

Regulatory and Economic Consequences of Empirical Uncertainty for Urban Stormwater Management

Marcus F. Aguilar

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Civil Engineering

Randel L. Dymond, Chair
Thomas J. Grizzard
Glenn E. Moglen
Kurt Stephenson

September 28th, 2016
Blacksburg, Virginia

Keywords: Urban Stormwater Management, Regulations, Uncertainty, Hydraulics and Hydrology

Regulatory and Economic Consequences of Empirical Uncertainty for Urban Stormwater Management

Marcus F. Aguilar

ABSTRACT

The responsibility for mitigation of the ecological effects of urban stormwater runoff has been delegated to local government authorities through the Clean Water Act's National Pollutant Discharge Elimination Systems's Stormwater (NPDES SW), and Total Maximum Daily Load (TMDL) programs. These programs require that regulated entities reduce the discharge of pollutants from their storm drain systems to the "maximum extent practicable" (MEP), using a combination of structural and non-structural stormwater treatment – known as stormwater control measures (SCMs). The MEP regulatory paradigm acknowledges that there is empirical uncertainty regarding SCM pollutant reduction capacity, but that by monitoring, evaluation, and learning, this uncertainty can be reduced with time. The objective of this dissertation is to demonstrate the existing sources and magnitude of variability and uncertainty associated with the use of structural and non-structural SCMs towards the MEP goal, and to examine the extent to which the MEP paradigm of iterative implementation, monitoring, and learning is manifest in the current outcomes of the paradigm in Virginia.

To do this, three research objectives were fulfilled. First, the non-structural SCMs employed in Virginia in response to the second phase of the NPDES SW program were catalogued, and the variability in what is considered a "compliant" stormwater program was evaluated. Next, the uncertainty of several commonly used stormwater flow measurement devices were quantified in the laboratory and field, and the importance of this uncertainty for regulatory compliance was discussed. Finally, the third research objective quantified the uncertainty associated with structural SCMs, as a result of measurement error and environmental stochasticity. The impacts of this uncertainty are discussed in the context of the large number of structural SCMs prescribed in TMDL Implementation Plans. The outcomes of this dissertation emphasize the challenge that empirical uncertainty creates for cost-effective spending of local resources on flood control and water quality improvements, while successfully complying with regulatory requirements. The MEP paradigm acknowledged this challenge, and while the findings of this dissertation confirm the flexibility of the MEP paradigm, they suggest that the resulting magnitude of SCM implementation has outpaced the ability to measure and functionally define SCM pollutant removal performance. This gap between implementation, monitoring, and improvement is discussed, and several potential paths forward are suggested.

Regulatory and Economic Consequences of Empirical Uncertainty for Urban Stormwater Management

Marcus F. Aguilar

GENERAL AUDIENCE ABSTRACT

Responsibility for mitigation of the ecological effects of urban stormwater runoff has largely been delegated to local government authorities through several Clean Water Act programs, which require that regulated entities reduce the discharge of pollutants from their storm drain systems to the “maximum extent practicable” (MEP). The existing definition of MEP requires a combination of structural and non-structural stormwater treatment – known as stormwater control measures (SCMs). The regulations acknowledge that there is uncertainty regarding the ability of SCMs to reduce pollution, but suggest that this uncertainty can be reduced over time, by monitoring and evaluation of SCMs. The objective of this dissertation is to demonstrate the existing sources and magnitude of variability and uncertainty associated with the use of structural and non-structural SCMs towards the MEP goal, and to examine the extent to which the MEP paradigm of implementation, monitoring, and learning appears in the current outcomes of the paradigm in Virginia.

To do this, three research objectives were fulfilled. First, the non-structural SCMs employed in Virginia were catalogued, and the variability in what is considered a “compliant” stormwater program was evaluated. Next, the uncertainty of several commonly used stormwater flow measurement devices were quantified in the laboratory and field, and the importance of this uncertainty for regulatory compliance was discussed. Finally, the third research objective quantified the uncertainty associated with structural SCMs, as a result of measurement error and environmental variability. The impacts of this uncertainty are discussed in the context of the large number of structural SCMs prescribed by Clean Water Act programs. The outcomes of this dissertation emphasize the challenge that uncertainty creates for cost-effective spending of local resources on flood control and water quality improvements, while successfully complying with regulatory requirements. The MEP paradigm acknowledged this challenge, and while the findings of this dissertation confirm the flexibility of the MEP paradigm, they suggest that the resulting magnitude of SCM implementation has outpaced the ability to measure and functionally define SCM pollutant removal performance. This gap between implementation, monitoring, and improvement is discussed, and several potential paths forward are suggested.

ACKNOWLEDGEMENTS

I am grateful for the deliberate commitment of family, friends and colleagues to the successful completion of this work – which I hope will contribute in even a small way, to better stewardship of the urban water environment. The most important of these, is my wife, Mary Katherine - An excellent wife, who can find? – I’ve found one. Thank you for your patient support and help. I am also extremely grateful for Dr. Dymond’s guidance though the writing of this work, and for his commitment to myself and to his students. He has become a great mentor to me, as I’ve witnessed his selfless leadership style, and healthy perspective on work, recreation, and family. Thank you for the countless hours you’ve invested in me, and for a terrific graduate school experience. The input, care, and thoughtful conversations of my advisory committee, and many friends and colleagues also helped shape and refine this work, including Hu Di, Paul Bender, Walter McDonald, Clay Hodges, David Poteet, Kyle Strom and many others. I am also thankful for parents who taught me the value of education, and who encouraged me through the writing of this dissertation. This work was funded by the Department of Civil and Environmental Engineering at Virginia Tech, through the Via Fellowship, and by the City of Roanoke, Virginia. I am grateful to the Via family for their wise stewardship, and to our colleagues at the City of Roanoke, who helped give legs to some of our research ideas. Each chapter in this dissertation is a stone, which collectively form a cairn of remembrance, that uncertainty isn’t hopelessness, and that a grain of wheat that falls to the earth bears much fruit – though it must first fall to the earth. I hope that I will read these papers one day, and remember the providential circumstances, manifest in the relationships mentioned above (and many others), by which they were written.

TABLE OF CONTENTS

Abstract.....	ii
General Audience Abstract.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Figures.....	ix
List of Tables.....	xiii
Attribution.....	xv
1. Introduction.....	1
1.1 Problem Statement.....	1
1.2 Dissertation Objectives.....	1
1.3 Research Approach.....	1
1.4 Dissertation Layout.....	2
2. Literature Review.....	4
2.1 Urbanization and Ecology.....	4
2.2 Urban Stormwater Management.....	5
2.2.1 Policy.....	5
2.2.2 Economics.....	6
2.2.3 Treatment.....	8
2.3 Uncertainty.....	9
3. Evaluation of Variability in Response to the NPDES Phase II Stormwater Program in Virginia.....	12
3.1 Abstract.....	12
3.2 Introduction.....	12
3.3 Methods.....	15
3.3.1 Data Development.....	15
3.3.2 Analysis of the MS4 Type.....	19
3.3.3 Analysis of Cultural, Environmental, and Economic Factors.....	20
3.3.4 Analysis of Other Factors.....	21
3.3.5 SCM Use Efficiencies.....	22
3.4 Results and Discussion.....	23

3.4.1	Value of a Quantitative SCM Database	23
3.4.2	The Effect of MS4 Type.....	25
3.4.3	The Effect of Stormwater Organizations, Consultants, and the Chesapeake Bay TMDL Requirements	30
3.4.4	Cultural, Environmental, Economic Factors	33
3.4.5	SCM Use Efficiencies	36
3.5	Conclusions.....	38
4.	Benchmarking Laboratory Observation Uncertainty for In-Pipe Storm Sewer Discharge Measurements	42
4.1	Abstract.....	42
4.2	Introduction.....	42
4.3	Instrumentation	45
4.3.1	Depth Measurement	47
4.3.2	Velocity Measurement	47
4.4	Laboratory Procedure.....	49
4.4.1	Benchmark Uncertainty.....	50
4.4.2	Bias Uncertainty.....	52
4.4.3	Precision Uncertainty	53
4.4.4	The Effect of Woody Debris	53
4.4.5	Combined Uncertainty	54
4.5	Field Procedure	55
4.5.1	Selection of a Field Location and Installation.....	55
4.5.2	Sensor Power Considerations.....	57
4.5.3	Application of Laboratory Uncertainty to Field Observations	58
4.6	Results and Discussion	58
4.6.1	Bias Uncertainty.....	59
4.6.2	Precision Uncertainty	63
4.6.3	The Effect of Woody Debris	64
4.6.4	Application to Field Observations	66
5.	Evaluation of Stormwater Control Measure Performance Uncertainty	73
5.1	Abstract.....	73

5.2 Background.....	73
5.3 Introduction.....	74
5.4 Methods.....	78
5.4.1 Environmental Uncertainty	81
5.4.2 Observation Uncertainty	83
5.4.3 Category Uncertainty	88
5.5 Results and Discussion	89
5.5.1 Environmental Uncertainty	90
5.5.2 Observation Uncertainty	93
5.5.3 Category Uncertainty	95
5.6 Conclusion	101
6. Conclusion	103
6.1 Outcomes	103
6.2 Broader Impacts	105
6.3 Future Work	106
References.....	108
Appendix A - Flow Sensor Information and Flume Rating.....	130
A.1 Ultrasonics	130
A.2 Acoustic Doppler Velocimeters (ADV).....	131
A.3 Sensor Communication and Power Specifications	133
A.4 Manometer and Venturi Measurements	134
Appendix B - Laboratory Methods for Phosphorus, Nitrogen, and Total Suspended Solids.....	136
B.1 Phosphorus	136
B.2 Nitrogen.....	139
B.3 Total Suspended Solids	139
Appendix C - Stormwater Control Measure Performance Uncertainty.....	141
C.1 Runoff Volume Reduction	141
C.2 Total Phosphorus Pollutant Removal.....	144
C.3 Total Nitrogen Pollutant Removal	150
C.4 Total Suspended Solids Pollutant Removal	154

Appendix D - References Used to Develop Stormwater Control Measure Performance

Uncertainty.....	159
D.1 Table of References	159
D.2 List of References	165

LIST OF FIGURES

Figure 2.1 – The estimated capital outlay cost to Virginia MS4 permit holders to achieve the requirements of the Chesapeake Bay TMDL Implementation Plan (IP). Estimates from (Virginia Senate Finance Committee, 2011) and (AMEC, 2012). VDOT – Virginia Department of Transportation..... 7

Figure 3.1 - The 59 reported Stormwater Control Measures (SCMs) in Program Plans for 2013 – 2018 permit years* 18

Figure 3.2 - SCMs used at a significantly different frequency across MS4 Types. The “Municipal” cluster includes cities (n = 21), counties (10), and towns (9), and the “Other” cluster includes federal/state facilities (18), local school boards (2), military installations (8), and a transport system (1). The number in each cell is the percentage of use within categories, also indicated by the cell shading. 27

Figure 3.3 - Tukey Box and Whisker plot demonstrating total number of SCMs reported, categorized by MS4 type. The “Other” category includes federal/state facilities, local school boards, military installations, and a transport system..... 28

Figure 3.4 - Similarity of Municipal MS4 Program Plans in reported stormwater organizations with bubble size representing relative size of each group, and number of reported membership (n) in black within each bubble..... 31

Figure 3.5 - Significant logistic correlations between factors and frequency of SCM use in Municipal MS4s at 95 and 99% confidence. The sign in parentheses represents the direction of the effect (β_1) based on the estimated log-odds 34

Figure 4.1 – Uncertainty associated with the observation of a measurand X , adapted from Coleman and Steele (1995)..... 44

Figure 4.2 – Model woody debris field used to evaluate sensor uncertainty with obstructions in storm sewer channels 54

Figure 4.3 – Sensor depth observations plotted against manual observations (A - C), and sensor velocity observations plotted against manometer derived velocity (D - E) for all trials with least-squares regression lines shown 61

Figure 4.4 – Sensor observation residuals, $X_{o,j} - X_{s,j}$ for depth (A – C) and velocity (D – E) with line of perfect agreement. (E) Sensor 2’s velocity residuals varied with Froude number; the least-squares regression line is shown 62

Figure 4.5 – Standard deviation of i sensor observations for each trial, Ss, j plotted against manometer derived Froude Number, Fro, j with the median value shown representing precision uncertainty 64

Figure 4.6 – (A) Peak flow and (B) total runoff volume calculated for nine storm events based on original and adjusted discharge observations from Sensor’s 1 and 2 with the upper and lower bounds of laboratory benchmarked uncertainty shown 68

Figure 4.7 – Adjusted (A) depth, (B) velocity, and (C) discharge time series for storm event on 4/16/2015 with observations from three depth sensors, two velocity sensors, and the upper and lower bounds of combined uncertainty from the laboratory. The adjusted discharge observations for this storm event had the highest Nash-Sutcliffe Efficiency of any event recorded. 69

Figure 5.1 - The organizational structure of i data points in j SCM source studies for an example SCM category k shown with dimensions of environmental and observation uncertainty. Category uncertainty is for this example category is shown as the dark gray background region. 80

Figure 5.2 - Uncertainty associated with removal capabilities for bioretention stormwater control measures for: (A) Runoff Volume Reduction (RR), (B) Total Phosphorus EMC Reduction (PR), (C) Total Nitrogen EMC Reduction (PR), and (D) Total Suspended Solids EMC Reduction (PR). Information on the right of each Figure is as follows: “n” represents the number of data points in each study, the bracketed number represents upper and lower uncertainty bounds of the Total Study Uncertainty ($u_{env} + u_{obs}$), and the text describes the type of information reported. Citations noted with an * use RR data from (Poresky et al., 2011) and PR data from the 12/20/2014 version of the International Stormwater BMP Database (WERF et al., 2014). 92

Figure 5.3 - Summary of environmental and observation uncertainty for all stormwater control measures (SCMs) studied in each of the four pollutants of concern: runoff volume (RR), total nitrogen (TN), total phosphorus (TP), and total suspended solids (TSS). Bar heights represent the median uncertainties for j SCMs studied in each group (shown in brackets), and whiskers represent 25th and 75th percentile values. 94

Figure 5.4 - Uncertainty in total mass load reduction (TR): $1 - (\text{effluent mass}/\text{influent mass})$, estimated based on the Runoff Reduction Method (Equation 3). VDEQ Level 1 and 2

designate the credit given by the VDEQ for the two design levels of the SCM (VDEQ, 2013); TSS credit is not specified by the VDEQ.....	97
Figure A.1 – The Massa M-300/95 mounted on the soffit of a concrete box culvert.....	130
Figure A.2 – Vectrino II Velocimeter mounted in the flume	133
Figure B.1 - Bias and Precision associated with various methods for determining total phosphorus in the laboratory. The median bias, precision, and overall uncertainty used in Chapter 5 are shown as red, blue, and black dashed lines respectively.....	137
Figure B.2 - Bias and Precision associated with various methods for determining total nitrogen in the laboratory.	139
Figure C.1 – Runoff Volume Reduction capabilities of Bioretention Cells	141
Figure C.2 – Runoff Volume Reduction capabilities of Compost Soil Amendment.....	141
Figure C.3 – Runoff Volume Reduction capabilities of Dry Swales.....	141
Figure C.4 – Runoff Volume Reduction capabilities of Grass Channels	142
Figure C.5 – Runoff Volume Reduction capabilities of Infiltration.....	142
Figure C.6 – Runoff Volume Reduction capabilities of Permeable Pavement.....	142
Figure C.7 – Runoff Volume Reduction capabilities of Rainwater Harvesting	142
Figure C.8 – Runoff Volume Reduction capabilities of Rooftop Disconnect	143
Figure C.9 – Runoff Volume Reduction capabilities of Vegetated Filter Strip.....	143
Figure C.10 – Runoff Volume Reduction capabilities of Vegetated Roof.....	143
Figure C.11 – Total Phosphorus pollutant removal capabilities of Bioretention cells	144
Figure C.12 – Total Phosphorus pollutant removal capabilities of Constructed Wetlands	144
Figure C.13 – Total Phosphorus pollutant removal capabilities of Extended Detention.....	145
Figure C.14 – Total Phosphorus pollutant removal capabilities of Filtering Practices	145
Figure C.15 – Total Phosphorus pollutant removal capabilities of Grass Channels	146
Figure C.16 – Total Phosphorus pollutant removal capabilities of Infiltration	146
Figure C.17 – Total Phosphorus pollutant removal capabilities of Permeable Pavement.....	147
Figure C.18 – Total Phosphorus pollutant removal capabilities of Rainwater Harvesting	147
Figure C.19 – Total Phosphorus pollutant removal capabilities of Vegetated Filter Strip.....	148

Figure C.20 – Total Phosphorus pollutant removal capabilities of Vegetated Roof. Results from Moran (2005) were excluded, as the soil media used composted cow manure to improve vegetation growth, leading to significant export of nutrients.	148
Figure C.21 – Total Phosphorus pollutant removal capabilities of Wet Ponds	149
Figure C.22 – Total Phosphorus pollutant removal capabilities of Wet Swales	149
Figure C.23 – Total Nitrogen removal capabilities of Bioretention	150
Figure C.24 – Total Nitrogen removal capabilities of Constructed Wetlands.....	150
Figure C.25 – Total Nitrogen removal capabilities of Extended Detention	150
Figure C.26 – Total Nitrogen removal capabilities of Filtering Practices	151
Figure C.27 – Total Nitrogen removal capabilities of Grass Channels	151
Figure C.28 – Total Nitrogen removal capabilities of Infiltration.....	151
Figure C.29 – Total Nitrogen removal capabilities of Permeable Pavement	151
Figure C.30 – Total Nitrogen removal capabilities of Rainwater Harvesting	152
Figure C.31 – Total Nitrogen removal capabilities of Vegetated Filter Strips.....	152
Figure C.32 – Total Nitrogen removal capabilities of Vegetated Roof. Results from Moran (2005) were excluded, as the soil media used composted cow manure to improve vegetation growth, leading to significant export of nutrients.	152
Figure C.33 – Total Nitrogen removal capabilities of Wet Pond	153
Figure C.34 – Total Nitrogen removal capabilities of Wet Swale.....	153
Figure C.35 – Total Suspended Solids removal capabilities of Bioretention	154
Figure C.36 – Total Suspended Solids removal capabilities of Compost Soil Amendment	154
Figure C.37 – Total Suspended Solids removal capabilities of Constructed Wetlands.....	154
Figure C.38 – Total Suspended Solids removal capabilities of Extended Detention	155
Figure C.39 – Total Suspended Solids removal capabilities of Filtering Practices.....	155
Figure C.40 – Total Suspended Solids removal capabilities of Grass Channels	156
Figure C.41 – Total Suspended Solids removal capabilities of Infiltration.....	156
Figure C.42 – Total Suspended Solids removal capabilities of Permeable Pavement	157
Figure C.43 – Total Suspended Solids removal capabilities of Rainwater Harvesting	157
Figure C.44 – Total Suspended Solids removal capabilities of Vegetated Filter Strips.....	157
Figure C.45 – Total Suspended Solids removal capabilities of Wet Ponds.....	158
Figure C.46 – Total Suspended Solids removal capabilities of Wet Swales	158

LIST OF TABLES

Table 3.1 - Six minimum control measures for stormwater management in Phase II MS4s from USEPA (1999)	13
Table 3.2 - Explanatory variables describing MS4 characteristics and their respective sources .	16
Table 4.1 – Classification, Manufacturer Specifications, and Cost of Sensors	46
Table 4.2 – The transformation functions used to adjust sensor depth (<i>ds</i>) and velocity (<i>Vs</i>) observations towards manual observations with and without debris, the bias uncertainty of the original observations and the adjusted observations, precision uncertainty, benchmark uncertainty, and combined uncertainty from Equation 4.7. Values of uncertainty are in +/- the unit of measurement.	60
Table 4.3 - Minimum, maximum and mean Nash-Sutcliffe efficiency (NSE) values for nine storm events using Sensor 1 as the “observed” data.....	67
Table 4.4 - Average Daily Volume Summary	70
Table 5.1 - Stormwater control measure (SCM) categories used in the Virginia Department of Environmental Quality’s Best Management Practice (BMP) Standards and Specifications (VDEQ, 2013). L1 and L2 designate Level 1 and 2 design criteria for each SCM category, where L2 requires increased size, depth, or a more restrictive soil specification. Sheet Flow to Conserved Open Space and Sheet Flow to Vegetated Filter Strip were treated as a single category for this study.	79
Table 5.2 - Uncertainties associated with various methods of estimating inflow and outflow volume for SCMs as a percent of the measurement. Where error distributions were available, the median was used. Where error was reported in units of the measurement, it was converted to a percentage based on a reasonable range of observations for SCM studies.	84
Table 5.3 - Analytical uncertainty associated with laboratory characterization of phosphorus, nitrogen, and total suspended solids using several methods. Observation uncertainty, u_{obs} is estimated as the probable error range of bias and precision uncertainty	86
Table 5.4 - Number of stormwater control measures (SCMs) for which data was collected for each pollutant of concern (POC), and the number of individual data points in these studies. Many studies reported results for multiple SCMs and/or multiple POCs.	90

Table A.1 – Manufacturer Specifications and Cost of Ultrasonic Sensors.....	130
Table A.2 - Manufacturer Specifications and Cost of Acoustic Doppler Sensors.....	131
Table A.3 - Sensor Communication and Power Specifications	134
Table D.1 - Studies reviewed for percent removal values for runoff volume reduction (RR), and reduction in EMC concentration as total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS).	160

ATTRIBUTION

Each of the following co-authors contributed to the manuscripts published, or in review, that comprise the work of this dissertation as credited below:

Randel L. Dymond, Ph.D., Associate Professor

Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA.

Co-author of Chapters 2, 3, and 4

Walter M. McDonald, Ph.D., Assistant Professor

Department of Civil, Construction, and Environmental Engineering, Marquette

University, Marquette WI.

Co-author of Chapter 3

1. INTRODUCTION

1.1 PROBLEM STATEMENT

Mitigation of the effects of urbanization on stormwater runoff quality and quantity presents a complex challenge to local government authorities, as the existing treatment paradigm set forth in the Clean Water Act (CWA) (*33 USC 26 - Federal Water Pollution Control Act, 1972*) – known as “maximum extent practicable” - leads to significant expenditures on stormwater control measures (SCMs) with unproven ecological benefit at large time and space scales. The objective of this dissertation is to enumerate the SCMs currently prescribed as the “maximum extent practicable” pollutant reduction practices, quantify the extent to which the pollutant removal benefits of these SCMs can be empirically measured, and explore the consequences of this empirical uncertainty on urban stormwater management.

1.2 DISSERTATION OBJECTIVES

In order to fulfill the overall research objective, this dissertation is framed as the following three research questions:

1. What non-structural SCMs are being used in small regulated entities in Virginia and why, and are these SCMs pursuant to the “maximum extent practicable” goal of the CWA’s stormwater program?
2. How uncertain are stormwater discharge measurements, and what are the regulatory implications of that uncertainty?
3. How uncertain are SCM pollutant removal efficiencies, and how does this affect the selection of the most cost-effective SCMs?

1.3 RESEARCH APPROACH

The first research question was addressed by reviewing the current Virginia outcomes of the section of the CWA which requires stormwater quality control programs of small local governments. Documents describing local government stormwater management programs, supplemented with interviews of stormwater authorities were used as descriptors of the non-structural, programmatic approaches that local governments in Virginia are using to manage stormwater quality and quantity. A correlational and observational approach was used, such that the initial hypotheses were added to iteratively, as patterns in the data emerged, and the results identify these patterns but were not designed to be predictive.

The second research question focused on the uncertainty associated with discharge measurements for stormwater flows, and included data collection in a laboratory flume and field culvert. Several hypotheses were formed regarding the drivers of uncertainty for the flow sensors tested, and the sensors were evaluated over a range of hydraulic conditions in the lab to test the validity of these hypotheses. Uncertainty was presented as \pm the absolute values of the measurements, and as percent deviations from benchmarks. Field uncertainty was estimated using the uncertainty bounds estimated in the laboratory, and field results were observational.

Finally, the third research question was addressed using a meta-analysis approach to the SCM monitoring literature. The hypothesis of this study – that SCM performance could be explained using a single categorical variable – was tested by gathering existing datasets from the literature, and developing uncertainty bounds for each SCM category. The outcomes of the meta-analysis are evaluated comparatively, but without formal statistical testing, as the data collected from source studies was not consistent enough to allow for hypothesis tests.

1.4 DISSERTATION LAYOUT

The dissertation is organized into the following Chapters, with associated Appendices noted:

Chapter 2 provides background literature supporting the need for the research questions in this dissertation. It starts with a broad overview of the field of urban stormwater management, which is followed by an overview of the current regulatory environment. Next, the economic implications of the stormwater regulations are discussed, along with an introduction to the current methods used to treat urban stormwater: SCMs, also known as “best management practices” (BMPs). Finally, a brief discussion of uncertainty in the water resources literature is provided, as an introduction to the subsequent chapters.

