load contributions from various landuse types to total loads of the upper Ballona Creek watershed, zeroth-order regularization and L-BFGS-B optimization techniques used. Relative form was used in the objective function to compensate strong contributions of high magnitude variables.

To account for the variability of the percent of each representative landuse types for the landuse monitoring catchments, which were ranged from 54% to 100%, weighting factors were introduced and optimized. Contributions of each landuse types to the total loadings of upper Ballona Creek watershed were adjusted by these weighting factors. For total Zn and TKN total load calculations, optimized weighting factors were closer to unity when the representative landuses for corresponding landuse types were closer to 100%. For example, the weighting factor of the Retail/Commercial landuse, which has the lowest percent representative landuse, had the greatest deviation from one. However, weighting factors for TSS did not show any tendency. Also, weighting factors did not show any relationships to the size of landuse monitoring catchments for the three pollutants.

7. Model predictions showed reasonable agreement with measured loadings for total Zn, TKN, and TSS without using site specific information. The RMS error of the optimized model without regularization for total Zn was 24 Kgs and only 9 Kgs with regularization, which is low compared to the mean discharge of 77 Kgs per storm event. The RMS error for TKN was 0.5 Kgs for the mean discharge of 2.9 Kgs. The RMS error for TSS was 25 Kgs for the mean discharge of 220 Kgs. However, measured TSS loads for the validation dataset were much lower than model predictions. This discrepancy might be caused by measurement errors or it might be an indication of other activities in the watershed, which would require further investigation to understand.

8. Highways and local roads were added as land types to the LADPW landuse types. In this way, the contributions of these two landuses can be separately quantified. Total Zn load for Dec. 2, 2001 storm event was analyzed as an example. For this storm event, the leverage (percent mass emitted per percent area) of highways was 6.6, which was greater than all the other landuses (the next highest was 4.1 for Transportation among LADPW landuse types). However, highways contributed only 4.3% of the total watershed load because of its small runoff (1.3%). The load contribution from local roads to the total watershed load was 52.4 %, even though local road area was only 10.8 % of the watershed area, and its runoff volume was only 21.6 % of the total runoff. The leverage for local roads was 4.8, which makes it the best candidate for BMP applications.

9. Among the monitored storm events from 1998 to 2003 storm seasons, total Zn load of the Feb. 9, 1999 storm event exceeded the TMDL by 18% and the Jan. 8, 2001 storm event exceeded the TMDL by over 70%. If the total zinc concentration of the Jan. 8, 2001 storm event is uniform over the upper Ballona Creek watershed, and over the duration of the storm event, BMPs with pollutant removal efficiencies higher than 44 % will be needed to treat the entire runoff to meet the required TMDL.

   Two subwatersheds with high leverages were identified and Austin sand filters were simulated for the storm event. By assuming typical total Zn removal efficiency of 45%, total Zn removal at the mouth of the watershed was 12.4 % with the treatment of 5 % of the upper Ballona Creek watershed area. If the same leverage and runoff per unit area are assumed for additional locations in the upper Ballona Creek watershed, about 20% of the watershed area (or equivalently 36 % of the runoff volume) would need to be treated to meet the TMDL. This would require about $240 million if one assumes Caltrans' predicted construction cost per unit volume of runoff.

The TMDL was met with the treatment of 68% of the local road area, which is equivalent to 7 % of the upper Ballona Creek watershed area. The construction cost for the infiltration trenches will be approximately $58 million, using Caltrans' construction costs, which is about one-fourth of the cost needed for the subwatershed approach.

10. Even though the saturated infiltration rate of upper Ballona Creek watershed is low (0.5 – 2.3 mm/hr), first flush of the Feb. 9, 1999 storm event can be captured by choosing appropriate modified infiltration trench area (for example, 31 m^2 per 500 m^2 local road drainage area). In turn, this will provide enough coverage to avoid the 18 % exceedence of the storm event. According to Han et al. (2006a), total metals, COD, and DOC are generally well correlated with mass fist flush for stormwater runoff from highways. Thus, high concentration of those pollutants can also be removed by capturing early portion of runoff from local roads with modified infiltration trenches.

