DISSERTATION

STORMWATER CONTROL MEASURE MODELING AND UNCERTAINTY ANALYSIS FOR

TOTAL MAXIMUM DAILY LOAD COMPLIANCE

Submitted by

Christopher C. Olson

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2017

Doctoral Committee:

    Advisor: Mazdak Arabi

    Larry Roesner
    Ken Carlson
    Stephanie Kampf

ABSTRACT


STORMWATER CONTROL MEASURE MODELING AND UNCERTAINTY ANALYSIS FOR

TOTAL MAXIMUM DAILY LOAD COMPLIANCE


Cities, counties and other stormwater management agencies throughout the United States face billions of

dollars of urban stormwater improvements each year to meet total maximum daily load (TMDL)

regulations.  In many cases, they will accomplish this by implementing stormwater control measures

(SCMs) that are designed to capture urban stormwater and remove pollutants before the stormwater is

discharged back to receiving waters.  A wide variety of SCMs exist, each with unique pollutant removal

performance and associated costs.

A critical aspect of TMDL projects is modeling alternative SCM implementation strategies to evaluate

which strategies offer the greatest opportunity of TMDL compliance at reasonable costs.  However,

current SCM modeling practice suffers from several deficiencies, particularly as it relates to modeling for

TMDL compliance.  One problem is that most SCM modeling studies do not incorporate uncertainty

analysis (UA), despite recommendations from the National Research Council (NRC) and others.  This is

generally due to a lack of knowledge for how to perform UA, lack of available models/algorithms that

include UA capabilities and/or perceptions by decision makers that UA will not affect the most cost-

effective decision.  Another problem is that SCM models are typically calibrated and operated on an

"event-basis" (assuming steady-state hydraulic conditions), whereas most watershed and receiving water

models operate dynamically.  This presents practical difficulties for modelers as they link watershed

models to SCM models to receiving water models for TMDL studies and can also affect decision making

as SCM model results are based on events and many TMDLs are subject to durations of hours, days,

months, etc.

This dissertation addresses those problems by providing new tools and knowledge that can improve SCM

modeling and decision making for TMDL compliance.  In Chapter 2 ("*Uncertainty Analysis of a*

*Stormwater Control Measure Model using Global Sensitivity Analysis and Bayesian Approaches*"), we compare different UA methods and use global sensitivity analysis to determine the most sensitive parameters in a new pollutant removal model. We conclude that an informal Bayesian approach (the Generalized Uncertainty Estimation Method) provides better estimates of SCM pollutant removal uncertainty compared to a formal Bayesian approach. We also show that the TSS removal in EDBs is most sensitive to the particle size distribution and particle density of solids in the runoff entering EDBs.

In Chapter 3 ("*Appraisal of Steady-State Stormwater Control Measure Pollutant Removal Models within a Dynamic Stormwater Routing Framework with Uncertainty Analysis*"), we evaluate the effects of applying three different event-based (steady-state) SCMs models to a dynamic modeling framework. The linear regression model produces almost identical outputs under both steady-state and dynamic conditions, however the modified Fair and Geyer (MFG) model and k-C* model both produce results that underestimate TSS pollutant removal by 20-90% at the median. Using those same three models, 5-95% percentile prediction intervals (PI) were also evaluated using Monte-Carlo (MC) and first-order variance estimation (FOVE) UA methods. The FOVE method generally produced smaller PIs compared to the MC method, however, the 95th percentile values generated from the dynamically-applied SCM models were closer to the 95th percentile values generated from the steady-state SCM models using MC.

In Chapter 4 ("*Selecting Stormwater Control Measures to Achieve Total Maximum Daily Loads: The Effects of Performance Measures and Uncertainty*"), we evaluate how incorporating UA into SCM modeling can affect decision making to achieve TMDL compliance. Using theoretical TMDL scenarios and three different TMDL compliance measures, our results show that the most cost-effective SCM design/implementation strategy can be different based on the decision maker's risk level, which can only be incorporated into the decision making process through the use of UA. This finding justifies the recommendations from the NRC and others that UA should be included all TMDL modeling studies.

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# Chapter 1 - Introduction

## 1.1    Background

Urban stormwater pollution is now the primary cause of impairment of thousands of water bodies throughout the United States (EPA 2007).  Once a receiving water is designated as "impaired", meaning that one or more water quality standards (WQS) are exceeded, regulations in Section 303(d) of the Federal Clean Water Act (33 CFR §1313) are triggered.  These regulations, better known as the Total Maximum Daily Load (TMDL) regulations, require states to proceed with the TMDL process of identifying target pollutant sources and allocating target pollutant reductions to those sources.  Pollutant reduction allocations may result in *numeric limits* being set on stormwater discharges that may be enforceable on timescales as short as daily (EPA 2006).

The estimated cost of compliance for TMDLs is astounding.  For example, the EPA estimated the national *average annual* costs of implementing TMDLs to be $900 million to $3.2 billion assuming cost-effective strategies were implemented (EPA 2001).   For the City of Los Angeles alone, recent studies have estimated the total cost to implement all of its TMDLs to be as high as $70 billion (LAWPD 2009).  These costs are now becoming reality for many municipalities as regulatory, political and social pressures to re-attain WQS can no longer be ignored.

To reduce the amount of stormwater pollutants discharged to receiving waters, municipalities and other agencies responsible for stormwater management (MS4s) implement stormwater control measures (SCMs).  A SCM (also commonly referred to as a best management practice [BMP]) is "a technique, measure or structural control that is used for a given set of conditions to manage the quantity and improve the quality of stormwater runoff in the most cost-effective manner" (EPA 1999).  A wide variety of SCMs exist, each with unique pollutant removal performance and associated costs.  A critical aspect of TMDL projects is modeling alternative SCM implementation strategies to evaluate which strategies offer the

greatest reliability of TMDL compliance at reasonable costs. One limitation with current SCM modeling methods is that the majority of stormwater monitoring data that may be used for model calibration is event-averaged, so naturally most models that simulate SCM pollutant removal are steady-state, event-based models. However, there are several problems with the use steady-state SCM models. First, they are not directly compatible with the dynamic models that are most commonly applied today to simulate watershed and receiving water models. Second, the duration of individual events is variable and subject to the modeler's selection of event definition criteria. This presents problems when trying to evaluate event-based modeling outputs for TMDL compliance since most TMDL criteria are specified over durations of days, weeks and months; not events.

The uncertainty associated with water quality modeling is widely acknowledged. To address uncertainty, the EPA requires all TMDLs to incorporate a "margin of safety" that "accounts for the uncertainty in the response of the waterbody to loading reduction" (EPA 2012). While strict interpretation of this definition suggests that uncertainty need only be considered in modeling of the receiving water, we believe that uncertainty also needs to be considered in the evaluation of load reduction strategies at achieving the load reduction; that is, can SCMs reliably meet the prescribed load reduction?

Accounting for uncertainty in the modeling process using formal, explicit uncertainty analysis (UA) methods can help to avoid some of the problems currently facing TMDL stakeholders; on one hand underinvestment in pollutant reduction actions resulting in non-compliance and on the other hand overinvestment resulting in wasting of funds (NRC 2001; Reckhow 1994; Walker 2003). Despite these benefits, UA is still rarely performed in practice (Dilks and Freedman 2004; NRC 2001). One reason for this is that UA is impractical for the average practitioner to perform using today's available models (Pappenberger and Beven 2006; Reckhow 1994). To our knowledge, none of the most common models used by stormwater modelers have "built in" UA capabilities; therefore modelers must develop their own UA capability, and that often requires additional programming efforts of which they might not have the knowledge or budget to perform. Additional reasons are that uncertainty is often assumed to be either too

large (Arabi et al. 2007; Dilks and Freedman 2004) or too small (Reckhow 2003) to be of practical use in decision making. Considering the mounting pressures of re-attaining WQS in TMDL-listed water bodies and the financial costs associated with achieving those goals, it is crucial that these reasons for neglecting UA be addressed.

## 1.2 Research Objectives and Expected Contribution

The overall goal of this research is to develop new tools and knowledge that will improve the state of practice of stormwater SCM modeling for TMDL compliance. This goal is expected to be accomplished through the following research objectives:

1. Improve understanding of SCM model uncertainties
2. Develop a model that simulates SCM hydraulics and pollutant removal processes dynamically
3. Investigate the difference in SCM model outputs generated under dynamic and steady-state hydraulic conditions
4. Compare different methods for performing uncertainty analysis of SCM models
5. Evaluate how incorporating uncertainty into SCM modeling can influence decision making under scenarios of TMDL compliance

The successful attainment of these objectives is expected to demonstrate the importance of considering uncertainty when evaluating different SCMs for achieving compliance with TMDLs, as many practitioners are either not aware of the benefits of performing UA or believe that the benefits are not worth the investment. In addition, a product of this research will be a new model that simulates the performance of a variety of SCMs *and* automatically simulates uncertainty for the user. These capabilities are not currently available in any SCM simulations tools that we are aware of. Finally, testing the model using data obtained from literature and the International BMP Database will assess the applicability of the model to estimate SCM performance and its uncertainty to unmonitored or unconstructed BMPs.

## 1.3    Dissertation Outline

This dissertation is comprised of three primary chapters that each address a separate research need, but in total provide a foundation for improving stormwater SCM modeling and decision making; particularly as those activities pertain to TMDL compliance.

In Chapter 2, we compare the performance of several different Bayesian uncertainty analysis methods to provide uncertainty estimates of a SCM pollutant removal model.   The SCM model is applied to 11 different extended detention basins using data obtained from the International BMP Database and the published literature.  The performance of the different Bayesian methods are evaluated based on the percentage of measured data that fall within the $5^{th}$-$95^{th}$ percentile intervals produced by each method.  In addition, the sensitivity of the SCM model parameters are evaluated using Sobol's Global Sensitivity Method.

In Chapter 3, we simulate three different SCM pollutant removal models under both steady-state and dynamic hydraulic conditions to evaluate how applying those models (which are typically calibrated and applied assuming steady-state conditions) to a dynamic modeling framework will affect the pollutant removal outputs.  Additionally, we compare the performance of the monte-carlo (MC) uncertainty propagation method (which is very computationally intensive under a dynamic modeling framework) to the first order variance estimation (FOVE) uncertainty propagation method (which is considerably less computationally intensive) to determine if the FOVE method could be applied for SCM pollutant removal uncertainty analysis within a dynamic modeling framework.

In Chapter 4, we apply dynamic SCM pollutant modeling with uncertainty analysis to a theoretical TMDL scenario to investigate how incorporating UA into stormwater SCM modeling might affect decisions regarding SCM implementation/design when decision makers are faced with TMDL compliance.  In an attempt to simulate "real-world" decisions, the implications of uncertainty in the

reliability and vulnerability of exceeding a TMDL are considered along with the costs of different SCM designs.

# Chapter 2: Uncertainty Analysis of a Stormwater Control Measure Model using Global Sensitivity Analysis and Bayesian Approaches

## 2.1    Introduction

The extended detention basin (EDB) is one of the most common types of stormwater control measures (SCMs) used to reduce the impacts of urbanization on receiving waters.  During runoff events, EDBs capture runoff and discharge it over an extended period of time (typically 24-72 hours) to reduce downstream peak flows and remove pollutants from runoff.  The primary mechanism by which EDBs remove pollutants is through settling of particulates, and SCM performance is often measured by their ability to remove total suspended solids (TSS) due to the tendency of many pollutants to attach to particles in stormwater (United States Environmental Protection Agency 1983; Urbonas and Stahre 1993).

Simulation models are increasingly being used in the design and implementation of SCMs and the performance validity of SCM models and uncertainties associated with the estimated pollutant removal effects should be evaluated using rigorous statistical approaches (National Research Council 2001; Shirmohammadi et al. 2006). One example of a model used to evaluate SCM pollutant removal is the pseudo-physical k-C* model (Wong et al. 2006).  Parameters k and C* are "lumped" calibration coefficients that represent various factors that affect particle settling such as particle size, specific gravity, settling velocity, water temperature, flow-through rate, and other factors (Park et al. 2011; Wong et al. 2006). Wong et al. (2006) underscore the importance of investigating the influence of each of these factors individually. Park et al. (2011) used the pseudo-physical k-C* model to estimate the uncertainty of EDB TSS effluent concentrations and found that the uncertainty in the areal removal rate (k) values completely explained all of the uncertainty in measured effluent concentrations.  However, since k is a lumped parameter, it cannot be determined which factors contribute more or less to uncertainty.  The Fair and Geyer (1954) model is another model that has been recommended by the United States Environmental Protection Agency (USEPA) (USEPA 1986) and the Water Environment Research

Foundation (WERF 2013) for modeling particle removal in "sedimentation" SCMs. To the authors' knowledge, no past studies have evaluated the uncertainty of this model for estimating the TSS effluent concentrations of EDBs.

Model predictions are influenced by uncertainties in model parameters and structure, as well as errors in inputs and measurements (Beck 1987). Mathematical models are approximate representations of real world processes, therefore, any model structure will invariably include uncertainties due to incomplete understanding and representation of the system under study. While some processes are not known or well-understood, pragmatic considerations may favor neglecting processes that are deemed insignificant in the development of models. Model parameters are time invariant constants and are assumed to remain constant over computational units of the model. Since mathematical models are resolved at spatial scales that are larger than the point scales, where first principles are valid, the uncertainty in model parameters propagates into model predictions. Finally, inputs forcings to models (i.e. meteorological, soil, and land use) and measurements (observed hydrological and water quality responses) can contain errors due to instrument resolution and/or sampling methods.

The Bayesian approach is widely used as a rigorous statistical approach to assess modeling uncertainties (e.g. Ajami et al. 2007; Arabi et al. 2007; Kavetski et al. 2006; Kuczera et al. 2006; Thiemann et al. 2001). The Bayesian approach generates posterior (updated) parameter distributions (PPDs) using prior estimates of parameter uncertainty and a likelihood function that quantifies how well model simulations generated from a set of model parameters fits observed data. The PPDs can then be propagated using a Monte Carlo approach to estimate the model output uncertainty.

Two types of Bayesian approaches, "formal" and "informal", are commonly used, which differ in the types of likelihood functions that characterize the model error structure. In the formal approach, the likelihood function is objectively selected based on the form (either known or assumed) of the model residuals. Residuals are the difference between observed and simulated values and, in the formal

approach, are assumed to represent model structure and/or measurement uncertainty. Various formal likelihood functions have been derived for specific residual forms (Schoups and Vrugt 2010; Smith et al. 2010; Vrugt 2016) and various transformations can been used to generate residuals that better fit formal likelihood functions (e.g. Bates and Campbell 2001; Kuczera 1983). As long as the actual model residuals *perfectly* fit the form of the likelihood function, the likelihood function quantifies the exact probability of a parameter set to reproduce the model outputs, and the uncertainty estimates (e.g. prediction intervals) generated adhere to statistical theory and represent the exact probability of a future event occurring (Stedinger et al. 2008). However, it is rare that the actual residuals perfectly fit the assumed residual form in a real problem, even after transformation.

Alternatively, informal Bayesian approaches, the most common of which is the Generalized Likelihood Uncertainty Estimation (GLUE) method (Beven and Binley 1992), allows for goodness-of-fit measures to be used as likelihood functions. Many different goodness-of-fit measures can be used with GLUE (Beven and Binley 1992; Beven and Freer 2001; Freer et al. 1996) and the choice is ultimately that of the modeler (i.e. subjective). A primary justification for using the GLUE approach is that is it difficult to determine the exact form of model residuals required in application of the "formal" approach without making strong and perhaps unjustifiable assumptions (Beven 2006; Beven and Freer 2001). One detriment of the GLUE approach, however, is that uncertainty estimates generated from GLUE do not adhere to statistical theory (Stedinger et al. 2008), and instead may only be considered quantiles of the model predictions (Beven and Freer 2001).

Both Bayesian methods have been used extensively to calibrate and estimate uncertainty of various models, however, the outcomes of applying each method have resulted in PPDs that generate different quantities of uncertainty estimates. Some studies (e.g. Dotto et al. 2012; Freni and Mannina 2012; Freni et al. 2009; Li et al. 2010) have shown that GLUE produces much wider uncertainty predictions than formal Bayesian approaches, while others(e.g. Hutton et al. 2014; Jin et al. 2010) have shown the opposite. Vrugt et al (2008) applied formal and informal methods to estimate streamflow uncertainty on

two different watersheds and showed that both methods produce very similar estimates of uncertainty intervals.

## 2.2     Goal and Objectives

The overall goal of this study is to enhance characterization of uncertainties associated with SCM pollutant removal performance.  Specific objectives are to: (i) examine the applicability of a modified Fair and Geyer model for simulating TSS effluent concentration discharges from EDBs; (ii) investigate the importance of model parameters in the overall uncertainty of model responses; and (iii) evaluate differences in model output uncertainty using formal and informal Bayesian approaches. While a few studies have evaluated the uncertainties associated with EDBs, none of those studies used a physically-based model that explicitly represents the various physical mechanisms that drive particle settling in EDBs.  Model output uncertainties can be estimated using Bayesian methods, however previous studies have shown that different Bayesian methods (e.g. formal versus informal) can produce different estimates of uncertainties.

## 2.3     Methods and Materials

A modified version of the Fair and Geyer model was used to simulate the effluent concentrations of TSS discharged from EDBs. Data and information about performance of EDBs were obtained from the International BMP Database (www.bmpdatabase.org). The method of Sobol global sensitivity analysis was used to characterize the influence of model parameters in the variability of model responses. Both formal and informal Bayesian uncertainty estimation methods were used to obtain the PPDs of the important model parameters and predictive uncertainties in evaluation of TSS pollutant removal in EDBs.

### 2.3.1     Modified Fair and Geyer (MFG) Particle Settling Model

The USEPA (United States Environmental Protection Agency 1986) suggested the Fair and Geyer (Fair and Geyer 1954) model (Eq. 1) as an appropriate model for simulating particle removal in stormwater detention basins under flow-through conditions:

$$R = 1 - \left(1 + \frac{vA}{nQ}\right)^{-n}$$

<div align="right">Eq. 1</div>

where $R$ is the fraction of particles removed, $v$ is particle settling velocity (m/s), $A$ is basin surface area

($m^2$), $Q$ is flow rate through the basin ($m^3$/s), and $n$ is hydraulic efficiency factor.

The Fair and Geyer model was originally developed for application to water/wastewater treatment

facilities where steady-state conditions apply and the existence of hydraulic "dead zones" (i.e. non-ideal

settling) within settling basins reduced particle settling efficiency from that predicted using Camp's

(Camp 1946) concept of ideal basins. Even though such conditions rarely exist in stormwater basins

because of the intermittent and highly variable nature of rainfall/runoff, basins that fill over a short period

of time and empty over an extended period of time (i.e. EDBs) can be reasonably assumed to be operating

at steady-state over the duration of an entire event.

In this study, a modified version of the Fair and Geyer model was used, referred to hereafter as the

modified Fair and Geyer model (MFG model). In the MFG model, the removal term ($R$) in Eq. 1 is

replaced with influent ($C_{in}$) and effluent ($C_{out}$) TSS concentrations and removal of particles is simulated

with different settling velocities. Incorporating multiple particle "bins" into the model allows for a more

realistic representation of the large variability in particle sizes found in stormwater runoff (Greb and

Bannerman 1997; Kim and Sansalone 2008; Roseen et al. 2011; Selbig and Bannerman 2011; United

States Environmental Protection Agency 1986). The MFG model calculates the effluent TSS

concentration ($C_{out}$) in an EDB with $K$ particle size bins as follows:

$$C_{out} = \sum_{k=1}^{K} C_{in} * psf_k * \left(1 + \frac{v_k A}{nQ}\right)^{-n}$$

<div align="right">Eq. 2</div>

where $C_{in}$ is influent TSS concentration (mg/L), $psf_k$ is fraction of particles in particle size bin $k$, and $K$ is

the number of particle size bins. This study used the five particle bins in recent recommendations by

WERF (Water Environment Research Foundation 2013), including particle size ($d_k$) ranges: less than 10μm, 11-30μm, 31-60μm, 61-100μm, greater than 100μm.

Particle size fraction (*psf*) is a function of the distribution of particle sizes in stormwater runoff, which vary temporally and spatially. To determine ranges of *psf*, lognormal distributions were fit to particle size distributions reported in the literature. The lognormal distribution parameters (lognormal mean $\mu_d$ and lognormal standard deviation $\sigma_d$) for the particle size distribution with the highest fraction of fine particles are 1.75 and 0.75, respectively (Greb and Bannerman 1997). For the particle size distribution with the highest fraction of course particles, $\mu_d$ and $\sigma_d$ are 4.99 and 1.15 (Kim and Sansalone 2008). For any value of $\mu_d$ and $\sigma_d$, *psf* values for each of the five particle size bins were obtained as follows:

$$psf_k = F(d_k|\mu_d, \sigma_d) - F(d_{k-1}|\mu_d, \sigma_d)$$

Eq. 3

where $F(x|\mu, \sigma)$ is nonexceedance probability of $d_k$ from the lognormal cumulative distribution function with parameters μ and σ.

Particle settling velocity (*v*) in m/s is estimated using Stoke's Law in Eq. 4:

$$v = g\frac{\rho_p - \rho_w}{18v_w}d^2$$

Eq. 4

where *g* is gravitational constant (i.e., 9.81 m/s$^2$), $\rho_p$ is particle density (kg/m$^3$), $\rho_w$ is water density (kg/m$^3$), *d* is particle diameter (m), and $v_w$ is water dynamic viscosity (Pa-s). Particle diameters were assumed to be 5μm, 20 μm, 45 μm, 80 μm and 100 μm for particle size bins 1 to 5, respectively.

The application of Stoke's Law requires assumptions of spherical particles, discrete (non-hindered, non-flocculated) settling and laminar flow conditions. This assumption should be corroborated when data on particle settling velocity is available.