Chapter 3 addresses research objective 1, by reviewing the outcomes of the NPDES stormwater program for small MS4s in Virginia. In this Chapter, 90 stormwater programs were reviewed for the non-structural SCMs reported, in accordance with the six minimum control measures (MCMs) outlined in the Phase II final rules. These SCMs are thought to reduce the discharge of pollutants from urban areas to the maximum extent practicable, and this research catalogues the measures reported, attempts to explain their use based on a number of socio-economic and environmental factors, and identifies management inefficiencies.

Chapter 4 addresses research objective 2, by focusing on the measurement uncertainty of a fundamental parameter in urban stormwater management: discharge. Several commonly used discharge sensors are tested in a laboratory flume for bias and precision uncertainty, and the sensors are then placed in a storm sewer pipe in Blacksburg, Virginia, where the laboratory-benchmarked uncertainty bounds are applied. The effects of uncertainty on the estimation of peak discharge, and total storm event runoff volume are demonstrated, and the regulatory consequences, namely the importance of observation uncertainty on TMDL models and SCM percent removal values are discussed. Information regarding the sensors used in this Chapter is included in Appendix A, along with a description of the development of the flume discharge rating curve using manometry.

Chapter 5 addresses research objective 3, by developing a framework for total uncertainty in SCM performance according to the percent removal performance metric. Total uncertainty is estimated for the 15 categories of SCMs defined by the Virginia stormwater regulations. Uncertainty results from the stochasticity of runoff processes, the challenge of accurately measuring flow and pollutant concentrations, and the aggregation of many SCMs of different composition into a single SCM category. The effects of this uncertainty on estimating cost-effectiveness of a SCM are demonstrated using an example SCM from Fairfax County, Virginia. The references used to develop the SCM uncertainty bounds are included in Appendix D, and the methods used to develop several of the laboratory uncertainty bounds are included in Appendix B. Comprehensive results of all SCMs are provided in Appendix C.

Chapter 6 summarizes the outcomes of the dissertation, examines the broader impacts of these outcomes for the field of urban stormwater management, and recommends potential directions for future research. This is the final Chapter of the dissertation body, and is followed by a References section for the entire body. Sources in the Appendices but not in the dissertation body are not included in this References section, but are in the relevant Appendix.

2. LITERATURE REVIEW

2.1 URBANIZATION AND ECOLOGY

In the 20th and 21st centuries, global population has shifted from approximately 10% to over 50% of the population dwelling in urban areas – defined by the United Nations as a contiguous territory inhabited at urban population density levels – this is projected to reach 66% by 2050. In the U.S., the urban population growth between 2014 and 2050 was projected as an additional 82 million urban-dwellers, amounting to 89% of the total projected U.S. population of 400 million in 2050 (United Nations, 2014). The dramatic shift of the population into cities has generally allowed for improved delivery of critical infrastructure (e.g. transportation, electricity) to a larger proportion of the total population, and cities also appear to generate increased rates of innovation and wealth due to the high density of social interaction (Bettencourt et al., 2007). These increased efficiencies allow for higher consumption in cities, with the cost being the exposure of the surrounding ecosystem to denser fluxes of waste generated by this consumption (Bettencourt and West, 2010; Costanza and Daly, 1992). The future of the city-ecosystem relationship is subject to a variety of speculative theories, ranging from the eventual collapse of the relationship due to the un-sustainability of the urban form (Kunstler, 2003), to an increasingly beneficial symbiosis, resulting from the scaling principles noted above (Glaeser, 2012).

Glaeser also argues that the densification of pollution in cities is a more efficient method of polluting than a more spatially distributed model, though this is not uniformly agreed upon. Whatever the outcome of this debate, theorists will benefit from the substantial knowledge base that has documented the large ecological footprints of cities – in some cases tens to hundreds of times the area occupied by the city itself (Folke et al., 1997). This urban footprint includes shifting land cover/use patterns within cities and in support of cities' resource needs, an associated shift in biogeochemical cycling (e.g. the dense output of CO₂ from urban areas), thermal effects on the climate at local and synoptic scales, hydro-modification, and altered biodiversity (Grimm et al., 2008). The magnitude and composition of each of these inter-related categories of the urban footprint have been explored extensively, and the ability to manage these effects varies. The focus of this dissertation is on hydro-modification – changes to the water environment affected by urbanization – and the limitations of our existing management paradigm to mitigate these changes.

2.2 URBAN STORMWATER MANAGEMENT

The effects of urbanization on the water environment were organized by Leopold (1968) into four related but separable categories: changes to (1) peak flow characteristics, (2) total runoff volume, (3) water quality, and (4) the hydrologic amenities (i.e. the overall appearance of a river or stream). In the four decades since Leopold's primer, the literature characterizing the changes to the water environment affected by land surface and runoff process modification due to urbanization has grown substantially. Changes to the hydrologic regime, are principally caused by the increased magnitude of impervious surfaces, and their efficient hydraulic connectivity to receiving waters through storm sewer pipes (Elmore and Kaushal, 2008; Kaushal and Belt, 2012; Walsh, Fletcher, et al., 2005). This hydrologic alteration has been referred to as the master (but not exclusive) variable driving geomorphic, ecological, and biochemical changes in the watershed (Brabec, 2009; Doyle et al., 2005; Paul and Meyer, 2001; Walsh, Roy, et al., 2005). Leopold's work was intended as a guidance document for land planning, as it was realized even then, that the management of runoff from urban areas is driven by land use. This connection between the modification of the land surface, and the downstream changes to the hydrologic regime, creates a politically, economically, and technically challenging problem.

2.2.1 *Policy*

Politically, stormwater pollution is an expression of the tragedy of the commons, whereby multiple (usually) private land owners are responsible for flooding, erosion, or degraded water quality that is sufficiently downstream to be invisible to the upstream land owners, or sufficiently diffuse so that responsibility is difficult to attribute (Hardin, 1968). A strong sentiment of private land ownership in the U.S., as exemplified for example, by the failed National Land Use Policy Act of 1974 (Nolon, 1996), makes attempts to mitigate stormwater effects on private land an especially sensitive issue. This means that urban stormwater pollution control has largely become the responsibility of local governments by way of several Clean Water Act programs.

Legal responsibility to mitigate the effects of urbanization has been delegated through the Clean Water Act's National Pollutant Discharge Elimination System Stormwater (NPDES SW) program and the Total Maximum Daily Load (TMDL) program (33 USC 26, §§ 402 and 303, respectively). The NPDES SW program created a permitting system requiring owners of

municipal separate storm sewer systems (MS4s, including cities, counties, towns, and other local government entities) to create and document stormwater management programs that work toward the NPDES SW program's objective: "reduce the discharge of pollutants to the maximum extent practicable (USEPA, 1990, p. 47994, 1999, p. 68752)."

This subjective definition of regulatory compliance was proposed because of its adaptability to innovation in stormwater control measures (SCMs, also commonly referred to as "best management practices", BMPs), but has since incorporated TMDL effluent limitations, known as waste load allocations (WLAs). These "pollution diets" for urbanized areas are based on state-level water quality and hydrologic monitoring and modeling, and their enforcement through the NPDES SW program has shifted the regulatory environment for MS4s toward a pseudo-numeric based system. As the MS4 permit provides statutory authority for TMDL enforcement in regulated urban areas, but not outside of these areas [40 CFR §122.44(d)(1)(vii)], significant regulatory responsibility for mitigating stormwater pollution has been delegated to MS4 entities (i.e. urban areas) (D. Owen, 2011). The cost of local government programs developed pursuant to the NPDES and TMDL requirements are passed along to private-property owners by means of taxes or stormwater utility fees (Kea and Dymond, 2016).

2.2.2 Economics

The recent proliferation of these stormwater utility fees (SWUs) underscores the substantial cost of stormwater treatment, as over 1,400 localities in the U.S. now use some variety of SWU to fund stormwater programs (Campbell et al., 2014). These fees have led to political push-back, again, largely because of the conceptual disconnection between land surface modifications on private land, and downstream ecological consequences (NAFSMA, 2006), and it is still not clear if the revenue generated by these fees will provide sufficient funding for local governments to prevent flooding, improve water quality, and meet regulatory requirements. The increased level-of-effort required by stormwater regulations was not supported by additional federal funding, as the required cost-benefit analysis for both Phase I and II of the NPDES SW program showed a net benefit to localities, and because the total estimated cost to municipalities, industry, and State/Federal authorities was \$14.5 million for Phase I (USEPA, 1990), and between \$848 and \$982 million for Phase II (Science Applications International Corporation, 1999a; USEPA, 1999).

It should be noted that these original estimates of the cost of the NPDES program did not include the cost associated with SCM implementation towards TMDL action plans. The cost to design, construct, operate, and maintain SCMs – i.e. the whole life cost – is a topic of current research (Hodges et al., 2016; S. Taylor et al., 2014), and while these costs are subject to a wide range of variability resulting from fluctuating material and labor costs, preliminary cost removal estimates for total phosphorus have been reported in the tens of thousands of dollars per pound of phosphorus removed (Houle et al., 2013; Nobles et al., 2014). The impact of these per pound estimates at a regional scale is, for example, demonstrated by the large estimated expenditures incurred by Virginia localities for the requirements of the Chesapeake Bay TMDL Watershed Implementation Plan. These estimates for individual localities, based only on the capital outlay required to meet the final WLA targets, range from \$7 million to \$2.1 billion by 2025 [Figure 2.1, (AMEC, 2012; Virginia Senate Finance Committee, 2011)].

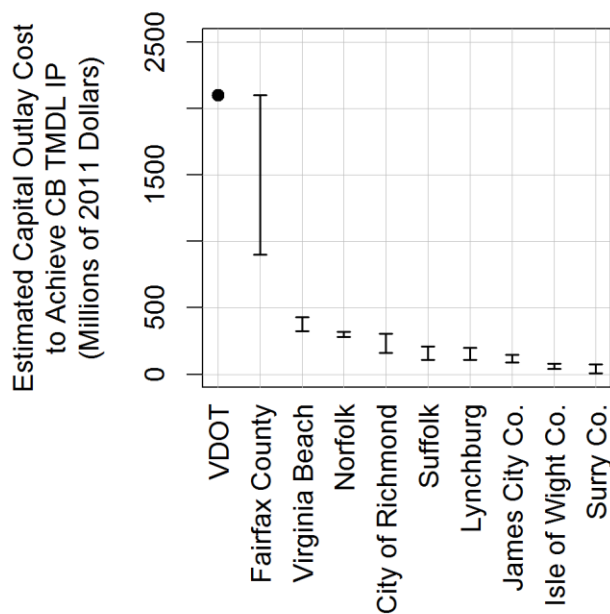


Figure 2.1 – The estimated capital outlay cost to Virginia MS4 permit holders to achieve the requirements of the Chesapeake Bay TMDL Implementation Plan (IP). Estimates from (Virginia Senate Finance Committee, 2011) and (AMEC, 2012). VDOT – Virginia Department of Transportation.

The large anticipated costs for urban stormwater management are further emphasized by the lack of conclusive evidence that the currently used treatment technologies – SCMs - will provide sufficient long-term pollutant removal capabilities across a watershed to meet water quality standards.

2.2.3 *Treatment*

The ability of passive treatment technologies to treat stormwater runoff is a matter of much current research, which can generally be divided into two categories: (1) studies that evaluate the effects of individual SCMs, and (2) studies that evaluate the effects of SCMs at a larger (e.g. sub-watershed or watershed) scale. In general, the local effects of SCMs over short time periods are better understood than the large-scale, long-term effects of SCMs on a watershed. In both cases, the observed or modeled removal capabilities are constrained by sensing technology, and the large amount of variability in climate, soils, slope, vegetation, construction quality, etc. (Barrett, 2008; Sample et al., 2012).

The literature on local effects of SCMs is large and growing, and began approximately in the 1970's when the only available SCM was what is now categorized as the dry detention pond, and the focus was finding the optimal configuration of detention pond size, volume, and outlet structure configuration to prevent excessive downstream flooding (e.g. McCuen, 1974). At this time, the desired function of the SCM was limited to flood control objectives, as the magnitude of the water quality effects of urbanization had not yet been quantified.

This paradigm began to shift in the 1980's with the publication of the Nationwide Urban Runoff Program (USEPA, 1983), and the subsequent data in the National Stormwater Quality Database (Pitt et al., 2004). It became clear at this time, that managing peak flow during runoff events would not be sufficient for protecting receiving waters, and it was noted that a peak-flow only paradigm would lead to excessive downstream channel erosion (McCuen and Moglen, 1988). Moreover, it has been noted more recently that local success at peak-flow management can actually have negative flood control consequences at a watershed scale (Emerson et al., 2005).

The inclusion of erosion and sediment control as a management objective pointed to the increased total volume of runoff as the primary driver – a principle noted by Leopold (1968) (see Chapter 2.2), though not statutorily recognized until the NPDES Phase II final rules (USEPA, 1999). During this period, it also became clear that the excess sediment in stormwater carried with it a suite of pathogens, organic chemicals, nutrients, and heavy metals that would present a human health risk, and degrade the quality of the aquatic ecosystem (Makepeace et al., 1995; Walsh, Roy, et al., 2005). This combination of outcomes, attributable to the diminished abstractive capacity of urban watersheds, and the subsequent increase in runoff volume, has led

to a newer paradigm that attempts to restore these services by means of smaller, more distributed SCMs that focus on infiltrating runoff (Burns et al., 2012).

Evaluative studies of the function of SCMs designed for volume-control began to appear in the mid to late 2000's (e.g. Emerson and Traver, 2008; Hunt et al., 2006), and the literature has grown substantially since then, though it is not reviewed here, as this is done in Chapter 5, and a table of available studies is provided in Appendix D. Overall, these SCMs appear to perform the services they are designed for, at the site for which they were designed, although data on their long-term performance is still not widely available (although see, Komlos and Traver, 2012; S. Taylor et al., 2014). Furthermore, the evaluation of the effects of SCMs at the watershed scale appear to be limited to modeling studies. No studies have been able to detect improvement in urban water quality as a result of SCM implementation, though this may be due to the implementation-response lag time (Meals et al., 2010), or the un-controllable confounding variables driving water quality in watersheds. However, there are ongoing studies to this end (e.g. Jastram, 2014), though the long-term effects are still inconclusive.

Another option that has been explored for stormwater treatment are non-structural or programmatic SCMs - any type of stormwater treatment that does not require the construction of a physical treatment facility. These SCMs form the six minimum control measures (MCMs) for stormwater quality control required of permitted entities in the NPDES SW Phase II Final Rules (USEPA, 1999), and a summary of non-structural SCMs is given in Chapter 3. These types of SCMs have been recommended because of their low capital requirements and flexibility in implementation (A. Taylor and Wong, 2002), though their effects on water quality are very difficult to measure (Dietz et al., 2004). As a result, the effects of non-structural SCMs on water quality, like their structural counterparts, have not been conclusively shown.

2.3 UNCERTAINTY

As described above, the effectiveness of structural and non-structural SCMs – the prescribed treatment methods from the Clean Water Act - at the removal of pollutants in stormwater runoff is subject to a high level of uncertainty. This is not unique in the hydrologic sciences, as the problem of uncertainty in monitoring and modeling watershed physical and biochemical processes has been noted since at least the 1970's (Bogardi and Szidarovsky, 1974). In general, uncertainty is categorized into:

- Natural Uncertainty – uncertainty that arises from the stochasticity of the underlying process, also referred to as “aleatory”, or “Type-A” uncertainty.
- Epistemic Uncertainty – uncertainty that is introduced by the limitations of observation or measurement, also referred to as “ignorance”, or “Type-B” uncertainty.

These two categories of uncertainty can be thought of separately, and in most cases there is some component of each type, though it can be difficult to separate the two in practice (Merz and Thielen, 2005). In the hydrologic sciences, these two uncertainties manifest in various ways, but two of the primary sources of uncertainty are “monitoring” or “observation” uncertainty, and “modeling” uncertainty.

Monitoring uncertainty is comprised of both natural and epistemic components, as the inability to perfectly measure a hydrologic process is a result of both the stochasticity of the process, and the limitations of observation. Several components of observation uncertainty, including precipitation, discharge, and the characterization of nutrients and sediment, are reviewed in meta-analyses by McMillan et al. (2012) and Harmel et al. (2006). Monitoring uncertainty is also important because of the effect that it has on the calibration of hydrologic models, shown, for example in Harmel and Smith (2007). However, the uncertainty associated with building hydrologic models also stems from the parametrization of the model (modeling uncertainty), especially important because multiple parameter combinations can produce the same model results – known as equifinality (Beven and Freer, 2001). This implies that the model, while providing the “correct” output, may not be faithfully representing the physical processes, and may therefore not be transposable to other uses (Klemeš, 1986). Modeling uncertainty in the hydrologic sciences is reviewed extensively in Beven (1993), and the importance in urban stormwater models in Dotto et al. (2012).

Uncertainty in reported measurements becomes a management issue when the measurements are used to make decisions or assessments based on these observations. For example, the designation of a water body as “impaired” based on the results of ambient water quality sampling pursuant to Sections 305(b) and 303(d) of the Clean Water Act, is highly subject to the environmental conditions during the collection of the grab samples (Smith et al., 2001). Once the impairment has been designated, a TMDL is generated (typically) using a hydrologic and water quality model – also subject to a significant amount of uncertainty due to

model parameterization, and error in the measurements to which the model is calibrated (Dilks and Freedman, 2004). The result of the impairment designation, and TMDL creation process, is that regulated entities in the TMDL watershed are then required to spend significant amounts of resources on the reduction of pollutants towards the TMDL endpoint (see Section 2.2.2), in contrast to the significant amount of empirical uncertainty behind the development of the endpoint, and the impairment which it is intended to address (Cooter, 2004; Houck, 2003).

3. EVALUATION OF VARIABILITY IN RESPONSE TO THE NPDES PHASE II STORMWATER PROGRAM IN VIRGINIA

3.1 ABSTRACT

Authorities of municipal separate storm sewer systems (MS4s) in small urbanized areas (population less than 100,000) are required to implement stormwater control measures (SCMs) to mitigate and reduce the impacts of urbanization on stormwater runoff under Phase II of the National Pollutant Discharge Elimination System's (NPDES) stormwater program. This sixteen year old policy has been challenged in its effectiveness in maintaining or improving water quality, but reviews are scarce because of the policy's subjective requirements, and because it governs MS4s across a wide variety of characteristics, objectives, and institutional capacity. This research models SCM selection as a function of these differences, thereby systematically evaluating the policy's outcome in its constituents. The results show that certain characteristics of an MS4 community significantly affect the selection of SCMs, suggesting that regulations may need to be refined to address distinct groups of MS4s. The results also reveal inefficiencies and underutilizations in the SCMs employed – a problem that could be resolved by effectively sharing strategies among permittees. Subsequent recommendations are provided for policy makers and stormwater authorities.

3.2 INTRODUCTION

A lawsuit settlement agreement reached in May 2010 with the Chesapeake Bay Foundation and others (Fowler et al. v. USEPA, 2010) compelled the U.S. Environmental Protection Agency (EPA) to propose a new national stormwater rule that will require additional pollutant load reductions from new and redeveloped sites, Municipal Separate Storm Sewer System (MS4) retrofits to reduce existing loads, and the expansion of the definition of an MS4. This proposal called for a comprehensive reform of the current National Pollutant Discharge Elimination System (NPDES) Phase I and II programs - the current regulatory mechanism for addressing stormwater runoff quality in urbanized areas (USEPA, 1999). As of summer 2015, the EPA has not submitted a proposal, stating that the agency has deferred regulatory action, instead opting to pursue incentive-based programs. This deferral has been contested in the U.S. Ninth Circuit Court of Appeals as an evasion of prior rulings requiring regulatory reform (EDC and NRDC v. USEPA, 2014).

To provide supporting data analysis for future regulations, this research examines the outcomes of the old NPDES program - specifically the Phase II component for urbanized areas of population less than 100,000 (as defined by the U.S. Census Bureau). The Phase II program began in 1999, and has since grown to include approximately 6,700 urbanized areas across the U.S. This research extends Supreme Court Justice Louis Brandeis’ abstraction of states as “laboratories of democracy (Brandeis, 1932)” to include local governments as microcosms of policy success and failure. In spite of the shortcomings of the current national stormwater rules (NRC, 2009), the EPA has the opportunity to use adaptive management (Doyle, 2012) in light of the outcomes of their old rules, and new priorities and technologies, in order to improve the forthcoming legislation.

The conditions for compliance with the existing Phase II policy require that MS4 permittees (MS4s) implement a combination of interventions [i.e. strategies or stormwater control measures (SCMs), also commonly known as stormwater best management practices (BMPs)] to control the quality of stormwater runoff in their jurisdictional areas. These interventions are organized into six categories, called Minimum Control Measures (MCMs, see Table 3.1), with the presumption that implementing the appropriate combination of interventions would reduce pollution from stormwater runoff to the maximum extent practicable in regulated urbanized areas (USEPA, 1999). The MS4s collate descriptions of their interventions into a single document called a Program Plan, which is submitted to the EPA designated stormwater authority (often the State environmental regulatory agency), where the MS4 is granted “compliance” status under a General Permit from the EPA. This permitting process varies in some instances; for a more detailed description, see USEPA (2014).

Table 3.1 - Six minimum control measures for stormwater management in Phase II MS4s from USEPA (1999)

Measure	Description
1	Public education and outreach on storm water impacts
2	Public involvement and participation
3	Illicit discharge detection and elimination
4	Construction site storm water runoff control
5	Post-construction storm water management in new development and re-development
6	Pollution prevention/good housekeeping for municipal operations

There have been two challenges to the Phase II rule regarding compliance. First, compliant status is predicated on the hypothesis that implementation by MS4 entities of the

proper combination of SCMs in sufficient quantity produces an improvement in water quality. However, the rule does not explicitly require water quality improvement by numeric criteria, obviating both the need for MS4 entities to monitor water quality, and the possibility of studying the longitudinal effects of the Phase II program (NRC, 2009; Wagner, 2005). The effects of some individual programmatic SCMs such as public education (Dietz et al., 2004), land cover controls (Horner et al., 2002), and regulations (Ahlman et al., 2005) have been tested, though the results are inconclusive, and studies do not cover the range of reported programmatic SCMs (A. C. Taylor and Fletcher, 2007). This, along with the long intervention-response lag times for SCMs (Bende-Michl and Hairsine, 2010) and the difficulty in measuring and reporting these data (Barrett, 2008) suggest that conditions for assessing the overall water quality effects of the Phase II Program are unfavorable at this time, and an attempt to do so is not included in this research.

The second challenge to the Phase II program, affirmed by a U.S. Court, is that the flexible nature of the MCMs, while allowing for management plans tailored to entity-specific conditions, also allows for a widely ranging level-of-effort that may still be considered compliant (EDC and NRDC v. USEPA, 2003). This variability in level-of-effort is not immediately evident, as Program Plans are normally submitted in a format (PDF documents) that is not conducive to analysis, also making it difficult to determine the salient factors driving this variability.

Several studies have attempted to organize Phase II response, and describe its variability using a variety of contextual factors. Galavotti et al. (2012) queried 471 regulated MS4 entities across the U.S. (249 Phase I and 222 Phase II) providing contextual factors such as a general description of the entities, percentage of SCM implementation, and some rationale behind their usage. Kaufman (1995) studied 46 local government stormwater plans as a function of several contextual factors, and found that mean household age, green staff index, total tax rate, and community type were significant determinants of overall stormwater plan quality. Similarly, Morison and Brown (2011) found that total population, recurrent income, median household income, percentage of population that has completed year 12 education, and an index of economic resources were significantly correlated to commitment to water sensitive urban design by municipal professionals around Melbourne, Australia. Finally, White and Boswell (2006), in a study of Phase II MS4 response in California and Kansas, found that contextual factors such as

percent high school education and median home value are significant determinants of Program Plan quality in both states.

The objective of this work is to inform future stormwater rulemaking and stimulate policy transfer by cataloging the stormwater interventions employed in response to Phase II of the NPDES policy in 90 small MS4s in the Commonwealth of Virginia, and by explaining the use of those SCMs using contextual factors. To that end, this paper presents a quantitative method to compare Phase II MS4 Program Plans based on the reported use of SCMs, and identifies patterns of SCM use as a function of MS4 characteristics and use efficiencies with other SCMs. These observations are synthesized to provide insight into how and why an SCM is used in certain MS4s but not others, in order to make recommendations to stormwater authorities concerning the merits of individual strategies, and the NPDES Phase II program as a whole. The two hypotheses tested are: (1) contextual factors affect the types of stormwater interventions a local government uses and (2) certain SCMs are complementary to others, as shown by the correlation of use between two or more SCMs across the studied MS4s.

3.3 METHODS

3.3.1 *Data Development*

To test these hypotheses, MS4 response to the Phase II policy and information regarding MS4 characteristics was collected and organized in a systematic fashion. The Program Plans describing stormwater management interventions for years 2013 - 2018 were used as descriptors of the response to the Phase II policy for each MS4. These plans were collected in PDF format from the Virginia Department of Environmental Quality (DEQ), the NPDES permitting authority in the Commonwealth. A comma separated values (CSV) file was created such that each MS4's Program Plan occupied a row, and for each SCM a new column was created and populated with a 1 or 0 based on the MS4's use (or not) of that SCM. This matrix of binary data was used as a numerical representation of MS4 response in Virginia, which was then supplemented by informal interviews with stormwater managers and permitting authorities to assure that the data and methods were consistent with implementation conditions.