APPENDICES

APPENDIX A. LANDUSE MONITORING SITES

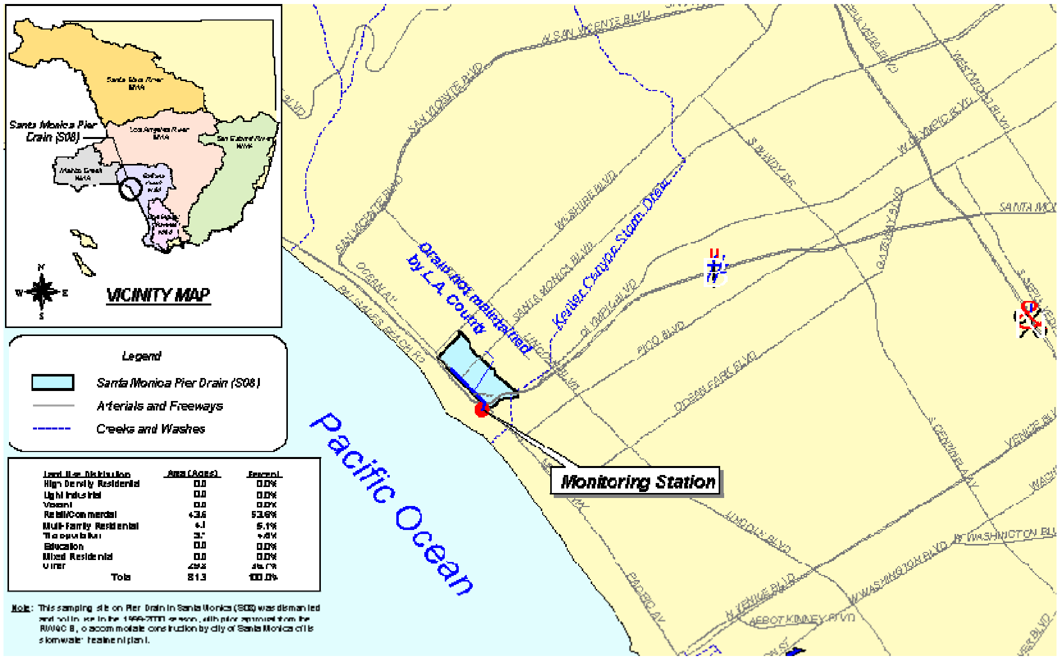


Figure A.1 Retail/Commercial landuse monitoring site at Appian Way, Santa Monica

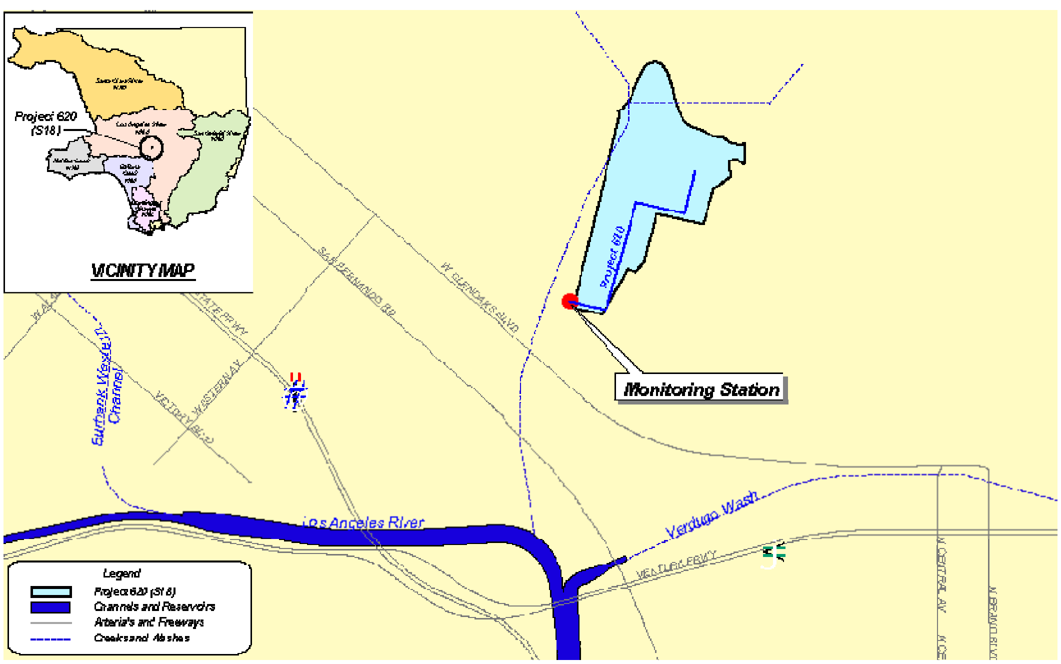
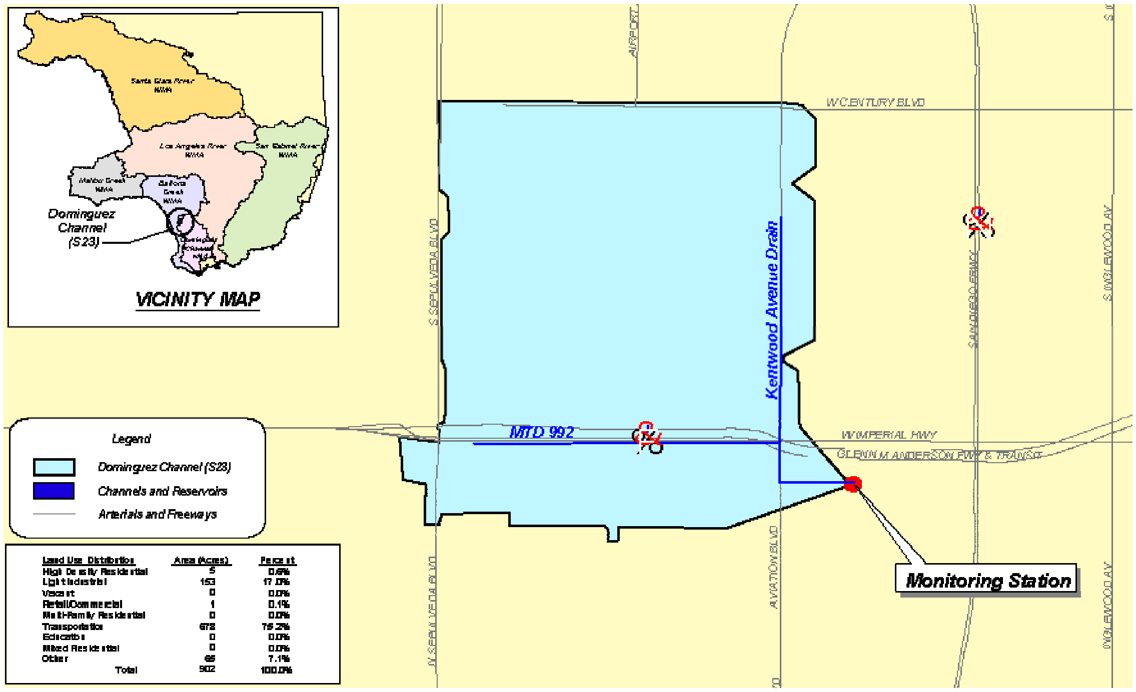
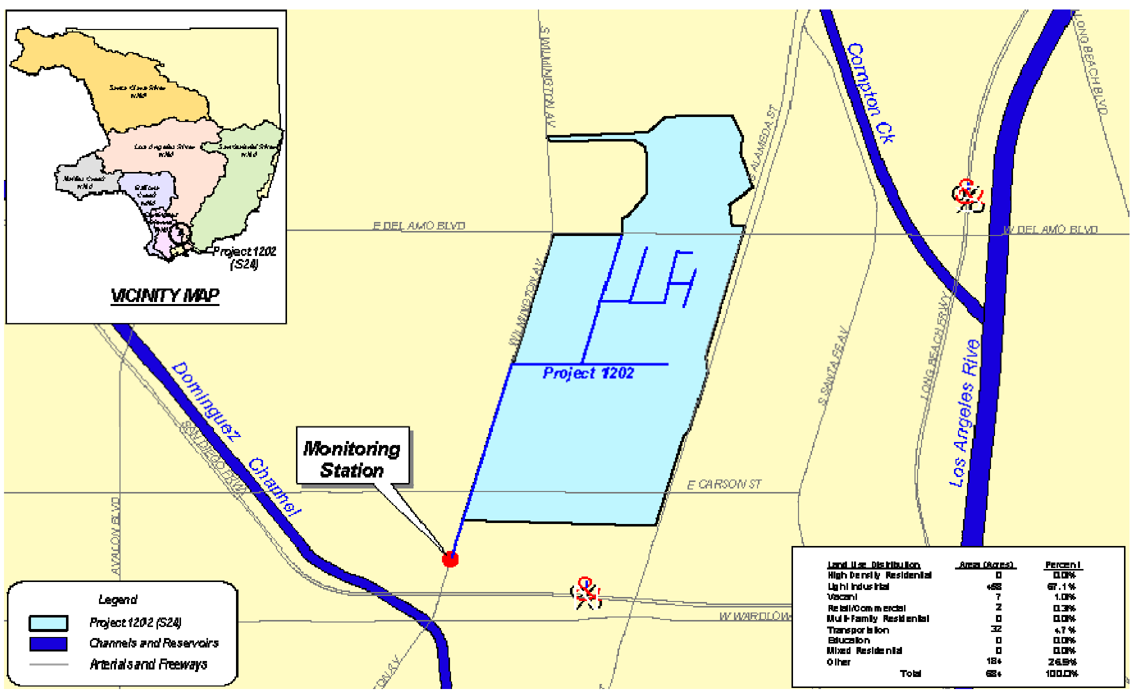


Figure A.2 High density single family residential landuse monitoring site at Glenwood Rd, Glendale



Land Use Distribution	Area (Acres)	Percent
High Density Residential	5	0.6%
Light Industrial	153	17.0%
Vacant	0	0.0%
Retail/Commercial	1	0.1%
Multi-Family Residential	0	0.0%
Transportation	678	75.2%
Education	0	0.0%
Mixed Residential	0	0.0%
Other	65	7.1%
Total	902	100.0%