Particle densities are assumed to range from 1000-2000 kg/m$^3$ (Karamalegos et al. 2005; Kayhanian et al. 2008) and is assumed independent of particle size. Water density ($\rho_w$ in kg/m$^3$) and water viscosity ($v_w$ in

Pa-s) can be estimated using Eq. 5 and Eq. 6, respectively, based on water temperature ($t_w$). These linear

regressions were generated for $t_w$ values ranging from 10-30 °C using published values in Potter and

Wiggert (1997).

$$\rho_w = -0.2894\, t_w + 1002.7 \qquad\qquad \text{Eq. 5}$$

$$\nu_w = (-0.0245\, t_w + 1.5704)/1000 \qquad\qquad \text{Eq. 6}$$

The hydraulic efficiency factor ($n$) reflects the amount of short-circuiting that exists within the basin and

conceptually can be viewed as the number of continuously stirred tank reactors (CSTR) in series. On one

hand, very inefficient basins may be modeled as a single CSTR within which pollutants are evenly

distributed vertically and horizontally upon entry into the EDB. On the other hand, very efficient BMPs

may be modeled as a larger number of CSTRs (e.g. > 5) to represent near plug flow conditions (Fair and

Geyer 1954; United States Environmental Protection Agency 1986). Measuring the hydrodynamic

conditions in existing basins requires tracer studies that can be difficult and expensive to perform. Persson

et al. (1999) simulated the hydrodynamic conditions of settling basins with various shapes and inlet/outlet

configurations in order to estimate hydraulic efficiency based on those factors. In this study, the range of

$n$ was limited to 1.3 to 2.1 to reflect typical design configurations of most EDBs.

Table 1: Summary of the MFG model parameters and ranges used in this study

| Parameter | Units | Lower | Upper | Median | Source(s) |
|-----------|-------|-------|-------|--------|-----------|
| $\mu_d$ | - | 1.5 | 5 | 3.25 | (1) to (5) |
| $\sigma_d$ | - | 0.75 | 1.15 | 0.95 | (1) to (5) |
| $n$ | - | 1.3 | 2.1 | 1.7 | (6) |
| $t_w$ | Celcius | 10 | 30 | 20 | (7) |
| $\rho_p$ | kg/m$^3$ | 1000 | 2000 | 1500 | (8) and (9) |

1) United States Environmental Protection Agency 1986
2) Greb and Bannerman 1997
3) Selbig and Bannerman 2011
4) Kim and Sansalone 2008
5) Roseen et al. 2011
6) Persson et al. 1999
7) Jones and Hunt 2010
8) Karamalegos et al. 2005
9) Kayhanian et al. 2008

### 2.3.2    Global Sensitivity Analysis (GSA)

The purpose of the GSA was twofold: (1) to evaluate the importance of the MFG model parameters in simulating TSS effluent concentration and removal in EDBs; and (2) to examine the structure of model residuals. GSA can indicate model parameters that most significantly influence model outputs. Parameters that are not important can be treated as constants (typically an average or median value), which reduces the dimensionality of the parameter space that must be explored during uncertainty analysis. The exploration of the model residual structure using systemic sampling in GSA can help formulate "statistically correct" likelihood functions for formal Bayesian uncertainty analysis (Stedinger et al. 2008).

The Sobol's SA method (Sobol 1990), a global SA method based on variance decomposition, was used in this study.  The benefits of using the Sobol's method are that it is model independent and provides estimates of both the first-order and total-order effects of random parameters.  The method is the most comprehensive approach to the implementation of variance-based global sensitivity methods (Saltelli 2008; Saltelli et al. 2000).

For a given function $Y=f(X_i,...,X_k)$, variance decomposition methods propose that the overall variance of the function outputs can be decomposed as (Saltelli et al. 2004):

$$V(Y) = \sum_i V_i + \sum_i \sum_{i>j} V_{ij} + \cdots + V_{12...k}$$

Eq. 7

where $V_i = V\big(E(Y|X_i)\big)$ is the variance of the expected value of the model output with respect to parameter $X_i$, and $V_{ij} = V\big(E(Y|X_i,X_j)\big) - V_i - V_j$ is the variance of the expected value of the model output with respect to parameters $X_i$ and $X_j$ together.

To compute the variance terms, Sobol suggested to redefine the function $Y=f(X)$ as:

$$f(x) = f_o + \sum_{i=1}^{k} f_i(x_i) + \sum_{i=1}^{k} \sum_{j>i}^{k} f_{ij}(x_i, x_j) + \cdots + f_{1,\ldots,k}(x_1, \ldots, x_k)$$

Eq. 8

where $X=\{X_i,\ldots,X_k\}$ is the parameter set re-scaled to the unit hyperspace [0,1]. Assuming that all parameters are independent and that Eq. 8 is square integrable, the total variance of the model output is:

$$V(Y) = \int_0^1 f^2(x_i) - f_0^2$$

Eq. 9

The partial variance of the model output attributable to each parameter and its interaction with other parameters is:

$$V_{i_1,\ldots,i_s} = \int_0^1 \cdots \int_0^1 f_{i_1,\ldots,s}^2(x_{i_1}, \ldots, x_{i_s}) dx_{i_1}, \ldots, dx_{i_s}$$

Eq. 10

where $1 \leq i_1 \leq \cdots \leq i_s \leq k$ and $s=1, \ldots, k$.

Solutions to the integrals in Eq. 9 and Eq. 10 are not tractable. Hence, Monte-Carlo based sampling techniques are used to computationally solve the equations. Sobol's quasi-Monte Carlo sampling method (Sobol' 1967) was used in this study for sampling the parameter space in an efficient manner to compute the sensitivity indices.

The variances in Eq. 9 and Eq. 10 can then be applied to compute Sobol's "first order effect", $S$, which represent the effects of parameters on the variance of the model output individually, with all other parameters set as constants:

$$S_i = \frac{V_i}{V}$$

Eq. 11

The "total effect" sensitivity indices, $S_T$ are computed as:

$$S_{Ti} = S_i + \sum_{j \neq i} \frac{V_{ij}}{V}$$

<div align="right">Eq. 12</div>

$S_T$ represent the effect of the variance of random parameters on the variance of the model output including interactions among model parameter. The larger the value of $S$ and $S_T$, the more sensitive the model output is to the parameter.

For this study, Sobol sampling and the calculation of $S_i$ and $S_{Ti}$ were conducted using the SimLab Version 2.2.1 software package (Joint Research Center of the European Commission 2004). The MFG model includes five uncertain random parameters (Table 1), hence, a total of 24,576 model evaluations were used to estimate the sensitivity indices.

### 2.3.3    Bayesian Uncertainty Analysis

*Formal Likelihood Function*

Bayesian methods are widely applied for parameter estimation, data assimilation, and uncertainty characterization due to their intrinsic ability to account for model input, parameter, structure and measurement uncertainty (Freer et al. 1996; Kuczera et al. 2006; Vrugt 2016). Using Bayesian formalism, the posterior of model parameter ($\theta$) can be derived by conditioning the model behavior on observed responses of the system (i.e. measurement) $Y: \{y_1, y_2, \dots, y_n\}$:

$$P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{\int P(Y|\theta)P(\theta)d\theta}$$

<div align="right">Eq. 13</div>

where $\hat{Y}: \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ denotes the model response, $P(\theta)$ and $P(\theta|Y)$ denote prior and posterior parameter distributions, respectively, and $P(Y|\theta) \equiv L(\theta|Y)$ is the likelihood function.

When information about prior parameter distributions are not available, uniform ("non-informative") distribution is assumed for $P(\theta)$:

$$P(\theta) = \prod_{\theta} \left( \frac{1}{b-a} \right)$$

<div align="right">Eq. 14</div>

where $a$ and $b$ denote lower and upper bounds for model parameters. Table 1 provides a summary of the information for the parameters of the MFG model.

For the formal approach, the likelihood function represents the form of the model residuals (known or assumed) that result from model structure and/or measurement uncertainty. Assuming residuals are independent and identically distributed (IID) following a normal distribution with mean zero and constant but unknown variance $\sigma_\varepsilon^2$, the likelihood function can be derived as (Box and Tiao 1992):

$$L(\theta|Y, \sigma_\varepsilon^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \, exp\left( \frac{[y_i - \hat{y}_i(\theta)]^2}{2\sigma_\varepsilon^2} \right)$$

<div align="right">Eq. 15</div>

The likelihood function in Eq. 15 is often transformed using a log transformation for simplicity and to enhance numerical stability, which results in the log-likelihood $\mathcal{L}(\theta|Y, \sigma_\varepsilon^2)$:

$$\mathcal{L}(\theta|Y, \sigma_\varepsilon^2) = -\frac{n}{2} log(2\pi) - n \, log(\sigma_\varepsilon^2) - \frac{1}{2} \sum_{i=1}^{n} \frac{[y_i - \hat{y}_i(\theta)]^2}{\sigma_\varepsilon^2}$$

<div align="right">Eq. 16</div>

In the current study, the observed data were obtained from the International BMP Database Version 3-24-13 (BMP Database) (www.bmpdatabase.org). The variance of model residuals was estimated using the best solution, i.e. the solution with the lowest sum of squared errors, from the 24,576 model evaluations from the Sobol's GSA results.

In most hydrologic and water quality modeling cases, model errors are rarely normally distributed or homoscedastic. To generate residuals that are normally distributed and homoscedastic, the Box-Cox transformation (Box and Cox 1964) is applied:

$$X_{bc} = \frac{X_i^\lambda - 1}{\lambda}$$

Eq. 17

where $X$ denotes the original responses, $X_{bc}$ represents the Box-Cox transformed values, and $\lambda$ is the Box-Cox transformation parameter value.

### *Informal Likelihood Function*

For the implementation of the Bayesian uncertainty analysis using informal likelihood functions, two different likelihood functions suggested by Freer et al. (1996) were used. The "efficiency criterion" likelihood measure, also known as Nash-Sutcliffe Coefficient of Efficiency, is expressed as:

$$L(\theta|Y) = \left(1 - \frac{\sigma_\varepsilon^2}{\sigma_Y^2}\right); \sigma_e^2 < \sigma_Y^2$$

Eq. 18

where $\sigma_\varepsilon^2$ and $\sigma_Y^2$ denote the variance of the residuals and observed data, respectively.

The "exponential" likelihood measure is similar to the efficiency criterion, but with a shape factor $Z$. Higher values of $Z$ result in giving greater weight to those parameter sets that better fit the observed data

$$L(\theta|Y) = exp\left(-Z\frac{\sigma_\varepsilon^2}{\sigma_Y^2}\right)$$

Eq. 19

Throughout the rest of this paper, the results pertaining the application of Eq. 18 will be referred to as "Informal-1", and the results pertaining to the application of Eq. 19 using $Z$ values of 1 and 2 will be referred to as "Informal-2" and "Informal-3", respectively. For all informal scenarios, parameter sets that resulted in likelihood values greater than 0 were retained and rescaled such that the sum of all likelihood values was equal to 1.

### 2.3.4 Experimental Dataset

Information on EDB design and influent/effluent EMCs were obtained for 11 different EDBs from the BMP Database.

The 11 EDBs were selected because they had sufficient information included in the BMP Database to estimate the surface area ($A$) and discharge rate ($Q$).

Table 2 displays the names, location and primary contributing watershed land use of the EDBs used in this study. The surface area used was the average of the "Water Quality Detention Surface Area" and "Water Quality Detention Bottom Area" values reported in the BMP Database for each EDB. The one exception was for the "Lexington Hills Detention Basin" where the surface area was calculated from the reported "Water Quality Detention Volume" assuming an average depth of 4 feet (1.22m). The EDB average discharge rate (Q) was estimated by dividing the "Water Quality Detention Volume" value by the "Brim-Full Drawdown Time" value.

A total of 158 pairs of influent/effluent TSS EMCs were retrieved from the Database, of which 20 pairs that showed effluent EMCs greater than influent EMCs were eliminated from the analysis because they are not representative of typical EDB conditions. The number of EMC pairs available for each EDB ranged from 6 ("Mountain Park Detention Basin") to 27 ("Orchard Pond"). 104 (75%) EMC pairs were randomly selected as the "training" dataset in the sensitivity and uncertainty analysis procedures. The remaining 34 EMC pairs were used as the "testing" dataset.

Table 2: Extended Detention Basin information obtained from the International BMP Database

| BMP Name | Location | Land Use | Surface Area (ha) | WQCV ($m^3$) | Drawdown Time (hrs) | Discharge Rate (L/s) |
|---|---|---|---|---|---|---|
| BMP12-Malcolm Brook | Valhalla, NY | n/a | 0.196 | 3614 | 33 | 30.4 |
| BMP13- West Lake Drive | Valhalla, NY | n/a | 0.104 | 862 | 43 | 5.63 |
| I-15/SR-78 | Escondido, CA | Road/Highway | 0.0895 | 1123 | 72 | 4.33 |
| I-5/Manchester | Encinitas, CA | Road/Highway | 0.0276 | 253 | 72 | 0.977 |
| I-605/SR-91 | Los Angeles, CA | Road/Highway | 0.00854 | 70 | 72 | 0.268 |
| Lexington Hills Detention Pond | Portland, OR | Residential | 0.0279[a] | 340 | 24 | 3.93 |
| Orchard Pond | Littleton, CO | Residential | 0.104 | 506 | 40 | 3.51 |
| Greenville Pond | Greenville, NC | Office/Commercial | 0.349 | 9572 | 75 | 35.43 |
| I-5/SR-56 | San Diego, CA | Road/Highway | 0.0695 | 391 | 72 | 1.51 |
| Mountain Park Detention Basin | Lilburn, GA | Residential | 0.296 | 2803 | 10 | 33.83 |
| I-5/I-605 EDB | Downey, CA | Road/Highway | 0.0543 | 365 | 72 | 1.41 |

a = estimated assuming average depth = 4 feet (1.22m)

## 2.4    Results and Discussion

### 2.4.1    Global Sensitivity Analysis Results

Figure 1 presents the results of Sobol's global sensitivity analysis. First order sensitivity indices represent the sensitivity of the parameter with all other parameters being held constant, while total order sensitivity indices represent parameter sensitivity considering interactions with other parameters. In both cases, higher values represent greater parameter sensitivity. The results show that parameters $\mu_d$ was the most sensitive parameter, while $\rho_p$ is the second-most important parameters. Low sensitivity indices for $\sigma_d$, $n$, and $t_w$ indicated that the MFG model TSS concentration outputs are not sensitive to changes in these parameters.
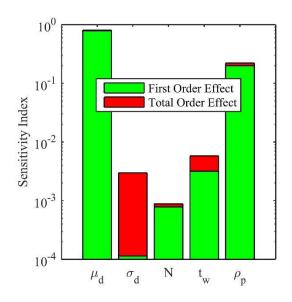
Figure 1: First and total order sensitivity indices for the MFG model parameters using Sobol's global sensitivity analysis method

For both $\mu_d$ and $\rho_p$, the sensitivity indices for total order effects are slightly greater than for first order effects. This indicates that the sensitivity of both parameters is only slightly affected by interactions with other parameters. In other words, model outputs are mainly affected by parameters $\mu_d$ and $\rho_p$, regardless of the values assigned to other model parameters. The sum of first effects and total effect indices were 0.981 and 1.035, respectively, indicating insignificant effects from interactions among model parameters on simulated TSS concentration responses.

Based on these results, the subsequent application of Bayesian methods and model output uncertainty estimation were performed assuming that $\sigma_d$, $n$, and $t_w$ values were constant and equal to the median value reported in Table 1.

### 2.4.2    Exploration of the MFG Model Error Structure and Properties

The analysis of the MFG model residuals indicated that residuals were not normally distributed and were heteroscedastic. Hence, the Box-Cox transformation was applied. Using the solutions from the Sobol's GSA, the optimal value of the Box-Cox transformation coefficient was estimated to be 0.14. Model residuals after transformation of both observed and model responses with $\lambda = 0.14$ followed a normal

distribution with mean and standard deviation -0.4 and 0.85, respectively, i.e. $\varepsilon \sim N(-0.4, 0.85)$. The

normality of model residuals after transformation of observed and simulated responses were corroborated

based on the Chi-Square Goodness-of-Fit test with a $p$-value of 0.69. Figure 2(a) shows the transformed

residuals against the theoretical normal probability distribution and Figure 2(b) shows the transformed

residuals plotted against the measured TSS effluent concentrations. For most of the data points (effluent

concentrations < 40 mg/L), the residuals appear homoscedastic; however, some bias is still apparent for

larger effluent concentrations.



Figure 2: Test of normality and homoscedasticity of the transformed residuals (Box-Cox
Transformation with λ=0.14): (a) Normal probability plot; (b) model residuals versus measured TSS
effluent concentration (mg/L).

### 2.4.3    Parameter Posterior Probability Distributions

Figure 3 shows the marginal posterior probability distributions (PPD) of parameters $\mu_d$ and $\rho_p$ for each of

the Bayesian scenarios. Summary statistics, including the 5th and 95th percentiles, median, and variance,

of the parameter distributions are provided in Table 3.

For both parameters $\mu_d$ and $\rho_p$, the PPDs generated using the formal Bayesian approach were much narrower in range (i.e. smaller variance) than those produced using the informal likelihood functions.



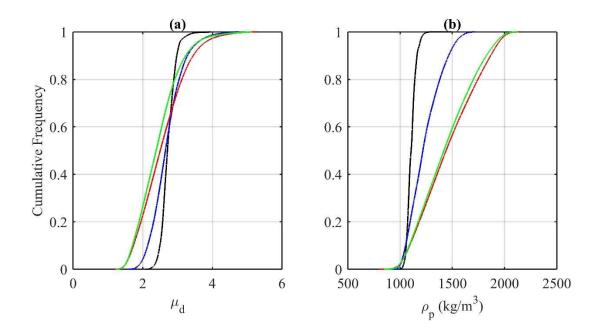Figure 3: Cumulative frequency plots of posterior parameter distributions (PPDs) for parameter $\mu_d$ (Panel a) and $\rho_p$ (Panel b) using formal Bayesian method (black line), Informal-1 method (blue line), Informal-2 method (red line) and Informal-3 method (green line).

Table 3: Statistics of the posterior parameter distributions generated using four different Bayesian methods.

| Bayesian Method | Parameter | 5th Percentile | 95th Percentile | Median | Variance |
|---|---|---|---|---|---|
| Formal | $\mu_d$ | 2.42 | 3.05 | 2.71 | 0.042 |
| | $\rho_p$ | 1048 | 1188 | 1106 | 1801 |
| Informal-1 | $\mu_d$ | 2.02 | 3.54 | 2.65 | 0.214 |
| | $\rho_p$ | 1026 | 1522 | 1218 | 23599 |
| Informal-2 | $\mu_d$ | 1.60 | 3.74 | 2.50 | 0.448 |
| | $\rho_p$ | 1039 | 1935 | 1440 | 80230 |
| Informal-3 | $\mu_d$ | 1.59 | 3.51 | 2.38 | 0.370 |
| | $\rho_p$ | 1038 | 1907 | 1419 | 73627 |

This is believed to be the result of the formal Bayesian method accounting for model structure and measurement uncertainties in the structure of the residuals (i.e. Eq. 15), while the informal methods incorporate those uncertainties implicitly into the estimation of parameter uncertainty. The Informal-1

scenario produced PPDs with much lower variance compared to both Informal-2 and Informal-3

scenarios, which suggests the use of the "efficiency criterion" likelihood measure (Eq. 18) is more

efficient as distinguishing parameter sets that better fit the observed data compared to the "exponential"

likelihood measure (Eq. 19). Results from Informal-2 and Informal-3 methods show that increasing the

value of the scaling factor (Z) from 1 to 2 resulted in slight decrease in PPD variance, as expected.

All PPDs for $\mu_d$ show the median value between 2.4 and 2.7. The $\mu_d$ parameter is the expected (mean)

value of a lognormal distribution that represents the distribution of particle sizes in urban runoff. Figure 4

shows particle size distributions associated with different values of $\mu_d$ when $\sigma_d$ is 0.95 (median value

from Table 1). Smaller values of $\mu_d$ represent particle size distributions with a higher fraction of smaller

particles (e.g. < 40µm) and higher values of $\mu_d$ represent distributions with a higher fraction of larger

particles. The value of 2.5 for $\mu_d$ for a lognormal distribution with $\sigma_d$ at 0.95 represents a particle size

distribution with approximately 42% of particles smaller than 10 µm, 40% of particles between 11 µm

and 40 µm, 13% of particles between 41 µm and 60 µm and the remaining 5% of particles being greater
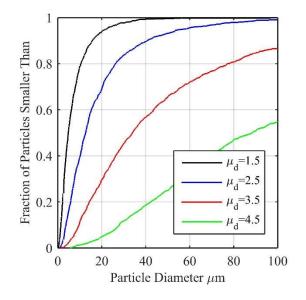
than 61 µm.



Figure 4: Relationship between $\mu_d$ parameter and particle size distribution computed assuming a lognormal distribution with $\sigma_d$ at 0.95.

The results of the Bayesian uncertainty analysis also highlighted a correlation structure between $\mu_d$ and $\rho_p$.

Figure 5 is a surface plot showing the relationship between likelihood values and different combinations

of $\mu_d$ and $\rho_p$ generated from the formal Bayesian method. Larger likelihood values (dark red color)

indicate parameter values that fit the measured dataset better than lower likelihood values (light red to

blue color). This plot shows a correlation between the two parameters which suggests the parameters are

not mutually independent. (Additional plots for informal scenarios are included in the supplemental

materials and show similar results). The correlation is inverse, where lower values of $\mu$ (i.e. higher

percentage of smaller particles) are associated with higher values of $\rho_p$, and vice versa.



Figure 5: Surface plot with 2-D contour lines showing relationship between the likelihood value
calculated from the formal Bayesian approach and values of $\mu_d$ and $\rho_p$.

### 2.4.4    Uncertainty Estimates of TSS Effluent Concentrations for EDBs

The posterior parameter distributions from all four Bayesian methods were used to estimate TSS effluent

concentrations from EDBs for each of the 34 "testing" events (Table S-1 in the supplementary materials

includes influent/effluent and BMP name for each of the testing events). For each testing event, 10,000

sets of values for $\mu_d$ and $\rho_p$ were drawn from their corresponding PPDs using Latin Hypercube Sampling

(McKay et al. 1979), and the MFG model outputs were obtained.  Parameters $\sigma_d$, $n$, and $t_w$ were held

constant at the median value shown in Table 1, while BMP-specific parameters ($A$, $Q$, $C_{in}$) were input

according to the information provided in Table 2 and Table S-1 (supplementary materials).

Figure 6 shows the results of the uncertainty analysis.  Panels a-d represent the results generated for each

of the Bayesian scenarios.  The range shown for each event represents the estimated 95% prediction

interval (P.I.) for TSS effluent concentrations. The circle symbols indicate the measured TSS effluent

concentrations obtained from the International BMP Database. Figure 7 summarizes the average 95% P.I.

and measured data inclusion rate for each of the Bayesian methods.