Explanatory variables (i.e. contextual factors, Table 3.2) were chosen based on precedence in the literature (Kaufman, 1995; White and Boswell, 2006), and the authors' experience with stormwater management in municipalities. Factorial data was not available for

non-traditional MS4s (of a type other than city, county, or town), and as it was hypothesized that the type of MS4 was a determinant of SCMs used (test described below), no further effort to estimate characteristic information for non-traditional MS4s was performed.

Table 3.2 - Explanatory variables describing MS4 characteristics and their respective sources

Variable	Source	Category
Total Population (count)	(U.S. Census Bureau, 2010) P1	Cultural
Median Age (yr.)	(U.S. Census Bureau, 2013) 5 yr. S0101	Cultural
Population age 25 and higher finished grade 12 (%)	(U.S. Census Bureau, 2013) 5 yr. S1501	Cultural
Total Jurisdictional Area (km ²)	(U.S. Census Bureau, 2009)	Environmental
Population Density (person/km ²)	(U.S. Census Bureau, 2009)	Environmental
Total Annual Precipitation (mm)	(NOAA, 2001)	Environmental
Proportion Urbanized Area (%)	(U.S. Census Bureau, 2009)	Environmental
12 Month Median Household Income (\$)	(U.S. Census Bureau, 2013) 5 yr. S1901	Economic
Home Ownership (%)	(U.S. Census Bureau, 2013) 5 yr. B25008	Economic
Real Estate Tax (\$/\$ appraised)	(University of Virginia, 2008)	Economic
MS4 Type (categorical)	Program Plans	Governmental
Consultant Use (Y/N)	Program Plans	Other
Involvement in Stormwater or Environmental Organization (Y/N)	Program Plans	Other
Chesapeake Bay MS4 (Y/N)	(USGS, 2008)	Other

Several of the explanatory variables were found to be significantly cross-correlated (e.g. Total Jurisdictional Area and Proportion of Urbanized Area, Pearson’s $R = 0.86$), but since models were built by iteratively testing a single explanatory variable against a single response, all explanatory variables were left in the analysis and results were individually reported.

The binary response data from the MS4 Program Plans was joined using place names with the characteristic information in Table 3.2 to create a matrix of 90 rows (representing each MS4), several columns of miscellaneous information reported in the Program Plans, and 59 columns representing each reported SCM (listed in Figure 3.1) with either a “1” or “0” in each cell. One MS4 submitted a stormwater pollution prevention plan, a different type of regulatory document that could not be used, so the total number of MS4s analyzed was 89. The uncertainties associated with this method are: (1) it is problematic to assure that reported SCMs have actually been implemented, and if so, how their implementation varies between MS4s, and (2) there may be SCMs not reported that have been implemented. The former is a monitoring

problem as audits can only be performed intermittently. The latter is an informational gap between stormwater managers and the suite of interventions that could be used or reported. Either way, a known source of uncertainty in this work is the potential discrepancy between the actions reported in a Program Plan, and what has actually been implemented

1. Public Education and Outreach

- Stormwater (SW) website [70]
- Hosts public seminars
- Print media
- Stormwater signage
- Inlet markings
- Uses TV/Radio commercials/news segments
- Children/minority SW education
- Phone/text/email
- Pet owner education
- Publication for high-risk businesses and construction sites
- Hazardous waste publication [18]

2. Public Participation/Involvement

- Posts program plan online [49]
- Program plan available on request
- Trash/stream cleanup day
- Citizen advisory committee
- Volunteer based SW education
- Adopt-a-stream program
- Adopt-a-drain program
- Adopt-a-highway program
- Recycling/composting/vegetation pickup
- SW pollution awareness day
- SW/environmental public survey [3]

3. Illicit Discharge (ID) Detection and Elimination

- Notifies downstream interconnected MS4s
- Dry weather outfall inspections [62]
- ID ordinance
- ID response procedure
- ID tracking/recording
- Stormwater maps (any variety)
- Stormwater maps (in GIS)
- Identify/inspect high-risk businesses
- Inspector ID training
- Visual outfall monitoring
- CCTV ID detection [2]
- Online ID reporting mechanism

4. Construction Site Runoff Control

- Tracks/reports land disturbing activities
- Requires SW permit for all construction activities
- Requires ESC plan for land disturbance $\geq 2,000$ ft² [74]
- Provides employee ESC training
- Inspects all construction sites
- Construction site buffers [3]
- Design/inspection/enforcement policy for construction SCMs
- Requires environmental protection plan

5. Post-Construction Runoff Control

- Published SW design manual
- Ordinance for SCM maintenance
- Promotes LID at new construction sites
- Plan review for post-construction SCMs
- Encourage pre-development hydrology in plan review
- Tracks SW facilities
- Performs post-construction inspections [78]
- Literature review for innovative SCMs
- Hydrologic/hydraulic modeling [1]

6. Pollution Prevention/Good Housekeeping

- Pollution control plan for municipal operations
- Hazardous waste collection day/location
- Nutrient/pesticide/herbicide application plan
- Uses street sweepers
- Pet waste bags and receptacles [12]
- Employee training [70]
- Municipal property evaluations, pollution strategies
- Storm sewer and SCM cleaning and maintenance

*Note: The number in brackets represents the frequency that an SCM was reported in 89 Program Plans. The most and least frequently reported SCM in each Minimum Control Measure are shown.

Figure 3.1 - The 59 reported Stormwater Control Measures (SCMs) in Program Plans for 2013 – 2018 permit years*

A database of CSV files was created in the R Statistical Software Environment (R Core Development Team, 2014) as R allowed for simple manipulation, sub-setting, and visualization of data, and because the necessary statistical tests could be performed iteratively over a large number of observations (89 MS4s) and variables (59 reported SCMs plus several other outcomes). As shown in Table 3.2, the variables in this study were both continuous (e.g. Total Population of the MS4) and categorical (e.g. MS4 Type and SCM use), so three suitable statistical methods were selected that could provide analysis of potential associations in the data, and a test for significance of those associations. These tests were first performed at a 95% confidence level, though this appeared to result in Type I error such that many of the statistically significant correlations had no practical explanation. It was therefore determined that a more rigorous confidence level of 99% should be used to reduce the possibility of false positives, and the results of testing at the 95% confidence interval are only mentioned to emphasize the absence of an effect that was hypothesized to exist.

3.3.2 *Analysis of the MS4 Type*

The Phase II regulations govern eight disparate forms of urban jurisdictions, including cities, counties, towns, colleges/universities, federal/state facilities (e.g. state owned hospitals, non-military defense centers), local school boards, military installations, and transport systems (i.e. departments of transportation). During analysis of the Program Plans, it was clear that the type of MS4 was an indicator of stormwater capabilities and objectives, though unclear how the types of MS4s could best be grouped, and to what extent MS4 type explains SCM use. To resolve this, contingency tables were created with rows of MS4 type and columns of SCM use (or not), which could then be tested for statistical association. For this test, and subsequent analysis of contingency tables, Fisher's Exact Test was chosen as it is capable of testing tables of dimensions greater than 2 by 2, and is robust against low expected frequencies. The original grounds for this test is described by Fisher (1970), with the evolution of the method described in Routledge (2005) and its implementation as a software algorithm in Mehta and Patel (1986).

In order to determine the best grouping of MS4s, all 59 SCMs were tested for dependence on the original eight MS4 type classifications, and the count of significant relationships based on Fisher's Test at $\alpha = 0.01$ was recorded. This count indicated the number of SCMs that were affected by MS4 type. These eight classifications were then progressively clustered into five,

three, and two groups and retested for the count of significant group (row) by SCM (column) combinations. The clustering with the highest count (i.e. largest effect on SCM usage) was determined to be the optimal grouping, and the use of the affected SCMs was noted as being contingent on MS4 type as described by that grouping.

This testing provided two pieces of valuable insight. First, it revealed how best to cluster the types of MS4s, and second, it showed which SCMs were particular to certain types of MS4s, but not others. As directive for new lawmaking, this will be helpful in determining how appropriate compliance baselines could be set based on inherent MS4 differences.

3.3.3 *Analysis of Cultural, Environmental, and Economic Factors*

Within MS4 groups, it was hypothesized that the ten factors presented in Table 3.2 within the cultural, environmental, and economic categories determined the use of SCMs; however information describing these contextual factors was only available for the 40 cities, counties, and towns in this study. Therefore, the method described in this section could only be performed for a subset of the 89 total MS4s. As the response variable for this portion of the analysis is the use of an SCM (or not), but the explanatory variable is continuous, a form of regression where the response is bounded by 0 and 1 was needed. Logistic regression allows for this, as the response is reported on the logit scale such that:

$$\text{logit}(P_i) = \beta_0 + \beta_1 x \quad \text{Equation 3.1}$$

and:

$$\text{logit}(P_i) = \log[P_i/(1 - P_i)] \quad \text{Equation 3.2}$$

where P_i is the probability of SCM use as a function of the explanatory variable x (from Table 3.2), and estimated parameters β_0 (intercept) and β_1 (slope) (Pampel, 2000). The right side of Equation 3.2 is known as the log-odds, and represents the log of the probability that an SCM will be used divided by the probability that it will not be used. A unit change in x will change the log-odds that an SCM is used by the magnitude and direction of β_1 , therefore β_1 is reported along with model significance. These models were created using R's generalized linear model function with a binomial family (Venables and Ripley, 2002). The null hypothesis is that the explanatory variable (i.e. the contextual factors from Table 3.2) has zero effect (i.e. $\beta_1 = 0$ on the logit scale) on the use of a given SCM. This test was iteratively applied to all combinations of factors and

SCMs at 99% significance, and a table of the significant combinations was produced with the β_1 and p -values.

These models are observational, and provide insight into why an SCM might be used in one locality but not another. They are not predictive, however, and should not be used as input-output tools for decision making. Rather, the models can be used to draw conclusions about the likelihood of SCM use under certain conditions, which will provide the grounds for discussion on the mechanistic link between factor and SCM use.

3.3.4 *Analysis of Other Factors*

The outcomes of the preceding analyses indicated that other factors might be consequential in the selection of SCMs, therefore additional hypotheses were formed. The first of these hypotheses was that an MS4s involvement in a stormwater organization (e.g. an environmental non-profit, regional commission, or other group intended to foster collaboration) would have some effect on the SCMs it chose to use. There were two different approaches to testing this.

The objective of the first method was to test for dependence of each SCM on involvement in any stormwater organization. This was done by performing Fisher's Test on a contingency table of organizational involvement (or not) on the rows, and use (or not) of each of the 59 SCMs on the columns. For these tests, only the 40 municipal type MS4s were used to avoid confounding variables with MS4 type, as only 7 non-municipal MS4s were members of stormwater organizations.

The objective of the second method was to measure the similarity of Program Plans within each reported organization, and compare this to the similarity of Program Plans from MS4 entities not involved in any organization. First, the nature of the organization and the degree to which MS4 membership could affect SCM selection was researched qualitatively using information in Program Plans and the organization's website. This provided further perspective on the effects of involvement, as some organizations are paid by members to create stormwater education and outreach programs, while others are voluntary initiatives with no contractual obligations. The MS4s were then clustered into nine groups – eight based on the group they reportedly belong to, and one group of the MS4s that did not report any organizational involvement. The “No Organizational Involvement Group” was thought to represent a baseline

of similarity between MS4 entities with no formal affiliation with each other. Within each group, the Program Plans were first compared to each other based on the SCMs they are, or are not performing using Fisher's Test, and the number of significantly similar Program Plan pairs (i.e. Program Plans with a significant proportion of similar SCM use) was reported as a proportion of the total possible pairs in each group. The results of this analysis, along with the degree to which the Program Plans were compositionally (i.e. format and organization) similar to each other, and the qualitatively assessed nature of the group was used to describe the degree of SCM collaboration within a group.

It was also hypothesized that MS4s that hired consultants to create their Program Plans would employ a different variety of SCMs. This was tested in the same manner as the first method for testing involvement in a stormwater organization – contingency tables were created with use (or not) of a consultant on the rows and use (or not) of each SCM on the columns, and tested for dependence using Fisher's Test.

Finally, as there are new (as of 2013) specific requirements to fulfill the Total Maximum Daily Load (TMDL) Waste Load Allocations (WLAs) for MS4 entities that discharge to streams located in the Chesapeake Bay watershed, it was hypothesized that these Program Plans would implement SCMs that specifically fulfill these requirements more frequently than MS4s not subject to these requirements. Although many of the MS4s in Virginia discharge to streams with other WLAs, the special conditions for MS4s within the Chesapeake Bay watershed require a specific accounting of Nitrogen, Phosphorus, and Total Suspended Solid load reductions as a function of the pervious and impervious acreage within the regulated area and predetermined loading rates (Section I.C Virginia Department of Environmental Quality, 2013). This special condition also explicitly requires a description of the MS4's nutrient management plan, construction site runoff controls, and post-construction runoff controls for MS4s in the Chesapeake Bay. The effects of these requirements were also evaluated using Fisher's Test.

3.3.5 SCM Use Efficiencies

The final hypothesis was that the use of certain SCMs were interdependent with the use of other SCMs or groups of SCMs; that if two or more SCMs are used together across a statistically significant proportion of the entities, that there was some underlying commonality in the implementation of the SCMs. For example, if SCM "A" and SCM "B" are used together by

a significant proportion of the MS4 entities, there may be a common resource (or other factor) that, if identified, could inform the remaining entities that employ SCM “A” but not “B” (or inversely), of a resource efficiency that is not being leveraged. To test this hypothesis, an exploratory clustering algorithm was used initially to group the SCMs based on frequency of concurrent use, and then these groupings were confirmed using Fisher’s Test.

To create initial groupings, the *K*-means clustering algorithm was used in the R environment which requires an input matrix and a number of centers (or groups) by which to cluster the SCMs. This algorithm sorts the SCMs into groups of varying sizes to maximize between group sum of squares (BSS) and minimize within group sum of squares (WSS). The basis for *K*-means clustering is found in Hartigan (1975), and its implementation as a computer algorithm in Hartigan and Wong (1979). As the number of groups, *n*, could be any value between two and 58, the selection of this parameter required some discretion. Initially the algorithm was run across the range of possible groups ($2 \leq n \leq 58$), and the resultant WSS/BSS was plotted against *n* (known as a scree plot). Typically, the scree plot will reveal a break in the plot at some value of *n*, indicating that further grouping will have an inconsequential improvement on the quality of the groups, but the scree plot for the SCMs did not reveal any such break. Therefore 20 groups ($n = 20$) was subjectively chosen on the basis that the average group size would be approximately 3 SCMs. This decision was confirmed by creating a function that tested the significance of the usage of each SCM paired with all other SCMs in the same group based on Fisher’s Test, and produced an output matrix of all *p*-values. In this way, the results of the *K*-means clustering algorithm could be evaluated at a range of values of *n*. This provided a confirmation of the groupings produced by the *K*-means clustering.

3.4 RESULTS AND DISCUSSION

3.4.1 *Value of a Quantitative SCM Database*

The basis for this research was the compliance database that was developed by parsing the 89 Program Plans into numerical values, allowing for quantitative analysis; an impractical task with SCMs reported in 89 separate PDF documents. The 59 SCMs shown in Figure 3.1, and their respective use frequencies, suggest a spectrum of intervention difficulty and effectiveness, but are all considered appropriate responses to the Phase II policy. Creation and sharing of this database might allow MS4 operators to learn about the interventions that other operators in their

state or region are employing to improve their own stormwater programs as the breadth of strategies reported was large, and strategy sharing was not highly reported.

Strategy sharing might be especially useful if operators were able to identify other MS4s with similar geographic and demographic characteristics as their own. This is supported by findings in Dolowitz et al. (2012) that although searching for stormwater innovations is occurring among local governments, this search is not rational and comprehensive, and that stormwater managers were unlikely to adopt strategies from municipalities of a different size than their own. As the range of total population for Phase II MS4s in Virginia (where census data was available) was 4,961 to 312,300, and the total jurisdictional area was 6.7 to 1,887.3 km², a repository that would allow a stormwater manager to peruse SCM use filtered by these sorts of metrics could augment collaboration between characteristically similar MS4s.

The population range also elicited some notable information about the Phase II program, as five Virginia Phase II MS4s serve populations greater than 100,000, the population threshold for Phase I of the NPDES stormwater program. Two of these MS4s were below this threshold as of the 1990 Decennial Census, the static population benchmark for Phase I designation (40 CFR §122.26), and although the populations of these localities now exceed this threshold, they will continue to be regulated under the Phase II Final Rule. As the population of US urban centers continues to increase (Alig et al., 2004), no further entities will be designated as Phase I MS4s, meaning that the Phase II Rule will become the de facto stormwater regulation for all urbanized areas (except those designated as Phase I in 1990) as it uses the most recent Decennial Census for designation. Two other MS4s have a significant proportion of their population served by a combined sewer, and were allowed to petition to have this population deducted from their total population, resulting in an MS4 service population below the threshold. Subsequently, designation for these MS4s was withheld until the Phase II Program began in 1999. The fifth Phase II above the population threshold was exempted in 2013 based on the determination that the extent of their storm sewer infrastructure was several discrete pipes draining government buildings; no “system” actually exists (J. Bauer, DEQ, personal communication, 2015). Nevertheless in the Program Plan submittal, the authorities of this entity insist that they will continue to fulfill permit requirements, and they are granted compliance; as such, their Program is included in this analysis.

A quantitative SCM database could also provide a permitting authority with the necessary information to create a baseline for granting or rejecting compliance status. In this study, all 90 MS4s had been granted compliance status, including the one MS4 that did not submit the correct document, and the total number of SCMs reported from MS4s that submitted Program Plans ranged from 6 – 35. The requirement for compliance is that an MS4 report yearly progress towards the interventions reported in their Program Plan, but there is no explicit definition of a “compliant” Program Plan. Therefore, MS4s can present a range of responses and still be within the regulatory framework. Although it may not be reasonable to create a single definition of compliance for all Phase II MS4s, a dynamic definition could be created as a function of MS4 characteristics – notably MS4 type – in order to maintain accountability across the variation of regulated entities.

Another ancillary benefit provided by an accounting of reported stormwater management activities is that it demonstrates the lack of use of measures that might reasonably be expected as outcomes of a policy created to improve water quality – namely water quality monitoring and modeling. The data confirms the National Research Council's (2009) assertion that monitoring and modeling “...might be the two weakest areas of the storm-water program...” as only one of 90 Program Plans report any variety of hydrologic/hydraulic modeling, and none report the measurement of any water quality or quantity parameters. Based on informal interviews, at least one MS4 is monitoring discharge in their storm sewers, though it was not reported in the Program Plan as monitoring is not required of MS4 entities.

A complete repository of Phase II responses is not available for the Commonwealth of Virginia, and at the national level, is only available as a summary of survey results in Galavotti et al. (2012) – though the EPA has proposed to move Phase II reporting to an electronic system (USEPA, 2013). The results of this study show the value that a national electronic reporting system could have if the system stored response data in a format that provides for accountability and analytical review.

3.4.2 The Effect of MS4 Type

The results of clustering MS4 types into progressively fewer groups provides direction for compliance monitoring and insight into how the MS4 types are functionally different. The eight original classifications of MS4 type were clustered into five, three, and two groups and

tested for the count of significant group and SCM combinations. The clustering with the largest effect on SCM usage was determined to be the optimal grouping. The SCMs that were dependent on MS4 type are reported in Figure 3.2, with the optimal clustering of MS4s into three groups: (1) municipal type MS4s including cities, counties, and towns; (2) higher education MS4s including colleges and universities; and (3) other MS4s including federal/state facilities, local school boards, military installations, and transport systems. Under this clustering scheme, 21 of the 59 total SCMs were found to be significantly dependent on MS4 type; grouping “colleges/universities” with “other” resulted in 20 significant SCMs. This suggests that the different types of MS4s may need to be separately regulated, as 21 SCMs are particular to certain types of MS4s but not others. This is further exemplified by an analysis of the affected SCMs.

Key				
90	Percentage of MS4s in Cluster Reporting SCM Use			
MCM	Stormwater Control Measure	Municipal (n = 40)	College/University (21)	Other (29)
1	Stormwater website	90	90	52
	Print media	93	67	62
	Uses TV/Radio commercials/news segments	50	14	7
	Children/minority SW education	65	10	21
	Pet owner education	45	5	0
	Publication for high-risk businesses and construction sites	53	19	14
2	Posts Program Plan online	78	52	24
	Adopt-a-stream program	23	5	0
	Recycling/composting/vegetation pickup	25	48	7
3	Illicit Discharge (ID) ordinance	88	38	38
	ID response procedure	80	38	31
	Stormwater maps (in GIS)	63	19	14
4	Provides employee ESC training	60	29	24
	Design/inspection/enforcement policy for construction SCMs	78	48	45
5	Published SW design manual	38	10	0
	Ordinance for SCM maintenance	85	67	48
6	Hazardous waste collection day/location	43	10	3
	Nutrient/pesticide/herbicide application plan	35	76	41
	Uses street sweepers	63	29	10
	Pet waste bags and receptacles	28	5	0
Other	Member of environmental organization	75	19	10

Figure 3.2 - SCMs used at a significantly different frequency across MS4 Types. The “Municipal” cluster includes cities (n = 21), counties (10), and towns (9), and the “Other” cluster includes federal/state facilities (18), local school boards (2), military installations (8), and a transport system (1). The number in each cell is the percentage of use within categories, also indicated by the cell shading.

Of the 21 SCMs contingent on MS4 type, 16 were performed by municipal type MS4s at a frequency that was significantly greater than College/University MS4s and Other MS4s. This suggests that the strategies customary to Phase II’s Minimum Control Measure format are more relevant to municipal MS4s than to non-traditional types, an effect that can also be seen in the disparity in total SCMs reported in Figure 3.3. The EPA addresses this problem to some degree in their MS4 Permit Improvement Guide (USEPA, 2010a) by suggesting that permit writers (i.e. States) account for this in the creation of their General Permits, although Virginia’s General

Permit has only a single caveat for non-traditional MS4s – the definition of “public” for public education and outreach (Virginia Department of Environmental Quality, 2013). While this is a helpful distinction to make, the disparity in SCM use across MS4 types suggests that a tiered but objective compliance baseline would allow for variation in institutional capacity across MS4 types, while maintaining a level of accountability for each MS4 type. For example, the results from Fisher’s Test show that municipal MS4s are considerably more likely to create a policy or ordinance that prohibits illicit discharges, publish a stormwater management design manual, and write an ordinance providing a mechanism for stormwater treatment facility maintenance. Non-municipal MS4s generally do not have the legal authority to create or enforce these policies and as such, it is impractical to require these regulation-based management techniques from non-traditional MS4s.

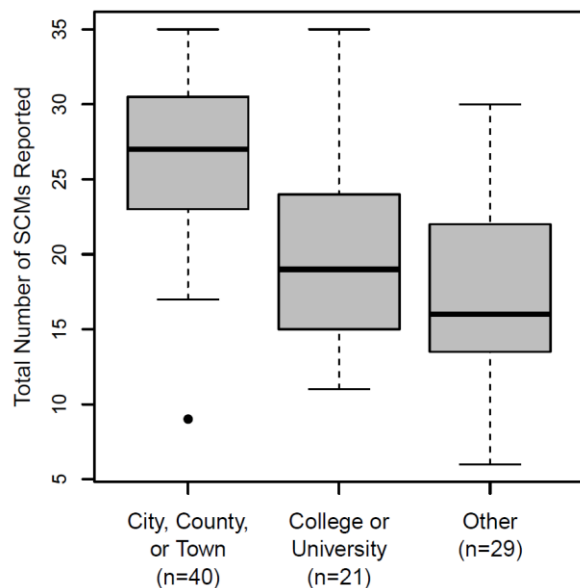


Figure 3.3 - Tukey Box and Whisker plot demonstrating total number of SCMs reported, categorized by MS4 type. The “Other” category includes federal/state facilities, local school boards, military installations, and a transport system

The results also showed that the creation of pet waste owner education programs, pet waste disposal stations, development of stormwater maps in GIS, and use of street sweepers were, by proportion, significantly less likely to be used by non-traditional MS4s. Educational programs (mostly signs, websites, and other advertising) informing pet owners of the stormwater impacts of pet waste could be used in certain non-municipal MS4s where a public audience could reasonably be reached – namely colleges and universities – but only one of the 21 colleges and universities in this study report the use of pet waste education. The same is true of the use of

pet waste disposal stations. These strategies are not relevant for all non-traditional MS4s (e.g. local school boards, hospitals, and departments of transportation), but could add value to Programs in MS4s with potential pet waste sources and target audiences.

The use of GIS to map stormwater infrastructure appears to be an SCM particular to municipal MS4s, as only 8 of 50 non-traditional MS4s use this SCM. It is not known if municipal GIS stormwater maps include infrastructure owned by other adjacent MS4 entities, or if mapping ends at jurisdictional boundaries, though informal interviews suggest that local government data is generally not continuous across political boundaries. If a watershed drains multiple MS4 entities, as they commonly do, it is only possible to completely understand the drainage system if storm sewer data from the different entities is available, and can be synthesized. This integration of data will likely require additional effort, as 28 of 50 non-traditional MS4s map stormwater infrastructure using a method other than GIS, and 14 do not report mapping at all.