Figure A.3 Transportation landuse monitoring site at 116th St., Torrance



Land Use Distribution	Area (Acres)	Percent
High Density Residential	0	0.0%
Light Industrial	458	67.1%
Vacant	7	1.0%
Retail/Commercial	2	0.3%
Multi-Family Residential	0	0.0%
Transportation	32	4.7%
Education	0	0.0%
Mixed Residential	0	0.0%
Other	184	26.9%
Total	684	100.0%

Figure A.4 Light industrial landuse monitoring site at Willmington Av.., Carson

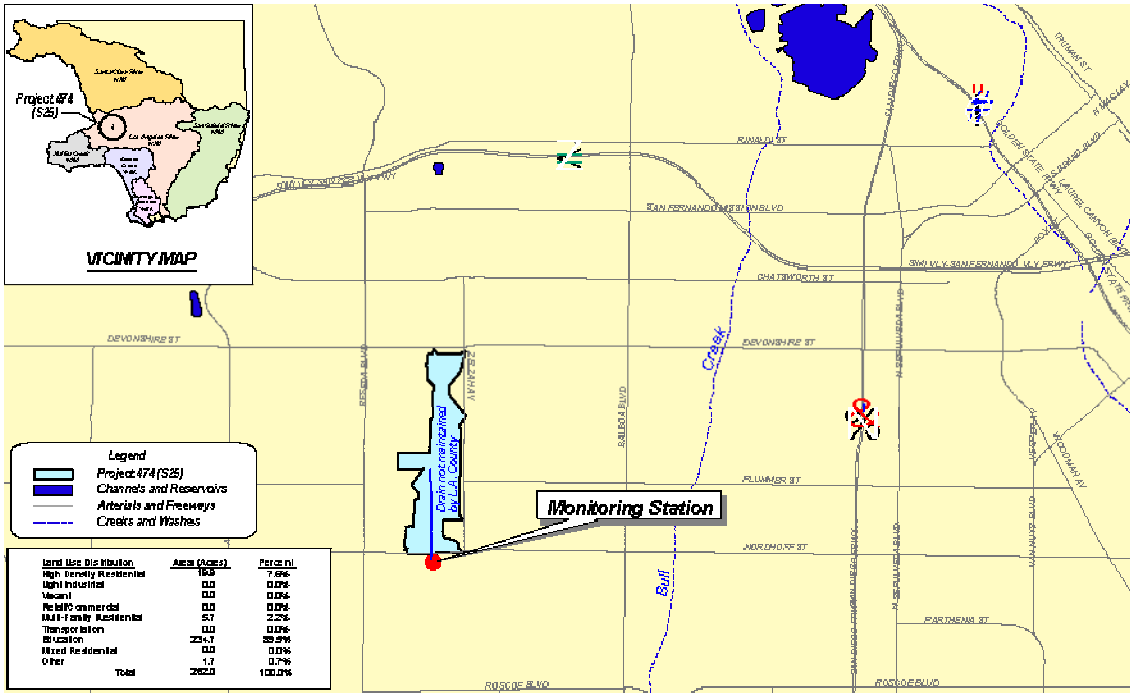


Figure A.5 Education landuse monitoring site at Nordhoff St.., Northridge

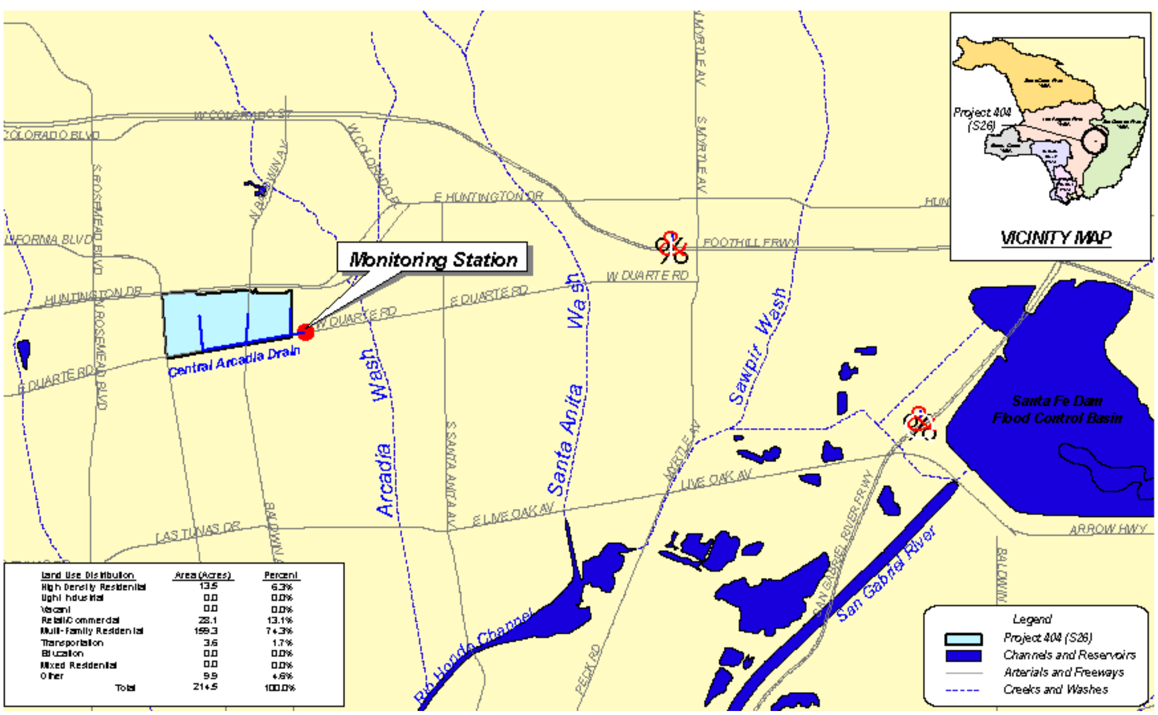


Figure A.6 Multi-family residential landuse monitoring site at La Cadena Ave.., Arcadia