The 95% P.I. from the formal Bayesian approach encompassed 17 (50%) of the measured TSS

concentration data.  7 (21%) of the measured data lie above and 10 (29%) of the measured data lie below

the 95% P.I.  Results from the informal Bayesian approaches Informal-1 and Informal-2 were virtually

identical, encompassing 32 (94%) of the measured data within the 95% P.I., while only 2 (6%) data points

lie above the 95% P.I.  Finally, 33 (97%) of the measured data fall within the 95% P.I. for the Informal-3

method, with only 1 data point lying above the 95% P.I.

The differences in the number of data points that fall within the uncertainty bands for the Bayesian and

GLUE approaches is due to the differences in each approach's methodology for producing PPDs.  As

discussed in the previous section, the formal Bayesian approach produced PPDs with considerably

smaller variance compared to all informal approaches.  PPDs with smaller variance mean less parameter

uncertainty, which results in less uncertainty of model outputs when parameter uncertainty is propagated

using a Monte Carlo (or similar) approach.

Another observation to note from Figure 6 is that the uncertainty bands are not constant for all events.

This suggests that the uncertainty of TSS effluent concentrations from EDBs is not simply a function of

uncertainty in parameters $\mu_d$ and $\rho_p$.  Figure 8 displays 95% P.I. against the measured TSS influent

concentration (mg/L) for all 34 testing events.  The four data series represent the results for the four

Figure 6: Uncertainty analysis results of EDB TSS effluent concentrations (mg/L) from 10,000 MC simulations for 34 testing events.  Blue error bars represent 5th and 95th percentiles of the MC simulations and red circles represent measured TSS effluent concentrations.  (a) Results from using PPDs generated using the formal Bayesian approach, (b) Results from using PPDs generated using the Informal-1 method, (c) Results from using PPDs generated using the Informal-2 method and (d) Results from using PPDs generated using the Informal-3 method.

Figure 7: Comparison of the average 95% PI spread (mg/L) and percent of 34 testing data points that fell within the 95% PI for each Bayesian scenario.

different Bayesian scenarios. Results show a strong and significant linear trends between the two variables, which demonstrates that uncertainty of TSS effluent concentration is also significantly affected by the TSS concentration entering EDBs. Linear trendlines were fit to each of the datasets and a linear regression equation was developed to provide a quantitative estimate of the relationship for each of the Bayesian scenarios ($R^2$ values for all regression equations were greater than 0.97). The formal Bayesian results produced uncertainty bandwidths equal to approximately 25% of the influent concentration. The Informal-1, Informal-2 and Informal-3 scenarios produced uncertainty bandwidths equal to approximately 51%, 57% and 58% of influent concentration, respectively.

27

Figure 8: Scatterplot of uncertainty bandwidth (95th percentile – 5th percentile) versus influent TSS concentration (mg/L) for the 34 testing events. Linear regression equations are provided in the figure legend. $R^2$-values for all regression equations were greater than 0.97.
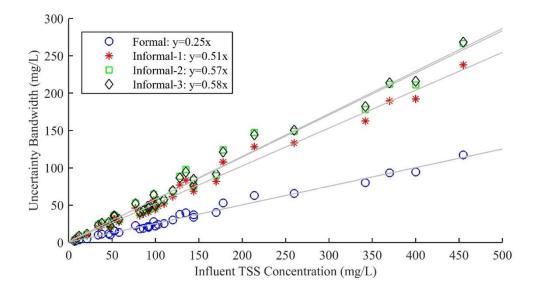
## 2.5    Summary and Conclusions

This study applied a modified version of the Fair and Geyer model (1954) to simulate the effluent

concentrations of TSS discharged from EDBs. A global sensitivity analysis was performed to identify the

most sensitive parameters of the model and uncertainty of the model outputs was estimated using both

formal and informal (i.e. GLUE) Bayesian methods. The model was applied to data obtained from the

International BMP Database; which included EDB design information for 11 different EDBs and 138

pairs of influent/effluent TSS EMCs. The following conclusions were determined as a result of this

study:

1.  When the influent TSS concentration is known, the most sensitive parameter of the modified Fair

    and Geyer model is $\mu_d$, which is the expected value of a lognormal distribution that represents the

    distribution of particle sizes (i.e. *psf*) in urban runoff. The second most sensitive parameter is

    particle density ($\rho_p$). Water temperature ($t_w$), standard deviation of the lognormal distribution

    representing particle sizes ($\sigma_d$) and the hydraulic routing parameter ($n$) were found to be relatively

    non-sensitive.

2. Posterior parameter distributions generated using the formal Bayesian approach had considerably smaller variance compared to all three informal scenarios. The Informal-1 scenario, which used the "efficiency criteria" likelihood measure, provided PPDs with smaller variance than both the Informal-2 and Informal-3 scenarios which used the "exponential" likelihood measure.

3. Prediction intervals ($5^{th}$ -$95^{th}$ percentiles) generated using the formal Bayesian method encompassed 50% of the testing dataset. The informal scenarios 1, 2 and 3 produced uncertainty bands that includes 94%, 94% and 97% of the testing dataset, respectively.

4. The width of prediction intervals produced for each testing event was found to be strongly ($R^2$ values greater than 0.97) linearly related to the (known) influent TSS concentration. Uncertainty bandwidth was found to be approximately 25%, 51%, 57% and 58% of the influent TSS concentration for the Formal, Informal-1, Informal-2 and Informal-3 methods, respectively.

5. Of the three informal methods, the Informal-1 scenario, which used the "efficiency criteria" likelihood value performed most efficiently by producing the smallest prediction intervals while still having 94% of the measured data fall within the uncertainty bands.

## Chapter 3: Appraisal of Steady-State Stormwater Control Measure Pollutant Removal Models within a Dynamic Stormwater Routing Framework with Uncertainty Analysis

### 3.1 Introduction

Stormwater runoff from urban areas can have detrimental impacts to receiving waters due to the increased discharge of runoff (volume and peak flows) and pollutants compared to undeveloped conditions. Stormwater control measures (SCMs) are constructed facilities designed to mitigate the impacts of urban stormwater by reducing the rate, volume and/or pollutant concentrations of runoff prior to being discharged downstream.

The literature is replete with studies that have quantified the effectiveness of SCMs through monitoring (e.g. Brown and Hunt 2011; Carleton et al. 2000; Comings et al. 2000; Greb and Bannerman 1997; Grizzard et al. 1986; Hathaway et al. 2009; Hunt et al. 2006; Martin 1988; Passeport et al. 2009; Stanley 1996; United States Environmental Protection Agency 1983; Winston et al. 2012). SCM effectiveness is typically reported as the difference in magnitudes of a particular metric (e.g. runoff volume, pollutant event mean concentration [EMC], pollutant load) entering (influent) and exiting (effluent) a SCM, and estimated benefits are often highly variable. Reasons for the large variability of BMP monitoring results include differences in: influent pollutant concentrations (Barrett 2008), influent particle characteristics (Greb and Bannerman 1997), location and climate (Barrett 2008), SCM design (Brown and Hunt 2011), and presentation/aggregation of results (Strecker et al. 2001), among others. The large variability of results presents challenges to extending SCM monitoring data directly to unmonitored SCMs.

Mathematical models offer an alternative method for predicting the performance of SCMs under varying environmental and design conditions. One challenge that modeler's face at the outset of a modeling project is the selection of an appropriate model. Model selection criteria may include (among others): availability of calibration/testing data, model temporal and spatial resolution compared to modeling

objectives, and the modeler's familiarity with the model (American Society of Civil Engineers and Water Environment Federation 1998; Engel et al. 2007; United States Environmental Protection Agency 2002; Urbonas 2007). The model selection criteria can conflict with one another when modeling SCM pollutant removal performance; especially for large-scale, planning-level modeling efforts such as those necessary for total maximum daily load (TMDL) allocation and implementation studies. Planning-level TMDL modeling, for example, may require the use of a watershed model to simulate pollutant loads discharged from the land surfaces, a SCM model to simulate pollutant load reductions from land surfaces under various SCM implementation scenarios, and a receiving water model to simulate fate and transport of pollutants discharged into the receiving water body. Watershed and receiving water modeling is often performed using dynamic models such as the EPA Stormwater Management Model (SWMM), USGS Hydrologic Simulation Program-Fortran (HSPF), EPA Water Quality Analysis Simulation Program (WASP) and others; which necessitates the use of dynamic SCM models in order to link dynamic inputs and outputs between various models.

Dynamic models resolve hydrologic and water quality processes in response to time-varying climatic inputs. However, most data available to calibrate SCM pollutant removal models are event-averaged values (e.g. event mean concentrations). Thus, most commonly-used SCM pollutant removal models are typically calibrated assuming steady-state conditions. The practical application of steady-state SCM pollutant removal models within a dynamic modeling framework requires the modeler to: 1) "pre-process" dynamic inputs to generate estimates of the influent pollutant EMC; 2) "post-process" the event-averaged SCM model outputs to generate dynamic inputs for the receiving water model; and 3) in some cases, perform separate dynamic hydrologic/hydraulic routing of influent flows to determine representative steady-state values of hydrologic/hydraulic flow metrics that are parameters in the steady-state SCM models. While not impossible tasks, these can be cumbersome and add to the cost and time resources needed to perform SCM modeling.

Also, the variability of SCM performance from monitoring studies underlines the importance of explicit and statistically rigorous analysis of modeling uncertainty. Uncertainty analysis (UA) techniques quantify the uncertainty of model responses as a result of uncertainty in input forcings, model parameters, model structure, and/or field/laboratory measurements. Compared to deterministic modeling, results of UA provide additional information that can enhance the decision making process; e.g. assessing the probability of an SCM discharging pollutants at a concentration/load greater than a regulatory threshold (e.g. Park et al. 2015; Park and Roesner 2012).

UA has been applied to a variety stormwater quality modeling exercises in the peer-reviewed literature (e.g. Avellaneda et al. 2009, 2010; Freni et al. 2008; Mannina and Viviani 2010; Park et al. 2011; Vezzaro et al. 2012). However the application of UA outside of academia remains relatively limited (Dilks and Freedman 2004; National Research Council 2001). One reason for this is that UA is impractical for the average practitioner to perform using current models (Pappenberger and Beven 2006; Reckhow 2003). To our knowledge, none of the most common models used by stormwater modelers have "built in" UA capabilities; therefore modelers must develop their own UA algorithms via computer programming efforts of which they might not have the knowledge or budget to perform.

Monte-Carlo (MC) and first-order variance estimation (FOVE) are two of the most common UA methods used (Shirmohammadi et al. 2006). MC methods involve realization of large numbers (typically thousands) of model responses with different combinations of model parameters. For a given set of model parameters, MC methods provide "exact" estimates of model output uncertainty as long as a sufficient number of model simulations are conducted to properly explore the model parameter space (Shirmohammadi et al. 2006; Tung et al. 2006). Thus, MC methods can be used to computationally evaluate the uncertainty estimates provided by other UA methods such FOVE (Bates and Townley 1988; Burges and Lettenmaier 1975; Melching and Anmangandla 1992; Yu et al. 2001). Practically, MC-based UA can be burdensome because of the time and resources necessary to not only run thousands of model

simulations (in some cases taking days/weeks of model run time), but also to develop code/algorithms to efficiently run thousands of model simulations and analyze thousands of model outputs.

Alternatively, FOVE methods can produce model output uncertainty estimates with a single model simulation. The FOVE method uses properties of variance to propagate the variance of random input parameters to estimate the variance of the model outputs. While the FOVE method is computationally efficient, the accuracy of the uncertainty estimates can decrease with increasing non-linearity of the model and increasing variance of the random input parameters (Morgan and Henrion 1990). FOVE has been applied to several models with relatively high nonlinearity including QUAL2E (Melching and Yoon 1996) and HSPF (Zhang and Yu 2004).

## 3.2 Study Goal and Objectives

The overall goal of this study is to evaluate the performance of new algorithms for dynamically-simulating the pollutant removal of SCMs with uncertainty. Two primary objectives have been identified for this study. The first is to investigate the effects of applying steady-state SCM pollutant removal models within a dynamic SCM routing framework. The results of this portion of the study are intended to reveal how accurate (or inaccurate) model responses (i.e. SCM effluent EMC concentrations) might be when commonly-used SCM pollutant removal models are applied within a dynamic routing framework. The second objective is to compare uncertainty estimates of the dynamic model responses using MC, prediction interval and FOVE UA methods. The results of this portion of the study are intended to reveal how the uncertainty methods that require less computational resources compare to more computationally-intensive MC methods.

## 3.3    Methods

### 3.3.1    Steady-State SCM Pollutant Removal Models

Three different SCM pollutant removal models were evaluated in this study.  The models were selected

primarily due to their prevalent use in other SCM modeling studies and recommendations as SCM

pollutant removal algorithms by the Water Environment Research Foundation (WERF) (2013).

### 3.3.2    Modified Fair and Geyer Model

The USEPA (1986) suggested the Fair and Geyer (Fair and Geyer 1954) model (Eq. 20) as an appropriate

model for simulating particle removal in stormwater detention basins under flow-through conditions.

$$R = 1 - \left(1 + \frac{vA}{nQ}\right)^{-n}$$
<div align="right">Eq. 20</div>

where $R$ denotes fraction of particles removed, $v$ denotes particle settling velocity (m/s), $A$ is basin surface

area ($m^2$), $Q$ is steady-state flowrate through basin ($m^3$/s), and $n$ represents hydraulic efficiency factor.

The Fair and Geyer model was originally developed for application to water/wastewater treatment

facilities where steady-state conditions apply and the existence of hydraulic "dead zones" (i.e. non-ideal

settling) within settling basins reduced particle settling efficiency from that predicted using Camp's

(1946) concept of ideal basins.  Even though such conditions rarely exist in stormwater basins because of

the intermittent and highly variable nature of rainfall/runoff, basins that fill over a short period of time

and empty over an extended period of time [i.e. extended detention basins (EDB)] can be reasonably

assumed to be operating at steady-state over the duration of an entire event.

In this study, we use a modified version of the Fair and Geyer model (herein referred to as the modified

Fair and Geyer model [MFG model]) (Eq. 2) that replaces the removal term ($R$) with influent ($C_i$) and

effluent ($C_e$) TSS concentrations and simulates removal of particles with different settling velocities.

Incorporating multiple particle "bins" into the model allows for a more representative analysis

considering the large distribution of particle sizes that are found in stormwater runoff (Greb and

Bannerman 1997; Kim and Sansalone 2008; Roseen et al. 2011; Selbig and Bannerman 2011; United States Environmental Protection Agency 1986). We follow recent recommendations by WERF (2013) to use five particle bins representing particle sizes in the following ranges; 2-10μm, 11-30μm, 31-60μm, 61-100μm, >100μm.

$$C_e = \sum_{k=1}^{K} C_i * psf_k * \left(1 + \frac{v_k A}{nQ}\right)^{-n}$$

Eq. 21

where $C_i$ denotes influent TSS concentration (mg/L), $C_e$ denotes effluent TSS concentration (mg/L), and $psf_k$ is fraction of particles in particle size bin $k$.

### 3.3.3 k-C* Model

The k-C* model was first proposed by Kadlec and Knight (1996) to model pollutant removal in wastewater treatment wetlands. The model uses a first-order decay coefficient ($k$, m/day) to reduce the influent concentration ($C_i$ in mg/L) towards a background or irreducible concentration ($C^*$ in mg/L) as a slug of pollutants moves through a treatment device. The model assumes steady-state and plug-flow conditions exist with the treatment device. The effluent concentration ($C_e$) is estimated as:

$$C_e = C^* + (C_i - C^*)e^{-k/q'}$$

Eq. 22

where $q'$ denotes hydraulic loading rate (m/day) which is equal to the average inflow rate (and/or discharge rate, assuming steady-state conditions) $Q$ in m³/s divided by the surface area ($A$) of the SCM, i.e. $q' = Q/A$.

Wong et al. (2006) recommended the k-C* model be adopted as a "unified approach" for simulating pollutant removal in SCMs and demonstrated the model's ability to be calibrated to measured data from several different types of SCMs for several different pollutants. It should be noted, however, that the calibration datasets in Wong et al. (2006) included spatially-variable measurements of the pollutants within the SCM. Park et al. (2011) calibrated the k-C* model to simulate TSS effluent discharged from EDBs using data obtained from the International BMP Database (BMP Database)

(www.bmpdatabase.org). The authors assumed a $C^*$ value based on the smallest TSS effluent EMC in the dataset, and calibrated $k$ assuming steady-state conditions. WERF (2013) included the k-C* model as a potential model for simulating pollutant removal in SCMs, but noted that the estimation of $k$ and $C^*$ parameters can be difficult using published SCM data (e.g. BMP Database).

### 3.3.4    Linear Regression Model

For a large number of SCMs and studies, the only information available for assessing SCM performance are measured values of influent and effluent pollutant EMCs. Recognizing the limitations of evaluating SCM performance using the percent removal metric (Strecker et al. 2001), Barrett (2005) suggested a new methodology for evaluating SCM pollutant removal performance by using linear regression of measured influent and effluent EMCs:

$$C_e = m * C_i + b + \varepsilon$$

Eq. 23

where $m$ is regression intercept, $b$ is regression slope, and $\varepsilon$ denotes residuals between the simulated and measured values of $C_e$. Barrett (2005) discussed several useful attributes of this model for evaluating SCM pollutant removal. One is that the intercept value ($m$) provides a reasonable approximation of the "irreducible minimum effluent concentration" that is frequently observed in SCM monitoring datasets. Another is that hypothesis testing can indicate whether the slope value is significantly different than zero. If it is not, then one may conclude that effluent concentrations are independent of influent concentrations for a particular BMP. Another benefit of this methodology it can be readily applied to data contained in the BMP Database for any type of BMP and pollutant (Barrett 2008).

### 3.3.5    SCM Pollutant Removal Model Parameters

Table 4 presents the mean parameter values and their uncertainty used in the pollutant removal models and UA. All of the parameter values and estimates of their uncertainty were obtained using data from the BMP Database, either as part of this study or others. Parameter values for the linear regression model were determined as part of this study using 137 pairs of influent/effluent TSS EMCs obtained from the

BMP Database. Those data and results of the linear regression analysis are presented in the supplemental materials. Parameter values for the MFG model are adapted from the results of Chapter 2 in this dissertation and parameter values for the k-C* model are adapted from results of Park et al. (2011).

Table 4: SCM pollutant removal parameter values and probability distributions

| Model | Parameter | Units | Mean | Variance | Probability Distribution Function |
|---|---|---|---|---|---|
| *Linear Regression* [1] | $m$ | - | 0.158 | - | - |
| | $b$ | mg/L | 14.95 | - | - |
| | $\varepsilon$ | mg/L | 0 | 276 | Normal |
| *MFG* [1] | $psf_1$ | - | 0.42 | 4.9E-2 | Normal |
| | $psf_2$ | - | 0.35 | 7.7E-3 | Normal |
| | $psf_3$ | - | 0.14 | 6.3E-3 | Normal |
| | $psf_4$ | - | 0.05 | 2.3E-3 | Normal |
| | $psf_5$ | - | 0.04 | 3.5E-3 | Normal |
| | $v_1$ | m/s | 6.29E-6 | 1.5E-11 | Normal |
| | $v_2$ | m/s | 1.01E-4 | 4.0E-9 | Normal |
| | $v_3$ | m/s | 5.10E-4 | 1.0E-7 | Normal |
| | $v_4$ | m/s | 1.61E-3 | 1.0E-6 | Normal |
| | $v_5$ | m/s | 2.52E-3 | 2.5E-6 | Normal |
| *k-C*$^{* [3]}$ | $k$ | m/day | 0.828 | 0.191 | Normal |
| | $C*$ | mg/L | 10 | 4 | Normal |
| Data Sources: [1] BMP Database Analysis (Appendix A) [2] Chapter 1 of this dissertation [3] Park et al (2011) | | | | | |

### 3.3.6    Dynamic Pollutant Routing

For dynamic modeling, we assume pollutants are routed through the SCM using a variable-volume, continuously stirred tank reactor (CSTR) model. This reactor model assumes that the pollutant concentration within the reactor is equal throughout (vertically and horizontally) and that the pollutant concentration discharged from the reactor is equal to the pollutant concentration within the reactor. The mass balance equation for this type of reactor can be written as:

$$\frac{\partial V \partial C}{\partial t} = Q_{i,t} \cdot C_{i,t} - Q_{e,t} \cdot C_{e,t}$$

Eq. 24

where $V$ denotes volume of runoff within the SCM (m³), $C$ is concentration of pollutant within the SCM (mg/L) (and also discharging from the SCM due to the assumption of a CSTR), $Q_i$ represents rate of runoff entering the SCM (m³/s), $Q_e$ represents rate of runoff discharging from the SCM (m³/s), $\frac{\partial V}{\partial t}$ represents rate of change of runoff stored within the SCM (m³/s), and $\frac{\partial C}{\partial t}$ denotes rate of change of pollutant concentration within the SCM in mg/L per second.  In typical reactor applications, Eq. 24 will also include additional terms describing the time-rate of pollutant removal within the reactor; however, such rates cannot be quantified for SCMs without intraevent pollutant concentration data.  Instead, we apply the steady-state pollutant removal equations (Eq. 2-Eq. 23) to runoff as it enters the reactor such that the value of $C_e$ terms in those equations become the $C_i$ term in Eq. 24.  In order to apply Eq. 2-Eq. 23 to this dynamic model, the values for $A$, $Q$ and $q'$ in those equations are computed at the timestep ($t$) that runoff enters the SCM, i.e. $A = A(t)$, $Q = Q(t)$, and $q'=q'(t)$.

Eq. 24 can be rewritten as a finite-difference approximation over a timestep interval Δt:

$$V_{t+1}C_{t+1} - V_tC_t = \frac{Q_{i,t} + Q_{i,t+1}}{2}\frac{C_{i,t} + C_{i,t+1}}{2}\Delta t - \frac{Q_{e,t} + Q_{e,t+1}}{2}\frac{C_t + C_{t+1}}{2}\Delta t \qquad \text{Eq. 25}$$

Eq. 25 can further be simplified if the dynamic model is operated over relatively small increments of Δt (i.e. on the order of minutes), where inflows and outflows are assumed to not change significantly over the timestep:

$$V_{t+1}C_{t+1} - V_tC_t = Q_{i,t}C_{i,t}\Delta t - Q_{e,t}C_t\Delta t \qquad \text{Eq. 26}$$

This assumption is justified under the conditions of this study because 1) the precipitation data used to generate the inflow hydrographs are aggregated on 1-hour increments, so runoff rates generated from the watershed model assume that precipitation is equal within each 1-hour increment; and 2) the inflow pollutographs are assumed constant throughout the entire storm event as is typically performed for planning-level studies.  (More discussion on these assumptions are provided in sections below).