Street sweeper use was reported by 9 of 50 non-traditional, and 25 of 40 municipal MS4s; a significant proportional difference. High capital cost, small relative size, and the need for dedicated street sweeper personnel are three possible reasons why non-traditional MS4s do not use this practice, though contracts or memoranda-of-understanding (MOU) between authorities may allow for some cost and resource sharing. These arrangements are impractical for certain entities, such as military installations, but could provide increased use efficiencies for municipal MS4s that already own sweepers, and the opportunity to use them in MS4 jurisdictions that would not otherwise be able to do so.

The disproportionate use of pet waste education and disposal stations, GIS stormwater mapping, and street sweeping in municipal MS4s over adjacent non-municipal MS4s, demonstrates the inconsistency between the need for hydrologic and water quality planning based on environmental boundaries such as a watershed divide (Ruhl, 1999), and MS4 operational capabilities constrained by jurisdictional boundaries (e.g. city or county limit). This inconsistency is known to produce negative economic and environmental externalities (Patterson et al., 2012; Roy et al., 2008), but may be ameliorated with template contracts and MOUs approved by the permitting authority and distributed with the General Permit.

Stormwater organizations could also foster this sort of collaboration (see section below), but the results indicate that non-traditional MS4s are significantly less likely to be involved in

these groups; only 7 of 50 non-traditional MS4s studied are involved in organizations, as compared to 30 of 40 municipal MS4s. This asymmetry may be an artifact of whether or not an entity has staff (or even a single staff member) dedicated specifically to stormwater management, as Program Plans from traditional MS4s were signed by engineers, public works directors, stormwater division chiefs, environmental administrators, and city/zoning/planning administrators, while non-traditional MS4s authorized all varieties of facilities managers, grounds supervisors, vice presidents, and environmental administrators to sign the Program Plans. The absence of non-traditional MS4s from stormwater organizations may therefore be attributed to the lack of a dedicated stormwater personnel, and while involvement in an organization does not conclusively mean that strategy sharing will occur, operators of non-traditional MS4s may not be leveraging the resources of their municipal MS4 neighbors simply because of a lack of inter-jurisdictional communication.

3.4.3 The Effect of Stormwater Organizations, Consultants, and the Chesapeake Bay TMDL Requirements

As only 7 non-traditional MS4s were members of stormwater organizations, only the 40 municipal MS4s were used as observations for testing the importance of stormwater organizations in order to reduce confounding variable effects with MS4 type. At 99% confidence, none of the SCMs were affected by involvement in an organization, and at 95% confidence, only three were affected: MS4s that were members of one or more stormwater organizations were more likely to post the Program Plan online, require a stormwater permit for all construction activities, and create a design, inspection, and enforcement policy for construction site control measures.

Although the involvement of MS4s in stormwater organizations had little effect on the selection of SCMs, the results showed that Program Plan similarity exists within organizations above the baseline level of similarity. MS4s reporting involvement in the same organization (and those not involved in any organization) were compared in a pairwise fashion, using Fisher's Test for each pair to determine if there was significant similarity between the SCMs reported in the two Program Plans. The number of significantly similar pairings (i.e. statistically indistinguishable Program Plans based on common reporting of SCMs) in each group, normalized by the total pairs in each group, represented the level of Program Plan similarity in a group. 24% of 55 pairs of municipal MS4s not participating in any group had significantly

similar Program Plans at $\alpha = 0.01$. This represents the baseline percentage of Programs Plans that are similar without any explicit inter-MS4 coordination. The level of similarity in each organization (i.e. the percentage of statistically similar program plans) is shown with respect to the baseline in Figure 3.4 with the relative size of the bubbles representing the size of each organization based on the number of members reported, n. Six of eight stormwater organizations had a higher proportion of statistically similar Program Plans above the group of non-associated MS4s, although three of the eight groups only had two members and therefore only one pairing (two of these three were below the baseline).

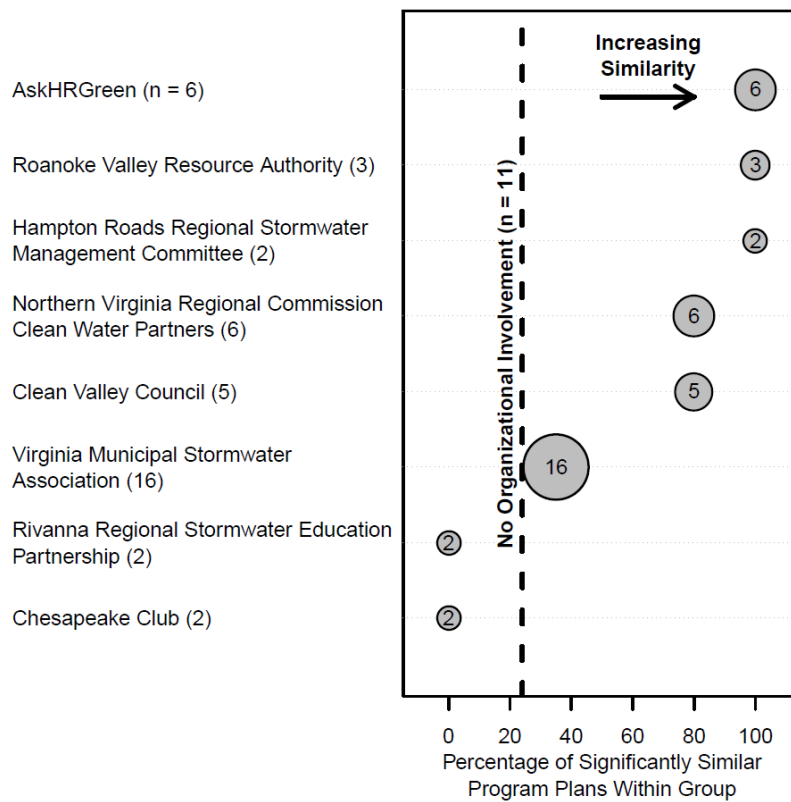


Figure 3.4 - Similarity of Municipal MS4 Program Plans in reported stormwater organizations with bubble size representing relative size of each group, and number of reported membership (n) in black within each bubble.

In order to relate statistical similarity to level of collaboration, qualitative observations about the composition of similar Program Plans, and the organizations themselves were made. Members of AskHRGreen had Program Plans that were statistically similar ($p < 2 \times 10^{-5}$), and all used the same tabular template to report SCMs, organized by MCM with columns for SCM Number, SCM Description, Measureable Goals, Metric, Responsible Party, Timeline, and Associated Documents. Most other MS4s organized SCMs into an outline format, and described

each action in a paragraph or sentence. AskHRGreen is a voluntary association that promotes general citizen environmental stewardship, and is an initiative of the Hampton Roads Planning District Commission created to “facilitate local government cooperation and state-local cooperation in addressing on a regional basis problems of greater than local significance” (“Regional Cooperation Act,” 1968). Members of the Clean Valley Council (CVC) described their SCMs in an outline format, and though less obviously similar than the AskHRGreen Program Plans, similarities in this organization did appear collaborative, as the CVC is contracted by its member MS4s to plan, implement, and document activities toward the fulfillment of MCMs 1 and 2. There were notable compositional similarities in the Program Plans from the Northern Virginia Regional Commission Clean Water Partner’s (another organization created under the “Regional Cooperation Act,” 1968) , notably in the four members that employed the same consultant to formulate their plans. Compositional similarities in the members of the Virginia Municipal Stormwater Association (VAMSA) were not obvious, perhaps because the geographical breadth of this organization is much larger than other organizations reported, suggesting that collaboration may also be a function of factors such as geographic proximity. It should also be noted that all reported members of the Roanoke Valley Resource Authority (actually a waste disposal authority) were also members of the Clean Valley Council, and all reported members of the Hampton Roads Regional Stormwater Management Committee were also members of AskHRGreen, as such, these subset organizations are not discussed further.

The test to determine effects of hiring a stormwater consultant showed that one SCM - construction site inspection - is significantly less likely for MS4s that hire consultants as opposed to those that do not at 99% confidence. Three other SCMs were found to be more likely to be used by MS4s with consultant help at 95% confidence, including the publication of a stormwater design manual, the distribution of informational material to businesses at high risk of illicit discharge, and identifying/inspecting those high risk businesses. These results agree with other findings that organizational membership, and the use of a consultant does not necessarily have an outcome on SCM selection (White and Boswell, 2006). However, it should be noted that three MS4s that hired the same consultant had statistically similar Program Plans, though it was not possible to determine if this was an effect of the consultant use, or their membership in a stormwater organization.

The hypothesis that MS4 entities within the Chesapeake Bay watershed would use certain SCMs tailored to the special requirements of the TMDL was rejected for all SCMs at 99% confidence. Nutrient management plans, a special requirement for Chesapeake Bay watershed MS4s for example, was reported by 4 of 11 non-Chesapeake Bay MS4s, and 22 of 78 MS4s within the Bay watershed – an insignificant proportional difference. As the special conditions require that the MS4 operator estimate the total existing load based on urban pervious and impervious area, it was hypothesized that Chesapeake Bay operators would use GIS for stormwater mapping, though this did not prove to be so at $\alpha = 0.01$ (it was the only SCM to fail rejection at $\alpha = 0.05$, however). It should be noted that these permit requirements are new, and while they do not appear to have an effect on Program Plans in this study, this may be due to the lag time between policy implementation and entity response.

3.4.4 Cultural, Environmental, Economic Factors

The results of testing SCM use as a function of the characteristic differences between the 40 municipal MS4s revealed that four of the 59 SCMs were dependent on one or more of the ten cultural, environmental, and economic (contextual) factors described in Table 3.2. The logistic regression models developed for this portion of the research were all bivariate in nature, with the explanatory variable (the ten contextual factors) hypothesized to determine the outcome variable (use or no-use of each of the 59 SCMs) at $\alpha = 0.01$. Seven combinations of factors and SCMs (out of a total of 590 possible combinations) were significant, and the effect of these characteristics is shown in Figure 3.5, with the sign in each cell representing the direction of each factor's effect on the use of the reported SCM (β_i), and the cell shading representing level of significance.

Key											
	(-)	p < 0.01, Factor and SCM are inversely related									
	(-)	p < 0.05, Factor and SCM are inversely related									
	(+)	p < 0.01, Factor and SCM are directly related									
	(+)	p < 0.05, Factor and SCM are directly related									
MCM	Stormwater Control Measure	Median Age (yr.)	Population Age > 25 finished grade 12 (%)	Total Population (count)	12 Month Median Household Income (\$)	Home Ownership (%)	Real Estate Tax (\$/\$ appraised)	Total Annual Precipitation (mm)	Total Jurisdictional Area (km2)	Proportion Urbanized Area (%)	Population Density (person/km2)
1	Stormwater website		(+)								
	Hosts public seminars				(-)	(-)					
	Uses TV/radio commercials/news segments				(+)						
	Children/minority SW education		(-)					(-)			
	Publications for high-risk businesses and construction sites	(+)						(+)			
2	Hosts trash/stream cleanup day		(-)		(-)						
	Volunteer-based SW education				(+)						
	Has recycling, composting, or vegetation pick-up program	(-)				(-)	(-)				
	Sponsors SW pollution awareness day		(+)								
3	Notifies downstream interconnected MS4s							(+)		(-)	
	Illicit discharge tracking/reporting							(+)			
	Online ID reporting mechanism									(+)	(+)
4	Requires ESC plan for land disturbing activities >= 2,000 ft ²				(-)						
	Provides employee ESC training										(-)
	Inspects all construction sites				(-)						(-)
5	Published a SW design manual			(+)							
	Plan review for post-construction SCMs	(-)			(-)	(-)		(-)			
	Performs post-construction inspections								(-)		
6	Uses street sweepers*	(-)	(-)		(-)				(-)	(+)	(+)
	Specifically created a Nutrient Management Plan										(+)
Other	Is located in the Chesapeake Bay Watershed				(+)						
	Used a consultant to help prepare their Program Plan										(+)

*With counties excluded from the analysis, none of these relationships were significant at 99%, though Total Jurisdictional Area, and Proportion Urbanized Area were still significant at 95%.

Figure 3.5 - Significant logistic correlations between factors and frequency of SCM use in Municipal MS4s at 95 and 99% confidence. The sign in parentheses represents the direction of the effect (β1) based on the estimated log-odds

The use of street sweepers was found to be significantly associated with the total area of an MS4, the density of urbanization and population, and percent home ownership. It was found, however, that in Virginia, the Department of Transportation (VDOT) owns and maintains

primary and secondary road right-of-way outside of incorporated places, therefore counties are not responsible for street sweeping. Only one of the ten counties reported using street sweepers, and their Program Plan states, “[The] County, with the cooperation of VDOT, has initiated a program to clean local roadways.” The effects on MS4 density and jurisdictional area (as well as the other factors) on the frequency of street sweeper use was re-evaluated with counties excluded, and the relationships were no longer significant at 99% confidence.

The use of seminars as a medium for education on the impacts of stormwater runoff was found to decrease in frequency as community 12 month income increased; a result that is counter-intuitive, but deserving of consideration given the low probability of Type II error. Although wealthier MS4s reported stormwater education programs less frequently, percentage of high school education did not affect the frequency of use of this SCM. Community wealth could not be related to more frequent reporting of any other SCM at 99%, contrasting Morison and Brown's (2011) findings that median household income increased commitment to innovative stormwater management. Further research is needed to explore the relationship between community wealth and the use of high-capital structural SCMs as an alternative to programmatic SCMs, as MS4s are not required to include an enumeration of structural stormwater treatment facilities in their Program Plans.

The frequency that children and minorities were targeted for stormwater education decreased significantly as the percentage of high school graduates in a community increased. This, along with the lack of correlation between this variable and the frequency of use of any other SCM at 99% was also a result contrary to the intuitive hypothesis that communities with higher educational attainment are more committed to high quality stormwater management programs (Morison and Brown, 2011; White and Boswell, 2006).

Frequency of use of an online system for reporting illicit discharges increased significantly with population density, though no other SCMs were affected by this factor at 99% confidence. This SCM may be more useful in densely urbanized areas, where the possibility of noticing and reporting the dumping of leaves, grass clippings, motor oil, and other illicit discharges is more salient.

Aside from the three statistically significant SCM – factor correlations noted, it does not appear that the SCMs employed by stormwater managers pursuant to the requirements of the MCM format are representative of the cultural, environmental, and economic characteristics of

their entities. The anecdotal assertion that wealthier, more populous, and educated communities produce high quality stormwater compliance was not supported by the data, as these communities largely report the same SCMs as their poorer, less populous and less educated counterparts.

3.4.5 *SCM Use Efficiencies*

Results from *K*-means clustering of 59 SCMs into 20 groups of sizes ranging from 1 – 13 provided two outcomes for discussion: (1) certain SCMs occurred together frequently and (2) certain SCMs could not be grouped with any other SCM because they were not frequently used with any other SCM. The first outcome suggests that performing all the SCMs within a group takes advantage of some efficiency, while the second outcome suggests exclusivity – that the resources needed for a strategy are unique, and cannot be easily shared to employ any other SCMs.

For example, construction site inspection and post-construction inspection were grouped together as 63 of 89 MS4s performed both actions, but dry weather outfall inspections was not performed synonymously with the other two (or any other SCM, for that matter). While outfall inspection may require separate knowledge and/or a higher level of effort than the other two, an adequate training program could prepare a single inspector or group of inspectors for all three duties. Training programs for construction site or illicit discharge inspectors were reported in 18 MS4s, but this training was not significantly correlated with construction, post-construction, illicit discharge, or high risk business inspection, though these SCMs were used in 70, 77, 62, and 19 MS4s respectively. Inspector training for construction and post-construction sites is available in the Commonwealth of Virginia, though training for illicit discharge and high risk business inspectors is not. Operators may find that preparation for the latter types of inspections is ancillary to the former training with only a marginal increase in effort. As it stands, inspections appear to take a high priority for operators, but training for inspectors does not.

The results of testing for SCM efficiency revisits the use of seminars on stormwater education, as this SCM was unique in its exclusivity; it did not consistently appear in Program Plans along with any other SCM. The primary resource required to host a seminar on stormwater education – an educator – may not be employable in any other reported SCM in this study except for special education programs for children and minorities. Fifty of the MS4s

perform one of the two of these SCMs, but only 13 perform both. This may be an oversight of a potential efficiency for an MS4 operator. Similarly, the hosting of trash or stream clean-up days was found to occur exclusive of all other SCMs as it may be contingent on the availability of a coordinator. While MS4s that are members of stormwater organizations would seem to have this resource, this SCM is not associated with organizational membership.

The creation of protocols for the handling of potential stormwater pollutants such as fertilizer and pesticides was hypothesized to occur with other policy controls such as the authorship of a hazardous waste control plan, ordinance for stormwater facility enforcement, stormwater design manual, construction site stormwater facility policy or illicit discharge policy. This hypothesis was not found to be true for Virginia MS4s, although several of these SCMs did appear together in Program Plans. As it was noted that some of these policy controls were so frequently used together but others were not, the use of these controls was tested for association with MS4 type. Six of these seven SCMs were contingent on the type of MS4, further reinforcing the variability in legal authority and subsequent authorship of stormwater related policies for different types of MS4s.

Several SCMs that require special infrastructure, including the use of phone, text, or e-mail to send stormwater information, and the creation of an online reporting mechanism for illicit discharges, were also found to be exclusive of all other SCMs. An MS4 would need an online electronic database system capable of interacting with both municipal employees and the general public. An SCM in this study that was hypothesized to take this role was the use of a stormwater GIS, though this SCM was not consistently used with either of the two SCMs mentioned.

The use of a stormwater GIS was also hypothesized to correlate with other actions such as illicit discharge tracking, the tracking of land disturbances, stormwater facility tracking, and hydrologic/hydraulic modeling. Although GIS is a tool that could facilitate the completion of these additional SCMs, within the 33 MS4s that reported the use of a GIS, none of these other SCMs were commonly performed. This suggests that the use of GIS in Virginia Phase II MS4s is primarily a means for visualizing underground stormwater infrastructure, a requirement of the Virginia General Permit (Virginia Department of Environmental Quality, 2013). This is a notable underutilization of a tool that could enable both stormwater and infrastructure management for stormwater operators (Aguilar and Dymond, 2014).

Finally, SCM concurrent use results showed that MS4s that used street sweepers also frequently performed storm sewer and structural BMP cleaning and maintenance, and implemented recycling, composting, or vegetation pickup programs. The use of these SCMs is of special interest because of the large amount of capital required to employ them, above the rest of the SCMs in this study. Their use describes a spectrum of ability to intervene using investment-heavy strategies, across which the Virginia MS4s lie. Thirteen MS4s performed all three of these SCMs: five cities, five colleges/universities, two towns, and a military installation. Counties did not commonly use any of these three SCMs, as only one county (of ten in this study) performed two of the three, and five performed only one. Although these SCMs could not unilaterally be correlated with urbanization and population density using logistic regression, as this data was only available for municipal MS4s, it appears that the use of these resource-intensive SCMs are subject to these densities via the proxy variable of MS4 type. This variable describes both environmental context and institutional capacity and it may be a subject of further research to determine how these two factors affect the use of the three resource-intensive SCMs mentioned here, as well as other SCMs of a structural nature.

3.5 CONCLUSIONS

Based on the results from the analysis of Virginia Phase II MS4s, the following conclusions and recommendations are made for policy makers and stormwater authorities:

- A quantitative response database as described in this research provides a forum for MS4 operators to look up the SCMs that other operators of characteristically similar entities are using, and provides a baseline for permitting authorities to grant or reject compliance status. Authorities could create a compliance baseline tailored to MS4 type or other factors that significantly explain Program Plan response. This supports the proposal for electronic reporting for Phase II MS4s, so long as the input to the electronic reports from stormwater operators is housed in a database that can be analyzed.
- MS4s reported between 6 and 35 SCMs in their Program Plans. No MS4s reported water quality or quantity monitoring, and only one reported hydrologic/hydraulic modeling, as these are not required components of the Phase II rule.
- MS4 type had a significant effect on the suite of SCMs that an operator chose to employ. When MS4s were grouped based on their type into: (1) cities/counties/towns, (2)

colleges/universities, and (3) federal/state facilities/local school boards/military installations/transport systems, 21 of 59 SCMs were used at significantly different frequencies between these groups. Most of these 21 SCMs were reported more frequently by municipal MS4s than the other two types, as they are not appropriate for the community, authority, or objectives of non-traditional MS4s. This should not preclude some sort of compliance baseline for non-traditional MS4s.

- Based on the results of the statistical testing, a qualitative comparison of the composition of the Program Plan documents, and a consideration of the scope of stormwater organization's authority and subsequent influence on Program Plan development, stormwater organization involvement increases the concurrent use of SCMs between MS4s above what normally occurs. This is presented in two notable fashions: (1) MS4s shared a Program Plan template and (2) several adjacent MS4s contracted an external entity to complete MCMs 1 and 2. This shows that collaboration between local entities provides some operational efficiencies for SCM implementation.
- Cooperation between adjacent regulated entities of different types does not occur. Where appropriate, strategy sharing could be facilitated by encouraging inter-organizational collaboration through organizations, or by creating boilerplate Memoranda of Understanding that allow for the deployment of programmatic SCMs across jurisdictional boundaries.
- Chesapeake Bay MS4s did not report any SCMs at a significantly different frequency than those not in the Chesapeake Bay, in spite of the special requirements of these MS4s in Virginia's General Permit, though it is possible that this is due to the time lag between permit implementation and MS4 response.
- MS4s that hired consultants to assist in Program Plan writing used construction site inspectors significantly less frequently, but otherwise used the same combinations of SCMs as MS4s that did not hire consultants. Although consultants did not affect the selection of SCMs, three MS4s that hired the same consultant had statistically similar Program Plans, suggesting that hiring a consultant provides some operational efficiency.
- Street sweepers were employed by 24 of 30 cities and towns, but by only 1 of 10 counties, because in Virginia, VDOT has maintenance authority for primary and secondary roads outside of incorporated areas. The one county that uses street sweepers created an agreement

with VDOT, demonstrating that inter-organizational cooperation could allow additional opportunities for SCMs.

- Municipal type MS4s with higher income were less likely to use stormwater education seminars, though the use of this SCM was not explained by educational attainment. Testing for concurrent use also showed that stormwater seminars were exclusive of all other SCMs, including similar educational programs, a possible oversight of resource use efficiency.
- Construction, post-construction, illicit discharge, and high risk business inspections were commonly used SCMs in Virginia, but were infrequently used in combination with each other. MS4 managers may be missing an operational efficiency, by employing some types of inspection but not others. Furthermore, although inspection was a frequently reported activity, training for inspectors was not, suggesting that inspectors may not be prepared to effectively evaluate a feature with respect to stormwater quality.
- The use of a trash or stream clean-up day; dissemination of stormwater information by phone, text, or e-mail; and the creation of an online reporting mechanism for illicit discharges were all SCMs that were exclusive to all other SCMs in this study. The availability of the necessary resources for these SCMs, ancillary to other interventions, did not increase the likelihood that these measures would be used, further describing efficiency losses in MS4 administration.
- 33 of 90 Virginia MS4s report the use of GIS in their Program Plans, though SCMs that could leverage the capabilities of GIS are not synonymously reported, suggesting an underutilization of this technology.
- The use of street sweepers; storm sewer and structural BMP maintenance; and recycling, composting, or vegetation pickup programs occurred frequently with each other, suggesting commonality in the MS4s that employed these SCMs. The common factor for these three resource-intensive SCMs appeared to be whether or not the entity was a Municipal MS4, as a proxy variable for both institutional capacity and environmental context, though it is unclear which of these is the stronger explanatory variable – a potential area for further research.

These conclusions synthesize the results of the evaluation of the NPDES Phase II stormwater final rule for MS4s in Virginia. Statistical hypothesis testing led to the conclusion that some contextual factors significantly affect the SCMs that MS4s use, although other factors such as

SCM cost should be evaluated in future research. Results also showed that that while concurrent use exists in the SCMs that operators choose, there are some notable missed resource efficiencies that could add value to stormwater programs with marginal effort. The extent to which the reported SCMs end in a positive outcome for water quality and hydrology is a matter for further research, as this will be the terminal measure of success of the Phase II rule. Until intervention-response data becomes more widely available (especially for non-structural interventions), alternative methods of regulatory evaluation will continue to be useful for developing guidance documents for permittees, and for informing the process of future stormwater policy making.