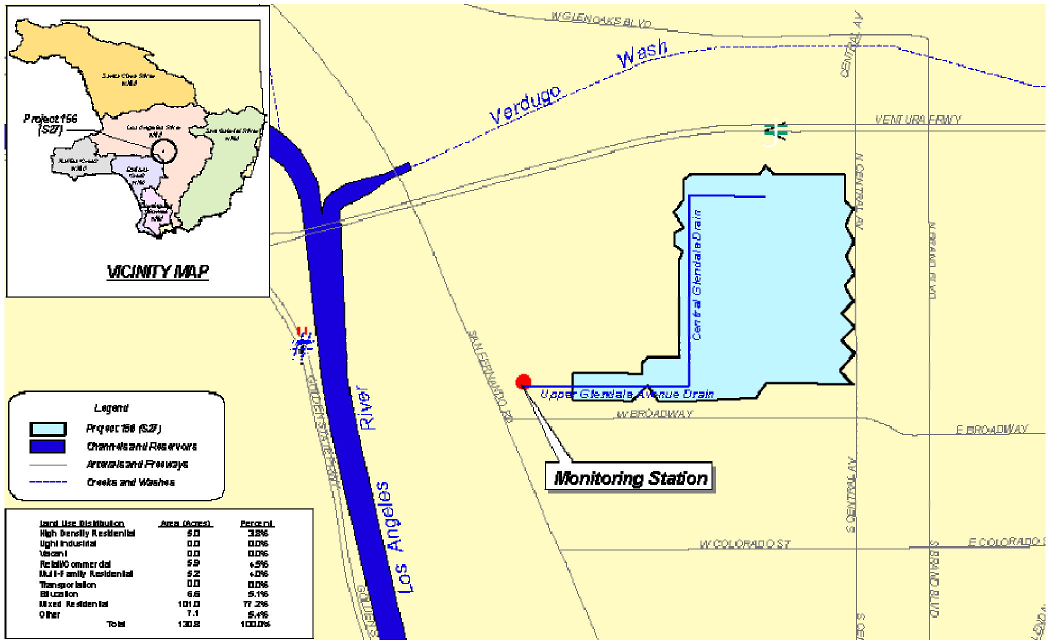


Figure A.7 Mixed residential landuse monitoring site at Concord St.., Glendale

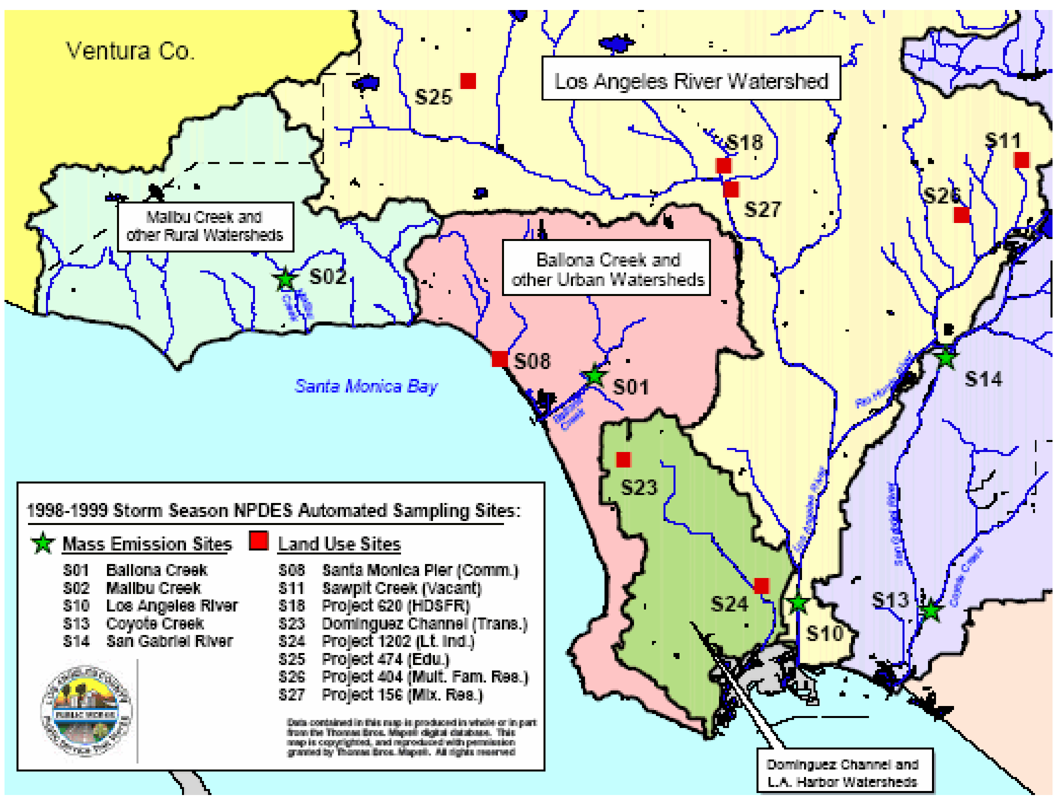


Figure A.8 Mass Emission and Landuse Sampling Sites (LADPW)

APPENDIX B. RELATIONSHIP OF LADPW AND SCAG LANDUSE CATEGOREIS

LADPW Landuse Types	SCAG Landuse Types
Educational Facilities	Special Care Facilities Pre-Schools/Day Care Centers Elementary Schools Junior or Intermediate High Schools Senior High Schools Colleges and Universities Trade Schools and Professional Training Facilities
High Density Single Family Residential	High-Density Single Family Residential Trailer Parks and Mobile Home Courts, High-Density Rural Residential, High-Density
Light Industrial	Manufacturing, Assembly, and Industrial Services Motion Picture and Television Studio Lots Research and Development Open Storage Mineral Extraction - Other Than Oil and Gas Mineral Extraction - Oil and Gas Wholesaling and Warehousing Electrical Power Facilities Water Storage Facilities Natural Gas and Petroleum Facilities Water Transfer Facilities Mixed Commercial and Industrial Mixed Urban Under Construction Marina Water Facilities
Mixed Residential	Low-Density Single Family Residential Mixed Multi-Family Residential Mixed Residential Rural Residential, Low-Density
Multifamily Residential	Duplexes, Triplexes and 2-or 3-Unit Condominiums and Townhouses Low-Rise Apartments, Condominiums, and Townhouses Medium-Rise Apartments and Condominiums High-Rise Apartments and Condominiums
Retail/Commercial	Low- and Medium-Rise Major Office Use High-Rise Major Office Use Skyscrapers Regional Shopping Center Retail Centers (Non-Strip With Contiguous Interconnected Off-Street) Modern Strip Development Older Strip Development

	Commercial Storage Commercial Recreation Hotels and Motels Attended Pay Public Parking Facilities Government Offices Fire Stations Major Medical Health Care Facilities Religious Facilities Other Public Facilities Other Special Use Facilities Communication Facilities
Transportation	Police and Sheriff Stations Non-Attended Public Parking Facilities Airports Railroads Freeways and Major Roads Park-and-Ride Lots Bus Terminals and Yards Truck Terminals Maintenance Yards
Vacant	Base (Built-up Area) Improved Flood Waterways and Structures Golf Courses Developed Local Parks and Recreation Developed Regional Parks and Recreation Undeveloped Regional Parks and Recreation Cemeteries Specimen Gardens and Arboreta Beach Parks Other Open Space and Recreation Irrigated Cropland and Improved Pasture Land Orchards and Vineyards Nurseries Horse Ranches Vacant Undifferentiated Water, Undifferentiated

APPENDIX C. RUNOFF COEFFICIENTS (BROWNE, 1990)