Finally, Eq. 26 can be rearranged to solve for $C_{t+1}$, which is the pollutant concentration within the SCM at the end of the simulation timestep and the concentration of pollutant discharging from the SCM during the next timestep:

$$C_{t+1} = \frac{V_t C_t + Q_{i,t} C_{i,t} \Delta t - Q_{e,t} C_t \Delta t}{V_{t+1}}$$

Eq. 27

Dynamic runoff routing through the SCM is based on the continuity equation (Eq. 28) where the value for $Q_e(t)$ is estimated using the storage-indication method (Viessman and Lewis 1996) and *a priori* knowledge of the stage-volume-discharge relationship of the SCM characterized as:

$$\frac{\partial V}{\partial t} = Q_{i,t} - Q_{e,t}$$

Eq. 28

### 3.3.7 Dynamic Watershed Modeling for Inputs to Dynamic SCM Algorithms

Inputs for the SCM models were generated using SWMM. SWMM is a widely-used dynamic stormwater model capable of simulating runoff and pollutant concentrations from urban areas. Runoff quality and quantity were simulated from a hypothetical 12.1 ha (30-acre) watershed in Fort Collins, Colorado, (Table 5) for three different precipitation events. All events had a total precipitation depth of 12.7 mm (0.5 inches), but different durations were used (Table 6). These represent actual events obtained from the hourly National Climatic Data Center (www.ncdc.noaa.gov) record of Gage 053005 recorded in Fort Collins, Colorado (Table 6). A total precipitation depth of 12.7 mm was selected because the stormwater drainage criteria for Fort Collins, CO requires that SCMs be designed to capture and treat the runoff that occurs from approximately 12.7 mm of precipitation (City of Fort Collins Colorado 2011). The three different storm durations were selected to evaluate if storm duration affects the performance of the dynamic algorithms.

Pollutant concentrations were simulated using the EMC method in SWMM. The SWMM EMC method applies a constant pollutant concentration to runoff. While it is possible to calibrate SWMM's buildup/washoff algorithms to generate intraevent variability of pollutant concentrations in runoff, it is

more common to apply the EMC method for planning-level studies due to lack of available intraevent

calibration data.  In this study, we applied TSS EMCs of 50 mg/L, 125 mg/L and 200 mg/L.

Table 5: SWMM subcatchment parameter values used to generate runoff from a synthetic watershed

| Parameter | Units | Value | Parameter | Units | Value |
|-----------|-------|-------|-----------|-------|-------|
| Area | ha | 12.1 | Dstore-Imperv | mm | 2.54 |
| Width | m | 2655 | Dstore_Perv | mm | 5.08 |
| Slope | % | 2 | % Zero-Imperv | % | 25 |
| Imperviousness | % | 75 | Horton (Max Infil. Rate) | mm/hr | 76.2 |
| N-Impev | - | 0.012 | Horton (Min Infilt. Rate) | mm/hr | 12.7 |
| N-Perv | - | 0.25 | Horton (Decay Constant) | 1/hr | 4 |

Table 6: Information about three different storm events evaluated during this study

| Event | Start Date (Time) | End Date (Time) | Total Precipitation Depth (in) | Event Duration (hrs) |
|-------|-------------------|-----------------|-------------------------------|----------------------|
| 1 | 7/5/2013 (1400) | 7/5/2013 (1500) | 0.5 | 1 |
| 2 | 8/5/1993 (1100) | 8/5/1993 (1700) | 0.5 | 6 |
| 3 | 5/20/1951 (1800) | 5/21/1951 (1800) | 0.5 | 24 |

Figure 9 show the hydrographs and pollutographs generated from SWMM for the three different

precipitation events.  These hydrographs and pollutographs were used as input to the dynamic SCM
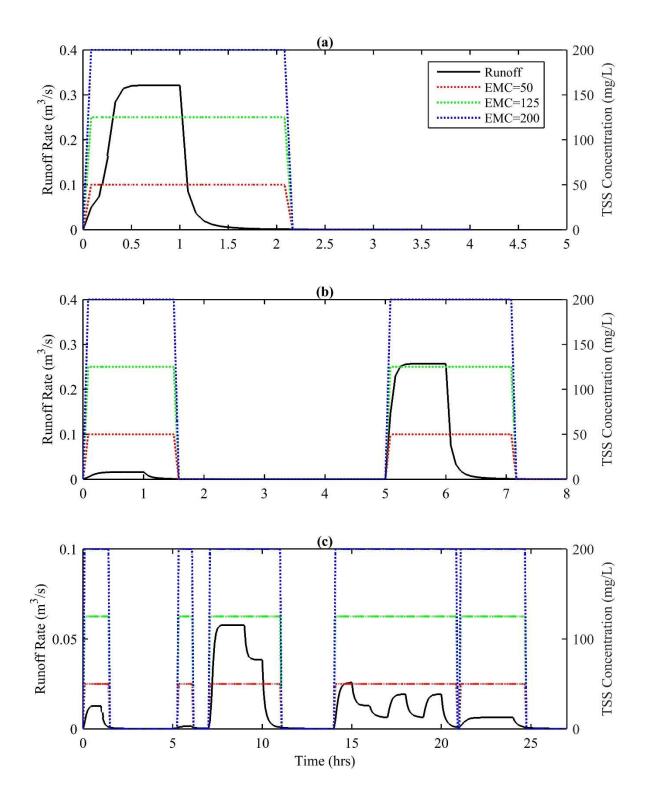
models.

Figure 9: Hydrographs and TSS pollutographs generated from SWMM and used as input to the SCM models for 1-hour precipitation event (Panel a), 6-hour precipitation event (Panel b) and 24-hour precipitation event (Panel c).

### 3.3.8 Extended Detention Basin Design

The objectives of this study were evaluated using an EDB SCM. The simulated EDB was designed according to design criteria published by the Urban Drainage and Flood Control District (UDFCD) (Urban Drainage and Flood Control District 2015), which describe methods for calculating the water quality capture volume (WQCV) and drawdown time (40 hours) of EDBs implemented along the Front Range of Colorado. We assumed the EDB has vertical side slopes and a WQCV depth of 0.91m. Table 7 presents the EDB design parameters used in the SCM simulations.

Table 7: Design parameters of the EDB evaluated in this study

| Parameter | Description | Units | Value | Notes |
|---|---|---|---|---|
| WQCV | Water Quality Capture Volume | $m^3$ | 923.7 | Computed according design criteria in UDFCD (2010). |
| $A$ | Surface area | $m^2$ | 1010 | Surface area is constant (assumed vertical side slopes) |
| $Q$ | Average discharge rate | $m^3/s$ | 6.41E-3 | Used in steady-state version of the MFG model |
| $q'$ | Average hydraulic loading rate | m/day | 0.549 | Used in steady-state version of the k-C* model |

SWMM was used to generate a stage-surface area-discharge table for the EDB assuming a single 102mm (4-inch) circular orifice is used to control the drawdown of the WQCV. The stage-surface area-discharge curve (Figure 10) was used as input to the dynamic SCM model.
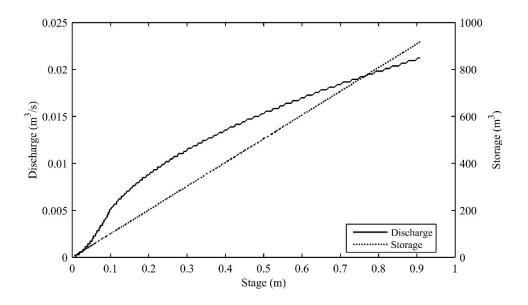
Figure 10: Stage-storage-discharge relationship for the simulated EDB

### 3.3.9    Aggregation of Dynamic Outputs

The dynamic model generates estimates of effluent flows, pollutant concentrations and uncertainty of the pollutant concentrations at every timestep of the simulation.  To compare the dynamic results to steady-state model results, the TSS effluent EMC and variance of the EMC were computed using Eq. 29 and Eq. 30, respectively:

$$EMC\ (C_e) = \frac{\sum Q_e(t) C_e(t)}{\sum Q_e(t)}$$

Eq. 29

$$VAR[EMC\ (C_e)] = \frac{\sum Q_e(t) VAR(C_e(t))}{\sum Q_e(t)}$$

Eq. 30

### 3.3.10   Uncertainty Analysis

Estimates of TSS effluent EMC uncertainty were generated using three different UA methods; MC, prediction interval for linear regression and FOVE.

### 3.3.10.1 Monte-Carlo Method with Latin Hypercube Sampling

Parameter uncertainty for the MFG and k-C* models was propagated using 1000 simulations of the steady-state and dynamic models. (Initial testing of higher numbers of simulations showed that 1000 simulations of both models were sufficient to generate stable results). Random parameter combinations were generated using the Latin Hypercube Sampling (LHS) method (McKay et al. 1979). The LHS method is a stratified sampling method that divides the multi-variate space into a specified number of equally-probable intervals from which one sample is generated randomly. The primary benefit of using the LHS sampling method is that it provides similar estimates of uncertainty to traditional random sampling methods, but using fewer model simulations. 5th and 95th percentile values were obtained for each scenario by sorting the 1000 model outputs lowest to highest and retrieving the value of the 50- and 950-ranked values, respectively.

### 3.3.10.2 First Order Variance Estimation Method

The FOVE method was also used to generate uncertainty estimates for all three pollutant removal models applied to the dynamic algorithms. The FOVE method propagates the variance of random parameters using a Taylor-series expansion of the model function to generate an estimate of the variance of the model output. Consider a function $Y = g(X)$, where $X$ is a series of random variables $X = \{X_1, X_2, \ldots, X_k\}$ and $X_o = \{\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k\}$ denotes the set of mean parameter values. The value of the function at $X_i = \{x_1, x_2, \ldots, x_k\}$ can be approximated using a first-order Taylor Series approximation (Morgan and Henrion 1990; Tung et al. 2006);

$$Y(X_i) = Y(X_o) + \sum_{j=1}^{k} \left[\frac{\partial Y}{\partial x_j}\right]_{X_o} (x_j - \bar{x}_j)$$

Eq. 31

According to properties of variance, the variance of function $Y = g(X)$ can be expressed as:

$$VAR(Y) = \sum_{j=1}^{k} \left[\frac{\partial Y}{\partial x_j}\right]_{X_o}^2 VAR(x_j) + 2 \sum_{i=1}^{n} \sum_{j>1}^{n} \left[\frac{\partial Y}{\partial x_i}\right]_{X_o} \left[\frac{\partial Y}{\partial x_j}\right]_{X_o} COV(x_i, x_j)$$

Eq. 32

The partial derivative terms for each pollutant removal model response with respect to each model parameter were determined analytically and are provided in Table 8. The probability distribution function, expected value, and variance of each random parameter assumed in this study is provided in Table 4. Additionally, it was assumed that all random parameters were independent and uncorrelated, hence the second term in the right-hand side of Eq. 32 was dropped from the final analysis.

Table 8: Partial derivatives used in the FOVE application of the SCM pollutant removal models

| Model | Parameter | Analytical Partial Derivative |
|---|---|---|
| Linear Regression | $\varepsilon$ | $\dfrac{\partial C_e}{\partial \varepsilon} = 1$ |
| MFG | $psf$ | $\dfrac{\partial C_e}{\partial psf} = \dfrac{C_i}{\left(1 + \dfrac{Av}{nQ}\right)^n}$ |
|  | $v$ | $\dfrac{\partial C_e}{\partial v} = \dfrac{-C_i * psf * A * \left(1 + \dfrac{Av}{nQ}\right)^{-n-1}}{Q}$ |
| k-C* | $C^*$ | $\dfrac{\partial C_e}{\partial C^*} = 1 - e^{-k/q'}$ |
|  | $k$ | $\dfrac{\partial C_e}{\partial k} = \dfrac{(C^* - C_i)e^{-k/q'}}{q'}$ |

Recalling that Eq. 27 is the finite-difference approximation of effluent pollutant concentrations for the dynamic algorithm, the values of both $C_t$ and $C_{i,t}$ are considered random parameters that affect the variance of $C_{t+1}$. The FOVE derivation of the dynamic pollutant routing model is:

$$VAR(C_{t+1}) = a^2 VAR(C_t) + b^2 VAR(C_{i,t}) + 2ab COV(C_t, C_{i,t})$$

Eq. 33

where $a = \dfrac{V_t - Q_{e,t}\Delta t}{V_{t+1}}$, $b = \dfrac{Q_{i,t}\Delta t}{V_{t+1}}$ and $COV(C_t, C_{i,t}) = r\sigma_{C_t}\sigma_{C_{i,t}}$ ($r$ is correlation coefficient between $C_t$ and $C_{i,t}$). We assumed $C_t$ and $C_{i,t}$ are perfectly and positively correlated (i.e. $r = 1$) and the 5th and 95th percentiles of $C_e$ were estimated assuming that $C_e$ was lognormally distributed.

### 3.3.10.3  Prediction Interval for Linear Regression Model

The prediction interval method (Walpole et al. 1998) is used estimate uncertainty intervals for the linear regression model.  The prediction interval provides the uncertainty of the dependent variable (i.e. effluent EMC) of a single, unknown event as a function of the independent variable (i.e. influent EMC) how well the best-fit linear regression parameters fit the data.  For an individual storm event with influent concentration $C_i$, the uncertainty of the predicted effluent concentration $C_e$ can be estimated using Eq. 34, which is the $(1-\alpha)100\%$ prediction interval (Walpole et al. 1998);

$$C_e \pm t_{\alpha/2,n-2}\, s \sqrt{1 + \frac{1}{n} + \frac{(C_i - \overline{C_i})^2}{\sum_{j=1}^{n}(C_{i,j} - \overline{C_i})^2}} \qquad \text{Eq. 34}$$

Where $t_{\alpha/2,n-2}$ denotes t-distribution value at $\alpha$ significance level with ($n$-2) degrees of freedom, $n$ is the number of data point pairs, $s$ represents standard error of the linear regression model, $\overline{C_i}$ denotes mean of the influent concentration data points, and $C_{i,j}$ denotes individual influent concentration data point. The $5^{\text{th}}$ and $95^{\text{th}}$ percentiles of $C_e$ were obtained using $\alpha=0.1$.

### 3.4  Results and Discussion

### 3.4.1  Deterministic Comparison of the three EDB Pollutant Removal Models

The results of deterministically modeling SCM effluent EMCs using the mean value of pollutant removal model parameters are shown in Table 9.  The linear regression model produced approximately the same (<1% difference) effluent EMCs for both the steady-state and dynamic models.  This was expected because the linear regression equation is only a function of the influent concentration, and the influent concentration was constant throughout the duration of the storm events.  These results demonstrate that linear regression models calibrated to EMC data can be applied to SCM models with dynamic routing without introducing additional error to the model results.  It is likely that the results may be different if the influent concentration varied throughout the duration of the event, however calibrating pollutant buildup/washoff models at the planning level is generally not possible due to lack of sufficient data.

Results of applying the MFG model showed that effluent EMCs are higher when applied to a dynamic model compared to a steady-state model. The dynamic model produced effluent EMCs that were approximately 50%, 38% and 22% higher than the steady-state model for storm durations of 1 hour, 6 hours and 24 hours, respectively. The effect of influent concentration on the difference between steady-state and dynamic effluent EMCs was considerably smaller (1-4%) than the effect of storm duration.

The dynamic k-C* model also produced much larger effluent EMCs compared to the steady-state model, with differences ranging from 27% to 92% higher. The largest difference (92%) was produced by applying an influent concentration of 200 mg/L during a 1-hour storm duration, while the smallest difference (27%) was produced by applying an influent concentration of 50 mg/L during a 24-hour storm duration. Overall, the differences between dynamic and steady-state model results increased with larger influent concentrations and decreased with longer storm durations.

Since the EDB was assumed to have constant surface area and influent concentration, the only difference between the steady-state and dynamic models is the value of discharge rate ($Q$). All other parameters being equal, a higher value of $Q$ will result in a higher effluent concentration for both the MFG and k-C* models. For the steady-state model, the value of $Q$ was assumed to be constant and equal to the WQCV divided by the design drawdown time (40 hours), an assumption also used by Park et al. (2011). The dynamic models considered that the discharge rate is not constant in an EDB, as the discharge rate increases with increasing depth of storage due to hydrostatic head on the outlet orifice. Figure 11 shows discharge rate generated from the dynamic model compared to the assumed discharge rate for the steady-state model for all three precipitation events tested. Clearly, the discharge rates simulated using the dynamic model are higher than the average discharge rate during a considerable portion of all storm events. Most importantly, the discharge rates are generally higher than the average discharge rate when the discharge hydrograph is rising, which represents conditions when runoff is entering the EDB. The dynamic models apply the pollutant removal equations instantaneously at the time that runoff enters the EDB, using the discharge rate computed for that timestep. When the instantaneous discharge rate is

larger than the average discharge rate when runoff enters the EDB, the MFG and k-C* models generate

higher effluent concentrations of pollutants.  Figure 11 also shows that the difference between the

instantaneous discharge rates and the average discharge rate is much lower for the 24-hour storm event

than for the 1-hour and 6-hour storm events.  This explains why the difference between the dynamic and

steady-state model results are lower for the 24-hour storm duration compared to the results of the 1-hour

and 6-hour storm durations.

Table 9: TSS Effluent EMCs generated using steady-state and dynamic applications of the SCM
pollutant removal models

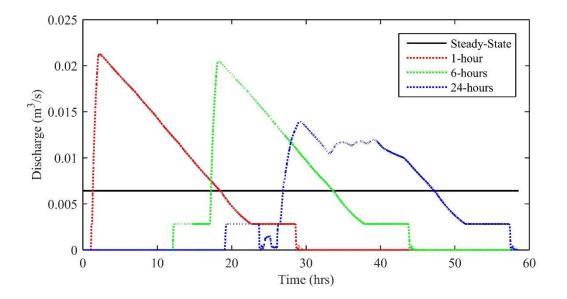| Model | Influent EMC | Effluent EMC | | |
|---|---|---|---|---|
| | | Steady-State | Dynamic | % Difference |
| Storm Duration = 1 hour | | | | |
| Linear Regression | 50 | 22.9 | 22.8 | -0.4 |
| | 125 | 34.7 | 34.6 | -0.3 |
| | 200 | 46.6 | 46.4 | -0.4 |
| Fair and Geyer | 50 | 10.0 | 14.9 | 49.0 |
| | 125 | 24.9 | 37.3 | 49.8 |
| | 200 | 38.9 | 59.7 | 53.5 |
| k-C* | 50 | 18.8 | 28.9 | 53.7 |
| | 125 | 35.4 | 64.3 | 81.6 |
| | 200 | 52.0 | 99.7 | 91.7 |
| Storm Duration = 6 hours | | | | |
| Linear Regression | 50 | 22.9 | 22.8 | -0.4 |
| | 125 | 34.7 | 34.6 | -0.3 |
| | 200 | 46.6 | 46.4 | -0.4 |
| Fair and Geyer | 50 | 10.0 | 13.6 | 36.0 |
| | 125 | 24.9 | 34.1 | 37.0 |
| | 200 | 38.9 | 54.6 | 40.4 |
| k-C* | 50 | 18.8 | 27.0 | 43.6 |
| | 125 | 35.4 | 59.0 | 66.7 |
| | 200 | 52.0 | 91.0 | 75.0 |
| Storm Duration = 24 hours | | | | |
| Linear Regression | 50 | 22.9 | 22.7 | -0.9 |
| | 125 | 34.7 | 34.4 | -0.9 |
| | 200 | 46.6 | 46.1 | -1.1 |
| Fair and Geyer | 50 | 10.0 | 12.0 | 20.0 |
| | 125 | 24.9 | 30.1 | 20.9 |
| | 200 | 38.9 | 48.2 | 23.9 |
| k-C* | 50 | 18.8 | 23.8 | 26.6 |
| | 125 | 35.4 | 49.8 | 40.7 |
| | 200 | 52.0 | 75.8 | 45.8 |

Figure 11: EDB discharge hydrographs for the three dynamically simulated rainfall/runoff events compared to average discharge rate used for steady-state simulations

## 3.4.2   Modeling Uncertainty in the Assessment of the Effects of EDBs

The 95% Prediction Interval (P.I.) TSS effluent concentrations are presented in Figure 12 for the MFG model, Figure 13 for the k-C* model, and Figure 14 for the linear regression model.  The marker on each figure represents the mean effluent TSS EMC while lines show the P.I.  Each panel represents the results for a specific influent TSS concentration (e.g. 50 mg/L, 125 mg/L or 200 mg/L).  For the MFG and k-C* models, three different UA results are presented for each storm duration (e.g. 1-hour, 6-hours or 24-hours).  The first interval for each storm duration category ("circle" marker) represents the results of 1000 Monte Carlo simulations of the respective steady-state models using the random parameter distributions shown in Table 4.  The second interval ("square" marker) represents the results of 1000 Monte Carlo simulations of the respective dynamic model. The third interval ("diamond" marker) represents the results of applying the FOVE method to the respective dynamic models.

49

For the Monte Carlo scenarios, the percentiles were obtained directly from the 1000 simulation outputs while the percentiles for the FOVE scenario were obtained assuming a lognormal distribution of the outputs. For the linear regression model, the left interval ("circle" marker) was generated using the steady-state model with uncertainty estimated using the prediction interval method and the right interval ("square" marker) was the result of applying the FOVE method to the dynamic model.

### 3.4.2.1 MFG Model

The width of the P.I.'s generated using Monte Carlo simulations were approximately the same for both the steady-state and dynamic MFG models. The primary difference between the results of those two scenarios is that the dynamic mean EMC value and uncertainty intervals were consistently higher. As discussed in the previous section, this shift in outputs from the dynamic model is likely attributed to the use of the instantaneous discharge rate to compute pollutant removal in the dynamic model. The difference between the steady-state and dynamic model results was smaller for larger duration storms, but the overall P.I. remained mostly unchanged.

Application of the FOVE method to the dynamic MFG model produced slightly narrower P.I. compared to the MC method under all combinations of influent concentration and storm duration. The 5th percentile values generated using the FOVE were all larger than those generated using the MC simulations and the 95th percentile values generated using the FOVE are all smaller than those generated using the MC simulations. The maximum *absolute* difference between FOVE- and MC-generated 95th percentile values was approximately 24 mg/L, which occurs with influent concentration at 200 mg/L and storm duration of 24 hours. The maximum *relative* difference between those values was approximately 21%, which occurs with influent concentration at 125 mg/L and storm duration of 24 hours. The width of P.I. generated by the FOVE method tended to decrease with increasing storm duration.