4. BENCHMARKING LABORATORY OBSERVATION UNCERTAINTY FOR IN-PIPE STORM SEWER DISCHARGE MEASUREMENTS

4.1 ABSTRACT

The uncertainty associated with discharge measurement in storm sewer systems is of fundamental importance for hydrologic/hydraulic model calibration and pollutant load estimation, although it is difficult to determine as field benchmarks are generally impractical. This study benchmarks discharge uncertainty in several commonly used sensors by laboratory flume testing with and without a woody debris model. The sensors are then installed in a field location where laboratory benchmarked uncertainty is applied to field measurements. Combined depth and velocity uncertainty from the laboratory ranged from $\pm 0.207 - 0.710$ in., and $\pm 0.176 - 0.631$ fps respectively, and when propagated and applied to discharge estimation in the field, resulted in field discharge uncertainties of between 13% – 256% of the observation. Average daily volume calculation based on these observations had uncertainties of between 58 and 99% of the estimated value, and the uncertainty bounds of storm flow volume and peak flow for nine storm events constituted between 31 - 84%, and 13 - 48% of the estimated value respectively. Subsequently, the implications of these observational uncertainties for stormwater best-management practice evaluation, hydrologic modeling, and Total Maximum Daily Load development are considered.

4.2 INTRODUCTION

The value of accurate discharge measurements in urban storm sewer systems was first recognized as rudimentary flood gaging stations appeared in, and upstream of urban areas in the 1970's (H. J. Owen, 1979), but has since multiplied with the inclusion of water quality management in the stormwater paradigm (Roy et al., 2008). Discharge measurements paired with constituent concentration data allows for the estimation of pollutant loads, now regulated in many urban areas by the intersection of the National Pollutant Discharge Elimination System's (NPDES) Municipal Separate Storm Sewer System (MS4) program and the Total Maximum Daily Load (TMDL) program (Sections 303 and 402 of the Clean Water Act, respectively). Stormwater managers must now show that their localities are reducing pollutant runoff to achieve limits called Waste Load Allocations (WLAs), and though discharge measurement is necessary for pollutant load estimation, explicit requirements for discharge monitoring are absent

from the MS4 and TMDL programs, and the regulation of discharge as a pollutant unto itself was prohibited by the U.S. District Court of Virginia (VDOT v. USEPA, 2013).

As the current regulatory environment does not require discharge monitoring, and may even disincentivize it (Wagner, 2005), only a small proportion of approximately 7,000 regulated MS4 entities (USEPA, 2014) monitor discharge. Nevertheless, there are certain localities that have developed monitoring programs either through relationships with the USGS (e.g. Hoogstraat, 2015; Jastram, 2014; Storms et al., 2015), as a department of the local or regional government (e.g. City of Austin, 2009), or as consulting contracts (e.g. Gauron, 2015).

The literature provides thorough guidance on the measurement of discharge in open channels (Turnipseed and Sauer, 2010; USBR, 2001; WMO, 2010), but MS4 permits ascribe the water quality effects of urban stormwater to the underground system's terminal pipe discharging into jurisdictional waters of the U.S. – known as an “outfall.” Furthermore, the treatment prescribed for urban stormwater pollution is a combination of programmatic measures and structural controls (Aguilar and Dymond, 2015) whose hydrologic and water quality benefit is yet unknown or uncertain (Barrett, 2008; A. C. Taylor and Fletcher, 2007). Detailed guidance addressing the nuances of monitoring storm sewer discharges from MS4 outfalls and stormwater best management practices (BMPs) is needed, and in particular, there is a need for characterization of the uncertainty associated with in-pipe sensor discharge measurements and its effects on the use of flow data for modeling and pollutant load estimation (Harmel and Smith, 2007). The type of uncertainty associated with sensor measurements is called “measurement” or “observation” uncertainty (McMillan et al., 2012) – the focus of this paper.

The term “uncertainty” should be distinguished from the term “error”, which is defined as the difference between the true value and measured value (measurand), which is not operationally helpful since true values are almost never known (Moffat, 1988). Rather, uncertainty is defined as “a parameter associated with the result of a measurement that characterizes the dispersion of values that could reasonably be attributed to the measurement (WMO, 2010).” The two components of observation uncertainty as defined in Coleman and Steele (1995), and applied to hydrologic measurements in Bertrand-Krawjewski and Muste (2008a) are (1) uncertainty due to bias, and (2) precision uncertainty (Figure 4.1). While bias and precision are typically thought of as error sources, they are defined as components of uncertainty in this paper, as the true value of the measurand is not known. Bias uncertainty is the

systematic difference between the mean of the observations and the benchmark value, while precision uncertainty is the random scatter of observations about the mean, conforming to some probability distribution, and generally described by a simple statistic such as the standard deviation.

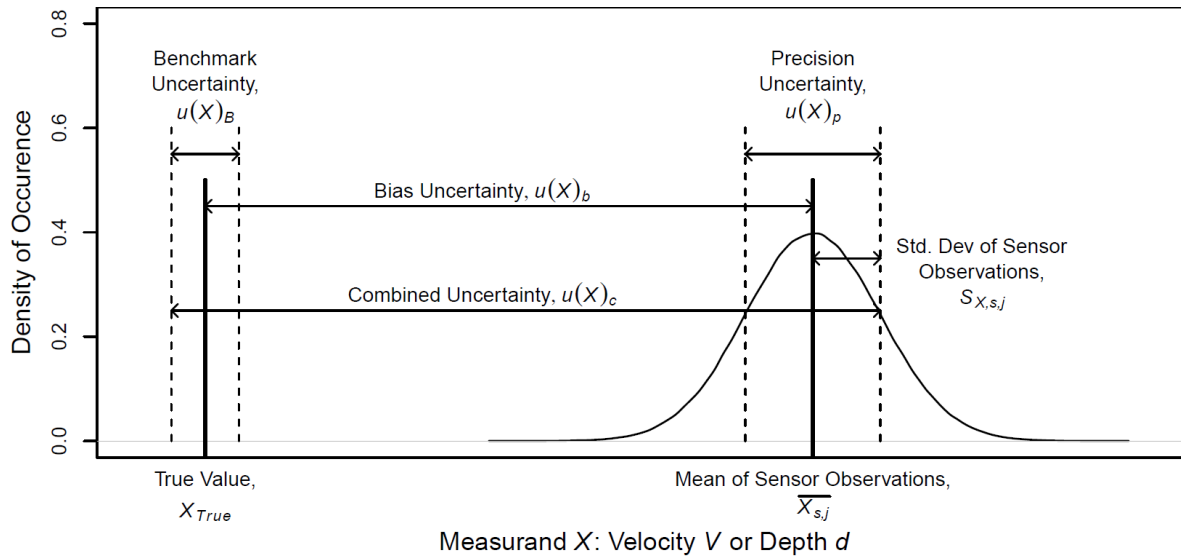


Figure 4.1 – Uncertainty associated with the observation of a measurand X , adapted from Coleman and Steele (1995)

As discharge benchmarks are generally not available, there are limited studies that attempt to quantify the components of discharge observation uncertainty with applications to storm sewer field measurements. McMillan et al. (2012) provide a meta-analysis of observation uncertainty for various types of hydrologic measurements, and Lee et al. (2014) apply a standardized uncertainty framework to river flow sensor observations, but neither provide specific information regarding the measurement of discharge in storm sewer pipes. McIntyre and Marshall (2008) and Rehmel (2008) partially fill this gap by comparing acoustic Doppler current profiler observations to the commonly used impeller current meter in nine storm sewer cross sections, and 43 USGS stations respectively, however no attempt was made to perform laboratory benchmarking in these studies. Maheepala et al. (2001) perform flume calibration of flow sensors that are then placed in storm sewer pipes and evaluated, but the procedure and results of laboratory work are not reported. Heiner and Vermeyen (2013) performed laboratory evaluations of nine sensors in a rectangular, circular, and trapezoidal channel, though laboratory constraints allowed comparisons at a limited number of discharge values, and the lab results were not applied to field measurements. The literature on discharge monitoring uncertainty lacks

the connection between laboratory benchmarking of sensor uncertainty, application of that uncertainty in the field, and the implications of uncertainty for stormwater monitoring, modeling, analysis, and decision making.

The purpose of this study is to benchmark the uncertainty associated with discharge measurements from several common sensors for their use in storm sewer monitoring and modeling. To do this, uncertainty is determined in the laboratory under controlled conditions, and with the effects of a woody debris model. Laboratory benchmarked uncertainty is then applied to field measurements, and finally the implications of observational uncertainty for urban storm water monitoring and modeling is discussed.

4.3 INSTRUMENTATION

To obtain flow measurements without structural devices (e.g. weirs and flumes), electronic sensors can be used that employ a variety of technologies to measure stage and velocity in open channels and pipes. The sensors used in this study are shown in Table 4.1, and the technologies employed are discussed in the following sections.

Table 4.1 – Classification, Manufacturer Specifications, and Cost of Sensors

Sensor	Depth				Velocity				Source	Cost (USD)
	Type ¹	Range ²	Precision (in.)	Accuracy	Type ¹	Range (fps)	Precision (fps)	Accuracy		
Massa M-300/95	US	12 in.–13 ft.	0.01	±0.1%	-	-	-	-	Massa (2014)	900
Global Water WL 705	US	4 in.–12 ft.	0.035	<±0.5%	-	-	-	-	Global Water (2014)	800
Teledyne ISCO 2150	PT	0.4 in.-10 ft.	0.012	±0.01 ft	ADV	-5 – 20	0.01	±0.1 fps ³	ISCO (2011)	3,000
FloWav PSA-AV	PT	0.9 in.-15 ft.	N/A	±0.25%	ADV	-5 – 20	N/A	±2%	FloWav (2013)	1,800
SonTek Argonaut SW	PT	7.2 in.-16 ft.	N/A	±0.1%	ADCP	-16 - 16	0.003	±1%	Xylem (2009)	5,000 ⁴
Nortek Vectrino II	-	-	-	-	ADCP	0.3 – 9.8	N/A	±0.5%	Nortek (2013)	10,000

¹Sensor types: ultrasonic (US), pressure transducer (PT), acoustic Doppler velocimeter (ADV), and acoustic Doppler current profiler (ADCP)

²Range has a different meaning for ultrasonics and pressure transducers: for ultrasonics, it represents the minimum and maximum distance between the sensor and water surface (blanking distance). For acoustic dopplers, range represents the minimum and maximum depth of water above the sensor.

³The reported accuracy of the ISCO 2150 is ±0.1 fps up to 5 fps, and ±2% between 5 and 20 fps

⁴The SonTek Argonaut SW has since been discontinued. The price shown is the cost of a SonTek IQ, a comparable unit

4.3.1 *Depth Measurement*

Sensors that employ ultrasonic (US) technology are mounted at the top of a pipe, and estimate distance to the water surface by dividing the return time-of-flight of an emitted high frequency sound wave by the velocity of that wave (Angrisani et al., 2009). A shortcoming of ultrasonic sensors is that they require a minimum distance between the sensor and water surface (known as a dead zone or blanking distance) above which the sensor is not able to take measurements (Table 4.1), constraining the number of potential installation sites. The US instruments tested in the laboratory were the Massa M-300/95 (reabeled as the Telog UT-33u/95) and Global Water WL705, known henceforth as the Massa and GW respectively. These sensors are similar in make, with the primary difference being that the GW includes a data logger that contains the battery power source, while the Massa must be connected to a separate logger for data collection and power.

Sensors that use pressure transducers (PTs) estimate depth using a submerged piezo-resistive chip that is exposed to water pressure and open to the atmosphere through a hose in the communication cable, such that the electrical signal from the chip can be calibrated to water depth. Depending on the sensor design, these electrical signals are processed within the device, or relayed to a recording and telemetry unit (RTU, also known as a data logger) for processing.

PTs in storm sewer applications have a number of shortcomings. Their installation at the bottom of a pipe affects the flow regime immediately around the sensor, subjects the pressure/communications hose to entanglement from debris, and subjects the transducer to burial from sediment. They are also subject to drift; a change in their output over a period of time that is not a function of the measured water level. To address these shortcomings, PTs must be routinely calibrated and maintained. The Teledyne ISCO 2150 ADV, FloWav PSA-AV, and SonTek Argonaut SW (known henceforth as the ISCO, FloWav, and SonTek, respectively) use PTs to sense depth.

4.3.2 *Velocity Measurement*

Acoustic Doppler current profilers (ADCPs) are a subset of acoustic Doppler velocimeters (ADV) – instruments that estimate velocity by continuously transmitting sound waves into a control volume, then measuring the Doppler frequency shift caused by reflection off particulates in the water. ADCPs provide multi-dimensional velocity measurements, while

ADV's (as defined in this study) integrate samples to provide a one-dimensional water velocity (Larrarte et al., 2008). Generally, these sensors also measure depth in order to estimate a discharge based on a form of the flow continuity equation that incorporates the observed frequency shift, as shown in McIntyre and Marshall (2008).

The ability of these sensors to accurately characterize velocity in a storm sewer pipe is contingent on several factors. First, the position of the sensor on the pipe bottom subjects them to sedimentation at low velocities, and debris accumulation during storm events, potentially leading to data loss or inaccuracy if the velocity transducer is blocked or excessive turbulence is created. Second, the velocity is estimated based on the magnitude of the frequency shift from suspended particles within the conical control volume; measurement error occurs if this volume is not representative of the cross sectional velocity (Bonakdari and Zinatizadeh, 2011), or if there is an uneven vertical distribution of sediment within the volume (McIntyre and Marshall, 2008; Nord et al., 2014). Further sources of error and considerations for field implementation are enumerated in Aguilar and Dymond (2014) and McIntyre and Marshall (2008).

The ADVs/ADCPs used in this study were the ISCO, FloWav, Sontek, and Nortek Vectrino II (known henceforth as the Nortek). The ISCO and FloWav are designed in a similar fashion; they provide a single, one-dimensional velocity measurement for a conical control volume and a single depth at each time interval. They are both bottom mounted, and both require an external data logger and power source. The SonTek has a similar setup, but is larger and measures velocity in two dimensions: parallel to the primary flow direction of the channel and normal to the top face of the sensor. The Nortek is the most dissimilar of the four velocimeters used in this study as it is usually top or side mounted, does not provide a depth measurement, and uses four receiving beams at 60 degrees from the central transmit beam to measure three-dimensional velocity. Due to its sensitivity to damage from external forces commonly found in storm flow situations, it would not be an ideal sensor for autonomous storm flow measurement.

All sensors shown in Table 4.1 were tested in the laboratory, though several of the sensors were only available for a limited amount of testing, and were not included in the results. Manufacturer names are censored from results, as the objective of this study is to evaluate observation uncertainty, not to recommend one sensor over another.

4.4 LABORATORY PROCEDURE

The laboratory setup was a 20 ft. long x 18 in. wide x 18 in. deep experimental channel that recirculates water with a user-controlled pump operated at variable speeds from 0 – 60 Hz at 0.1 Hz increments. To provide a direct benchmark with which to compare sensor observations, the flume provided a venturi–differential manometer configuration with an established rating curve derived from Bernoulli’s equation, flow continuity, and hydrostatic pressure. The venturi reduces flow from a diameter of 6.068” to 3.040”, and has a discharge efficiency coefficient, $C_d = 0.987$. The manometer measures the head difference in the venturi (Δh) to the thousandth of a foot. The equation relating discharge, Q (cfs) with Δh (ft.) is:

$$Q = C_d A_2 \sqrt{\frac{2g\Delta h}{1-r^2}} \quad \text{Equation 4.1}$$

where A_2 is the cross sectional area of the contraction in the venturi in ft^2 , g is acceleration due to gravity (32.2 ft/s^2), r is the contraction cross section area divided by the entrance cross section area (0.251), and C_d is the discharge efficiency coefficient (0.987). Information regarding the calibration of this venturi meter is from Engineering Laboratory Design (1999). In this paper, the flow as measured by the venturi-manometer configuration is referred to as “manometer discharge”.

For each of seven experiment days, the sensors were installed in the flume channel as specified by the manufacturer, and several “tests” were performed where the system was set to an initial channel depth with the pump off (i.e. zero discharge). For each test, the pump was run at increasing frequencies from 0 – 60 Hz, usually at 5 Hz increments, and the system was allowed to reach steady-state at each pump frequency. Steady-state channel depth and manometer discharge were maintained for at least four minutes – referred to as a “trial”, and annotated as j . Depth and manometer discharge were recorded for each trial, and an observation i was taken by each sensor, s , mounted in the channel every minute, referred to as $d_{s,i}$ or $V_{s,i}$ for $i = 1 \rightarrow n$ depth and velocity observations in trial j respectively.

Summary statistics were performed on each trial such that i observations were aggregated into j trials for $j = 1 \rightarrow N$ trials. The mean and standard deviation of sensor depth and velocity observations i in trial j were calculated and annotated as $\overline{X_{s,j}}$ and $S_{X,s,j}$ respectively, where X represents the measurands depth, d and velocity, V , and S represents the standard deviation of i observations of the measurand. The following section describes how $\overline{X_{s,j}}$ and $S_{X,s,j}$ were used to

benchmark observation bias and precision uncertainty respectively, and how the uncertainty associated with the benchmark itself was defined. Uncertainty associated with a measurand X will be annotated as $u(X)_k$ where k is one of the three components of uncertainty defined for this study: bias $u(X)_b$, precision $u(X)_p$, and benchmark $u(X)_B$; and where $u(X)_c$ is the combined uncertainty of the three components (Figure 4.1).

Measurements taken outside of manufacturer reported sensor limitations (e.g. minimum depth or blanking distance, see Table 4.1) were removed from analysis, as the sensors were not expected to operate outside of specified conditions. Other erroneous measurements noted during sensor data post-processing were marked and censored from analysis; this will be discussed further in the Results section.

4.4.1 *Benchmark Uncertainty*

The uncertainty associated with the benchmark measurements ($u(X)_B$) was calculated for the laboratory, as it provided a measure of the extent to which the depth and velocity could be measured given a set of laboratory tools. This is referred to as epistemic uncertainty – the limit to what can be known about a system; its counterpart – aleatory or “natural” uncertainty – is discussed later in this section (definitions from Merz and Thielen, 2005). It was not possible to evaluate benchmark uncertainty in the field, as the depth and velocity were not directly observable, but it was assumed that the instruments used for observation of hydraulic parameters in the field would be at least as uncertain as those used in the laboratory. As a result, laboratory benchmark uncertainty is applied to field measurements as a conservative means of describing the upper limit of what can be directly observed in the field.

The uncertainty associated with the discharge benchmark based on the 0.001 ft. head increments applied to the manometer equation (Equation 4.1) was ± 0.013 cfs. The pump frequency-initial channel depth-discharge relationship was created by filling the flume to an initial depth in the channel, running the variable speed pump from 0 – 60 Hz, and estimating discharge based on manometer readings. The initial depth in the flume was then incrementally increased, and the process was repeated. Non-linear least squares regression was performed separately for each initial channel depth to relate manometer discharge to pump frequency based on a power-law relationship. The uncertainty associated with these rating curves was defined as the maximum standard error of the residuals (RSE) of these power-law models (0.01 cfs) where:

$$RSE = \sqrt{\frac{\sum_{i=1}^n (X_{o,i} - X_{p,i})^2}{n - 2}} \quad \text{Equation 4.2}$$

and $X_{o,i}$ is the observation of the value by direct measurement, $X_{p,i}$ is the predicted value from the model, and n is the number of observations ($56 \leq n \leq 92$). The RSE is in the same units as the measurement and provides a metric of the difference between modeled values and observed values at all discharges. The total uncertainty associated with the manometer reading and frequency-discharge rating curve was calculated based on the probable error range defined in Bertrand-Krawjewski and Muste (2008a), Harmel and Smith (2007), and Sauer and Meyer (1992):

$$u^2(y) = \sum_{i=1}^N u^2(x)_i \quad \text{Equation 4.3}$$

where $u(y)$ is the combined uncertainty of N independent components of uncertainty, $u(x)_i$, and for the manometer and rating curve, the combined discharge benchmark uncertainty, $u(Q)_B$ was ± 0.016 cfs.

Depth, $d_{o,i}$ was measured in the channel with a depth gauge attached to the plexi-glass channel wall perpendicular to the water surface for a range of depths between zero and twelve inches, with a benchmark uncertainty, $u(d)_B = \pm 0.0625$ in. Benchmark velocity, $V_{o,i}$ was estimated based on the manometer discharge observation divided by cross sectional flow area (as calculated by $d_{o,i}$). The uncertainty in the velocity benchmark is the propagated uncertainty in depth and discharge benchmarks based on the principle from Coleman and Steele (1995):

$$u^2(y) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) \quad \text{Equation 4.4}$$

where $u(y)$ is the uncertainty of a value calculated as a function of x_i , and f is that function. The application of this principle to the estimation of velocity uncertainty as a function of depth, discharge, and their respective uncertainties is:

$$u(V)_B = \sqrt{\left(\frac{\partial V}{\partial Q} \right)^2 u(Q)^2 + \left(\frac{\partial V}{\partial d} \right)^2 u(d)^2 + \left(\frac{\partial V}{\partial w} \right)^2 u(w)^2} \quad \text{Equation 4.5}$$

where V is velocity in fps (estimated as $V = Q/wd$), d is channel depth in ft., w is channel width in ft. (constant and relatively certain) and Q is discharge in cfs. The benchmarked velocity

uncertainty, as it is a derived value, varies with manually observed depth and manometer discharge, such that the maximum $u(V)_B$ achieved in the laboratory was ± 0.272 fps as Froude number, $Fr \rightarrow 1$. Although it was possible to calculate the velocity benchmark uncertainty in the laboratory, it was not possible to do so in the field, as depth and discharge were not directly observable in the field. As such, the bootstrapped upper 95% confidence interval of the velocity benchmark uncertainty was used, such that $u(V)_B = 0.056$ fps.

4.4.2 Bias Uncertainty

The bias uncertainty was defined as the difference between the benchmarked value ($X_{o,j}$) and the mean of i sensor observations for each trial, j ($\overline{X_{s,j}}$), and was calculated as the standard error of the observation residuals using the equation for RSE:

$$u(X)_b = \sqrt{\frac{\sum_{i=1}^N (X_{o,j} - \overline{X_{s,j}})^2}{N-2}} \quad \text{Equation 4.6}$$

First, all $\overline{X_{s,j}}$ were plotted against their respective manual observation, $X_{o,j}$ with a line of perfect agreement (1:1 line) for reference. Then, observation residuals ($X_{o,j} - \overline{X_{s,j}}$) were plotted against channel depth, velocity, and Froude number (Fr) to determine the effects of these parameters on the residuals. The 1:1 plots and residual plots were visually inspected, and it was determined how best to model sensor observations as a function of manual observations, and thereby transform sensor observations to better fit manual observations.

Models were built for $X_{o,j}$ as a function of $\overline{X_{s,j}}$ so that sensor observations could be transformed to better fit manual observations, and so this transformation could be applied in the field. These transformation functions were created in three ways depending on the transformation necessary to meet homoscedasticity and autocorrelation assumptions of linear modeling in the adjusted (i.e. transformed) mean observations ($\overline{X_{s,j,adj}}$): (1) linear least-squares regression (2) non-linear least-squares regression (3) addition of observation residual models (as a function of Fr) to sensor observations. These transformation functions were applied to the original sensor observations to adjust them towards the line of perfect agreement with direct observations, and to remove heteroscedasticity and autocorrelation. Subsequent linear models were then built for the manual observations as a function of these adjusted sensor observations

$(\overline{X_{s,j,adj}})$, and the RSE of this adjusted model was reported as the adjusted bias uncertainty $(u(X)_{b,adj})$.

The date of testing was later included in these models as a categorical variable to determine if sensor depth observations drifted with time, but only for the purpose of commenting on lab conditions, as direct depth observations were not measured in the field. All linear and non-linear least-squares regression modeling was performed in the R software language using the “lm” and “nls” functions respectively (R Core Development Team, 2014).

4.4.3 Precision Uncertainty

Precision uncertainty (Figure 4.1) is defined as the random scatter of sensor observations about the mean value of these observations due to electrical limitations of the sensor, and environmental stochasticity [i.e. aleatory uncertainty (Merz and Thielen, 2005)]. Testing in the laboratory was designed to minimize the effects of environmental conditions, although the very presence of the bottom mounted ADVs and ADCPs created non-uniformities in the channel. It was not possible to control for this, and was deemed acceptable as the same effects would be present in the field installation.

This random scatter was evaluated for N trials at steady-state channel depth and velocity as the standard deviation of $n > 3$ sensor observations ($S_{X,s,j}$). These standard deviations were calculated for trials across a range of depths and velocities in the channel, and it was hypothesized that precision uncertainty would increase as $Fr \rightarrow 1$, as the flow regime in the channel was least stable under critical flow conditions. To test this hypothesis, $S_{X,s,j}$ was plotted as a function of the manually observed depth, velocity, and Froude number, and the plots were visually inspected for trends. If no trend was present, the precision uncertainty $(u(X)_p)$ was calculated as the median of all $S_{X,s,j}$, as the median is robust against non-normality and outliers.

4.4.4 The Effect of Woody Debris

Previous experience with storm sewer discharge monitoring, and the results in McCuen et al. (2014) has shown that storm sewer systems are subject to the accumulation of trash and debris, especially in large pipes and at locations where open channels enter daylighted culverts, as there are no preventative measures for removing these detritus before they enter the storm sewer system. Although it was not possible to build an exact scale model of any particular debris

jam, the principles from Gippel et al. (1996) were used to approximate a woody debris jam that might exist in a storm sewer pipe.