Landuse	A			B			C			D		
	0-2%	2-6%	6%+	0-2%	2-6%	6%+	0-2%	2-6%	6%+	0-2%	2-6%	6%+
Cultivated	0.08	0.13	0.16	0.11	0.15	0.21	0.14	0.19	0.26	0.18	0.23	0.31
land	0.14	0.18	0.22	0.16	0.21	0.28	0.2	0.25	0.34	0.24	0.29	0.41
Pasture	0.12	0.2	0.3	0.18	0.28	0.37	0.24	0.34	0.44	0.3	0.4	0.5
	0.15	0.25	0.37	0.23	0.34	0.45	0.3	0.42	0.52	0.37	0.5	0.62
Meadow	0.1	0.16	0.25	0.14	0.22	0.3	0.2	0.28	0.36	0.24	0.3	0.4
	0.14	0.22	0.3	0.2	0.28	0.37	0.26	0.35	0.44	0.3	0.4	0.5
Forest	0.05	0.08	0.11	0.08	0.11	0.14	0.1	0.13	0.16	0.12	0.16	0.2
	0.08	0.11	0.14	0.1	0.14	0.18	0.12	0.16	0.2	0.15	0.2	0.25
Residential												
Lot size	0.25	0.28	0.31	0.27	0.3	0.35	0.3	0.33	0.38	0.33	0.36	0.42
1/8 acre	0.33	0.37	0.4	0.35	0.39	0.44	0.38	0.42	0.49	0.41	0.45	0.54
Lot size	0.22	0.26	0.29	0.24	0.29	0.33	0.27	0.31	0.36	0.3	0.34	0.4
1/4 acre	0.3	0.34	0.37	0.33	0.37	0.42	0.36	0.4	0.47	0.38	0.42	0.52
Lot size	0.19	0.23	0.26	0.22	0.26	0.3	0.25	0.29	0.34	0.28	0.32	0.39
1/3 acre	0.28	0.32	0.35	0.3	0.35	0.39	0.33	0.38	0.45	0.36	0.4	0.5
Lot size	0.16	0.2	0.24	0.19	0.23	0.28	0.22	0.27	0.32	0.26	0.3	0.37
1/2 acre	0.25	0.29	0.32	0.28	0.32	0.36	0.31	0.35	0.42	0.34	0.38	0.48
Lot size	0.14	0.19	0.22	0.17	0.21	0.26	0.2	0.25	0.31	0.24	0.29	0.35
1 acre	0.22	0.26	0.29	0.24	0.28	0.34	0.28	0.32	0.4	0.31	0.35	0.46
Industrial	0.67	0.68	0.68	0.68	0.68	0.69	0.68	0.69	0.69	0.69	0.69	0.7
	0.85	0.85	0.86	0.85	0.86	0.86	0.86	0.86	0.87	0.86	0.86	0.88
Commercial	0.71	0.71	0.72	0.71	0.72	0.72	0.72	0.72	0.72	0.72	0.72	0.72
	0.88	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.9	0.89	0.89	0.9
Streets	0.7	0.71	0.72	0.71	0.72	0.74	0.72	0.73	0.76	0.73	0.75	0.78
	0.76	0.77	0.79	0.8	0.82	0.84	0.84	0.85	0.89	0.89	0.91	0.95
Open space	0.05	0.1	0.14	0.08	0.13	0.19	0.12	0.17	0.24	0.16	0.21	0.28
	0.11	0.16	0.2	0.14	0.19	0.26	0.18	0.23	0.32	0.22	0.27	0.39
Parking	0.85	0.86	0.87	0.85	0.86	0.87	0.85	0.86	0.87	0.85	0.86	0.87
	0.95	0.96	0.97	0.95	0.96	0.97	0.95	0.96	0.97	0.95	0.96	0.97

- First row of each entry gives runoff coefficients for storm recurrence intervals less than 25 years; second row gives runoff coefficients for storm recurrence intervals of 25 years or more.

APPENDIX D. Cumulative EMCs of Landuse Types for 1994-2000 Storm Seasons (LADPW)

Pollutants	Unit	HDR	Light Industrial	Vacant	Retail/ Commercial	Multi-family Residential	Transportation	Education	Mixed Residential
Total Cu	µg / L	15.30	31.04	9.12	34.77	12.23	51.86	21.49	17.33
Dissolved Cu	µg / L	8.44	20.22	n/m	14.60	6.75	32.68	12.80	11.52
Total Pb	µg / L	9.59	14.87	n/m	11.53	5.13	9.08	4.53	8.70
Total Zn	µg / L	80.35	565.60	38.81	238.53	134.88	279.45	123.69	184.85
Dissolved Zn	µg / L	39.11	460.19	n/m	164.12	75.36	203.89	65.97	125.83
TSS	mg / L	104.65	229.37	164.68	67.40	46.35	75.35	103.02	69.06
Total P	mg / L	0.39	0.44	0.11	0.41	0.19	0.44	0.31	0.26
Dissolved P	mg / L	0.29	0.28	0.06	0.30	0.16	0.36	0.27	0.20
TKN	mg / L	2.80	3.07	0.81	3.37	1.86	1.81	1.62	2.70
NO3-N	mg / L	1.04	0.86	1.11	0.58	1.73	0.75	0.63	0.71
Oil & Grease	mg / L	1.36	1.87	n/m	3.65		3.19		n/m

- HDR: High Density Single Family Residential
- Blank cells: No data available.
- n/m: Not meaningful, not enough data above detection limit.