Interestingly, the P.I. values were nearly the same for the FOVE and steady-state MC results for the 1-hour duration storm and it appears that the narrower intervals generated using the FOVE method

somewhat compensated for the dynamic model's tendency to predict higher effluent concentrations compared to the steady-state model.  However, the difference between those values increased with longer storm durations, with the FOVE method generating upper P.I. values approximately 20% smaller for a 24-hour storm duration.
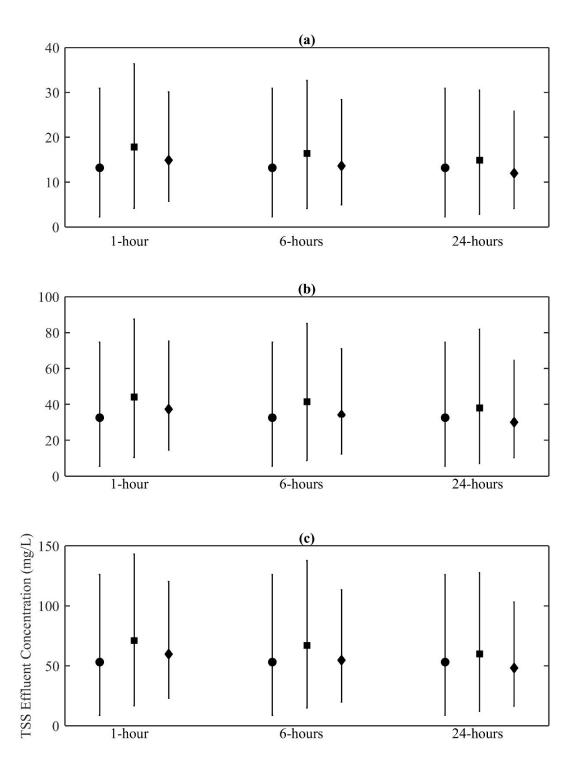
Figure 12: Uncertainty analysis results for the MFG model with influent concentration at 50 mg/L TSS (Panel a), 125 mg/L TSS (Panel b) and 200 mg/L TSS (Panel c). Markers indicate the mean effluent EMC and error bars represent the 5th and 95th percentiles. Above each storm duration, the "circle" marker indicates results from MC simulations of the steady-state MFG model, the "square" marker indicates results from MC simulations of the dynamic MFG model and the "diamond" marker indicates results from the FOVE method applied to the dynamic MFG model.

### 3.4.2.2    k-C* Model

MC simulations of the dynamic k-C* model produced noticeably narrower uncertainty intervals compared to MC simulations of the steady-state k-C* model.  However, most of the difference in uncertainty results between those applications is in the estimation of the lower percentiles; as the 95th percentile values are similar under all conditions.  The maximum *absolute* difference between 95th percentile values is approximately 13 mg/L, which occurs with influent concentration at 200 mg/L and storm duration of 1 hour.  The maximum *relative* difference between those values is approximately 7%, which occurs with influent concentration at 125 mg/L and storm duration of 1 hour.  This is an important finding when considered within the context of UA, as decision makers may be interested in planning for "worse-case" scenarios to reduce the probability that discharged pollutants exceed a regulatory threshold.  The 95th percentile value, for example, has a 5% chance of being exceeded under the simulated conditions.  Thus, one may conclude that applying the dynamic k-C* model with MC methods produces estimates of higher percentile values that are reasonably close to the steady-state model with MC methods.

The FOVE method generally produces narrower uncertainty intervals compared to the MC method when applied to the dynamic model.  One exception is for the case of influent concentration of 50 mg/L and storm duration of 24 hours where the FOVE uncertainty estimates are slightly wider.  Similar to the results generated using MFG model, the FOVE method produced lower values of 95th percentiles and higher values of 5th percentiles.  The maximum *absolute* difference between 95th percentile values is approximately 33 mg/L and the maximum *relative* difference between those values is approximately 20%, which both occur with influent concentration of 200 mg/L and storm duration of 24 hours.  In general, the FOVE uncertainty intervals tend to increase in width with increasing storm duration.
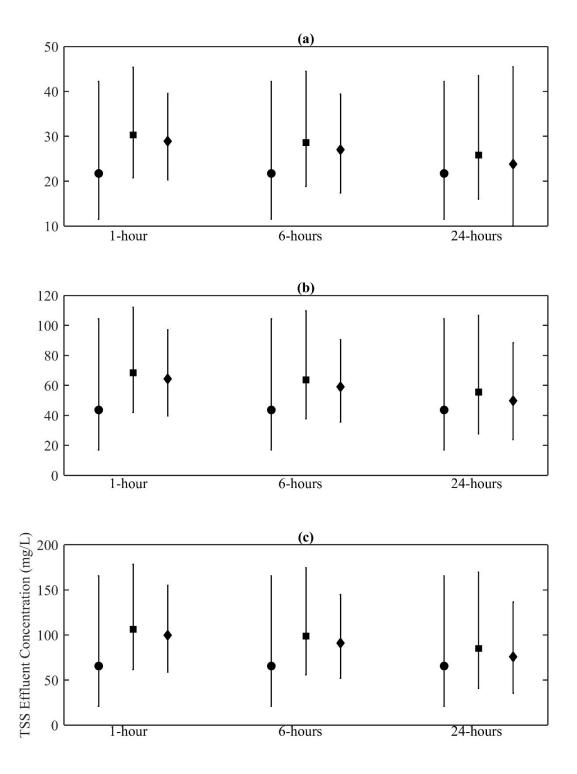
Figure 13: Uncertainty analysis results for the k-C* model with influent concentration at 50 mg/L TSS (Panel a), 125 mg/L TSS (Panel b) and 200 mg/L TSS (Panel c). Markers indicate the mean effluent EMC and error bars represent the 5th and 95th percentiles. Above each storm duration, the "circle" marker indicates results from MC simulations of the steady-state MFG model, the "square" marker indicates results from MC simulations of the dynamic MFG model and the "diamond" marker indicates results from the FOVE method applied to the dynamic MFG model.

### 3.4.2.3 Linear Regression Model

Overall, the FOVE method produces narrower uncertainty intervals than the prediction interval for all combinations of influent concentration and storm durations, however the primary effect of the smaller uncertainty intervals is the estimate of the 5th percentile value. The 95th percentile estimates are very similar for both methods; with the maximum *absolute* difference between 95th percentile values being approximately 4 mg/L (for influent concentration of 200 mg/L, regardless of storm duration) and the maximum *relative* difference being approximately 5% (for storm durations of 1 hour, regardless of influent concentration). As discussed in the previous section, the estimates of higher percentile values may be more important to decision makers as they consider how to plan and implement pollutant reduction strategies with low probability of failure; thus the application of the FOVE method within a dynamic modeling framework can likely be used without significantly affecting decision making under uncertainty.

Figure 14: Uncertainty analysis results for the linear model with influent concentration at 50 mg/L TSS (Panel a), 125 mg/L TSS (Panel b) and 200 mg/L TSS (Panel c).  Markers indicate the mean effluent EMC and error bars represent the 5th and 95th percentiles.  Above each storm duration, the "circle" marker indicates results from the prediction interval method of the steady-state model and the "square" marker indicates results from the FOVE method applied to the dynamic model.

### 3.5    Conclusions and Recommendations

This study evaluated how the application of three different SCM pollutant removal models, calibrated to available data under the assumptions of steady-state conditions, performed when applied in a dynamic modeling framework.  The primary objectives of the study were to; 1) compare estimates of SCM effluent EMCs generated from the steady-state and dynamic models and 2) compare estimates of SCM effluent EMC uncertainty generated using the FOVE method (applying with the dynamic modeling framework) and "more accurate" MC and prediction interval methods.  *It is important to reiterate that the outputs of this study are not intended to represent **absolute** estimates of effluent EMCs and their uncertainty, but instead compare how applying the same models under steady-state and dynamic conditions affect the estimates of those values in a **relative** manner.*  The following conclusions were drawn from the results of this study.

1.  Linear regression SCM pollutant removal models can be applied within a dynamic modeling framework and produce very similar estimates of SCM effluent EMCs as the calibrated steady-state model.

2.  The application of the MFG and k-C* SCM pollutant removal models within a dynamic modeling framework produce higher outputs of SCM effluent EMCs compared to the steady-state application of those models.  The MFG model may overpredict effluent EMCs by approximately 20-50% and the k-C* model may overpredict effluent EMCs by approximately 25-90%; depending on influent concentration and storm duration.  For both models, the overprediction of the dynamic effluent EMCs was reduced as the duration of the storm event increased.  This finding suggests that the results of applying these models will vary based on location and the general duration of storm events.  For example, applying these models to areas with high-intensity, short duration precipitation events (e.g. the Front Range of Colorado) will result in higher effluent EMCs compared to areas with low-intensity, long duration precipitation events (e.g. Pacific Northwest); with all else being equal.

3. The FOVE method for estimating SCM effluent EMC uncertainty produces narrower uncertainty intervals (defined by 5th and 95th percentiles) compared to "more accurate" MC and prediction interval methods in almost all scenarios. However, the differences in estimates of the 95th percentile values among different UA methods were generally much smaller than the overall uncertainty interval width generated by each UA method. For example, the maximum absolute difference between the FOVE- and MC-generated 95th percentiles was 33 mg/L (using the k-C* model, influent concentation = 200 mg/L and 24-hour storm duration). This difference is relatively small, however, considering the overall uncertainty interval width was approximately 129 mg/L and 111 mg/L for the MC and FOVE method, respectively. In addition, in some cases, the 95th percentile estimates generated using the FOVE were much closer to estimates generated using MC simulations of the steady-state model compared to MC simulations of the dynamic models. Thus, we conclude that using the FOVE method within the dynamic modeling framework described in this study produces uncertainty estimates that are similar enough to MC methods to warrant their use; especially considering the significant time and resources saved by performing a single model simulation compared to hundreds or thousands required by MC simulations.

Several assumptions and limitations of the study scope may limit the application of these conclusions to alternative modeling efforts. The limitations are discussed below along with recommendations for future studies to improve the knowledge base of SCM modeling.

1. Models were tested on only three storm events, all with total precipitation depth of 0.5 inches and durations of 1-hour, 6-hours and 24-hours. Future work could better quantify the effects of total precipitation depth (i.e. runoff volume) and intensity (i.e. runoff rate) on the performance of the dynamic models tested in this study.

2. This study was limited to evaluating TSS effluent concentrations discharged from EDBs. While EDBs are very common throughout the US, there are many other types of SCMs that similar

studies could be performed on. Wet ponds and constructed wetlands, for example, tend to operate

under conditions that are closer to steady-state due to having designed permanent pools where

runoff is stored for extended periods of time. Thus, the results of simulating those types of SCMs

within a dynamic modeling framework may produce outputs much closer to the steady-state

model results.

3. All uncertain (random parameters) were considered independent and uncorrelated in the UA.

   This is a common assumption in UA due to lack of available data to quantify correlation among

   different parameters. Future studies could look to identify correlations between particle size and

   particle density for application of the MFG model and/or correlations between k and C* for

   application of the k-C* model to potentially improve uncertainty estimates.

4. All random parameters were assumed to be normally distributed and outputs from the FOVE

   method were assumed to be lognormally distributed. These assumptions were made primarily

   due to lack of available data and to decrease complexity of the overall modeling efforts. The

   differences in uncertainty estimates generated using the different UA methods are likely to

   change with different random parameter distributions.

**Chapter 4: Selecting Stormwater Control Measures to Achieve Total Maximum Daily**

**Loads: The Effects of Performance Measures and Uncertainty**

### 4.1    Introduction

In the US, impaired waterbodies are listed under Section 303(d) of the Clean Water Act, also known as the "total maximum daily load" (TMDL) regulation.  The TMDL process generally involves the following steps (USEPA 2008); 1) listing the waterbody as "impaired", 2) identifying the sources of the pollutant causing the impairment 3) determining how much each source needs to reduce the amount of pollutant it discharges 4) evaluating and implementing pollution control practices to achieve the required pollutant discharge reductions and 5) continued monitoring of the water body to determine if the selected pollution control practices are achieving the TMDL goals.  (If monitoring indicates that goals are not being achieved, then additional pollution control practices may be required).

The United States Environmental Protection Agency (USEPA 2015) reports over 10,000 waterbodies impaired by pathogens, over 7,000 impaired by nutrients and almost 6,000 waterbodies impaired by sediment; with urban stormwater identified as a source of those pollutants in many of those impaired waterbodies (USEPA 2008).  Regulated stormwater communities (herein referred to as "decision makers") that discharge to these impaired waterbodies must implement stormwater control measures (SCMs) to achieve the required pollutant load reduction.  Faced with a myriad of options of different types of SCMs and designs to choose from, decision makers often rely on mathematical modeling of SCM performance to aid in decision making.

In a review of the TMDL program, the National Research Council (NRC) suggested that TMDL modeling studies include explicit quantification of modeling uncertainties to reduce the potential for overestimating necessary pollutant reductions (and increasing costs) and underestimating pollutant reductions (and increasing the duration of waterbody impairment) (National Research Council 2001).  Specifically, it is

recommended that uncertainty analysis be used to estimate the margin of safety component of the TMDL, which in effect increases the load reduction requirements. Multiple studies have demonstrated how to apply uncertainty analysis to watershed and receiving water models to quantify a MOS (e.g. Franceschini and Tsai 2008; Jia and Culver 2008; Langseth and Brown 2011; Zhang and Yu 2004). However, once a TMDL has been established, additional modeling must be performed to determine what pollution control practices (e.g. SCMs) can be implemented to achieve the TMDL. Studies of both agricultural best management practices (e.g. Arabi et al. 2007) and stormwater SCMs (e.g. Avellaneda et al. 2010; Park and Roesner 2012) have shown that pollutant removal can be highly uncertain and Shirmohammadi et al. (2006) recommended that uncertainty be considered in the modeling of TMDL pollution control implementation plans. In addition, the NRC (National Research Council 2001) suggested that "equivalent" pollution control actions implemented to achieve a TMDL must meet the applicable water quality standard while also considering the uncertainty of those actions.

Despite various reports and studies demonstrating the importance of considering uncertainty in the TMDL modeling process, uncertainty analysis is still rarely performed in practice (Dilks and Freedman 2004; Pappenberger and Beven 2006). We posit that in order for UA to be performed routinely in practice, it must be demonstrated that incorporating uncertainty into the decision making process could result in different (hopefully better) decisions compared to decisions made based on deterministic modeling results. For example, Krzysztofowicz (2001) discusses a real-world example of the consequences of failing to report confidence values on modeling results; when decision makers in Grand Forks, North Dakota built temporary flood control structures based on deterministic modeling results and the structures failed when the actual peak flood stage exceeded the deterministic results. He notes that the decision makers were lead to believe that the flood stage forecast was "certain", and may have made alternative decisions (e.g. constructing taller structures) if they knew the uncertainty associated with the forecast.

The objective of this study is to evaluate whether incorporating uncertainty into the decision making process for SCM implementation can result in different decisions. Specifically, we evaluate different

designs of the extended detention basin (EDB) under a theoretical situation where SCMs must achieve numeric limits on pollutant load discharges, as could be the case if the receiving water has a TMDL. SCM performance is evaluated using three different metrics and uncertainty of the performance metrics is introduced as a result of uncertainties in the pollutant removal processes of the EDB.

## 4.2 Methods

The system used in this study consists of a typical (but "synthetic") urban watershed that discharges runoff into a single EDB. Runoff from the watershed is temporarily detained in the EDB and pollutants are removed from the runoff water column prior to being discharged downstream into a receiving water. The system is simulated using a dynamic, continuous simulation approach over a 19 year period.

### 4.2.1 Watershed Modeling – Generation of Runoff Quantity and Quality Timeseries

Runoff from a 12.1 hectare (30 acre) theoretical watershed was generated using the USEPA Stormwater Management Model (SWMM) Version 5.1 (United States Environmental Protection Agency 2015b). The SWMM model was run for the period of 1/1/1995-12/31/2013 using 15-minute precipitation data obtained from the National Climatic Data Center (www.ncdc.noaa.gov) for Gage 053005 in Fort Collins, Colorado (Although this is a 19 year period of record, the gage was down for most of 1999 and from July 2001-July 2003, so the actual simulated number of years is closer to 16 years). The subcatchment parameters used in the SWMM5 model are presented in Table 5 and Figure 15 shows the timeseries of runoff rates from the watershed over the simulation period.

Table 10: SWMM subcatchment parameter values used to generate runoff from a synthetic watershed

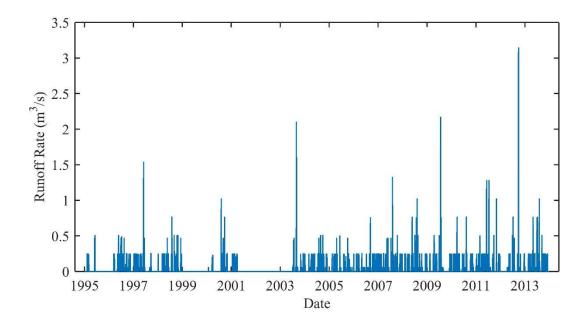| Parameter | Units | Value | Parameter | Units | Value |
|---|---|---|---|---|---|
| Area | ha | 12.1 | Dstore-Imperv | mm | 2.54 |
| Width | m | 2655 | Dstore_Perv | mm | 5.08 |
| Slope | % | 2 | % Zero-Imperv | % | 25 |
| Imperviousness | % | 75 | Horton (Max Infil. Rate) | mm/hr | 76.2 |
| N-Impev | - | 0.012 | Horton (Min Infilt. Rate) | mm/hr | 12..7 |
| N-Perv | - | 0.25 | Horton (Decay Constant) | 1/hr | 4 |

Figure 15: Runoff rate timeseries generated for the theoretical watershed using SWMM5

A timeseries of total suspended solids (TSS) concentrations in the runoff was generated by assigning a random TSS concentration to each separate runoff event. Individual runoff events were separated by at least 6 hours of zero runoff and random TSS concentrations were selected from a lognormal distribution with expected value ($\mu = 3.73$) and standard deviation ($\sigma = 1.31$). The distribution parameters were obtained by fitting a lognormal distribution to runoff TSS event mean concentrations (EMCs) obtained from the International BMP Database Version 12/31/2014 (www.bmpdatabase.org). Figure 16 shows the cumulative distribution frequencies of the EMCs obtained from the BMP Database and the EMCs applied to the runoff timeseries generated from the watershed. Figure 17 shows the timeseries of TSS concentration in runoff.
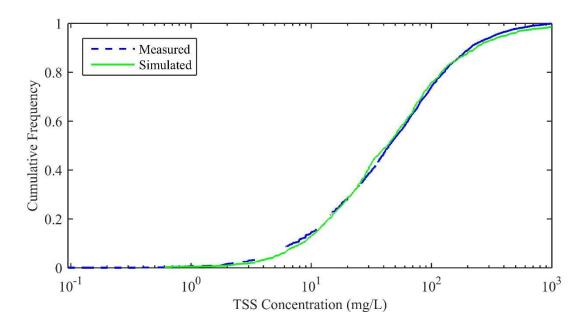
Figure 16: Cumulative distribution frequencies of TSS concentrations in runoff obtained from the International BMP Database (dashed line) and simulated from the study watershed (solid line)
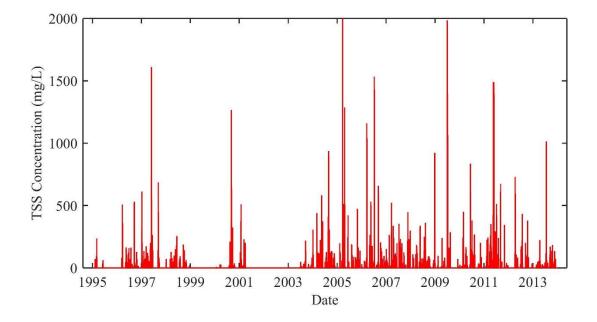


Figure 17: Runoff TSS concentration (mg/L) timeseries generated for the watershed

### 4.2.2 Extended Detention Basin Modeling

#### 4.2.2.1 Dynamic Flow Routing

EDBs were simulated dynamically using computer algorithms developed from principles of mass balance and continuity.

Runoff is routed through EDBs using a finite difference approximation of the continuity equation (Eq. 35)

$$V_{t+1} = V_t + \frac{Q_{i,t} + Q_{i,t+1}}{2}\Delta t - \frac{Q_{e,t} + Q_{e,t+1}}{2}\Delta t - Q_{b,t}\Delta t$$

Eq. 35

Where $V$ = volume of runoff stored in the EDB (m³), $Q_i$ is the rate of runoff entering the EDB (m³/s), $Q_e$ is the rate of water leaving the EDB through the control pipe/orifice (m³/s), $Q_b$ is the rate of water leaving the EDB as "overflow" (m³/s) and $t$ is the time since the start of the simulation (s). At each timestep, $Q_{e,t+1}$ is determined using the storage-indication method (Viessman and Lewis 1996) from a user-input table of the stage-volume-discharge relationship of the EDB and $Q_{b,t}$ is calculated using Eq. 36.

$$Q_{b,t} = max\left(0, V_t + \frac{Q_{i,t} + Q_{i,t+1}}{2}\Delta t - \frac{Q_{e,t} + Q_{e,t+1}}{2}\Delta t - V_T\right)$$

Eq. 36

Where $V_T$ = maximum storage volume of the EDB (m³).

#### 4.2.2.2 Pollutant Removal

Pollutant removal is simulated using a modified version of the Fair and Geyer model (Fair and Geyer 1954) (Eq. 37).

$$C_e = \sum_{j=1}^{5} C_i * psf_j * \left(1 + \frac{v_j A}{nq}\right)^{-n}$$

Eq. 37

Where $C_i$ = influent TSS concentration (mg/L), $C_e$ = effluent TSS concentration (mg/L), $psf$ = fraction of particles in particle size bin $j$, $v$ = particle settling velocity (m/s) of particles in particle size bin $j$, $A$ = basin surface area (m²), $n$ = hydraulic routing factor for non-ideal basins and $q$ = flow rate through the basin (m³/s). The Fair and Geyer model has been used in a number of reports and studies for modeling

stormwater detention basins (e.g. Chen and Adams 2006; Krishnappan et al. 1999; USEPA 1986) under the assumption that detention basins operate reasonably close to steady-state conditions when the basins fill over a short period of time and empty over a much longer period of time. The modified version of the model we use simulates the removal of five different particle sizes to provide a better representation of the distribution of particles found in urban stormwater runoff (Greb and Bannerman 1997; Kim and Sansalone 2008; Roseen et al. 2011; Selbig and Bannerman 2011; USEPA 1986). We follow recent recommendations by The Water Environment Research Foundation (Water Environment Research Foundation 2013) to use five particle bins representing particle sizes in the following ranges; 2-10µm, 11-30µm, 31-60µm, 61-100µm, >100µm.