A modular debris field was constructed from 3/4” and 1/8” balsa wood dowels mounted to a 1/2” plywood sheet cut to fit the width of the flume channel so that the field could easily be installed and removed from the channel (Figure 4.2). The previously described testing procedure was used to benchmark sensor bias and precision uncertainty with debris set in the channel at a fixed distance upstream of each sensor, so that a comparison could be made with the observation uncertainty without debris in the channel. Two days of experimentation were performed where tests were run without debris in the channel, and the same tests were repeated with debris in the channel.



Figure 4.2 – Model woody debris field used to evaluate sensor uncertainty with obstructions in storm sewer channels

The presence of the debris model was first used to subset data to determine $u(X)_b$ and $u(X)_p$ of the sensors with and without debris in the channel, and then as a presence/absence categorical variable to determine if the presence of debris had a significant effect ($p < 0.05$) on model specification or coefficients.

4.4.5 Combined Uncertainty

The combined observation uncertainty for a measurand was calculated as the probable error range (Equation 4.3) of adjusted bias, precision, and benchmark uncertainty:

$$u(X)_c = \sqrt{u(X)_{b,adj}^2 + u(X)_p^2 + u(X)_B^2} \quad \text{Equation 4.7}$$

To evaluate the relative importance of depth and velocity sensor uncertainty on total discharge uncertainty, the principal of error propagation (Equation 4.4) from Bertrand-Krawjewski and Muste (2008a) and Coleman and Steele (1995) was used, such that:

$$u(Q) = \sqrt{u(V)^2(wd)^2 + u(d)^2(Vw)^2} \quad \text{Equation 4.8}$$

as $Q = Vwd$ for both the flume channel and field installation.

4.5 FIELD PROCEDURE

4.5.1 Selection of a Field Location and Installation

As the cost for purchase, operation, and maintenance of instrumentation for water measurements may limit the number of monitoring stations, it is important that installation sites are carefully selected. Certain locations might be considered based on institutional knowledge or regulations, though if a city maintains a storm sewer GIS database, this can be helpful in systematically identifying potential monitoring locations.

The location for the installation of sensors in this study was selected by (1) querying a storm sewer GIS database described in Aguilar and Dymond (2014) and (2) consulting the municipal authority (known as the MS4 operator) responsible for the storm sewer system in which the sensors were to be placed. The GIS was used for a preliminary selection of sites in the MS4, using a series of logical queries. First, all reinforced concrete pipes were selected, as open channels can have unstable cross-sections, and other materials (e.g. corrugated metal) can produce non-uniformities in the flow. As large fluctuations in depth, velocity, and discharge were desired for this experiment, preference was given by descending pipe size and upstream watershed area. A practical consideration was the accessibility of the sensors for maintenance, as manufacturers recommend weekly maintenance until sufficient understanding of environmental conditions is obtained (ISCO, 2011), and many storm sewer pipes require confined space entry. Locations where the storm sewer system daylights into open streams serve as highly accessible monitoring locations, but are more susceptible to entrance and exit turbulence.

A short-list of potential monitoring locations was created by querying these criteria in the GIS, which was then presented to the MS4 operator. The operator provided insight into locations that might make valuable study areas – for example where future development was

proposed, or where flooding complaints had occurred. The inclusion of this institutional knowledge added a layer of validation to the GIS exercise that would have otherwise been missed.

The final site selected was a 70 in. wide x 49 $\frac{3}{4}$ in. tall rough bottom rectangular reinforced concrete pipe at 0.85% slope, draining a 490 acre large watershed near downtown Blacksburg, Virginia. The site was selected as it drained an area that was larger and more urbanized than the other potential sites, and there was future infill development proposed within the contributing watershed. The rectangular pipe was accessed through a manhole in an upstream junction box, which was built as a transition from two parallel 60 in. diameter reinforced concrete pipes upstream of the junction box. The junction box had artificial low-flow channels formed into the bottom to transition baseflow from the two upstream circular pipes, to a single downstream rectangular pipe.

Two ADVs and one US (referred to as Sensor 1, Sensor 2, and Sensor 3 respectively) were installed along the center line of the downstream rectangular pipe at 4 ft., 11.5 ft., and 18.5 ft. downstream of the end of the junction box respectively, as constrained by the lengths of the communication cables. All sensors were mounted to the pipe using concrete screws, and the communication cables were fastened to the pipe bottom, side, or top using plastic clamps to assure that debris would not be caught on the cables. Sensors 2 and 3 were connected to a single RTU, and Sensor 1 used a separate RTU; both of these were hung from a $\frac{3}{4}$ in. steel all-thread hanger in the manhole above the junction box. Sensors were left in the field location for 2.5 months, recording nine storm events in that time period.

There were several shortcomings of the site. First, the sensors may have been subject to non-uniformities in the flow produced by the transition from the upstream circular pipes and junction box – especially Sensor 3 (the ultrasonic), as it was closest to this transition. As sensors must generally be tethered to an RTU, site access and hydraulic stability are trade-offs unless additional cable lengths can be acquired. The site also did not allow direct observation of the hydraulic conditions during storm events, as this would have been dangerous to the observer. Nevertheless, the site was chosen because it drained the largest area, had the same cross sectional shape as the laboratory channel, and had the best access given the two previous constraints.

4.5.2 *Sensor Power Considerations*

During laboratory and field testing, it was noted that a sensor's power requirements are important considerations. The sensors were designed for field use, and were powered by batteries, which in some cases could be recharged using solar panels (Sensors 2 and 3). However solar recharge was not available in the lab or field setting, and the testing was power intensive as samples were needed at one minute time intervals. Sensor 1 conserved power using a feature that allowed observations at two time intervals given a velocity or depth threshold; it was set to record measurements at 1 minute intervals when depth was above 2 in., and 15 minute intervals otherwise. This option was not available for Sensors 2 and 3.

The manufacturer supplied battery power was insufficient for long term testing of Sensors 2 and 3, so in the laboratory, the data logger for these sensors was connected to a laboratory DC power supply set at a constant 12V. As the sensors were installed in a manhole covered junction box, there was no practical way to use solar recharge, so instead two 12V, 21 A-h lead-acid batteries were purchased so that one battery could power the sensors while the other recharged.

To determine the battery rating necessary for Sensors 2 and 3, the total load from the sensors was estimated using the DC power supply in the laboratory. As the sensors required 15 seconds of approximately 0.20 Amps (A) for measurements every minute, and the floating load was approximately 0.001 A for the remaining 45 seconds, the total current required for one minute was $0.20 A \times 0.25 + 0.001 A \times 0.75 = 0.051 A$. As it was desired that batteries would last in the field for two weeks without replacement with an additional 10% margin (370 hrs), it was determined that a battery rated at $0.051 A \times 370 h = 19 A - h$ or greater be purchased. Above this rating, the weight of the battery became limiting, as both RTUs would need to be hung from the $\frac{3}{4}$ in. steel all thread hanger in the junction box manhole.

Similar computations could be performed for other sensor-RTU combinations, though many sensors have proprietary battery mounts, making it difficult to use different power configurations. Sensor 1, for example, used two 6V lantern batteries mounted inside the RTU in a manner that would have been difficult to use other battery supplies. However, the power demands of Sensor 1, while not directly measured, appeared to be significantly smaller than Sensors 2 and 3, as the lantern batteries were only replaced once in the 2.5 months of field use.

4.5.3 Application of Laboratory Uncertainty to Field Observations

Time series data for the three depth and two velocity sensors installed in the field were downloaded and integrated into a single time series so that analysis could be more easily performed. First, the laboratory defined transformation equations were applied to the field data to determine if these equations reconciled the observations from different sensors to each other. Second, the combined uncertainty of the depth and velocity measurements was applied to the field data using Equation 4.8 to determine the total uncertainty of discharge measurements for each sensor. Finally, the level of agreement between the sensors was assessed for each recorded storm event using a variation of the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970):

$$E_{a,b} = 1 - \frac{\sum_{t=1}^T (Q_a^t - Q_b^t)^2}{\sum_{t=1}^T (Q_a^t - \bar{Q}_a)^2} \quad \text{Equation 4.9}$$

where $E_{a,b}$ is the NSE coefficient representing the goodness-of-fit of sensor b 's discharge observations (Q_b^t) with respect to sensor a 's discharge observations (Q_a^t) and compared to sensor a 's mean discharge \bar{Q}_a . As the NSE coefficient is commonly used to assess the predictive power of hydrological models by comparing simulated and observed hydrographs, it is used in this paper to compare the hydrographs observed using two different flow sensors in order to comment on how well models should fit measured data. The NSE coefficient can range between $-\infty$ and 1, however values greater than 0.5 are generally deemed acceptable for model simulations (Moriassi et al., 2007).

4.6 RESULTS AND DISCUSSION

The results of all laboratory experiments were first plotted as time-series in order to visually inspect the agreement between sensor and manual observations, and several problems were noted in doing so. First, Sensor 1's velocity transducer reported zero values (though not consistently) at manometer derived velocities of 0.3 fps and less, suggesting a minimum detection threshold. Zero velocities and depths were also reported by both Sensor 1 and 2 at random times in the experiments, contrary to channel conditions and serial measurements. These erroneous values were censored from analysis.

It was also noted that at the end of trials when the pump was turned off and velocity in the channel rapidly dropped to zero, Sensor 1's velocity sensor continued to report the previously reported value (i.e. the transducer appeared to be "stuck") until the velocity changed significantly

from zero, though it was not possible to directly quantify what allowed the sensor to resume taking observations. These values were also censored from analysis, as this phenomenon was unlikely to occur in the field, since the rate of change of velocity in the field was limited to the ascending and descending limbs of storm hydrographs.

As the sensors tested did not provide any sort of error flagging, data post-processing was necessary to assure that uncertainty benchmarking and field data analyses were not skewed by erroneous data. In general, data validation conformed to recommendations in Bertrand-Krawjewski and Muste (2008b), and it should be noted that as the number of measurements increases, data validation and post-processing can become a limiting factor for large-scale monitoring programs.

This section presents the results of the laboratory benchmarking of bias, precision, and combined uncertainty; the effects of woody debris therein, and the application of laboratory benchmarking to observations in the field. Although benchmark uncertainty is a component of combined uncertainty, benchmark uncertainty is not discussed further in this section, as its estimation and outcomes are included in the laboratory procedure.

4.6.1 *Bias Uncertainty*

Uncertainty due to bias was defined as the standard error of the residuals (RSE), $X_{o,j} - \overline{X_{s,j}}$ for j trials. This was first calculated based on the RSE of the unadjusted $\overline{X_{s,j}}$, then models were built to transform these values to $\overline{X_{s,j,adj}}$, and the adjusted RSE was calculated. The purpose of this adjustment was to determine the relationship between sensor observations and laboratory benchmarks, so that this relationship could be applied to sensor observations in the field. The transformation functions, unadjusted bias uncertainty, and adjusted bias uncertainty are shown in Table 4.2.

Table 4.2 – The transformation functions used to adjust sensor depth (d_s) and velocity (V_s) observations towards manual observations with and without debris, the bias uncertainty of the original observations and the adjusted observations, precision uncertainty, benchmark uncertainty, and combined uncertainty from Equation 4.7. Values of uncertainty are in +/- the unit of measurement.

	Sensor	Type	Debris	Number of Trials	Transformation Function	Unadjusted Bias Uncertainty	Adjusted Bias Uncertainty	Precision Uncertainty	Benchmark Uncertainty	Unadjusted Combined Uncertainty	Adjusted Combined Uncertainty
Depth (in.)	1	PT	N	465	$(d_s + 0.106)/1.038$	0.744	0.707	0.031	0.0625	0.747	0.710
			Y ⁺	108	$(d_s + 1.27)/1.12$	0.863	0.198	0.040	0.0625	0.866	0.211
	2	PT	N	349	$(d_s + 0.176)/1.063$	0.347	0.279	0.141	0.0625	0.380	0.319
			Y	132	$(d_s + 0.156)/1.053$	0.444	0.393	0.173	0.0625	0.481	0.434
	3	US	N	353	$(d_s - 0.006)/0.993$	0.192	0.192	0.044	0.0625	0.207	0.207
			Y [*]	102	$(d_s - 0.568)/0.864$	0.394	0.404	0.085	0.0625	0.408	0.418
Velocity (fps)	1	ADV	N	287	$0.875 \times V_s^{0.968}$	0.314	0.159	0.050	0.056	0.323	0.176
			Y ^{**}	92	$0.692 \times V_s^{0.912}$	0.641	0.166	0.059	0.056	0.646	0.185
	2	ADV	N	349	$V_s - 3.42 + 4.65 \times Fr^{1/2}$	1.44	0.588	0.221	0.056	1.458	0.631
			Y [†]	132	$(V_s - 0.390)/0.956$	0.479	0.338	0.179	0.056	0.514	0.387

[†]Debris significantly affects y-intercept, but not slope

^{*}Debris significantly effects y-intercept and slope

[†]Debris changes the form of the transformation function

^{**}Debris significantly effects nls parameters

Figure 4.3 shows the unadjusted and adjusted observations, $\overline{X_{s,j}}$ with 60% transparency, and the associated least-squares regression lines demonstrating improved agreement after transformation. Bias uncertainty was reported as the RSE of the adjusted and unadjusted regression lines as shown in Table 4.2. Observations from the three depth sensors had a linear relationship with the manual observations (Figure 4.3, A-C), though the velocity sensors did not (Figure 4.3, D-E); Sensor 2's velocity observations did not appear to have any relationship to manual observations, and as such, no least-squares regression line for the unadjusted observations is shown (Figure 4.3E).

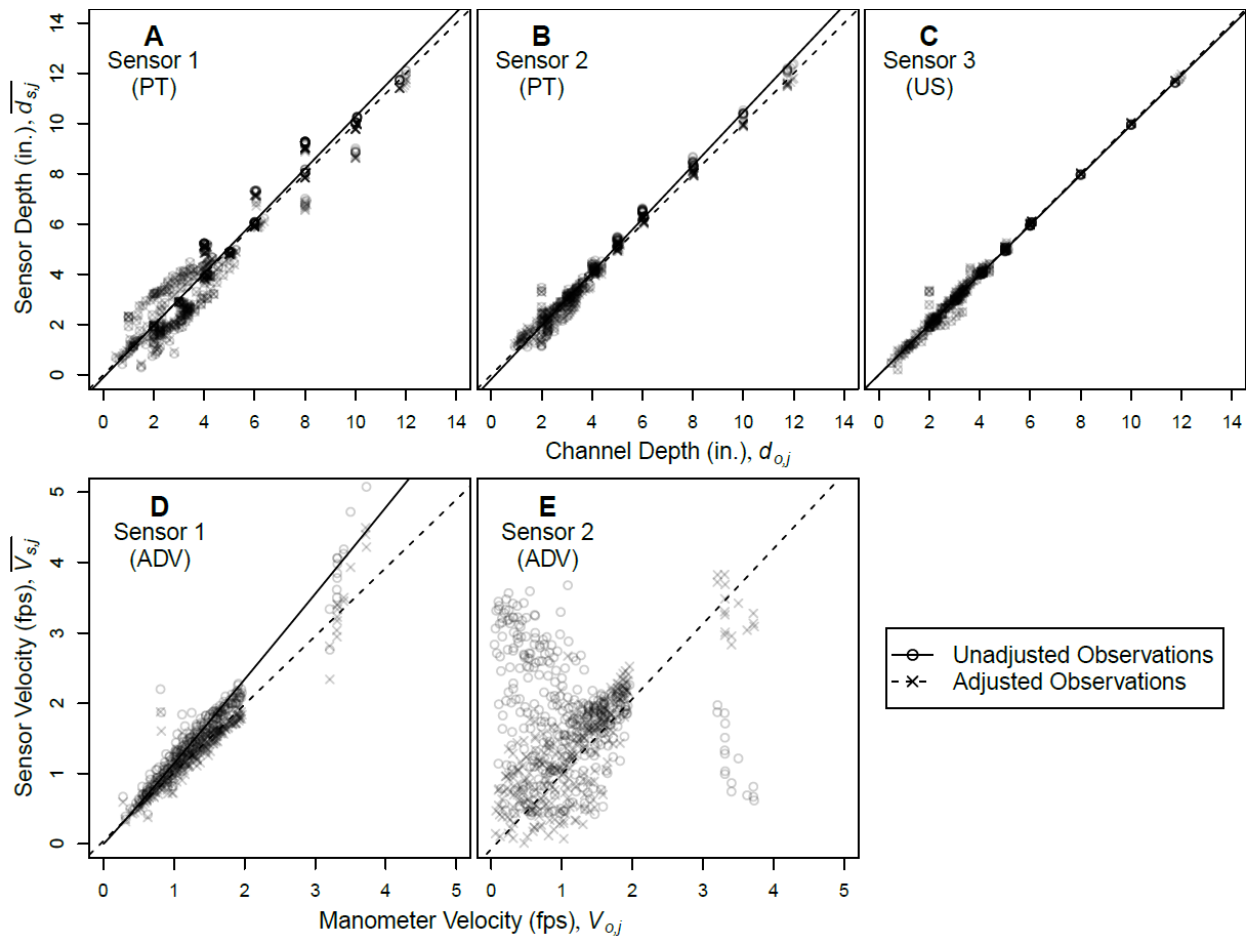


Figure 4.3 – Sensor depth observations plotted against manual observations (A - C), and sensor velocity observations plotted against manometer derived velocity (D - E) for all trials with least-squares regression lines shown

Sensor observation residuals were then plotted as a function of manually observed depth, velocity, and Froude number to determine how these parameters affected residuals. Froude number (Fr) plots are shown in Figure 4.4, as Fr represents the mixed effect of depth and velocity on residuals. The magnitude of the residuals of the three sensor's depth observations

were invariant with manually observed depth (Figure 4.4, A – C), so a linear regression was used to transform $\overline{d_{s,j}}$ to vary about the line of perfect agreement. Sensor 1’s velocity observations expressed a non-linear relationship with manometer derived velocity, and as such a non-linear least-squares regression model was created to transform Sensor 1’s velocity observations (Figure 4.3D). Sensor 2’s velocity observations did not appear to have any correlation with the manometer derived velocity (Figure 4.3E), although observation residuals appeared to vary systematically with Fr (Figure 4.4E). To adjust Sensor 2’s velocity observations, a linear least-squares regression model was developed for the observation residuals as a function of Fr (Figure 4.4E), and the adjusted velocity observations were calculated as the sum of the unadjusted observations and the output of this linear model built for observation residuals as a function of Froude number.

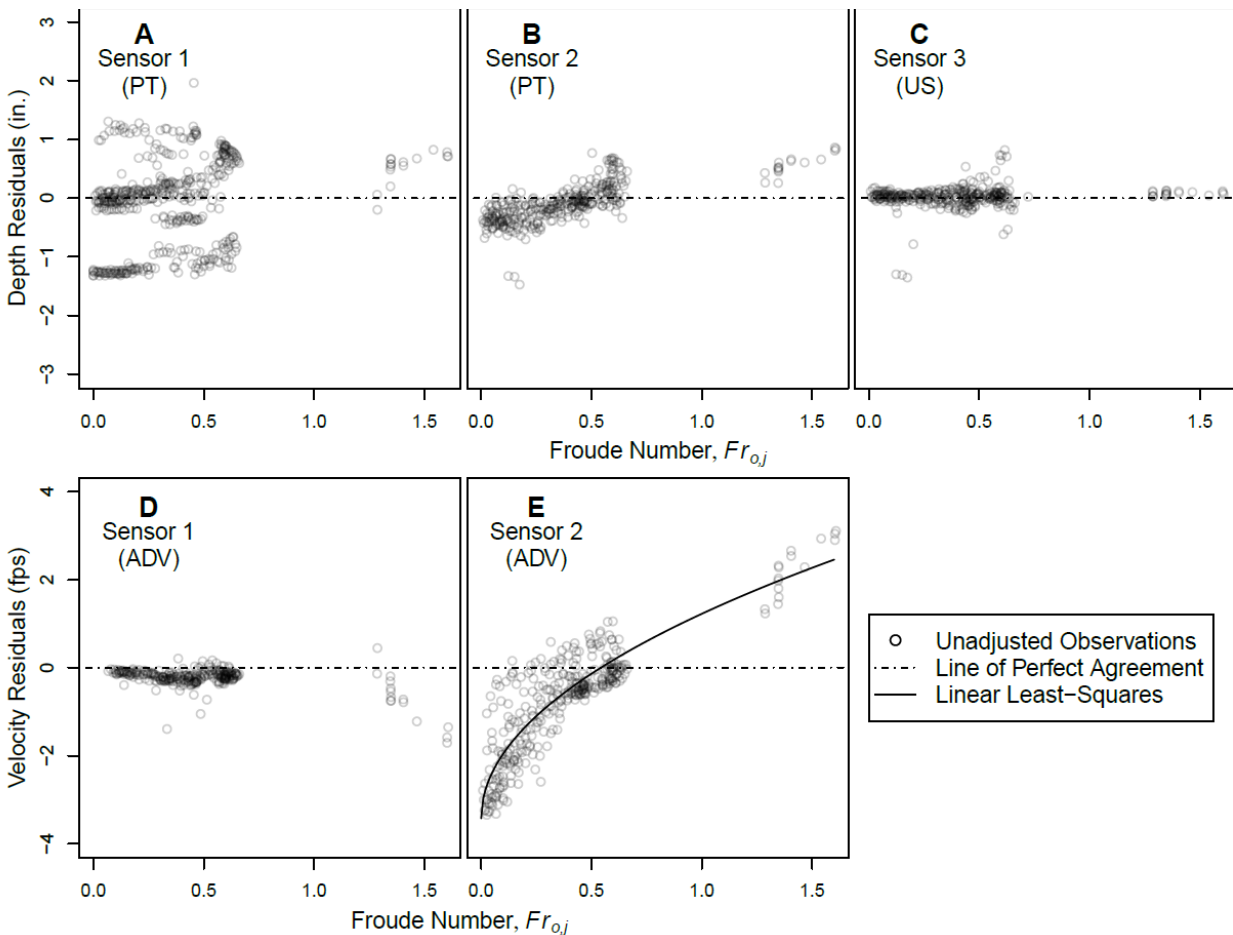


Figure 4.4 – Sensor observation residuals, $X_{o,j} - \overline{X_{s,j}}$ for depth (A – C) and velocity (D – E) with line of perfect agreement. (E) Sensor 2’s velocity residuals varied with Froude number; the least-squares regression line is shown

After initial experiments, it was noted that Sensor 1's pressure transducer (PT) systematically overestimated depth, as the sensor vs. manual depth observation line had a slope of approximately 1, but the y-intercept was 1.07 in. (i.e. the sensor systematically overestimated depth by 1.07 in.). This was thought to be caused either by sensor drift, or a mounting configuration in a stainless steel bracket with silicone caulk that prevented the sensor's pressure transducer from being properly exposed to water pressure. The sensor was re-mounted in an improved bracket without caulk, and calibrated by placing the sensor in a known depth of water, and setting the sensor's depth observation to the known depth. This shifted the intercept of the depth observations model by 0.74 in. towards the actual values. This was the only manual calibration or modification to any sensor that was performed, though Sensor 1's depth observations continued to shift between experiments, with a range of y-intercepts from -1.12 to +1.07 in. When the date of the experiment was included as a categorical variable to shift the y-intercept for different dates, the unadjusted bias uncertainty was reduced to ± 0.276 in.

Sensor 2 and 3's depth observations also showed significant shifts between experiments (Sensor 3's to a lesser degree), though there was no exogenous factor that could be found to cause this shift. The range of y-intercepts (i.e. measurement drift) for different experiment dates for Sensor 2's PT ranged from -0.20 to 0.71 and Sensor 3's ultrasonic (US) transducer from -0.02 to 0.11. As it was not possible to apply this information to the field, the overall y-intercept was used to adjust field measurements, and the sensor drift noted is included in bias uncertainty.

4.6.2 Precision Uncertainty

Sensor precision uncertainty was evaluated as the standard deviation of $n > 3$ sensor observations ($S_{X,s,j}$) when depth and velocity in the channel were at steady-state. It was hypothesized that precision uncertainty would increase as $Fr \rightarrow 1$ from either direction, as sensors aggregate many measurements into a single reported value of depth and velocity under these non-uniform conditions. Plots of $S_{X,s,j}$ as a function of $Fr_{o,j}$ at these steady-state conditions are shown in Figure 4.5, but were inconclusive to this end. It was not possible to evaluate $S_{X,s,j}$ at $0.72 < Fr < 1.28$, as the channel entered gradually varied flow conditions in this range, and the varying depth and velocity along the length of the channel prevented comparison to manual observations. As a result of this gap in the data, precision uncertainty is reported as the median of all $S_{X,s,j}$ for each sensor, with the caveat that near critical flow, sensor precision

uncertainty is likely to increase based on the principles of the measurement devices, but not substantiated by experimental results.