APPENDIX E. VBA PROCEDURE FOR STORM CHARACTERISTICS CALCULATION

```
Option Explicit

' 1. First record of calculated ADD is not correct because it doesn't have prior rainfall records.
' 2. This program is ONLY applicable to the National Climate Data Center (NCDC) hourly rainfall data.
'     It is assumed that the NCDC hourly rainfall data is imported to a table in a Access database.
' 3. Results are arranged by storm event start dates.

'LIMIT_RAINFALL: An event with below LIMIT_RAINFALL will be neglected when calculating ADD.
'                        [in]
'LIMIT_DRY:          Under LIMIT_DRY hours of no rainfall records during a storm event will be
considered
'                        as a continuous event. ( [hours] < 1 day)
'FILENAME:           An Access database name.
'TABLE:                A rainfall data table name in the FILENAME database.
'OUTTABLE:          A table name where calculated ADDs are to be written.
'LOCATION:           Location of the database
'OVERWRITE:        Controls whether the OUTTABLE will be overwritten
'                        or new results will be added to the table.
Private Const LIMIT_RAINFALL = 0.04
Private Const LIMIT_DRY = 6
Private Const FILENAME = "48092"
Private Const TABLE = "48092"
Private Const OUTTABLE = TABLE & "_ADD"
Private Const LOCATION = ".\"
Private Const OVERWRITE = True


Private Sub calculate_Characteristics()
'This procedure calculates storm characteristics from the TABLE.

 'Open the featureclass
 Dim pWorkspace As IWorkspace
 Dim pTable As ITable
 Call getDatabase(pWorkspace)
 Call getTable(pWorkspace, TABLE, pTable)

 If pTable Is Nothing Then
   MsgBox "Please provide a '" & TABLE & "' table in the '" & LOCATION & "' !"
   End
 End If

 'Get a field of the "Year" column
 Dim X As Integer
 X = pTable.FindField("YEAR")

 'Get the first row of the rainfall table
 Dim j As Long                          'Index of pTable row
 Dim pRow As IRow                       'Row of rainfall table
 Dim pRowBuffer As IRowBuffer           'Row of output table
```

```
Dim dStartDate As String             'ADD start date
Dim dStopDate As String              'ADD stop date
j = 1

'To store the calculated ADDs
Dim pOutTable As ITable
Dim pCursor As ICursor
If OVERWRITE Then
  Call createTable(pWorkspace, pOutTable)
Else
  Call getTable(pWorkspace, OUTTABLE, pOutTable)
End If
Set pCursor = pOutTable.Update(Nothing, False)
Set pCursor = pOutTable.Insert(True)

'For all rows in the rainfall table, calculate ADD.
Dim dRows As Long
Dim EventStarted As Boolean
Dim dDate
Dim dTmp As Double
Dim dEventRain As Double
Dim dDuration As Double
Dim dIntensity As Double
Dim dADD As Double
Dim dTmpADD As Double
Dim dDry As Double
Dim dMin As Long
Dim i As Long
dRows = pTable.RowCount(Nothing)
EventStarted = False
Set pRow = pTable.GetRow(j)
dStartDate = pRow.Value(X + 1) & "/" & pRow.Value(X + 2) & "/" & pRow.Value(X)
dMin = LIMIT_RAINFALL * 100

Do While j < dRows

  'For 24 hour rainfall records, search a rainfall event greater than the LIMIT_RAINFALL.
  i = 1
  Do While i < 25
    dTmp = pRow.Value(X + 4 * i)

    'For a greater than zero rainfall record.
    If 0 < dTmp And dTmp < 99999 Then
      dEventRain = dEventRain + dTmp
      dDuration = dDuration + dDry + 1
      dDry = 0                        'To reset for dry periods less than or equal to LIMIT_DRY
                                      'after adding it to the current rainfall duration.
      If dIntensity < dTmp Then       'Hourly maximum rainfall intensity.
        dIntensity = dTmp
      End If

      If Not EventStarted Then
        EventStarted = True
```

```
    dStopDate = pRow.Value(X + 1) & "/" & pRow.Value(X + 2) & "/" & pRow.Value(X)
    dTmpADD = (i - 1) / 24
  End If

'For zero rainfall record after a storm event started.
ElseIf EventStarted Then
  dDry = dDry + 1                                'Count the dry hours after a rainfall event.

  'End of a storm event. It is confirmed since LIMIT_DRY hours are passed without a rainfall record.
  If dDry > LIMIT_DRY Then

    'Under LIMIT_RAINFALL events are ignored.
    If dEventRain > dMin Then

      'Add a calculated ADD and rainfall data to the pOutTable.
      Set pRowBuffer = pOutTable.CreateRowBuffer
      pRowBuffer.Value(1) = dStopDate                'Start date of the current rainfall event,
                                                     'which is also ADD stop date.
      dDate = DateDiff("d", dStartDate, dStopDate)   'dDate contains ADD except hours.
      dADD = dADD + dTmpADD                          'dTmpADD is added only
                                                     'when the event is actually ended.
      If dStartDate = dStopDate Then
        pRowBuffer.Value(2) = Format(Abs(1 - dADD), "###0.00")              '[days]
      Else
        pRowBuffer.Value(2) = Format(dADD + CDbl(dDate) - 1, "###0.00")    '[days]
      End If
      pRowBuffer.Value(3) = dEventRain / 100         '[in]
      pRowBuffer.Value(4) = dDuration                '[hours]
      pRowBuffer.Value(5) = dIntensity / 100         '[in/hours]
      pRowBuffer.Value(6) = LIMIT_RAINFALL           '[in]
      pRowBuffer.Value(7) = LIMIT_DRY                '[hours]
      pRowBuffer.Value(8) = pRow.Value(1)            'Station ID
      pCursor.InsertRow pRowBuffer

      'Reset ADD parameters only when above LIMIT_RAINFALL event occured.
      'Assume here that the rest of the hours in a day as part of next storm event ADD.
      dADD = (24 - (i - 1 - LIMIT_DRY)) / 24
      dStartDate = pRow.Value(X + 1) & "/" & pRow.Value(X + 2) & "/" & pRow.Value(X)

    End If    '(If dEventRain > dMin Then)

    'Reset for the next storm event.
    dEventRain = 0
    dDuration = 0
    dDry = 0
    dIntensity = 0
    EventStarted = False

  End If    '(If dDry > LIMIT_DRY Then)

End If    '(ElseIf EventStarted Then)

i = i + 1
```

```
    Loop

    j = j + 1
    Set pRow = pTable.GetRow(j)
  Loop

  Set pWorkspace = Nothing
  Set pTable = Nothing
  Set pRow = Nothing
  Set pRowBuffer = Nothing
  Set pOutTable = Nothing
  Set pCursor = Nothing

End Sub



Private Sub getDatabase(pWorkspace As IWorkspace)

  Dim FullFileName As String
  FullFileName = Dir(LOCATION + FILENAME + ".mdb")

  Dim pWorkspaceFactory As IWorkspaceFactory
  Set pWorkspaceFactory = New AccessWorkspaceFactory

  'Open the specified database
  If FullFileName <> "" Then
    Set pWorkspace = pWorkspaceFactory.OpenFromFile(FullFileName, 0)
  Else
    MsgBox "Rainfall data file is not found."
    End
  End If

  Set pWorkspaceFactory = Nothing

End Sub



Private Sub getTable(pWorkspace As IWorkspace, TableName As String, pTable As ITable)
'Gets a pointer to the TableName table.

  Dim pEnumDataset As IEnumDataset
  Dim pDataset As IDataset
  Set pEnumDataset = pWorkspace.Datasets(esriDTTable)

  'Check whether the named table already exists
  Do
    Set pDataset = pEnumDataset.Next
    If Not pDataset Is Nothing Then
      If pDataset.Name = TableName Then
        Exit Do
      End If
```

```
    End If
  Loop While Not pDataset Is Nothing

  If pDataset Is Nothing Then
    Call createTable(pWorkspace, pTable)
  Else
    Dim pFeatureWorkspace As IFeatureWorkspace
    Set pFeatureWorkspace = pWorkspace
    Set pTable = pFeatureWorkspace.OpenTable(pDataset.Name)
  End If

  Set pEnumDataset = Nothing
  Set pDataset = Nothing
  Set pFeatureWorkspace = Nothing

End Sub



Private Sub createTable(pWorkspace As IWorkspace, pTable As ITable)
'Creates storm characteristics table with TABLE name and '_ADD' extension.

  'Create new Fields collection
  Dim pFields As IFields
  Dim pFieldsEdit As IFieldsEdit
  Set pFields = New Fields
  Set pFieldsEdit = pFields
  pFieldsEdit.FieldCount = 9

  'Create an ID field
  Dim pField As IField
  Dim pFieldEdit As IFieldEdit
  Set pField = New Field
  Set pFieldEdit = pField
  With pFieldEdit
    .Name = "ID"
    .AliasName = ""
    .Type = esriFieldTypeOID
  End With
  Set pFieldsEdit.Field(0) = pField

  'Copy field names:
  Dim FieldNames(9) As String
  FieldNames(1) = "Date"
  FieldNames(2) = "ADD"
  FieldNames(3) = "Rainfall"
  FieldNames(4) = "Duration"
  FieldNames(5) = "Intensity"
  FieldNames(6) = "LimitRainfall"
  FieldNames(7) = "LimitDry"
  FieldNames(8) = "COOPID"

  'Create data Fields
```

```
  Dim i As Long
  i = 1
  Do While i < 9
    Set pField = New Field
    Set pFieldEdit = pField
    pFieldEdit.Name = FieldNames(i)
    If i = 1 Then
      pFieldEdit.Type = esriFieldTypeString
    Else
      pFieldEdit.Type = esriFieldTypeDouble
    End If
    Set pFieldsEdit.Field(i) = pField
    i = i + 1
  Loop

  'If the named table already exists, delete the table first
  Dim pEnumDataset As IEnumDataset
  Dim pDataset As IDataset
  Set pEnumDataset = pWorkspace.Datasets(esriDTTable)

  Do
    Set pDataset = pEnumDataset.Next
    If Not pDataset Is Nothing Then
      If pDataset.Name = OUTTABLE Then
        Exit Do
      End If
    End If
  Loop While Not pDataset Is Nothing

  If Not pDataset Is Nothing Then   'Delete the existing table
      pDataset.Delete
  End If

  Dim pFeatureWorkspace As IFeatureWorkspace
  Set pFeatureWorkspace = pWorkspace
  Set pTable = pFeatureWorkspace.createTable(OUTTABLE, pFields, Nothing, Nothing, "")

  Set pFields = Nothing
  Set pFieldsEdit = Nothing
  Set pField = Nothing
  Set pFieldEdit = Nothing
  Set pEnumDataset = Nothing
  Set pFeatureWorkspace = Nothing

End Sub
```