### 4.2.2.3 Dynamic Pollutant Routing

A variable-volume, continuously stirred tank reactor (CSTR) model is used to route pollutants through EDBs. This model assumes that the pollutant concentration within the EDB is distributed evenly (vertically and horizontally) and that the pollutant concentration discharged from the EDB is equal to the pollutant concentration within the EDB. The mass balance equation for this type of reactor can be written as:

$$\frac{\partial V \partial C}{\partial t} = Q_{i,t} C_{i,t} - Q_{e,t} C_t \qquad \text{Eq. 38}$$

Where $\frac{\partial V}{\partial t}$ = rate of change of runoff stored within the EDB (m³/s), $\frac{\partial C}{\partial t}$ = rate of change of pollutant concentration within the SCM (mg/L/s), $C_i$ = pollutant concentration entering the SCM (mg/L), and $C$ = concentration of pollutant within the SCM (mg/L) (and also discharging from the EDB due to the assumption of a CSTR). In typical reactor applications, Eq. 24 will also include additional terms describing the time-rate of pollutant removal within the reactor; however such rates cannot be quantified for EDBs without intraevent pollutant concentration data. Instead, we apply the modified Fair and Geyer model (Eq. 37) to runoff as it enters the SCM such that the value of $C_e$ in Eq. 37 becomes the $C_i$ term in

Eq. 24. The values for $A$ and $q$ in Eq. 37 are computed at the timestep ($t$) that runoff enters the SCM [i.e. $A = A(t)$, $q = q(t)$].

Eq. 24 can be rewritten as a finite-difference approximation over a timestep interval $\Delta t$;

$$V_{t+1}C_{t+1} - V_tC_t = \frac{Q_{i,t} + Q_{i,t+1}}{2}\frac{C_{i,t} + C_{i,t+1}}{2}\Delta t - \frac{Q_{e,t} + Q_{e,t+1}}{2}\frac{C_t + C_{t+1}}{2}\Delta t \qquad \text{Eq. 39}$$

Eq. 25 can further be simplified if the dynamic model is operated over relatively small increments of $\Delta t$ (i.e. on the order of minutes), where inflows and outflows are assumed to not change significantly over the timestep;

$$V_{t+1}C_{t+1} - V_tC_t = Q_{i,t}C_{i,t}\Delta t - Q_{e,t}C_t\Delta t \qquad \text{Eq. 40}$$

This assumption is justified under the conditions of this study because; 1) the precipitation data used to generate the inflow hydrographs are aggregated on 15-minutes increments, so runoff rates generated from the watershed model assume that precipitation is equal within each interval and 2) the inflow pollutographs are assumed constant throughout the entire storm event as is typically performed for planning-level studies. Finally, Eq. 26 can be rearranged to solve for $C_2$, which is the pollutant concentration within the EDB at the end of the simulation timestep and the concentration of pollutant discharging from the EDB during the next timestep;

$$C_{t+1} = \frac{V_tC_t + Q_{i,t}C_{i,t}\Delta t - Q_{e,t}C_t\Delta t}{V_{t+1}} \qquad \text{Eq. 41}$$

### 4.2.3    EDB Design and Cost Estimation

Most EDB design criteria specify methods for determining the required storage volume of the EDB and drawdown time. For example, the Urban Drainage and Flood Control District (UDFCD) stormwater design criteria (Urban Drainage and Flood Control District 2015) suggests water quality EDBs to have a water quality capture volume (WQCV) computed from Eq. 42 and a brim-full drawdown time of 40 hours.

$$WQCV = a(0.91i^3 - 1.19i^2 + 0.78i)$$

<div align="right">Eq. 42</div>

Where WQCV has units of watershed-inches, $a$ = drain time coefficient whose value is 1.0 for EDBs and $i$ = imperviousness of the watershed (%/100). Others, such as the Colorado Department of Transportation (California Department of Transportation 2010), use methods outlined in the WEF/ASCE Manual of Practice (Water Environment Federation and American Society of Civil Engineers 1998) to compute the WQCV and allow for drawdown times ranging from 24-96 hours.

In this study, we evaluated 15 EDB designs with different combinations of storage volume and drawdown time. The WQCV for the watershed described in Table 5 was calculated to be 924 m$^3$ using Eq. 42. (The location of the watershed is Fort Collins, Colorado; which has adopted the UDFCD design criteria for sizing EDBs). We also evaluated four other storage volumes that represent 0.5*WQCV, 0.75*WQCV, 1.5*WQCV and 2*WQCV. For each storage volume, three different design drawdown times were evaluated: 24 hours, 40 hours and 72 hours. All EDBs were simulated with a maximum storage depth of 0.91 m (3 feet) and vertical side walls. Table 11 shows the design volume, surface area and drawdown time for each of the 15 scenarios evaluated.

Table 11: Characteristics of 15 different EDB designs evaluated

| Scenario | Volume (m³) | Surface Area (m²) | Drawdown Time (hrs) | Construction Cost ($) |
|----------|-------------|-------------------|---------------------|-----------------------|
| 1 | 462 | 505 | 24 | $53,777 |
| 2 | 462 | 505 | 40 | $53,777 |
| 3 | 462 | 505 | 72 | $53,777 |
| 4 | 693 | 758 | 24 | $63,938 |
| 5 | 693 | 758 | 40 | $63,938 |
| 6 | 693 | 758 | 72 | $63,938 |
| 7 | 924 | 1010 | 24 | $74,099 |
| 8 | 924 | 1010 | 40 | $74,099 |
| 9 | 924 | 1010 | 72 | $74,099 |
| 10 | 1386 | 1515 | 24 | $94,421 |
| 11 | 1386 | 1515 | 40 | $94,421 |
| 12 | 1386 | 1515 | 72 | $94,421 |
| 13 | 1848 | 2020 | 24 | $114,742 |
| 14 | 1848 | 2020 | 40 | $114,742 |
| 15 | 1848 | 2020 | 72 | $114,742 |

Practical SCM implementation plans must consider both the cost and performance of SCMs. For each EDB design, the cost of implementation is estimated using an EDB construction cost equation (Eq. 43) reported in UDFCD (2013)

$$\$23,897 + \$31.43(V)$$

Eq. 43

### 4.2.4 Uncertainty Analysis

The uncertainty analysis conducted in this study focused on the EDB pollutant removal performance as a result of uncertainties in the distribution of sizes and settling velocities of particles in urban stormwater runoff. These parameters of the modified Fair and Geyer model were found be most sensitive in the study included as Chapter 1 in this dissertation.

Uncertainty of EDB pollutant removal performance was evaluated by simulating each EDB scenario 1000 times using a Monte Carlo approach. Each simulation contained a different combination of the uncertain parameters shown in Table 12; with random parameter combinations being selected and applied using a Latin Hypercube (McKay et al. 1979) sampling method.

Table 12: EDB pollutant removal parameter values and their uncertainty

| Parameter | Units | Mean | Variance | Distribution |
|---|---|---|---|---|
| $psf_1$ | - | 0.42 | 4.9E-2 | Normal |
| $psf_2$ | - | 0.35 | 7.7E-3 | Normal |
| $psf_3$ | - | 0.14 | 6.3E-3 | Normal |
| $psf_4$ | - | 0.05 | 2.3E-3 | Normal |
| $psf_5$ | - | 0.04 | 3.5E-3 | Normal |
| $v_1$ | m/s | 6.29E-6 | 1.5E-11 | Normal |
| $v_2$ | m/s | 1.01E-4 | 4.0E-9 | Normal |
| $v_3$ | m/s | 5.10E-4 | 1.0E-7 | Normal |
| $v_4$ | m/s | 1.61E-3 | 1.0E-6 | Normal |
| $v_5$ | m/s | 2.52E-3 | 2.5E-6 | Normal |

### 4.2.5 SCM Performance Measures for TMDL Compliance

We evaluate SCM performance using three different measures; percent load reduction, reliability of achieving a TMDL and vulnerability of TMDL compliance. The percent load reduction metric (Eq. 44) is

a commonly reported metric in SCM monitoring studies. From a decision maker's perspective, a SCM that produces a greater load reduction would be preferred over a different SCM that produces a lower load reduction.

$$LR = \frac{L_W - L_D}{L_W} * 100$$

<div align="right">Eq. 44</div>

Where $L_R$ = pollutant load reduction (%), $L_W$ = pollutant load discharged from the watershed (kg) and $L_D$ = pollutant load discharged from the SCM (kg).

Decision makers may also be interested how often a TMDL is achieved and/or the magnitude by which the TMDL is exceeded (when it is exceeded). Hashimoto et al (1982) refers to these as "reliability" and "vulnerability", respectively, as measures of evaluating water resource system performance. The reliability measure (Eq. 45) can be useful in cases where the TMDL allows for a certain probability of exceedances, but does not specify a limit on the magnitude of the exceedance. The vulnerability measure (Eq. 46) can be useful when both the frequency and magnitude of TMDL exceedances are important.

$$Reliability = P(L_D < TMDL) = \frac{m}{N}$$

<div align="right">Eq. 45</div>

Where $m$ = number of days that $L_D$ is less than the TMDL and $N$ = total number of days in the simulation.

$$Vulnerability = \sum_{n=1}^{N} \left[ \max(0, L_D - TMDL)_n * \frac{1}{N} \right]$$

<div align="right">Eq. 46</div>

Where the term $(L_D - TMDL)$ represents the severity of the exceedance and $\frac{1}{N}$ represents the probability of that exceedance occurring. The *max* term ensures that only those days in which $L_D$ is greater than TMDL are included in the vulnerability calculation.

## 4.3    Results and Discussion

### 4.3.1    Variability and Uncertainty of Daily Discharged Pollutant Loads from EDBs

The modeling approach used in this study produced 1000 timeseries of TSS loads discharged from each

EDB design.  For each timeseries, the TSS load discharged over a 24-hour period (starting and ending at

0000 hrs) was calculated and retained for analysis.  The simulation period included a total of 6,940 days

with runoff being generated from the watershed (and entering the EDBs) during 697 of those days.  The

number of days that EDBs discharged TSS loads are presented in Table 13.

Table 13: Number of days that runoff discharged from each EDB design scenario

| Scenario | Drawdown Time (hrs) | Number of Days Discharging | Scenario | Drawdown Time (hrs) | Number of Days Discharging |
|----------|---------------------|----------------------------|----------|---------------------|----------------------------|
| 1 | 24 | 873 | 9 | 72 | 1171 |
| 2 | 40 | 1056 | 10 | 24 | 819 |
| 3 | 72 | 1449 | 11 | 40 | 887 |
| 4 | 24 | 843 | 12 | 72 | 1151 |
| 5 | 40 | 986 | 13 | 24 | 795 |
| 6 | 72 | 1314 | 14 | 40 | 859 |
| 7 | 24 | 811 | 15 | 72 | 992 |
| 8 | 40 | 908 | | | |

For each day of the simulation, the 1000 daily TSS loads that were generated from the Monte Carlo

simulations were ranked lowest to highest to facilitate calculation of different percentiles of model

outputs.  Figure 18 presents the cumulative probability (Y-axis) of the EDB design for Scenario 1

discharging a TSS load equal to or less than a particular value (X-axis).  Three different percentiles are

presented, 5th, 50th (i.e. median) and 95th.  Under a typical deterministic modeling approach, only the

median model output will be generated and reported (assuming median values of model input parameters

are used in the simulation).  With this information alone, the decision maker would understand that the

EDB would have approximately 92% probability of discharging a daily TSS load below 0.1 kg/ha, for

example.  The variability of discharged TSS loads along the x-axis is the result of variability in the runoff

rate, runoff duration and TSS concentration entering the EDB that occurs over the 19 year continuous

simulation. However, by definition, median model outputs have a 50% probability of being exceeded (aka a 50% "risk level"), so the actual cumulative probability of discharging a daily TSS load less than 0.1 kg/ha is some amount less than 92%. A more conservative estimate of the cumulative probability can be provided using the 95th percentile model output values, which correspond to a risk level of only 5%. Using this information, the decision maker could be provided with the following two scenarios to consider: 1) "Using the EDB design in Scenario 1, we are 50% confident that the EDB will discharge a daily TSS load less than 0.1 kg/ha, 92% of the time" and 2) "Using the EDB design in Scenario 1, we are 95% confident that the EDB will discharge a daily TSS load less than 0.1 kg/ha, 91% of the time". The difference between the two statements may be very important to a decision maker depending on the severity of consequences of violating the TMDL. Throughout the rest of this section, we report both the median (50% risk level) and 95th percentile (5% risk level) values for the different BMP performance metrics to compare how decisions might be affected by estimates of model output uncertainty.
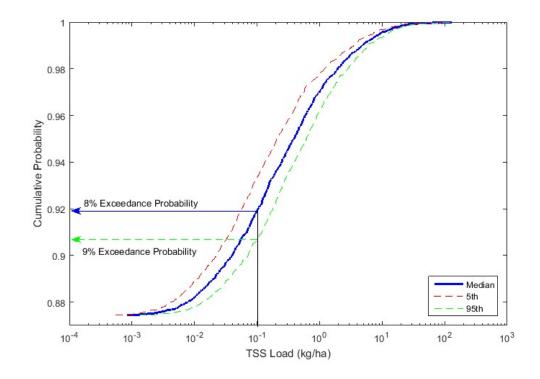


Figure 18: Median, 5th and 95th percentile empirical cumulative distribution function curves of daily TSS loads (kg/ha) discharged from EDB design for Scenario 1.

72

### 4.3.2    Performance Metrics and Uncertainty

In this section, we present and discuss the percent load reduction, reliability and vulnerability values computed for each EDB design.

### 4.3.2.1    Percent Load Reduction

Figure 19 shows the percent load reduction computed for each EDB design.  The total TSS load discharged from the watershed (and entering the EDBs) was 2,681 kg/ha.  The different EDB designs produced percent load reductions ranging from 39-72% for median outputs and 23-44% for 95th percentile outputs.  The results show that TSS load reduction generally increased with larger EDB storage volumes.  This is because greater EDB storage volume allows for more runoff to be captured and treated without overflowing/bypassing the EDB with little to no treatment.  Among EDBs with the same storage volume, the effect of increasing the drawdown time had variable effects.  For the smallest EDB storage volume (represented by Scenarios 1-3), increasing the drawdown time decreased the load reduction.  In this case, higher drawdown times likely increased the volume of runoff that overflowed the EDB because larger volumes of runoff from previous days remained stored in the EDB when additional runoff occurred 24-72 hours later.  For Scenarios 4-6 (storage volume = 0.75*WQCV), the effect of drawdown time had very little effect on the load reduction.  For EDBs with storage volumes greater than or equal to the WQCV (Scenarios 7-15), increased drawdown times resulted in increased load reduction among EDBs with the same storage volume.  Such findings are expected since greater particle residence time within the EDB will allow for more particles to settle out using the MFG model.  The EDB design with the greatest load reduction (Scenario 15) had the largest storage volume (1,848 m$^3$; 2*WQCV) and the longest drawdown time of 72 hours.  Scenario 12 had the second greatest load reduction using an EDB with storage volume of 1,386 m$^3$ (1.5*WQCV) and also a 72 hour drawdown time.  These trends hold true for both the median and 95th percentile model outputs.
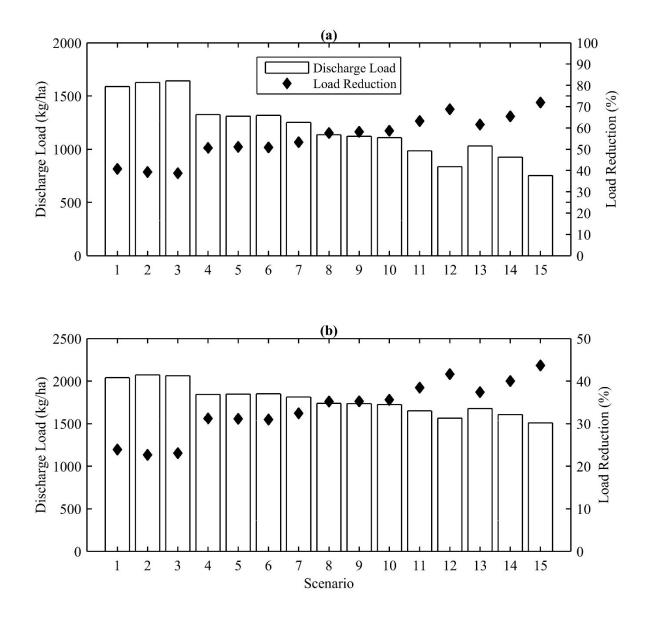
Figure 19: Discharged TSS load and percent load reduction for (a) median model outputs and (b) 95th percentile model outputs for 15 different EDB designs over the 19-year period of simulation. The TSS load discharged from the watershed (and entering the EDBs) was 2,681 kg/ha.

### 4.3.2.2 Reliability

Figure 20 shows the reliability values computed for each EDB design, under situations where the TMDL is 0.01 kg/ha and 1 kg/ha. A TMDL of 0.01 kg/ha represents a more stringent requirement (i.e. lower allowable discharge/higher load reduction) compared to a TMDL of 1.0 kg/ha. Reliability of achieving the 0.01 kg/ha TMDL ranged from 82-90% for median outputs and 80-89% for 95th percentile outputs for the different EDB designs. The EDB design for Scenario 13 produced the highest reliability using both median outputs and 95th percentile outputs; while the EDB design for Scenario 3 produces the lowest reliability at both risk levels. For each EDB design, the difference in reliability values between the median and 95th percentile outputs is relatively small compared to the difference in reliability values across EDB designs. Interestingly, the EDBs with 24-hour drawdown times performed better than those with longer drawdown times (and same storage volume). The greater reliability for these EDBs appears to be the result of a fewer number of days during which the EDBs discharge. Table 13 shows that the EDBs with 24-hour drawdown time discharge a considerably smaller number of days than the EDBs with longer drawdown times. Therefore, if a storm event discharges a high pollutant load from the watershed, and the EDB cannot reduce the pollutant load below the TMDL, the EDBs with 24-hour drawdown times will only violate the TMDL for one day, while the EDBs with longer drawdown times may violate the TMDL for up to 3 days.

The results for a TMDL of 1.0 kg/ha are markedly different from those for a TMDL of 0.01 kg/ha. Reliability of achieving the 1 kg/ha TMDL was higher than for the 0.01 kg/ha TMDL with median outputs ranging from 97-97.9% and 95th percentile outputs ranging from 95.9-96.7%. The highest reliability from median outputs was produced by the EDB design for Scenario 12, while the highest reliability based on 95th percentile outputs was produced by the EDB design for Scenario 13. The lowest reliabilities were produced by Scenarios 1 and 9 for median and 95th percentile outputs, respectively. Among all EDB designs, the median reliability values vary by less than 1% as do the 95th percentile

reliability values. However, for the same EDB, the difference between the median and 95<sup>th</sup> percentile

reliability values range from 1-2%.

It is worth noting that relatively small differences in reliability values can indicate profound and practical

impacts. For example, a 1% decrease in reliability means approximately 3-4 additional days per year that

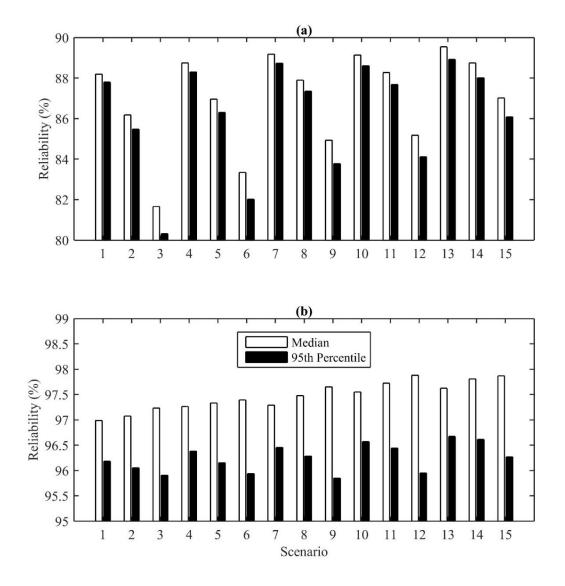the TMDL is exceeded for the system simulated in this study.



Figure 20: Reliability of 15 different EDB designs for achieving a TMDL of (a) 0.01 kg/ha and (b) 1.0 kg/ha

### 4.3.2.3 Vulnerability

Recall that the vulnerability measure considers both the probability of exceeding a TMDL and the magnitude of each exceedance. Unlike the previous performance measured, lower vulnerability values are preferred as they indicate a fewer number of exceedances, lower magnitudes of exceeding the TMDL, or both compared to larger vulnerability values.

Figure 21 shows the vulnerability values computed for each EDB design, under situations where the TMDL is 0.01 kg/ha and 1 kg/ha. Vulnerability associated with exceeding a TMDL of 0.01 kg/ha ranges from 0.11-0.24 for median outputs and 0.22-0.3 for 95th percentile outputs. EDB Scenario 15 produces the lowest vulnerability for both median and 95th percentile model outputs. EDB Scenario 3 produces the highest vulnerability for median outputs and EDB Scenario 2 produces the highest vulnerability based for 95th percentile outputs. There is little difference in vulnerability values among EDBs with the same storage volume, which suggests that drawdown time has little effect on this metric; although the effect is more noticeable with median outputs for EDBs with larger (greater than the WQCV) storage volumes. Generally, the difference in vulnerability metrics between the median and 95th percentile outputs for each EDB are greater than the differences of vulnerability among the different EDB designs.

For a TMDL of 1 ka/ha, vulnerability values range from 0.07-0.18 for median outputs and 0.15-0.23 for 95th percentile outputs. EDB Scenario 15 produces the lowest vulnerability for both median and 95th percentile model outputs and EDB Scenario 3 produces the highest vulnerability for median outputs; same as for the TMDL of 0.01 kg/ha. However, Scenario 1 produces the highest vulnerability for 95th percentile outputs compared to Scenario 2 for the median outputs. The effects of drawdown time on vulnerability are slightly greater under this TMDL compared to the 0.01 kg/ha and the differences in vulnerability among the EDB designs is similar to the difference in vulnerability between the median and 95th percentile outputs for each EDB.