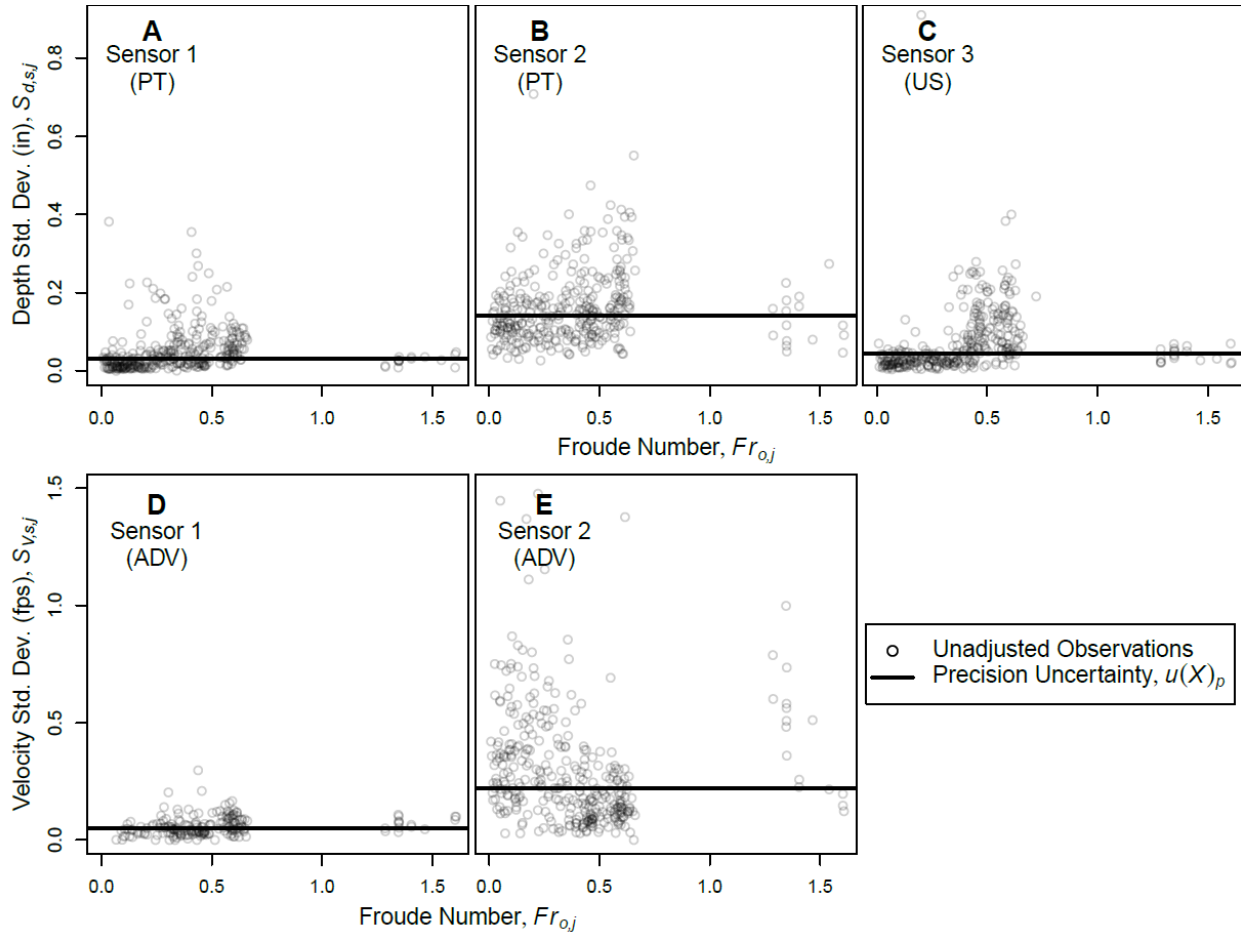


Figure 4.5 – Standard deviation of i sensor observations for each trial, $S_{s,j}$ plotted against manometer derived Froude Number, $Fr_{o,j}$ with the median value shown representing precision uncertainty

4.6.3 The Effect of Woody Debris

The effect of the presence of a woody debris field in a storm sewer on the uncertainty of sensor observations was tested by placing a balsa wood model designed to approximately simulate a debris jam in the flume at a set distance upstream from the sensors. The effects of the debris jam were evaluated by assessing how the presence of debris affected observation residuals, and if the presence of debris had a significant impact on sensor uncertainty. The significance of the debris field, transformation functions, and associated uncertainties are shown in Table 4.2.

The bias uncertainty associated with Sensor 1’s depth sensor significantly increased with upstream debris, though the transformation function adjusted these observations so that

$u(d)_{b,adj}$ was significantly smaller than the “no debris” model, since there were only two days of debris testing, and there was minimal sensor drift between these two days. The presence of debris jam in the channel resulted in insignificantly higher bias and precision uncertainties for Sensor 2’s depth transducer, although the debris had a significant effect on Sensor 3’s depth observations, as Sensor 3 overestimated at channel depths below three inches. These overestimations at low values had high leverage for the transformation model, and as such, the bias uncertainty of the adjusted model is actually worse than the original model.

Sensor 1’s velocity transducer was significantly less certain with debris in the channel, with unadjusted bias uncertainty increasing by 104% over the unadjusted no-debris bias uncertainty due to systematic overestimation, but when adjusted, bias uncertainty was nearly that of the sensor with no debris in the channel. The presence of debris in the channel had the opposite effect on Sensor 2’s velocity observation uncertainty – it significantly decreased both bias and precision uncertainty. There was a more visible agreement between Sensor 2’s $\overline{V_{s,j}}$ and $V_{o,j}$ with debris in the channel, though it was not possible to determine why this was the case.

The precision uncertainty for all other transducers increased with debris in the channel, an effect that would be expected given the decrease in flow uniformity. However the decrease in precision was small relative to the increase in bias due to debris, suggesting that while non-uniformities in the flow do not significantly reduce sensor observation repeatability, they make it more difficult for the sensor to report average measures of depth and velocity that are characteristic of the hydraulic conditions.

Combined Uncertainty

The precision, bias, and benchmark uncertainty was combined using the probable error range (Equation 4.7) to arrive at a total laboratory uncertainty for each of the sensors for application to the field (Table 4.2). Sensor 1 had the highest $u(d)_c$ of 0.710 in. of all the depth sensors when there was no debris in the channel, however it was apparent that this was due to sensor drift between experiment dates. When the data was controlled for sensor drift, the unadjusted bias uncertainty was reduced to 0.276 in., but this factor could not be included in the transformation function because it was not possible to determine how to shift Sensor 1’s depth measurements in the field.

Sensor 3 (an ultrasonic) had better bias uncertainty than both the pressure transducers, but the precision was approximately the same as Sensor 1's. As a result, Sensor 3 may provide a more reliable depth measurement than the other two, as the bias was low and minimal transformation was required, however when the water surface in the pipe is < 12 in. from Sensor 3's transducer, observations from one of the pressure transducers will need to be used. This is also true of low depth conditions when there is debris in the channel, though it is generally not possible to identify these conditions without direct observation.

Sensor 1's velocity bias and precision was lower than Sensor 2's, and as a result had lower combined uncertainty with and without debris in the channel. It did not appear that there was any relationship between Sensor 2's observations and manually observed velocity (Figure 4.3E) until the residuals were plotted against depth, velocity, and Froude number (Figure 4.4E), but it was not clear why the residuals of Sensor 2's observations would vary with these parameters. With debris in the channel, there was at least a visible relationship between Sensor 2 observations and manual velocity observations, but with marginally better precision.

4.6.4 *Application to Field Observations*

The sensors were deployed in the field for 2.5 months (March 12 – May 29, 2015) during which time they recorded 9 separate storm events. The transformation equations were applied to the sensor data to determine the adjusted discharges, and the combined uncertainties from Table 4.2 were used with Equation 4.8 to estimate the propagated uncertainty of discharge measurements. The transformation equation for Sensor 2's velocity requires a manually observed Froude number, but since there was no way to directly observe the depth or velocity in the field, the Froude number calculated using Sensor 1's adjusted observations was used as a surrogate.

While the transformation equations reduced the uncertainty in the field measurements, the NSE of the observed hydrographs showed only minor improvements for 4 of the 9 storms and an overall decrease in the average discharge NSE across all storms. This lack of improvement in the agreement by adjusting sensor discharge estimates was largely attributed to the influence that adjusted depth measurements have on the discharge, which showed only a marginal improvement in the average NSE (+0.1) for the PT and were worse for the US (Table 4.3). Although the adjusted velocity readings had an improved fit in 7 out of 9 storms, the velocity transformation function is inherently biased through the input of the adjusted Sensor 1

Froude number and would not be useful in most cases where there is not a second velocity measurement device.

Table 4.3 - Minimum, maximum and mean Nash-Sutcliffe efficiency (NSE) values for nine storm events using Sensor 1 as the “observed” data.

Storm	Depth				Velocity		Discharge	
	PT Unadj	PT Adj	US Unadj	US Adj	ADV Unadj.	ADV Adj.	Original	Adjusted
Min	-21.88	-21.26	-8.73	-9.69	-1.65	-2.82	-2.70	-15.71
Mean	-3.05	-2.95	-3.14	-3.51	-0.37	0.07	-0.49	-2.04
Max	0.55	0.55	-0.60	-0.66	0.42	0.77	0.55	0.44

These results do not support the hypothesis that adjusting field depth and velocity data using laboratory benchmarked values will result in better agreement between the sensors. One possible reason for this is that the physical conditions in the laboratory were restricted to the capabilities of the system and the power of the pump. Therefore, the depth and velocities simulated in the experimental flume may not be completely representative of field depth and velocity combinations. Another reason could be that the flow conditions between the two sensor locations vary during storm events, where limitations due to the site selection, such as the upstream pipe transition and junction box, result in non-uniformities in the flow. The following results should be viewed within the context of these field limitations.

The level of agreement between the sensors during storm events was first evaluated using the NSE for the unadjusted and adjusted time series (Table 4.3). The NSE of the depth, velocity and flow was computed for each storm event using Equation 4.9, with Sensor 1 as the observed data and Sensors 2 and 3 as the simulated data. The average NSE was less than zero for all permutations of adjusted and unadjusted depth, velocity and discharge, except for the adjusted velocity, though this agreement is a result of the use of Sensor 1’s depth and velocity observations for transformation.

The overall level of disagreement between the sensors is substantial, as a hydrologic model is considered to be acceptably calibrated when modelled hydrographs match observed hydrographs with $NSE \geq 0.5$. Only the unadjusted discharge data from the 3/15/2015 storm met this criteria ($NSE = 0.55$). These results suggest that when calibrating a stormflow model to observed discharges measured by an ADV, a modeler should consider whether the uncertainty in the flow measurement data merits a lowering of the acceptable NSE value.

The storm peak, volume, and their uncertainties were also evaluated using the unadjusted and adjusted observations. Figure 4.6 shows the peak flow, volume, and the range of cumulative uncertainty estimated using each of the sensor's observations for the nine storm events, and illustrates the effect of combined uncertainty on hydrograph characteristics that are commonly used in hydrologic modeling. For example, for the storm event that occurred on 3/26/2015, the unadjusted peak flow uncertainty varied between 7.1 and 24.3 cfs for the two sensors, and the unadjusted volume uncertainty for the storm event on 4/25/2015 ranged between 53,000 cf and 554,000 cf. Figure 4.6 also illustrates the effect that the transformation functions had on the range of uncertainty, as adjustments reduced the magnitude of peak and volume uncertainty for individual sensors for all storms. Overall, the uncertainty bounds constituted 13 - 48% of unadjusted peak flow, 10 - 25% of adjusted peak flow, 31 - 84% of unadjusted volume, and 25 - 41% of adjusted volume.

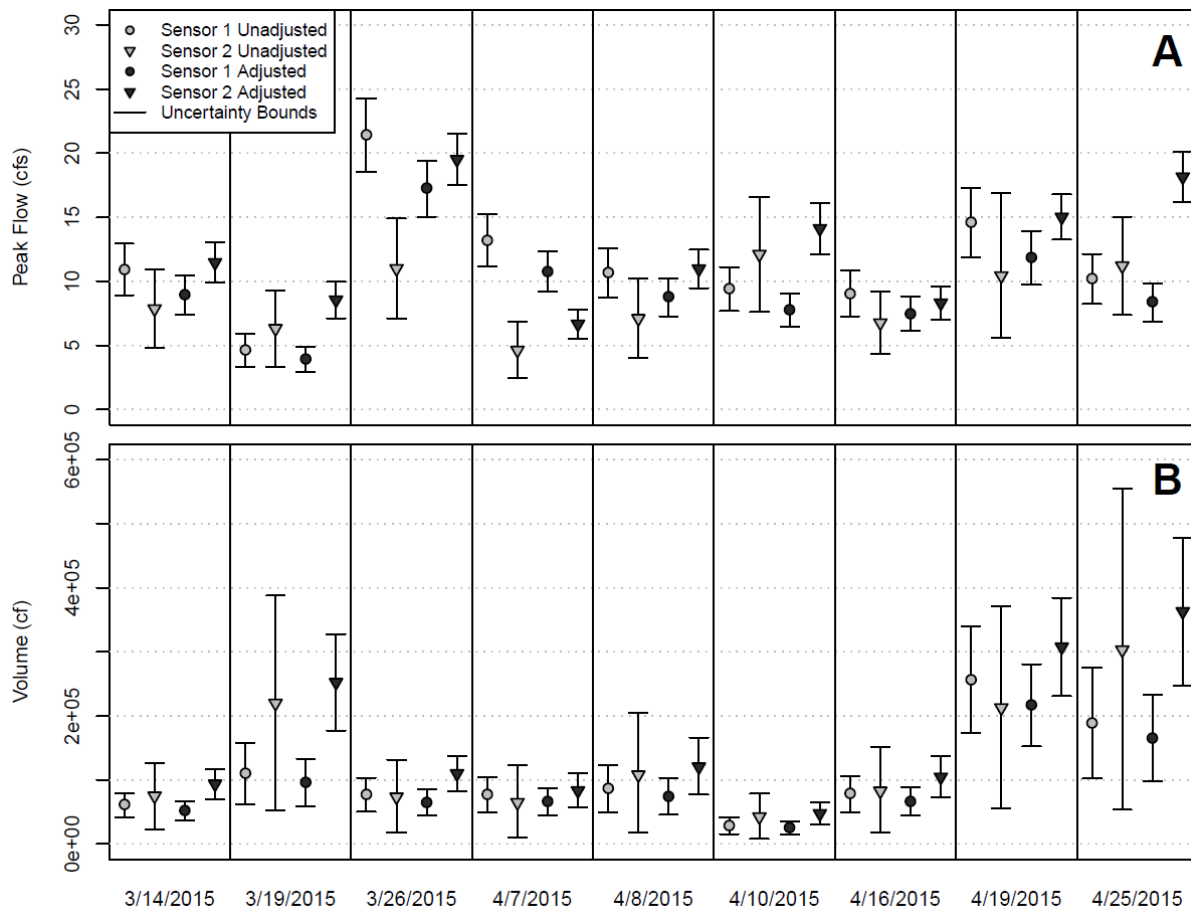


Figure 4.6 – (A) Peak flow and (B) total runoff volume calculated for nine storm events based on original and adjusted discharge observations from Sensor's 1 and 2 with the upper and lower bounds of laboratory benchmarked uncertainty shown

It should be noted that the uncertainties applied to the field data are the combined uncertainties developed in the laboratory without the woody debris field. While it is not known if and to what extent debris existed or affected the sensors during these events, it can be assumed that the field flow conditions are more turbulent than in the laboratory; therefore, the application of laboratory benchmarked uncertainties are conservative. This may be the reason that for some storms, the peak and volume uncertainty bounds for the two sensors do not overlap.

The effect of uncertainty between the sensors over the duration of a storm is further illustrated in Figure 4.7. This figure represents the adjusted depth, velocity, and discharge, and the maximum upper and lower uncertainty bounds of the storm with the best agreement between sensors (4/16/2015). Although the sensors had the highest level of statistical agreement for this storm event, and adjustments to the storm data reduced the peak and volume uncertainty of the sensors by an average of 11% and 28% respectively, the stage, velocity and discharge still exhibit a large range of uncertainty over the duration of the storm.

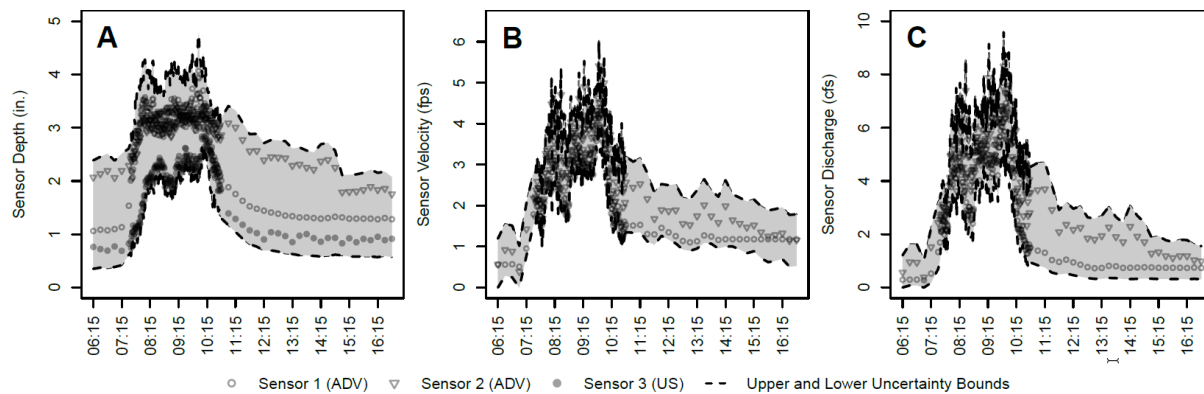


Figure 4.7 – Adjusted (A) depth, (B) velocity, and (C) discharge time series for storm event on 4/16/2015 with observations from three depth sensors, two velocity sensors, and the upper and lower bounds of combined uncertainty from the laboratory. The adjusted discharge observations for this storm event had the highest Nash-Sutcliffe Efficiency of any event recorded.

The demonstrated extent of discharge uncertainty has implications for watershed modeling and calibration. For example, when calibrating a watershed model, the behavior and performance of the model is evaluated through the comparison of simulated and observed storm peaks and volumes. These uncertainties suggest that as with the NSE, the peak and volume results from a simulation should be viewed within the context of the uncertainty of the data. One way to apply these results could be to use the range of uncertainty illustrated in Figure 4.7 to calibrate a watershed model, where the simulated hydrograph would need to fall within the upper and lower uncertainty bounds of the discharge time series (Harmel and Smith, 2007).

Data from the field observations were also evaluated to determine the magnitude of uncertainty associated with average daily flow volumes over the entire study period as a percentage of the adjusted and unadjusted observations (Table 4.4). This shows a considerable amount of uncertainty, with the unadjusted uncertainty bounds constituting between 63 and 99% of average daily volume. However, the magnitude of these values may be affected by the baseflow stage and velocity conditions at the field site which, during dry periods could be less than the combined uncertainty. To test this hypothesis, the average volumes and uncertainties using only discharge data during storm events were computed, and it was found that this resulted in an average decrease of 28% of the uncertainty as a percentage of the average volume.

Table 4.4 - Average Daily Volume Summary

	Unadjusted		Adjusted		Uncertainty as % of Average	
	Average Volume (cf)	Uncertainty (\pm cf)	Average Volume (cf)	Uncertainty (\pm cf)	Unadjusted	Adjusted
Sensor 1	70,364	44,646	62,681	36,122	63%	58%
Sensor 2	147,823	145,772	138,768	97,336	99%	70%
Sensor 3	65,660	53,440	74,783	53,108	81%	71%

The average daily volumes of the unadjusted and adjusted equations vary widely between the three sensors. For example, Sensor 2's unadjusted average daily volume is 2.1 times greater than Sensor 1's, and its adjusted volume is 2.2 times greater. As with the stormflow measurements, these results suggest that although the uncertainty is reduced from the unadjusted to the adjusted data, the adjustments to the stage and velocity measurements do not produce more agreeable average volumes over the study period. These results illustrate the impact that uncertainty can have on calculations and subsequent decisions that are dependent upon flow volumes, such as the estimation of pollutant loads, or the volume reduction capabilities of a stormwater BMP.

CONCLUSIONS

The propagation of the laboratory benchmarked depth and velocity uncertainty to unadjusted field discharge measurements resulted in discharge uncertainties of between 13% and 107% of the observation during storm events and baseflow conditions respectively for Sensor 1, and between 34% and 256% for Sensor 2. McIntyre and Marshall (2008) report ADV uncertainty as $\pm 20\%$ of the observation, and Heiner and Vermeyen (2013) report deviations

of -54.6% to +62.4% of venturi meter discharges for all ADVs and ADCPs tested, with some of the sensors producing considerably better results. These values contrast with the manufacturer reported discharge uncertainties of $\pm 2.4\%$ to 26% for Sensor 1 based on Equation 4.8, and $\pm 2\%$ for Sensor 2 based on the probable error range.

This study was constrained to the use of three depth sensors and two velocity sensors in the field, but discharge measurements with lower uncertainties than those reported in this study appear to be a possibility; further experimentation is needed to benchmark the relationship between laboratory and in-pipe uncertainty using additional sensors. Furthermore, ultrasonics and pressure transducers are used widely as depth ranging devices for structural discharge measurement techniques (e.g. weirs and flumes); further research is needed to determine the effects of depth observation uncertainty on discharge estimation using these devices. It should not necessarily be inferred that the technology underlying the sensors used in this study is not suitable for measuring storm sewer discharge, but rather that conditions in storm sewers provide challenges to accurate water flow measurement that laboratory results may not be able to account for.

This was evident in the lack of agreement between depth observations in the field, even after the transformation equations from the laboratory were applied. It was thought that these adjustments would reconcile sensor observations to each other in the field, but the transformation had little effect, and in some cases caused further disagreement. This could be an artifact of the hydraulic conditions in the rectangular culvert caused by the upstream transition from dual circular culverts, though this was not possible to determine. Storm sewer operators who choose to use autonomous flow sensors in difficult (or dangerous) to access locations will have the same problems verifying sensor measurements with direct observations. As such, the balance between ease-of-access and hydraulic stability should be considered during planning so additional equipment can be purchased if necessary.

The implications of the uncertainties reported in this research for stormwater science vary based on how a sensor is used. An inflow vs. outflow study of a stormwater BMP would need to consider the uncertainty of the instruments used while determining if there was a statistical difference between inflow and outflow volume, peak flow, and pollutant load (if measured). If a sensor is used as the baseline for model calibration and validation, there are two important implications. First, uncertainty of the sensor should be considered when deciding how “good,” a

goodness-of-fit metric needs to be (Dotto et al., 2014; Harmel et al., 2010). Second, if the model is being used to test the effect of a treatment in a pre- vs. post- BMP implementation (i.e. longitudinal) or a treatment vs. control (i.e. paired watershed) study, the uncertainty of the sensor may subsume any effects that the treatment is hypothesized to have.

For urban stormwater management, the importance of observation uncertainty in BMP studies and hydrologic modeling is compounded in the TMDL program, where numeric effluent limits called waste load allocations (WLAs) are developed for MS4 jurisdictions based on calibrated hydrologic models, and stormwater BMPs with a regulator-defined reduction efficiency are prescribed as the means by which these WLAs are met. Margins of safety as percentages added to TMDLs (Dilks and Freedman, 2004), and volume reduction percentages ascribed to BMPs (Geosyntec Consultants and Wright Water Engineers, 2011) are relatively inconsequential if the instruments used for TMDL model calibration and BMP testing have discharge observation uncertainties as reported in this study and others. As such, both the precedent for stormwater regulatory compliance in urban areas, and the compliance methods therein are contingent on discharge measurements that are attributed a level of certainty that the results of this study suggest may not be appropriate.

5. EVALUATION OF STORMWATER CONTROL MEASURE PERFORMANCE UNCERTAINTY

5.1 ABSTRACT

The percent removal metric is widely used in engineering practice for stormwater control measure (SCM) selection and design, as it has been adopted by regulatory agencies, and because it is straightforward to use for load reduction calculations. Various studies have shown that the metric is poorly-suited to the episodic and stochastic nature of stormwater flows, but the extent of uncertainty introduced by the use of the metric has not been quantified. This study evaluates SCM performance uncertainty for total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS) load reduction in 15 regulator-defined SCM categories, based on the existing SCM monitoring literature. The width of category uncertainty bounds range from 14% to 171%, with median widths of 82%, 62%, and 71% for TP, TN, and TSS respectively. It was unclear if many of the SCM categories would increase or reduce loading from the contributing drainage area, and additional data did not appear to reduce uncertainty. Bounds of TP load reduction in bioretention were 2 to 98%; economic consequences of this range are explored using an example bioretention cell. The challenge of making SCM spending decisions with large uncertainty bounds, and alternate methods for SCM performance characterization are discussed.

5.2 BACKGROUND

The hydrologic changes to urban streams observed for over four decades (Leopold, 1968), and the subsequent effects on riparian geomorphology, ecology, and biochemistry, have been attributed to the magnitude of impervious surfaces in contributing watersheds (Brabec, 2009), and their efficient connectivity to receiving waters through storm sewer drainage systems (Elmore and Kaushal, 2008; Walsh, Roy, et al., 2005). The legal responsibility to manage and mitigate these effects was delegated to local governments through the Clean Water Act's (CWA's) National Pollutant Discharge Elimination System (NPDES) and Total Maximum Daily Load (TMDL) programs (33 USC 26, §§ 402 and 303, respectively). TMDL studies model the sources of pollutants in impaired water bodies, and recommend structural stormwater control measures (SCMs) as the remedial action for the impairment. This recommendation becomes prescriptive in urban areas by the requirement of a TMDL "Action Plan" in annual NPDES

municipal separate storm sewer system (MS4) permit reports, wherein permittees must report the number and type of SCMs constructed toward TMDL requirements.