REFERENCES

Ackerman, D. and Schiff, K., Modeling Stormwater Mass Emissions to the Southern California Bight, Journal of Environmental Engineering, 129(4), 2003.

Barco, J., Wong, K.M., and Stenstrom, M.K., Calibration of the US EPA SWMM Model for a Large Urban Catchment, Journal of Hydraulic Engineering, ASCE, Submitted.

Brezonik, P.L. and Stadelmann, T.H., Analysis and Predictive Models of Stormwater Runoff Volumes, Loads, and Pollutant Concentrations from Watersheds in the Twin Cities Metropolitan Area, Minnesota, USA, Water Research, 36, 2002.

Browne, F.X., Stormwater Management, in Standard Handbook of Environmental Engineering, ed. by R.A. Corbitt, McGraw-Hill Inc., New York, 1990.

Brun, S.E. and Band, L.E., Simulating Runoff Behavior in an Urbanizing Watershed, Computers, Environment and Urban Systems, 24, pp 5-22, 2000.

Byrd, R.H., Lu, P., Nodedal, J., and Zhu, C., A Limited Memory Algorithm for Bound Constrained Optimization, Technical Report NAM-08, Dept. of Electrical Engineering and Computer Science, Northwestern University, 1994.

California Department of Transportation (Caltrans), BMP Retrofit Pilot Program Final Report, Division of Environmental Analysis, Sacramento, CA, 2004.

Cannon, A.J., and Whitfield, P.H., Downscaling Recent Stream-flow Conditions in British Columbia, Canada using Ensemble Neural Network Models, Journal of Hydrology, 259, pp.136-151, 2002.

Choe, J.S., Bang, K.W., and Lee, J.H., Characterization of Surface Runoff in Urban Areas, Water Science and Technology, 45(9), 2002.

Choi, K., Ball, J.E., Parameter Estimation for Urban Runoff Modeling, UrbanWater, 4, pp. 31-41, 2002.

Clar, M.L., Barfield, B.J., O'Connor, T.P., Stormwater Best Management Practice Design Guide Volume 3, Basin Best Management Practices, EPA/600/R-04/121B, National Risk Management Research Laboratory, Cincinnati, Ohio, 2004.

Cook, R.D., Detection of Influential Observations in Linear Regression, Technometrics, 19, pp. 15-18, 1977.

Cook, R.D. and Weisberg, S., Residuals and Influence in Regression, Chapman & Hall, New York, 1982.

DeVantier, B.A. and Feldman, A.D., Review of GIS Application in Hydrologic Modeling, Journal of Water Resources Planning and Management, ASCE, 119(2), 1993.

Donigian, A.S. and Huber, W.C., Modeling of Nonpoint Source Water Quality in Urban and Non-urban Areas, EPA 600/3-91/039, Office of Research and Development, Athens, GA., 1991.

Driver, N.E. and Tasker, G.D., Techniques for Estimation of Storm-Runoff Loads, Volumes, and Selected Constituent Concentrations in Urban Watersheds in the United States, USGS Open-File Report 88-191, Denver, CO., 1988.

Geman, S., Bienenstock, E., and Doursat, R., Neural Networks and the Bias/Variance Dilemma, Neural Computation, 4(1), 1992.

Han, Y.H., Lau, S-L, Kayhanian, M., and Stenstrom, M.K., Correlation Analysis among Highway Stormwater Runoff Pollutants and Characteristics, Water Science and Technology, 53(2), pp.235-243, 2006a.

Han, Y.H., Lau, S-L, Kayhanian, M., and Stenstrom, M.K., Characteristics of Highway Stormwater Runoff, Water Environment Research, 2006b (In Press).

Hansen, P.C., Analysis of Discrete Ill-posed Problems by Means of the L-Curve, SIAM Review 34(4), 1992.

Hansen, P.C. and O'Leary, D.P., The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems, SIAM J. Sci. Comput., 14(6), 1993.

Harrington, B.W., Design and Construction of Infiltration Trenches in Design of Urban Runoff Quality Control, American Society of Civil Engineers, 1989.

Hodgson, M.E., Hensen, J.R., Schmidt, R., Schill, S., and Davis, B., An Evaluation of LIDAR- and IFSAR-derived Digital Elevation Models in Leaf-on Conditions with USGS Level 1 and Level 2 DEMs, Remote Sensing of Environment, 84, 2003.

Kim, L-H, Lau, S-L, Kayhanian, M., and Stenstrom, M.K., A New Modeling Approach for Estimating First Flush Metals Loading, Paper 11, Theme 4, Proceedings of the IWA 7th Specialized Conference on Diffuse Pollution and Basin Management," Dublin, Ireland, Sept 17-22, 2003.

Kim, L-H, Kayhanian, M., Zoh, K.D., and Stenstrom, M.K., Modeling of Highway Stormwater Runoff, Science of Total Environment, Article in Press, 2005.

Lau, S-L, Bay, S., and Stenstrom, M.K., Assessment of Storm Drain Sources of Contaminants to Santa Monica Bay, Vol. VI, Toxicity of Wet Weather Urban Runoff, UCLA-ENG-99, pp. 1-E16, 1999.

Lee, J.H. and Bang, K.W., Characterization of Urban Stormwater Runoff, Water Research, 34(6), pp. 1773-1780, 2000.

Maier, H.R. and Dandy, G.C., Neural Networks for the Prediction and Forecasting of Water Resources Variables: A Review of Modeling Issues and Applications, Environmental Modeling & Software, 15, pp. 101-124, 2000.

Martin, P.H., LeBoeuf, E.J., Dobbins, J.P., Daniel, E.B., and Abkowitz, M.D., Interfacing GIS with Water Resource Models: A State-of-the-art Review, Journal of the Americal Water Resources Association, 41(6), pp. 1471-1487, 2005.

McCoy, J., Johnston, K., Kopp, S., Borup, B., Willison, J., and Payne, B., ESRI Software 9.0 Library – Using ArcGIS Spatial Analyst, pp. 141-148, ESRI, Redlands, CA, 2004.