Lastly, a visual comparison of the vulnerability figures and the discharged load bar graphs in Figure XX show similar trends. This suggests that the overall pollutant load reduction of each BMP design has a greater influence on the vulnerability metric compared to reliability.
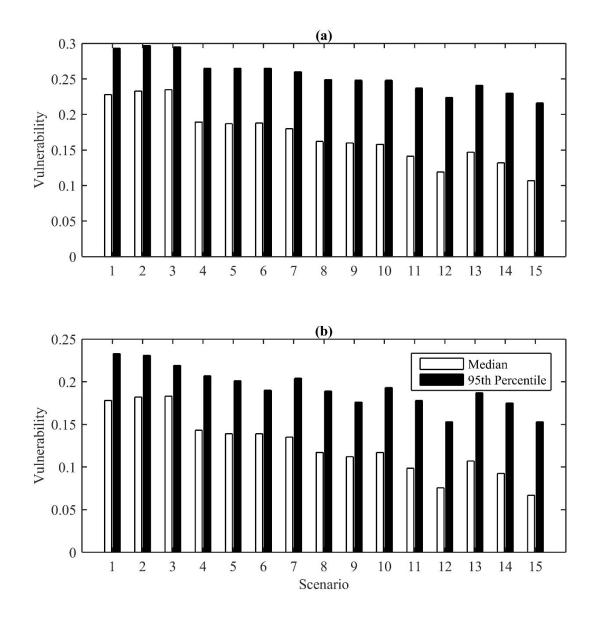


Figure 21: Vulnerability of 15 different EDB designs for achieving a TMDL of (a) 0.01 kg/ha and (b) 1.0 kg/ha

### 4.3.3    Cost Effective Decision Making

In this section, we present EDB performance metrics, their uncertainty, and cost estimates to represent the information that a decision maker might use while deciding which EDB design to implement to achieve the required pollutant load reduction. Using theoretical situations where the decision maker faces a constraint of achieving a certain EDB performance metric or a limited budget, we compare the decision that would be made when the decision maker is willing to accept a 50% risk level versus a 5% risk level of not achieving the desired metric.

Figure 22, Figure 23 and Figure 24 show the relationship between estimated construction costs for each EDB design and the percent load reduction, reliability and vulnerability, respectively. On Figure 22 and Figure 23 the right-most points represent "non-dominated" (i.e. "optimal") solutions that fall along a pseudo-Pareto optimal curve. For Figure 24, the optimal solutions are the left-most points because lower vulnerability values are preferred. Using these figures, one can determine the most cost-effective solution that achieves either a required performance metric value or a budget limitation. For example, consider the goal is to achieve a minimum of 35% load reduction. If the decision maker is willing to accept a 50% risk that the reported percent load reduction is exceeded, they would use Figure 22(a) to determine that all EDB designs meet that goal. Rationally, the decision maker should select the EDB design for Scenario 1 as it achieves the highest percent load reduction among the least costly solutions. However, if the decision maker is risk-averse and prefers to only accept a 5% risk that the goal is exceeded, they would find Scenarios 8 and 9 are the least costly solutions that provide at least 35% percent load reduction using Figure 22(b). In this case, the consequence of accepting less risk is an increase in costs. Alternatively, a decision maker may be constrained by budget and has the goal of selecting a design that maximizes (for percent load reduction and reliability) or minimizes (for vulnerability) the performance metric. Consider a situation with a TMDL of 1 kg/ha, a $60,000 maximum budget and a goal of maximizing reliability. Median outputs (Figure 23c) show that Scenario 3 is the most cost-effective EDB design, whereas 95th percentile outputs (Figure 23d) show that Scenario 1 is most cost-effective. These two examples prove

that decisions on SCM implementation can be affected by the uncertainty of SCM performance and the

decision maker's risk level, however that is not the case in all situations.
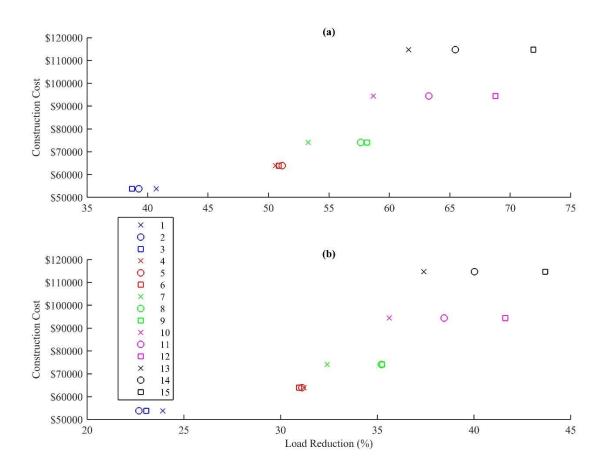


Figure 22: Construction cost versus TSS load reduction for 15 different EDB designs based on (a) median model outputs and (b) 95th percentile model outputs
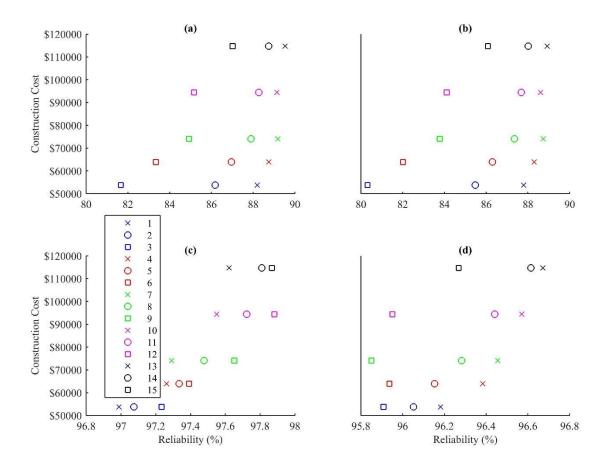
Figure 23: Cost-reliability curves for 15 different EDB designs and (a) TMDL = 0.01 kg/ha and 50% risk, (b) TMDL = 0.01kg/ha and 5% risk, (c) TMDL = 1.0 kg/ha and 50% risk and (d) TMDL = 1.0 kg/ha and 5% risk.
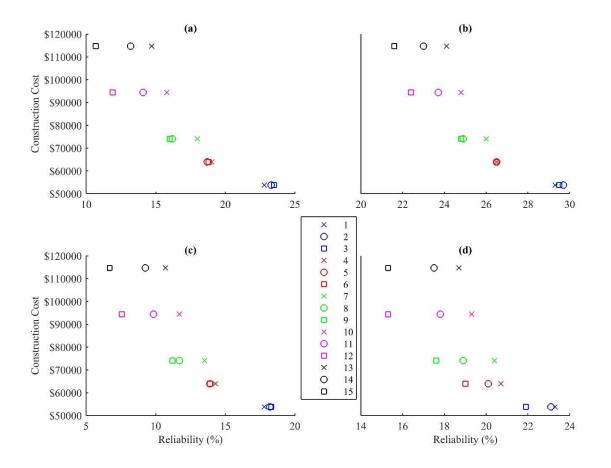
Figure 24: Cost-vulnerability curves for 15 different EDB designs and (a) TMDL = 0.01 kg/ha and 50% risk, (b) TMDL = 0.01kg/ha and 5% risk, (c) TMDL = 1.0 kg/ha and 50% risk and (d) TMDL = 1.0 kg/ha and 5% risk.

Following the same procedure described in the previous examples, we determined the optimal solution for many different values of required performance metrics or maximum cost and the results are shown in Table 14-Table 16. For percent load reduction, the optimal EDB designs are different for the 50% risk level and 5% risk level for all potential percent load reduction constraints. However, decisions made under a maximum cost constraint would be the same regardless of the accepted risk level. The results using reliability as a metric show that differences in decisions are affected by the TMDL value. Under the more severe TMDL (0.01 kg/ha) risk levels do not result in different decisions except for one case, however the optimal EDB designs are affected by risk for the less severe TMDL of 1 kg/ha. Optimal

decisions using vulnerability are affected by risk level when the constraint is a maximum vulnerability value, but not when budget is the constraint (except in one case).

Table 14: Most cost-effective EDB designs (by scenario number) per required percent load reduction and maximum budget constraints, for different decision maker's risk levels.

| Percent Load Reduction | Risk Level | | Maximum Cost | Risk Level | |
|---|---|---|---|---|---|
| | 50% | 5% | | 50% | 5% |
| 35% | 1 | 8,9 | $60,000 | 1 | 1 |
| 40% | 1 | 12 | $70,000 | 4,5,6 | 4,5,6 |
| 45% | 4,5,6 | n/a | $90,000 | 9 | 8,9 |
| 50% | 4,5,6 | n/a | $100,000 | 12 | 12 |
| 55% | 9 | n/a | $120,000 | 15 | 15 |
| 60% | 12 | n/a | | | |
| 65% | 12 | n/a | | | |
| 70% | 15 | n/a | | | |
| 75% | 15 | n/a | | | |
| n/a – indicates that load reduction cannot be achieved by any scenario | | | | | |

Table 15: Most cost-effective EDB designs (by scenario number) per required reliability and maximum budget constraints, for different decision maker's risk levels.

| Reliability | Risk Level | | Maximum Cost | Risk Level | |
|---|---|---|---|---|---|
| | 50% | 5% | | 50% | 5% |
| TMDL = 0.01 kg/ha | | | | | |
| 80% | 1 | 1 | $60,000 | 1 | 1 |
| 81% | 1 | 1 | $70,000 | 4 | 4 |
| 82% | 1 | 1 | $90,000 | 7 | 7 |
| 83% | 1 | 1 | $100,000 | 7 | 7 |
| 84% | 1 | 1 | $120,000 | 13 | 13 |
| 85% | 1 | 1 | | | |
| 86% | 1 | 1 | | | |
| 87% | 1 | 1 | | | |
| 88% | 1 | 1 | | | |
| 89% | 7 | 4 | | | |
| TMDL = 1 kg/ha | | | | | |
| 96% | 3 | 1 | $60,000 | 3 | 1 |
| 96.2% | 3 | 4 | $70,000 | 6 | 4 |
| 96.4% | 3 | 7 | $90,000 | 7 | 7 |
| 96.6% | 3 | 13 | $100,000 | 12 | 10 |
| 96.8% | 3 | n/a | $120,000 | 12 | 15 |
| 97% | 3 | n/a | | | |
| 97.2% | 3 | n/a | | | |
| 97.4% | 8 | n/a | | | |
| 97.6% | 9 | n/a | | | |
| 97.8% | 12 | n/a | | | |
| n/a – indicates that reliability cannot be achieved by any scenario | | | | | |

Table 16: Most cost-effective EDB designs (by scenario number) per required vulnerability and maximum budget constraints, for different decision maker's risk levels.

| Vulnerability | Risk Level | | Maximum | Risk Level | |
|---|---|---|---|---|---|
| | 50% | 5% | Cost | 50% | 5% |
| TMDL = 0.01 kg/ha | | | | | |
| 0.11 | 15 | n/a | $60,000 | 1 | 1 |
| 0.13 | 12 | n/a | $70,000 | 5 | 4,5,6 |
| 0.15 | 12 | n/a | $90,000 | 9 | 9 |
| 0.17 | 9 | n/a | $100,000 | 12 | 12 |
| 0.19 | 5 | n/a | $120,000 | 15 | 15 |
| 0.21 | 5 | n/a | | | |
| 0.23 | 1 | 12 | | | |
| 0.25 | 1 | 9 | | | |
| TMDL = 1 kg/ha | | | | | |
| 0.07 | 15 | n/a | $60,000 | 1 | 3 |
| 0.09 | 12 | n/a | $70,000 | 5,6 | 6 |
| 0.11 | 12 | n/a | $90,000 | 9 | 9 |
| 0.13 | 9 | n/a | $100,000 | 12 | 12 |
| 0.15 | 5,6 | n/a | $120,000 | 15 | 15 |
| 0.17 | 5,6 | 12 | | | |
| 0.19 | 1 | 6 | | | |
| n/a – indicates that reliability cannot be achieved by any scenario | | | | | |

## 4.4 Conclusions and Recommendations

This study aimed to evaluate whether decision makers might make different decisions of SCM implementation plans for TMDL compliance if they understood the uncertainty of SCM performance. A simplified system of a single watershed draining to a single EDB was simulated for a 19-year period of record, with uncertainty of the EDBs' pollutant removal performance evaluated using 1000 Monte Carlo simulations of the system. A total of 15 different EDB designs were simulated and pollutant load reduction performance was computed using three different metrics; percent load reduction, reliability of achieving a TMDL and vulnerability when the TMDL is exceeded. The effect of uncertainty and the decisions maker's risk level was evaluated by comparing median model outputs, 95th percentile model outputs and estimated construction costs of the different EDB designs.

Our results showed that uncertainty and the decision maker's risk level affected EDB design decisions in some circumstances, but not others. Under situations where the decision maker must achieve a particular

performance metric value, the most cost-effective EDB design for 50% risk level and 5% risk level was different in nearly all cases where the performance metric was percent load reduction or vulnerability. Using the reliability metric, the effect of uncertainty and decision maker's risk level on EDB design decisions appears to be sensitive to the actual TMDL value; as differences in decisions were shown for a less-stringent TMDL of 1 kg/ha, but not for a more-stringent TMDL of 0.1 kg/ha. In several of the scenarios evaluated, we also showed that a certain performance metric could be achieved at the 50% risk level, but at the 5% risk level none of the EDB designs could achieve the performance metric.

We also considered how cost constraints might affect decisions, as available budgets can be the limiting factor in selecting pollutant control practices, irrespective of if they actually meet TMDL targets. In most situations, the most cost-effective EDB design below certain budget thresholds was the same regardless of the decision maker's risk level. The usefulness of uncertainty under these conditions is not in selecting the most cost-effective EDB design, but in reporting the EDB's expected performance at different risk levels. This can help stakeholders, regulators and the public understand the probability of achieving the performance metrics under various budget scenarios and perhaps persuade additional funding for improved performance.

Our results justify the suggestions from NRC (National Research Council 2001) and Shirmohammadi et al. (2006) that uncertainty analysis be included not only in the development of a MOS for TMDLs, but also in the evaluation of pollutant control practices to achieve the TMDL. We demonstrate several situations where a decision maker informed with the uncertainty of the expected EDB performance metrics would make different decisions based on their willing to assume different levels of risk and/or faced with budget constraints, which could ultimately lead to less (wasteful) expenditure of funds on less cost-effective practices and/or improved probability of achieving a TMDL in a timely manner.

It is important to note the results of this study are based on a simplified watershed-BMP system and various assumptions necessary to complete a "synthetic" TMDL study in a timely manner, and we

recommend future studies follow the approach outlined in this study to determine if decisions are affected in a "real life" SCM implementation planning study.  Such a study could improve our understanding of SCM performance metrics and their uncertainty when measured over a larger watershed-scale, using multiple (perhaps hundreds) of individual SCMs.  We also acknowledge there are many other factors that can affect SCM implementation plans beyond performance and costs and the use of multi-criteria decision analysis tools can be used to synthesize all important factors for the decision maker.

# Chapter 5: Summary, Conclusions and Recommendations

This research resulted in the development of new tools and knowledge that can improve modeling of SCMs for TMDL compliance.

First, a new SCM pollutant removal model (modified Fair and Geyer) was evaluated using data from International BMP Database. Results of global sensitivity analysis showed that the removal of TSS in EDBs was most sensitive to the particle size distribution and particle density. In addition, the SCM model parameter uncertainty was estimated using both formal and informal (i.e. Generalized Likelihood Uncertainty Estimation) Bayesian methods. The results showed that the formal Bayesian method generated considerably smaller posterior probabilities compared to the informal approaches, which resulted in only 50% of the testing dataset falling within prediction intervals using the formal method compared to 94-97% for the informal approaches. Of the informal methods evaluated, the "efficiency criteria" likelihood measure performed most efficiently. These results demonstrate that the modified Fair and Geyer model can reasonably simulate the TSS removal performance of EDBs and the uncertainty of that performance when parameters are calibrated using the informal GLUE Bayesian method.

Second, three different SCM models were evaluated in a modeling framework that linked a dynamic watershed model (SWMM) to the SCM models. The SCM models were calibrated to data from the International BMP Database assuming steady-state hydraulic conditions for each event, then applied to the dynamic modeling framework with variable hydraulic conditions dictated by runoff generated from the dynamic watershed model. The linear regression model generated the same pollutant removal results under both steady-state and dynamic conditions, which was expected since pollutant removal is only a function of the influent pollutant concentration. However both the MFG and k-C* models underestimated pollutant removal by 20-90% under dynamic modeling conditions. This is because pollutant removal is inversely related to the discharge rate of the SCM and modeling results showed periods of time during

runoff events when the SCM discharge rates were higher than the average discharge rate assumed in steady-state modeling. The magnitude of pollutant removal estimation was larger for short (e.g. 1 hour) duration storms and smaller for long (e.g. 24 hour) duration storms.

Uncertainty of the SCM models was also evaluated using MC and FOVE approaches. In general, the FOVE method generated prediction intervals that were smaller than the MC method, with 95th percentile outputs generally being 5-20% lower using FOVE compared to MC. However, the smaller prediction intervals generated by the FOVE method partially compensated for the lower pollutant removal generated by the MFG and k-C* models under dynamic modeling conditions, such that the 95th percentile outputs generated using MC and steady-state assumptions were very similar to those generated using FOVE and dynamic modeling.

Finally, this study demonstrates the importance of incorporating uncertainty analysis into SCM modeling projects, as the final decision regarding SCM design/implementation may be dependent on the decision maker's risk level. The cost and performance of 15 different EDB designs was examined with respect to their ability to achieve TMDL compliance metrics such as percent removal, reliability and vulnerability. First, the effect of EDB performance uncertainty was demonstrated by comparing the median (50th percentile) metric value to the 95th percentile metric value. In all scenarios there is a noticeable difference between the two values, however the differences were generally greater for the percent load reduction and vulnerability metrics and less for reliability. Differences between median and 95th percentile reliability and vulnerability metrics were also affected by the TMDL magnitude, with greater uncertainty associated with higher (less stringent) TMDLs. These findings were translated into a theoretical decision making where TMDL compliance metrics and their uncertainty were evaluated against the costs of implementing different EDB designs. The results showed that, in some cases, the most cost-effective decision was different using the median metric values compared to the 95th percentile metric values. The use of different metric percentiles represents the decision maker's risk level. Not only did the uncertainty analysis reveal that optimal decisions can be affected by the decision maker's risk level, it also revealed

that some EDB designs could not achieve the required TMDL compliance threshold at a 5% risk level, when modeling showed they could achieve the same compliance threshold at the 50% risk level.

This research may also serve as a foundation for future research in the field of SCM modeling and uncertainty analysis. These studies were conducted using a single type of SCM (the extended detention basin), for a single pollutant (total suspended solids) and only considering the uncertainty of the pollutant removal mechanisms within SCMs (uncertainty of the hydraulics/hydrology was not considered). Future studies may apply similar experiments to assess the uncertainty performance of other types of SCMs (e.g. bioretention, permeable pavement, etc.), other pollutants (e.g. nitrogen, phosphorus, etc.) and include the uncertainty of the hydraulics/hydrology of the SCMs. Such studies should continue to use data from the International BMP Database as it provides large datasets from a variety of locations.

Lastly, a research project that takes the tools and knowledge presented here and applies them to a real-world TMDL modeling/decision-making situation perhaps could make the greatest impact on the state of practice if it could be demonstrated that 1) using the dynamic SCM models with FOVE uncertainty analysis saved considerable time/resources compared to the traditional SCM modeling/uncertainty analysis approaches and 2) that decision makers changed their SCM design/implementation plans based on uncertainty of SCM performance. Such a research project may also incorporate uncertainty analysis into the watershed model to evaluate how SCM performance and uncertainty compares to the uncertainty of pollutant concentrations and runoff rates generated from watershed models such as SWMM. As shown in Chapter 2 where 25-58% of the uncertainty of SCM pollutant removal was due to the magnitude of the influent pollutant concentration, it may be that the pollutant load discharged from SCMs is more sensitive to watershed model parameters than to SCM model parameters. This would then inform the modeler to ignore uncertainty of the SCM model and focus future research on understanding uncertainty of urban watershed models.

# References

Ajami, N. K., Duan, Q., and Sorooshian, S. (2007). "An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction." *Water Resources Research*, 43(1), n/a-n/a.

American Society of Civil Engineers, and Water Environment Federation. (1998). *Urban Runoff Quality Management - WEF Manual of Practice No. 23, ASCE Manual and Report on Engineering Practice No. 87*.

Arabi, M., Govindaraju, R. S., and Hantush, M. M. (2007). "A probabilistic approach for analysis of uncertainty in the evaluation of watershed management practices." *Journal of Hydrology*, JOUR, 333(2–4), 459–471.

Avellaneda, P., Ballestero, T. P., Roseen, R. M., and Houle, J. J. (2009). "On Parameter Estimation of Urban Storm-Water Runoff Model." *Journal of Environmental Engineering*, American Society of Civil Engineers, 135(8), 595–608.

Avellaneda, P., Ballestero, T., Roseen, R., and Houle, J. (2010). "Modeling Urban Storm-Water Quality Treatment: Model Development and Application to a Surface Sand Filter." *Journal of Environmental Engineering*, American Society of Civil Engineers, 136(1), 68–77.

Barrett, M. E. (2005). "Performance Comparison of Structural Stormwater Best Management Practices." *Water Environment Research*, Water Environment Federation, 77(1), 9.

Barrett, M. E. (2008). "Comparison of BMP Performance Using the International BMP Database." *Journal of Irrigation and Drainage Engineering*, ASCE, 134(5), 556–561.

Bates, B. C., and Campbell, E. P. (2001). "A Markov Chain Monte Carlo Scheme for parameter estimation and inference in conceptual rainfall-runoff modeling." *Water Resources Research*, 37(4), 937–947.

Bates, B. C., and Townley, L. R. (1988). "Nonlinear, discrete flood event models, 3. Analysis of prediction uncertainty." *Journal of Hydrology*, 99(1–2), 91–101.

Beck, M. B. (1987). "Water quality modeling: A review of the analysis of uncertainty." *Water Resources Research*, American Geophysical Union, 23(8), 1393.

Beven, K. (2006). "A manifesto for the equifinality thesis." *Journal of Hydrology*, 320(1–2), 18–36.

Beven, K., and Binley, A. (1992). "The future of distributed models: Model calibration and uncertainty prediction." *Hydrological Processes*, 6(3), 279–298.

Beven, K., and Freer, J. (2001). "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology." *Journal of Hydrology*, 249(1–4), 11–29.

Box, G. E. P., and Cox, D. R. (1964). "An Analysis of Transformations." *Journal of the Royal Statistical Society*, 211–252.

Box, G. E. P., and Tiao, G. C. (1992). *Bayesian inference in statistical analysis*. Wiley.

Brown, R. A., and Hunt, W. F. (2011). "Impacts of Media Depth on Effluent Water Quality and Hydrologic Performance of Undersized Bioretention Cells." *Journal of Irrigation and Drainage Engineering*, American Society of Civil Engineers, 137(3), 132–143.

Burges, S. J., and Lettenmaier, D. P. (1975). "PROBABILISTIC METHODS IN STREAM QUALITY MANAGEMENT." *Journal of the American Water Resources Association*, 11(1), 115–130.

California Department of Transportation. (2010). *Stormwater Quality Handbook - Project Planning and Design Guide*.

Camp, T. R. (1946). "Sedimentation and the Design of Settling Tanks." *Transactions of the American Society of Civil Engineers - Volume III*.

Carleton, J. N., Grizzard, T. J., Godrej, A. N., Post, H. E., Lampe, L., and Kenel, P. P. (2000). "Performance of a Constructed Wetlands in Treating Urban Stormwater Runoff." *Water Environment Research*, JOUR, Water Environment Federation, 72(3), 295–304.

Chen, J., and Adams, B. J. (2006). "Analytical Urban Storm Water Quality Models Based on Pollutant Buildup and Washoff Processes." *Journal of Environmental Engineering*, 132(10), 1314.

City of Fort Collins Colorado. (2011). "Stormwater Criteria Manual." Fort Collins, CO.

Comings, K. J., Booth, D. B., and Horner, R. R. (2000). "Storm Water Pollutant Removal by Two Wet Ponds in Bellevue, Washington." *Journal of Environmental Engineering*, American Society of Civil Engineers, 126(4), 321–330.

Dilks, D. W., and Freedman, P. L. (2004). "Improved Consideration of the Margin of Safety in Total Maximum Daily Load Development." *Journal of Environmental Engineering*, American Society of Civil Engineers, 130(6), 690–694.

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and Deletic, A. (2012). "Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling." *Water research*, 46(8), 2545–58.

Engel, B., Storm, D., White, M., Arnold, J., and Arabi, M. (2007). "A Hydrologic/Water Quality Model Applicati1." *Journal of the American Water Resources Association*, Blackwell Publishing Ltd, 43(5), 1223–1236.

Fair, G. M., and Geyer, G. C. (1954). *Water Supply and Wastewater Disposal*. John Wiley and Sons, Inc, Hoboken, NJ.

Franceschini, S., and Tsai, C. W. (2008). "Incorporating Reliability into the Definition of the Margin of Safety in Total Maximum Daily Load Calculations." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 134(1), 34–44.

Freer, J., Beven, K., and Ambroise, B. (1996). "Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach." *Water Resources Research*, 32(7), 2161–2173.

Freni, G., and Mannina, G. (2012). "Uncertainty estimation of a complex water quality model: The influence of Box–Cox transformation on Bayesian approaches and comparison with a non-Bayesian method." *Physics and Chemistry of the Earth, Parts A/B/C*, 42–44, 31–41.

Freni, G., Mannina, G., and Viviani, G. (2008). "Uncertainty in urban stormwater quality modelling: the

effect of acceptability threshold in the GLUE methodology." *Water research*, 42(8–9), 2061–72.

Freni, G., Mannina, G., and Viviani, G. (2009). "Urban runoff modelling uncertainty: Comparison among Bayesian and pseudo-Bayesian methods." *Environmental Modelling & Software*, 24(9), 1100–1111.

Greb, S. R., and Bannerman, R. T. (1997). "Influence of Particle Size on Wet Pond Effectiveness." *Water Environment Research*, 69(6), 1134–1138.

Grizzard, T. J., Randall, C. W., Weand, B. L., and Ellis, K. L. (1986). "Effectiveness of Extended Detention Ponds." *Urban Runoff Quality - Impacts and Quality Enhancement Technology*, B. R. Urbonas and L. A. Roesner, eds., American Society of Civil Engineers, New York, NY.

Hashimoto, T., Stedinger, J. R., and Loucks, D. P. (1982). "Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation." *Water Resources Research*, 18(1), 14–20.

Hathaway, J. M., Hunt, W. F., and Jadlocki, S. (2009). "Indicator Bacteria Removal in Storm-Water Best Management Practices in Charlotte, North Carolina." *Journal of Environmental Engineering*, American Society of Civil Engineers.

Hunt, W. F., Jarrett, A. R., Smith, J. T., and Sharkey, L. J. (2006). "Evaluating Bioretention Hydrology and Nutrient Removal at Three Field Sites in North Carolina." *Journal of Irrigation and Drainage Engineering*, American Society of Civil Engineers, 132(6), 600–608.

Hutton, C. J., Kapelan, Z., Vamvakeridou-Lyroudia, L., and Savić, D. (2014). "Application of Formal and Informal Bayesian Methods for Water Distribution Hydraulic Model Calibration." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 140(11), 4014030.

Jia, Y., and Culver, T. B. (2008). "Uncertainty Analysis for Watershed Modeling Using Generalized Likelihood Uncertainty Estimation with Multiple Calibration Measures." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 134(2), 97–106.

Jin, X., Xu, C.-Y., Zhang, Q., and Singh, V. P. (2010). "Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model." *Journal of Hydrology*, 383(3–4), 147–155.

Joint Research Center of the European Commission. (2004). "SIMLAB V2.2 - Simulation Environment for Uncertainty and Sensitivity Analysis."

Jones, M. P., and Hunt, W. F. (2010). "Effect of Storm-Water Wetlands and Wet Ponds on Runoff Temperature in Trout Sensitive Waters." *Journal of Irrigation and Drainage Engineering*, American Society of Civil Engineers, 136(9), 656–661.

Kadlec, R. H., and Knight, R. L. (1996). *Treatment Wetlands*. CRS Press LLC, Boca Raton, Fla.

Karamalegos, A. M., Barrett, M. E., Lawler, D. F., and Malina, J. F. (2005). *Particle Size Distribution of Highway Runoff and Modification Through Stormwater Treatment*.

Kavetski, D., Kuczera, G., and Franks, S. W. (2006). "Bayesian analysis of input uncertainty in hydrological modeling: 2. Application." *Water Resources Research*, 42(3), n/a-n/a.

Kayhanian, M., Rasa, E., Vichare, A., and Leatherbarrow, J. E. (2008). "Utility of Suspended Solid Measurements for Storm-Water Runoff Treatment." *Journal of Environmental Engineering*, American Society of Civil Engineers, 134(9), 712–721.

Kim, J.-Y., and Sansalone, J. J. (2008). "Event-based size distributions of particulate matter transported during urban rainfall-runoff events." *Water Research*, JOUR, 42(10–11), 2756–2768.

Krishnappan, B., Marsalek, J., Watt, W., and Anderson, B. (1999). "Seasonal size distributions of suspended solids in a stormwater management pond." *Water Science and Technology*, 39(2), 127–134.

Krzysztofowicz, R. (2001). "The case for probabilistic forecasting in hydrology." *Journal of Hydrology*, 249(1–4), 2–9.

Kuczera, G. (1983). "Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty." *Water Resources Research*, 19(5), 1151–1162.

Kuczera, G., Kavetski, D., Franks, S., and Thyer, M. (2006). "Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters." *Journal of Hydrology*, 331(1), 161–177.

Langseth, D. E., and Brown, N. (2011). "Risk-Based Margins of Safety for Phosphorus TMDLs in Lakes." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 137(3), 276–283.

Li, L., Xia, J., Xu, C.-Y., and Singh, V. P. (2010). "Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models." *Journal of Hydrology*, 390(3–4), 210–221.

Los Angeles Watershed Protection Division. (2009). *Water Quality Compliance Master Plan for Urban Runoff*.

Mannina, G., and Viviani, G. (2010). "An urban drainage stormwater quality model: Model development and uncertainty quantification." *Journal of Hydrology*, 381(3–4), 248–265.

Martin, E. H. (1988). "Effectiveness of an Urban Runoff Detention Pond-Wetlands System." *Journal of Environmental Engineering*, American Society of Civil Engineers, 114(4), 810–827.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics*, 21(2), 239–245.

Melching, C. S., and Anmangandla, S. (1992). "Improved First-Order Uncertainty Method for Water-Quality Modeling." *Journal of Environmental Engineering*, 118(5), 791.

Melching, C. S., and Yoon, C. G. (1996). "Key Sources of Uncertainty in QUAL2E Model of Passaic River." *Journal of Water Resources Planning and Management*, 122(2), 105.

Morgan, M. G., and Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press.

National Research Council. (2001). *Assessing the TMDL Approach to Water Quality Management*. National Academy Press, Washington, DC.

Pappenberger, F., and Beven, K. J. (2006). "Ignorance is bliss: Or seven reasons not to use uncertainty analysis." *Water Resources Research*, American Geophysical Union, 42(5), W05302.

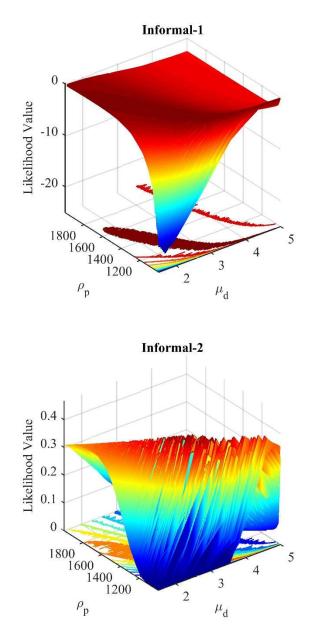Park, D., Kang, H., Jung, S. H., and Roesner, L. A. (2015). "Reliability analysis for evaluation of factors

affecting pollutant load reduction in urban stormwater BMP systems." *Environmental Modelling & Software*, 74, 130–139.
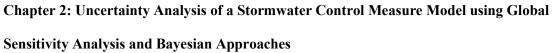
Park, D., Loftis, J. C., and Roesner, L. A. (2011). "Performance Modeling of Storm Water Best Management Practices with Uncertainty Analysis." *Journal of Hydrologic Engineering*, 16(4), 332.

Park, D., and Roesner, L. A. (2012). "Evaluation of pollutant loads from stormwater BMPs to receiving water using load frequency curves with uncertainty analysis." *Water research*, 46(20), 6881–90.

Passeport, E., Hunt, W. F., Line, D. E., Smith, R. A., and Brown, R. A. (2009). "Field Study of the Ability of Two Grassed Bioretention Cells to Reduce Storm-Water Runoff Pollution." *Journal of Irrigation and Drainage Engineering*, American Society of Civil Engineers, 135(4), 505–510.

Persson, J., Somes, N. L. G., and Wong, T. H. F. (1999). "Hydraulics Efficiency of Constructed Wetlands and Ponds." *Water Science and Technology*, 40(3), 291–300.

Potter, M. C., and Wiggert, D. C. (1997). *Mechanics of Fluids*. Prentice Hall, Inc, Upper Saddle River, NJ.

Reckhow, K. (1994). *Importance of scientific uncertainty in decision making*. *Environmental Management*, Springer New York, 161–166.

Reckhow, K. H. (2003). "On the Need for Uncertainty Assessment in TMDL Modeling and Implementation." *Journal of Water Resources Planning and Management*, 129(4), 245–246.

Roseen, R. M., Ballestero, T. P., Fowler, G. D., Guo, Q., and Houle, J. (2011). "Sediment Monitoring Bias by Automatic Sampler in Comparison with Large Volume Sampling for Parking Lot Runoff." *Journal of Irrigation and Drainage Engineering*, American Society of Civil Engineers, 137(4), 251–257.

Saltelli, A. (Andrea). (2008). *Global sensitivity analysis : the primer*. John Wiley.

Saltelli, A. (Andrea), Chan, K. (Karen), and Scott, E. M. (2000). *Sensitivity analysis*. Wiley.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley and Sons, Inc.

Schoups, G., and Vrugt, J. A. (2010). "A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors." *Water Resources Research*, 46(10), n/a-n/a.

Selbig, W. R., and Bannerman, R. T. (2011). *Characterizing the size distribution of particles in urban stormwater by use of fixed-point sample-collection methods*.

Shirmohammadi, A., Chaubey, I., Harmel, R. D., Bosch, D. D., Munoz-Carpena, R., Dharmasri, C., Sexton, A. M., Arabi, M., Wolfe, M. L., Frankenberger, J., Graff, C., and Sohrabi, T. M. (2006). "Uncertainty in TMDL Models." *Transactions of the ASABE*, 49(4), 1033–1049.

Smith, T., Sharma, A., Marshall, L., Mehrotra, R., and Sisson, S. (2010). "Development of a formal likelihood function for improved Bayesian inference of ephemeral catchments." *Water Resources Research*, 46(12), n/a-n/a.

Sobol', I. . (1967). "On the distribution of points in a cube and the approximate evaluation of integrals." *USSR Computational Mathematics and Mathematical Physics*, 7(4), 86–112.
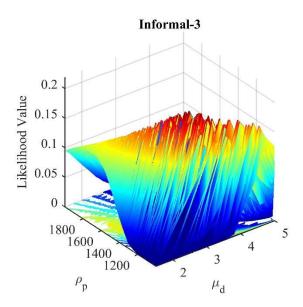
Sobol, I. M. (1990). "On Sensitivity Estimation for Nonlinear Mathematical Models." *Matematichesko Modelirovanie*, 2(1), 112–118.

Stanley, D. W. (1996). "Pollutant Removal by a Stormwater Dry Detention Pond." *Water Environment Research*, JOUR, Water Environment Federation, 68(6), 1076–1083.

Stedinger, J. R., Vogel, R. M., Lee, S. U., and Batchelder, R. (2008). "Appraisal of the generalized likelihood uncertainty estimation (GLUE) method." *Water Resources Research*, 44(12), n/a-n/a.

Strecker, E. W., Quigley, M. M., Urbonas, B. R., Jones, J. E., and Clary, J. K. (2001). "Determining Urban Storm Water BMP Effectiveness." *Journal of Water Resources Planning and Management*, ASCE, 127(3), 144–149.

Thiemann, M., Trosset, M., Gupta, H., and Sorooshian, S. (2001). "Bayesian recursive parameter estimation for hydrologic models." *Water Resources Research*, 37(10), 2521–2535.

Tung, Y.-K., Yen, B.-C., and Melching, C. S. (2006). *Hydrosystems Engineering Reliability Assessment and Risk Analysis*. McGraw Hill, New York, NY.

United States Environmental Protection Agency. (1983). *Results of the Nationwide Urban Runoff Program*.

United States Environmental Protection Agency. (1986). *Methodology for Analysis of Detention Basins for Control of Urban Runoff Quality*. Washington, DC.

United States Environmental Protection Agency. (1999). *Preliminary Data Summary of Urban Storm Water Best Management Practices*.

United States Environmental Protection Agency. (2001). *The National Costs of the Total Maximum Daily Load Program (Draft)*. Washington, DC.

United States Environmental Protection Agency. (2002). *Guidance for Quality Assurance Project Plans for Modeling. EPA QA/G-5 M. Report EPA/240/R-02/007*. Washington, DC.

United States Environmental Protection Agency. (2006). *Establishing TMDL "Daily" Loads in Light of the Decision by the U.S. Court of Appeals for the D.C. Circuit in Friends of the Earth, Inc. vs. EPA et al., No. 05-5015 (April 25-2006) and Implications for NPDES Permits*.

United States Environmental Protection Agency. (2007). *Total Maximum Daily Loads with Stormwater Sources: A Summary of 17 TMDLs*. Washington, DC.

United States Environmental Protection Agency. (2008). *TMDLs to Stormwater Permits Handbook*.

United States Environmental Protection Agency. (2012). "Glossary: Total Maximum Daily Loads (303d)." <http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/glossary.cfm>.

United States Environmental Protection Agency. (2015a). "National Summary of Impaired Waters and TMDL Information."
<https://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T>.

United States Environmental Protection Agency. (2015b). "Stormwater Management Model."

Urban Drainage and Flood Control District. (2013). *BMP-REALCOST User's Manual and Documentation*. Denver, Colorado.

Urban Drainage and Flood Control District. (2015). "Urban Storm Drainage Criteria Manual." Denver, Colorado.

Urbonas, B. R. (2007). "Stormwater Runoff Modeling - Is It As Accurate As We Think?" *Urban Runoff Modeling: Intelligent Modeling to Improve Stormwater Management*, Engineering Conferences International, Arcata, CA.

Urbonas, B. R., and Stahre, P. (1993). *Stormwater - Best Management Practices and Detention for Water Quality Drainage and CSO Management*. Prentice Hall, Inc, Englewood Cliffs, NJ.

Vezzaro, L., Eriksson, E., Ledin, A., and Mikkelsen, P. S. (2012). "Quantification of uncertainty in modelled partitioning and removal of heavy metals (Cu, Zn) in a stormwater retention pond and a biofilter." *Water research*, 46(20), 6891–903.

Viessman, W., and Lewis, G. L. (1996). *Introduction to Hydrology*. HarperCollins, New York, NY.

Vrugt, J. A. (2016). "Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation." *Environmental Modelling & Software*, 75, 273–316.

Vrugt, J. A., ter Braak, C. J. F., Gupta, H. V., and Robinson, B. A. (2008). "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?" *Stochastic Environmental Research and Risk Assessment*, 23(7), 1011–1026.

Walker, W. W. (2003). "Consideration of Variability and Uncertainty in Phosphorus Total Maximum Daily Loads for Lakes." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 129(4), 337–344.

Walpole, R. E., Myers, R. H., and Myers, S. L. (1998). *Probability and Statistics for Engineers and Scientists*. Prentice Hall, Inc, Upper Saddle River, NJ.

Water Environment Federation, and American Society of Civil Engineers. (1998). *Urban Runoff Quality Management*. (L. A. Roesner, B. R. Urbonas, and W. C. Pisano, eds.).

Water Environment Research Foundation. (2013). "Linking BMP Systems Performance to Receiving Water Protection - BMP Performance Algorithms."

Winston, R. J., Hunt, W. F., Kennedy, S. G., Wright, J. D., and Lauffer, M. S. (2012). "Field Evaluation of Storm-Water Control Measures for Highway Runoff Treatment." *Journal of Environmental Engineering*, American Society of Civil Engineers, 138(1), 101–111.

Wong, T. H. F., Fletcher, T. D., Duncan, H. P., and Jenkins, G. A. (2006). "Modelling urban stormwater treatment—A unified approach." *Ecological Engineering*, 27(1), 58–70.

Yu, P.-S., Yang, T.-C., and Chen, S.-J. (2001). "Comparison of uncertainty analysis methods for a distributed rainfall–runoff model." *Journal of Hydrology*, 244(1–2), 43–59.

Zhang, H. X., and Yu, S. L. (2004). "Applying the First-Order Error Analysis in Determining the Margin of Safety for Total Maximum Daily Load Computations." *Journal of Environmental Engineering*, 130(6), 664.

**Chapter 2: Uncertainty Analysis of a Stormwater Control Measure Model using Global Sensitivity Analysis and Bayesian Approaches**

**Informal-3**

| Event No. | BMP Name | Influent Concentration (mg/L) | Effluent Concentration (mg/L) |
|---|---|---|---|
| 1 | BMP12 | 178 | 48 |
| 2 | BMP12 | 128 | 25 |
| 3 | BMP12 | 52 | 24 |
| 4 | BMP12 | 34 | 7 |
| 5 | BMP12 | 38 | 23 |
| 6 | BMP13 | 144 | 50 |
| 7 | BMP13 | 455 | 114 |
| 8 | BMP13 | 46 | 6 |
| 9 | I15SR78 | 260 | 68 |
| 10 | I15SR78 | 370 | 70 |
| 11 | I15SR78 | 120 | 30 |
| 12 | I5Manchester | 400 | 33 |
| 13 | I5Manchester | 170 | 62 |
| 14 | I605SR91 | 110 | 19 |
| 15 | I605SR91 | 93 | 22 |
| 16 | LexHills | 77 | 53 |
| 17 | LexHills | 214 | 60 |
| 18 | LexHills | 53 | 22 |
| 19 | OrchardPond | 91 | 49 |
| 20 | OrchardPond | 21 | 11 |
| 21 | OrchardPond | 342 | 24 |
| 22 | OrchardPond | 144 | 10 |
| 23 | OrchardPond | 21 | 9 |
| 24 | OrchardPond | 58 | 23 |
| 25 | OrchardPond | 7 | 6 |
| 26 | OrchardPond | 103 | 26 |

| 27 | GreenvillePond | 98 | 26 |
|----|----------------|-----|-----|
| 28 | I5SR56 | 48 | 12 |
| 29 | MountainPark | 12 | 7 |
| 30 | MountainPark | 135 | 74 |
| 31 | I5I605 | 100 | 34 |
| 32 | I5I605 | 82 | 36 |
| 33 | I5I605 | 92 | 22 |
| 34 | I5I605 | 86 | 12 |

**Chapter 3: Appraisal of Steady-State Stormwater Control Measure Pollutant Removal**

**Models within a Dynamic Stormwater Routing Framework with Uncertainty Analysis**

*Linear Regression Analysis of EDB TSS Data from International BMP Database*

The figure below shows the linear regression analysis of 137 pairs of influent/effluent TSS EMCs

obtained from the BMP Database.  All data were collected from EDBs located within the United States.



The table below is the linear regression analysis output performed using Excel.  Values of the standard

error of the regression and sum of squares of the residuals were used in the prediction interval and FOVE

uncertainty analysis methods.

| SUMMARY OUTPUT | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0.670057439 | | | | | | | |
| R Square | 0.448976971 | | | | | | | |
| Adjusted R Square | 0.444895319 | | | | | | | |
| Standard Error | 16.62139936 | | | | | | | |
| Observations | 137 | | | | | | | |
| | | | | | | | | |
| ANOVA | | | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | | | |
| Regression | 1 | 30389.47889 | 30389.47889 | 109.9988 | 3.42773E-19 | | | |
| Residual | 135 | 37296.57374 | 276.2709166 | | | | | |
| Total | 136 | 67686.05263 | | | | | | |
| | | | | | | | | |
| | Coefficients | Standard Error | t Stat | P-value | Lower 95% | Upper 95% | Lower 95.0% | Upper 95.0% |
| Intercept | 14.94133088 | 2.302692253 | 6.488635578 | 1.52E-09 | 10.38731419 | 19.49534757 | 10.38731419 | 19.49534757 |
| 99.13 | 0.158044485 | 0.01506903 | 10.48803293 | 3.43E-19 | 0.12824258 | 0.18784639 | 0.12824258 | 0.18784639 |