McPherson, T.N., Burian, S.J., ASCE, A.M., Stenstrom, M.K., ASCE, F., Turin, H.J., Brown, M.J., and Suffet, I.H., Trace Metal Pollutant Load in Urban Runoff from a Southern California Watershed, Journal of Environmental Engineering, 131(7), 2005.

Meyer, S.P., Salem, T.H., and Labadie, J.W., Geographic Information Systems in Urban Stormwater Management, Journal of Water Resources Planning and Management, ASCE, 119(2), 1993.

Portland Bureau of Environmental Services, Siskiyou Curb Extension Flow Test Report, http://www.portlandonline.com/bes, Sustainable Stormwater Management Planning Group, Oregon, 2004.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., Numerical Recipes in Fortran 77: The Art of Scientific Computing, Cambridge University Press, New York, pp. 795-799, 1986.

Rahman, M. and Salbe, I., Modeling Impacts of Diffuse and Point Source Nutrients on the Water Quality of South Creek Catchment, Environment International, 21(5), 1995.

Raid, S., Mania, J., Bouchaou, L., and Najjar, Y., Rainfall-Runoff Model using an Artificial Neural Network Approach, Mathematical and Computer Modeling, 40, pp. 839-846, 2004.

Rawls, W.J., Brakensiek, D.L., and Saxton, K.E., Estimation of Soil Water Properties, Transactions ASAE, 25(5), pp. 1316-1320, 1328, 1982.

Reich, Y. and Barai, S.V., Evaluating Machine Leaning Models for Engineering Problems, Artificial Intelligence in Engineering, 13, pp. 257-272, 1999.

Reich, Y. and Barai, S.V., A methodology for Building Neural Networks Models from Empirical Engineering Data, Artificial Intelligence, 13, pp. 685-694, 2000.

Reinelt, L.E., Velikanje, J., Bell, E.J., Development and Application of a Geographic Information System for Wetland/Watershed Analysis, Computers, Environment and Urban Systems, 15, pp. 239-251, 1991.

Saez, P.B. and Rittmann, B.E., Model Parameter Estimation using Least Squares, Water Resources Bulletin, 26(6), 1992.

Scarselli, F. and Tsoi, A.C., Universal Approximation using Feed-forward Neural Networks: A Survey of Some Existing Methods, and Some New Results, Neural Networks, 11(1), 1998.

Schueler, T.R., Controlling Urban Runoff: A Practical Manual for Planning and Designing Urban BMPs, Department of Environmental Programs, Metropolitan Washington Council of Governments, Washington, D.C., 1987.

Seiker, H. and Klein, M., Best management Practices for Stormwater Runoff with Alternative Methods in a Large Urban Catchment in Berlin, Germany, Water Science and Technology, 38(10), 1998.

Shea, C., Grayman, W., Darden, D., Males, R.M., and Sushinsky, P., Integrated GIS and Hydrologic Modeling for Countywide Drainage Study, Journal of Water Resources Planning and Management, ASCE, 119(2), 1993.

Shen, J., Parker, A., and Riverson, J., A New Approach for a Windows-based Modeling System based on a Database-supporting Architecture, Environmental Modeling & Software, Article in Press, 2004.

Shinya, M., Tsuchinaga, T., Kitano, M., Tamada, Y., and Ishikawa, M., Characterization of Heavy Metals and Polycyclic Aromatic Hydrocarbons in Urban Highway Runoff, Water Science and Technology, 42(7-8), 2000.

Stenstrom, M.K. and Lau, S., Assessment of Storm Drain Sources of Contaminants to Santa Monica Bay, Volume VI. Toxicity of Wet Weather Urban Runoff, Southern California Coastal Water Research Project, 1998.

Stenstrom, M.K. and Lee, H., Final Report: Industrial Storm Water Monitoring Program Existing Statewide Permit Utility and Proposed Modifications, pp 1-79, 2005.

Tong, S.T.Y. and Chen, W., Modeling the relationship between Landuse and Surface Water Quality, Journal of Environmental Management, 66, pp 377-393, 2002.

US EPA, Better Assessment Science Integrating Point and Nonpoint Sources, BASINS 3.0 User's Manual, EPA-823-B-01-001, Office of Water, Washington, D.C., 2001.

US EPA, Results of the Nationwide Urban Runoff Program, Volume I, Final Report, NTIS PB84-185552, U.S. Environmental Protection Agency, Washington, D.C., 1983.

US EPA, Storm Water Technology Fact Sheet, Infiltration Trench, EPA 832-F-99-019, Office of Water, Washington, D.C., 1999a.

US EPA, Storm Water Technology Fact Sheet, Sand Filters, EPA 832-F-99-007, Office of Water, Washington, D.C., 1999b.

US EPA, Statistical Support Document for Proposed Effluent Limitations Guidelines and Standards for the Metal Products and Machinery Industry, EPA 821-B-00-006, Office of Water, Washington, D.C., 2000a.

US EPA, Storm Water Phase II Compliance Assistance Guide, EPA 833-R-00-002, Office of Water, Washington, D.C., 2000b.

US EPA, The National Water Quality Inventory, 1996 Report to Congress, EPA 841-R-97-008, Office of Water, Washington, D.C., 1998.

Wang, X., Integrating GIS, Simulation Models, and Visualization in Traffic Impact Analysis, Computers, Environment and Urban Systems, 29, 2005.

Weld, W.F., Cellucci, A.P., Coxe, T., Struhs, D.B., and O'Donnell, A., Stormwater Management Volume Two: Stormwater Technical Handbook, MA Dep. of Environmental Protection and MA Office of Coastal Zone Management, 1997.

Wischmeier, W.H. and Smith, D.D., Predicting Rainfall-Erosion Losses: A Guide to Conservation Planning, Agricultural Handbook No. 537, U.S. Dept. of Agriculture, Agricultural Research Service, 1978.

Woodward Clyde Consultants and Psomas and Associates, Evaluation of Landuse Monitoring Stations, Report prepared for Los Angeles County Department of Public Works, Alhambra, CA, August 1996.

Wong, K.M., Strecker, E.W., and Stenstrom, M.K., GIS to Estimate Stormwater Pollutant Mass Loadings, Journal of Environmental Engineering, ASCE, 123(8), 1997.

Wu, J.S., Allan, C.J., Saunders, W.L., and Evett, J.B., Characterization and Pollutant Loading Estimation for Highway Runoff, Journal of Environmental Engineering, 124(7), 1998.

Zeiler, M., Exploring ArcObjects Vol. I-Applications and Cartography, Vol. II-Geographic Data Management, ESRI, Redlands, CA, 2001.

Zimmerman, D., Pavlik, C., Ruggles, A., and Armstrong, M.P., An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting, Mathematical Geology, 31(4), 1999.

Zhu, C., Byrd, R.H., Lu, P., and Nocedal, J., L-BFGS-B – FORTRAN Subroutines for Large-scale Bound Constrained Optimization, http://www.ece.northwestern.edu/~nocedal/lbfgsb.html, Dept. of Electrical Engineering and Computer Science, Northwestern University, 1994.