The required cost-benefit analysis of the NPDES stormwater program showed a net benefit, with *ex ante* estimates of the cost of the program to municipalities, industry, and State/Federal authorities of \$14.5 million for large and medium MS4 entities (USEPA, 1990), and between \$848 and \$982 million for small MS4 entities (Science Applications International Corporation, 1999b; USEPA, 1999). It should be noted that this did not include the cost associated with TMDL Action Plans, and estimates of the SCM capital, operations, and maintenance investment needed to demonstrate progress toward TMDL effluent limitations for MS4 permittees has been estimated for several individual entities in the hundreds of millions, or even billions (AMEC, 2012; Virginia Senate Finance Committee, 2011).

This economic impact has been further demonstrated by the national proliferation of stormwater utility fees and taxes, designed to create a steady funding source for stormwater programs (Kea and Dymond, 2016). Local government authorities must now decide how to best spend these resources on SCMs, and mandate or incentivize the construction of SCMs on private development, in order to reduce pollutant loading to meet their waste load allocation (WLA) specified TMDL reports. These decisions are frequently based on an estimated cost-effectiveness of SCMs – the cost of an SCM divided by the ability of the SCM to reduce pollutant mass loading on an annual basis (e.g. cost per pound of total phosphorus removed). As substantial economic throughput depends on SCM cost-effectiveness, it is important to understand and quantify the uncertainty associated with the metric. This study focuses on the uncertainty in the denominator of the cost-effectiveness metric: the mass load removal effectiveness of SCMs.

5.3 INTRODUCTION

The need to estimate SCM cost-effectiveness has led to the widespread practice of estimating the mass “treated” by an SCM as the influent mass loading from the contributing drainage area (CDA) multiplied by the “percent removal” of the SCM. The percent removal metric has long been used for waste water treatment processes (e.g. Nesbitt, 1969), but its application for estimating pollutant of concern (POC) removal in stormwater fluxes is newer, and

frequently attributed to the Center for Watershed Protection's (CWP's) Runoff Reduction Method (RRM) (Battiata et al., 2010; CWP, 2008).

The RRM estimates a total mass load reduction percentage (TR), based on two components of POC removal – “runoff reduction” (RR), and “pollutant removal” (PR). RR is the magnitude of runoff volume abstracted (i.e. infiltrated, evaporated, or transpired) by an SCM as a proportion of the influent volume, estimated as:

$$RR = 1 - \frac{(\sum_{i=1}^n Q_i \Delta t_i)_{out}}{(\sum_{i=1}^n Q_i \Delta t_i)_{in}} \quad \text{Equation 5.1}$$

where Q_i is the average volumetric flow rate over Δt_i , the time interval at which $i = 1 \rightarrow n$ individual flow measurements are taken. RR was intended to provide an estimate of the volume abstracted over an annual duration. PR was intended to characterize the SCM's biogeochemical nutrient and sediment removal capacity as a proportion of the influent event mean concentration (EMC), estimated as:

$$PR = 1 - \frac{\left(\frac{\sum_{i=1}^n C_i Q_i \Delta t_i}{\sum_{i=1}^n Q_i \Delta t_i} \right)_{out}}{\left(\frac{\sum_{i=1}^n C_i Q_i \Delta t_i}{\sum_{i=1}^n Q_i \Delta t_i} \right)_{in}} \quad \text{Equation 5.2}$$

where C is a mass concentration, and the numerator and denominator are the EMC_{out} and EMC_{in} respectively. PR is intended to provide a flow-averaged estimate of the reduction in POC concentration across an SCM.

RR and PR values for each SCM category were developed by the CWP, by synthesizing the results of individual SCM studies (referred to as “source studies”) into a single document referred to in this study as “the RRM” (CWP, 2008). In general, these source studies monitor an SCM, and report RR and PR (or the individual components needed to calculate RR and PR) for either (1) individual storm events, (2) monthly, yearly, or study duration totals, or (3) a measure of central tendency and/or dispersion of (1) or (2). The RRM uses these individual RR and PR values to ascribe collective RR and PR values to SCM categories (e.g. “Infiltration”, “Retention Pond”) based on expert judgement of a conservative central tendency of the results of all source studies in a given category. The RR and PR values ascribed to each SCM category are then used as input for Equation 5.3, defining Total Reduction (TR), also developed as part of the RRM:

$$TR = RR + [PR \times (1 - RR)]$$

Equation 5.3

the intention of TR was to ascribe an overall nutrient and sediment load reduction to each SCM category based on hydrologic (RR) and biogeochemical (PR) removal capabilities of total phosphorus (TP) and total nitrogen (TN). Total suspended solids (TSS) percent removal values were not ascribed as part of the RRM, but were later developed in (Simpson and Weammert, 2009).

The RRM is used in this study, as it is codified in Virginia's 2013 Stormwater Management Program Regulations (9 VAC 25-870-65), and because its TR values and calculations are the standard practice for engineering design calculations and regulatory compliance. Design engineers use the RRM to demonstrate that the SCMs proposed on a new or re-development project will sufficiently prevent excessive flooding, erosion, and export of nutrients and sediment. Local governments use the TR values from the RRM to estimate how much progress has been made toward their TMDL Waste Load Allocation (WLA) (VDEQ, 2015), and to determine which SCMs will be the most cost-effective given those requirements (e.g. Kotula and Capiella, 2013).

There are other states that use the RRM and associated TR values for SCM selection and design (e.g. ARC et al., 2016; WVDEP, 2012), but the percent removal metric is used even more widely in stormwater regulations than the RRM. For example, the TR values in the RRM are the basis for another literature review by (Simpson and Weammert, 2009), which produced the SCM efficiencies used in the Chesapeake Bay model (USEPA, 2010b), and several other Chesapeake Bay states (and otherwise) use the percent removal metric in guidance manuals (e.g. NJDEP, 2004; PDEP, 2006). The regulatory acceptance of the RRM and the percent removal metric, along with the significant cost of SCM construction and operation, emphasizes the importance of understanding the limitations of the method and percent removal values, resulting from SCM performance uncertainty.

The value of the RRM is its straightforward application for engineering calculations, as the method can be used for any configuration of SCM. It also incorporates both runoff volume abstraction, and the reduction in concentration provided by an SCM, to estimate a mass load reduction, which more faithfully represents stormwater treatment processes than the concentration reduction paradigm typical in the waste water treatment literature. The simplicity

of the method, and its inclusion of both RR and PR also provides regulators with an equitable method to track progress toward TMDL WLAs.

The shortcomings of the RRM are mostly a result of the method's black-box approach to stormwater treatment. The principal of ascribing a single value of TR to a given category of SCM (e.g. permeable pavement) does not account for the large variability in that type of SCM's performance, attributable to three sources. First, the TR performance of SCMs is subject to the stochasticity of the influent POC concentration, and precipitation intensity, duration, frequency, and inter-event time, referred to as "environmental uncertainty". Second, the measurement of influent and effluent flow and concentration in an SCM is subject to uncertainty resulting from the various measurement methods employed, referred to as "observation uncertainty". Finally, these two sources of uncertainty are compounded when multiple individual SCMs of varying design, age, and condition are categorized as a single SCM type, referred to as "category uncertainty". Although these three sources of uncertainty are commonly acknowledged in source-studies, and even in literature reviews, variability and uncertainty disappear in the final regulator-defined values, and in their widespread use in stormwater engineering practice.

These uncertainties inherent to the RRM are at odds with its ease of use and subsequent widespread acceptance, which is further heightened by the scale of the current and projected expenditures on SCM construction, tracking, and maintenance in the Chesapeake Bay and other TMDL watersheds (Copeland, 2012; Virginia Senate Finance Committee, 2011). This tension necessitates a quantification of the uncertainties associated with the regulator-defined percent removal metrics for each SCM category, as it is not clear if the large amounts of resources spent on SCMs will result in POC removal according to these defined values. Therefore, the purpose of this study is to systematically evaluate and present the uncertainty associated with the regulator-defined SCM removal efficiencies used in Virginia. To do this, a review of source studies is performed, that includes the literature used for the TR values in the RRM, plus newer studies published since 2008. The results of these studies as RR or PR are systematically catalogued, and measures of observation and environmental uncertainty are ascribed to each source study. Finally, observation and environmental uncertainty is combined across source studies which evaluated SCMs of the same category to ascribe values of SCM category uncertainty.

5.4 METHODS

In order to estimate the components of uncertainty, 163 SCM source studies comprising 308 unique SCMs or laboratory simulations of SCMs were gathered and organized in an *R* database (R Core Development Team, 2014). All studies used for the development of the RRM (CWP, 2008) were reviewed, as were studies published after the document was released in 2008. These studies included peer-reviewed journal papers, conference papers, USGS reports, other technical reports, and theses/dissertations. In addition, volume reduction data and pollutant reduction data were downloaded from the 12/30/2014 version of the International Stormwater BMP Database (ISWBMPDB), and included in the *R* database. Source documentation for ISWBMPDB data were gathered so that important details of the study design could be evaluated, though in certain cases the data from the ISWBMPDB had not been documented elsewhere, and is referred to as “unpublished data”.

Each source study was reviewed, and the available results were extracted and entered into the *R* database along with important details pertaining to the monitoring methods used at the influent and effluent to each SCM. In addition, different studies calculate RR and PR over different durations (e.g. storm event, month), and present data as either the individual data points, or as summary statistics of the individual data points – this was also recorded. The highest resolution of data available from each source study was extracted, and is referred to as “source study data points”, i for each SCM studied, j . Each SCM studied was classified into $k = 15$ SCM categories defined by the Virginia DEQ’s BMP Standards and Specifications (VDEQ, 2013 and Table 5.1), though this was a general classification, as it was not possible to determine if the design and construction of each SCM conformed to the DEQ’s specifications (it is unlikely that they would). Sheet flow to conserved open space and sheet flow to vegetated filter strips were treated as a single category, as it was difficult to categorize source studies into one of these two sub-categories.

Stormwater Control Measure Category	Ann. Runoff Vol. Red. (%RR)		TP EMC Removal (%PR)		TN EMC Removal (%PR)		TP Load Removal (%TR)		TN Load Removal (%TR)	
	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2
Rooftop (Impervious Surface) Disconnection ^A	50	25	0	0	0	0	50	25	50	25
Sheet flow to:										
Conserved Open Space ^A	75	50	0	0	0	0	75	50	75	50
Vegetated Filter Strip ^B	50	50	0	0	0	0	50	50	50	50
Grass Channels ^A	20	10/20 ^C	15	15	20	20	32	24/32 ^C	36	28/36 ^C
Soil Compost Amendment	Increases credit for other SCM types		0	0	0	0	Increases credit for other SCM types			
Vegetated Roof	45	60	0	0	0	0	45	60	45	60
Rainwater Harvesting	≥ 90		0		0		≥ 90		≥ 90	
Permeable Pavement	45	75	25	25	25	25	59	81	59	81
Infiltration	50	90	25	25	15	15	63	93	57	92
Bioretention	40	80	25	50	40	60	55	90	64	90
Dry Swale	40	60	20	40	25	35	52	76	55	74
Wet Swale	0	0	20	40	25	35	20	40	25	35
Filtering Practices	0	0	60	65	30	45	60	65	30	45
Constructed Wetland	0	0	50	75	25	55	50	75	25	55
Wet Ponds	0	0	45/50 ^D	65/75 ^D	20/30 ^D	30/40 ^D	45/50 ^D	65/75 ^D	20/30 ^D	30/40 ^D
Extended Detention Pond	0	0	15	15	10	10	15	31	10	24

^ALevel 1 and 2 values are for use in hydrologic soil group (HSG) A/B or C/D respectively

^BLevel 1 and 2 values are for use in HSG A or B/C/D respectively

^CThe larger value can be obtained if compost soil amendments are used

^DThe lower value is for SCMs that are influenced by groundwater

Table 5.1 - Stormwater control measure (SCM) categories used in the Virginia Department of Environmental Quality's Best Management Practice (BMP) Standards and Specifications (VDEQ, 2013). L1 and L2 designate Level 1 and 2 design criteria for each SCM category, where L2 requires increased size, depth, or a more restrictive soil specification. Sheet Flow to Conserved Open Space and Sheet Flow to Vegetated Filter Strip were treated as a single category for this study.

After reviewing these studies, it was found that the sources of uncertainty could be defined as follows: (1) “Environmental Uncertainty”, u_{env} is generated by the stochastic nature of runoff processes and constituent concentrations in stormwater flows, and the subsequent variability in SCM performance. Environmental uncertainty is also referred to in the literature as “aleatory uncertainty” - variability arising from the underlying stochasticity of natural processes (Merz and Thielen, 2005). (2) “Observation Uncertainty”, u_{obs} is introduced by the measurement techniques used to quantify runoff volume and POC concentrations. This includes uncertainty associated with field sampling, sample preservation and storage, and laboratory characterization. Observation uncertainty is also referred to in the literature as “epistemic uncertainty” - incomplete knowledge due to the limitations of measurement (Merz and Thielen, 2005).

The two sources of uncertainty associated with the study of an individual SCM are then propagated when j SCMs are assumed to be compatible in their results and conclusions, and classified into a single SCM category, k . This is referred to as “Category Uncertainty”, u_{cat} , and is not a new source of uncertainty, but rather the product of distilling the performance of SCMs of varying age and design into a single category-based performance metric (RR or PR). Figure 5.1 demonstrates the combination of uncertainties for an example SCM category, such that environmental and observation uncertainty are shown as horizontal dimensions for each source study, and category uncertainty for the example SCM type is shown as the dark gray background region.

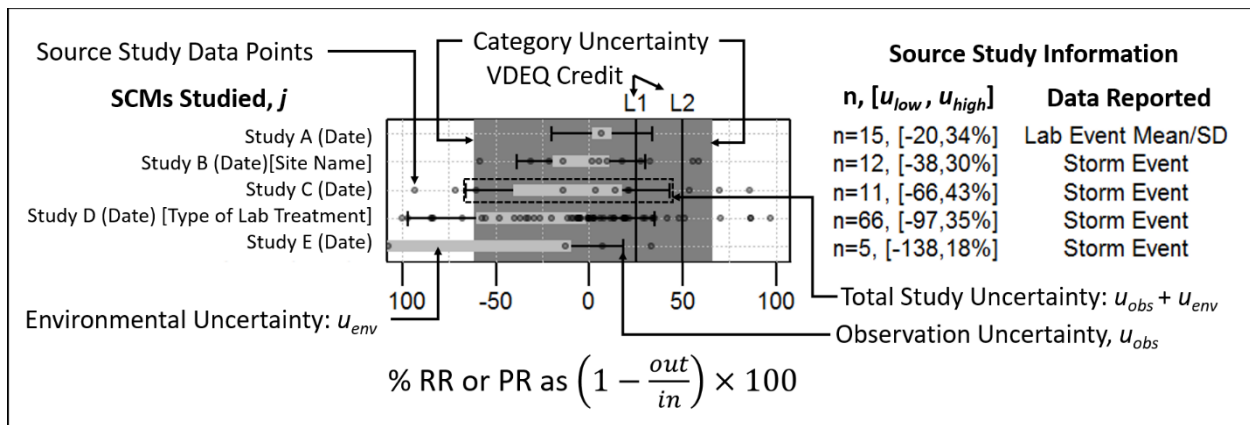


Figure 5.1 - The organizational structure of i data points in j SCM source studies for an example SCM category k shown with dimensions of environmental and observation uncertainty. Category uncertainty for this example category is shown as the dark gray background region.

For each SCM studied, u_{env} was estimated as the dispersion of individual data points in the study, and u_{obs} was estimated by attributing uncertainty to the monitoring methods reported. The u_{obs} was added bi-directionally to u_{env} for each study, and this sum was used to estimate u_{cat} for each SCM category and POC: runoff reduction (RR) and pollutant removal (PR) for total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS). Finally, the upper and lower RR and PR u_{cat} bounds were input into Equation 5.3 to estimate the uncertainty associated with TR – the overall mass load removal capability of each SCM category. The methods used to estimate u_{env} , u_{obs} , and u_{cat} are detailed in the following sections, but prior to this, two important considerations are noted.

First, the use of the RRM as the basis for estimating SCM uncertainty in this study should not be taken as an acknowledgement that it is the most effective or scientifically defensible method for quantitatively describing SCM pollutant removal dynamics. Rather, we use the RRM because of its regulatory adoption as an engineering design and analysis tool, and its subsequent economic importance. Second, the purpose of this study is to evaluate the uncertainty associated with this existing SCM design methodology, and as a result, it was necessary to estimate uncertainty in a manner that was compatible with the RRM. This required the estimation and reporting of uncertainty as \pm uncertainty bounds. This approach to uncertainty is still well-accepted in the literature (see e.g. Coz et al., 2012; K. Lee et al., 2014), and may be simpler to understand and apply in engineering practice (Willink and White, 2012), although it does not completely capture the probabilistic dynamics of uncertainty. Each \pm uncertainty bounds actually represents a probability distribution describing the range of observed values (Granato, 2014), however a probabilistic approach to uncertainty did not fit with the RRM's deterministic framework. Uncertainty bounds as \pm percentages of an actual value are estimated in the following sections, but discussion of underlying distributions are given where appropriate.

5.4.1 Environmental Uncertainty

When a given SCM is studied, the RR and PR values for i storm events (or other time durations) monitored in the study vary with the influent POC concentration, the precipitation characteristics of the monitored storm, and other random events in the SCM and contributing drainage area (CDA) (Koch et al., 2014). This is referred to as “Environmental Uncertainty”, u_{env} , and this section describes the sources of this uncertainty for RR and PR respectively.

Runoff reduction, RR, is the volume of water that is infiltrated, evaporated, or evapotranspired by an SCM as a proportion of the influent runoff volume from the CDA (Equation 5.1). The generation of runoff from a CDA is a joint probability distribution of precipitation depth, intensity (depth divided by duration), and inter-event time duration (Guo and Gao, 2015), and SCMs are designed to retain and abstract runoff up to a given precipitation frequency. Therefore, the RR of an SCM varies from 100% for small, frequent storm events up to a threshold inflow volume and rate from the CDA, at which point RR decreases to asymptotically approach zero. RR therefore takes a non-normal distribution, bounded by $[-\infty, 100\%)$, with a high density at 100% and a long left tail; this is shown theoretically in (Wild and Davis, 2009) and empirically in (Carmen et al., 2016; Davis, 2008; Winston et al., 2016). Inter-event time adds randomness to this distribution, if an SCM's infiltration capacity is affected by latent soil-moisture from previous events (Mangangka et al., 2015).

The mechanics of RR are different for several SCMs that are not designed for runoff volume reduction. Detention ponds are typically designed to alter peak flow and time-to-peak, but not volume reduction, and therefore do not normally provide RR benefit (Emerson et al., 2005). Perennially wet practices – retention ponds, wetlands, and wet swales – are also typically designed for attenuation and other pollutant removal functions, but not volume reduction, and in many cases it is not possible to accurately measure all hydrologic inputs and outputs to these SCMs (Geosyntec Consultants and Wright Water Engineers, 2011); as a result these SCMs are not included in the analysis of RR uncertainty.

The measured pollutant removal (PR) capacity of an SCM is a function of the influent concentration (McNett et al., 2011), which can vary based on the inter-event pollutant build up, and depth of previous rainfall events (H. Lee et al., 2004). This is especially true as influent concentrations approach the irreducible concentration – the concentration below which no further POC concentration reduction is possible (Granato, 2014; Schueler and Holland, 2000). Another source of uncertainty is the non-linear dynamics of POC transport and deposition from the inflow to the outflow points of the SCM. High EMCs observed at the influent may be retained in the SCM for a given storm event, leading to a high PR, but these retained POCs may be re-suspended and transported to the effluent during the subsequent event, leading to a lower PR (e.g. Gain, 1996; Yu et al., 2001). An appropriate probability density function for PR would also be bounded by $[-\infty, 100\%)$, with a high density between 0 and 100% and a long left tail.

As RR and PR values take the form of continuous, left-skewed distributions bounded by $[-\infty, 100\%)$, but these distributions have not been well specified, a robust measure of dispersion – the interquartile range (IQR) – was used to describe u_{env} for each individual SCM studied. The IQR was chosen because of its insensitivity to outliers, handling of asymmetric distributions, and precedence in the literature (Geosyntec Consultants and Wright Water Engineers, 2014; Helsel and Hirsch, 1994). The upper and lower bounds of u_{env} are defined as the first and third quartile of all RR or PR values in a given study.

Individual RR and PR values for storm events were not provided in all source studies, but were reported as measures of location and scale (e.g. mean and standard deviation) in 40% of all the studies reviewed. The metric used to summarize the range of RR and PR values varied; the most commonly reported summary statistics were means and medians of storm events, each comprising 10% of the studies reviewed. Other reported statistics included range and standard deviation, and in several studies RR and PR were estimated based on a month or year duration, not an individual storm event. These varying measures of location and scale cannot be compared directly to the IQR of storm events, but the u_{env} for these studies was set based on the best available statistic for the study, and the incompatibility of reported statistics across studies was incorporated in the estimation of category uncertainty.

5.4.2 *Observation Uncertainty*

The estimation of RR and PR for a given SCM study requires that runoff volume and/or EMC be measured at the influent and effluent of an SCM, or at a treatment (SCM) and control drainage area. The various hydrometric methods used to measure volume and EMC introduce uncertainty to the study results, and the procedures used to estimate this observation uncertainty, u_{obs} for each study is discussed in this section.

In order to calculate the runoff volume and EMC for a given time duration, observations of flow rate and concentration must be taken at intervals at the influent and effluent points to the SCM. Flow rate can be measured using any number of techniques, but is constrained by the (typically) piped configuration of the influent and effluent. The methods reported for monitoring flow rate in the 170 source studies reviewed are summarized in Table 5.2, along with values of uncertainty from the literature. The uncertainty value used for this study is the median of values from the various sources cited.

Table 5.2 - Uncertainties associated with various methods of estimating inflow and outflow volume for SCMs as a percent of the measurement. Where error distributions were available, the median was used. Where error was reported in units of the measurement, it was converted to a percentage based on a reasonable range of observations for SCM studies.

Measurement Type	Literature ± Uncertainty (%)	±Uncertainty Used (%)	Reference(s)
Tipping Bucket Rain Gage			
As a flow meter (no wind effects)	4 - 8%	6%	(Sieck et al., 2007)
For precipitation measurement (wind effects included)	5 – 15%	10%	(Neff, 1977)
Depth Measurements			
Ultrasonic Transducer	0.4 – 2.2%	1.3%	(Aguilar et al., 2016; Harmel et al., 2006)
Bubbler	0.14 – 5.3%	0.45%	(Dedrick and Clemmens, 1984; Harmel et al., 2006; Kirby, 1991)
Pressure Transducer	5.6 – 7.3%	6.4%	(Aguilar et al., 2016)
Float Recorder	1.4 – 2%	1.7%	(Harmel et al., 2006)
Conversion of Depth to Discharge (excluding depth and precipitation observation uncertainty)			
V-notch Weir - construction and discharge coefficient uncertainty	0.5 - 2.2%	2%	(Herschly, 1995; Krueger et al., 2010)
Standing Wave Flume	4%	4%	(Herschly, 1995)
Stage – Discharge relationship	6 – 20%	13%	(Harmel et al., 2006)
Manning’s Equation w/stable, uniform channel; accurate “n”	15%	15%	(Harmel et al., 2006)
NRCS Curve Number Equation	8 – 28%	18%	(Boughton, 1989; D’Asaro and Grillone, 2012)
Other Discharge Gages			
Acoustic Doppler Velocimeter discharge uncertainty (combined depth and velocity uncertainty)	8.4 - 13%	10%	(Aguilar et al., 2016; Heiner and Vermeyen, 2013; Maheepala et al., 2001)
Mechanical Flow Meters for Pressure Pipes	2 – 3%	2.5%	(AWWA, 2012)

Several of the methods shown in Table 5.2 are individual components of discharge measurement, and require two of the methods shown to estimate discharge. A commonly used example of this, is the combination of a depth gage and a primary device (e.g. weir or flume). For measurement methods that require multiple techniques, the probable error range (PER) was used to combine the uncertainty of the constituent techniques such that:

$$u(y) = \sqrt{\sum_{i=1}^N u^2(x)_i} \quad \text{Equation 5.4}$$

where $u(y)$ is the combined uncertainty of N independent components of uncertainty, $u(x)_i$ - in this case, the components of discharge measurement (Bertrand-Krawjewski and Muste, 2008a; Sauer and Meyer, 1992).

The measurement of nutrient and sediment EMC in the influent and effluent of SCMs introduces several additional components of observation uncertainty. The components of EMC measurement were defined as: (1) sample collection, (2) sample preservation and storage, and (3) laboratory analysis. Uncertainty associated with the field (or laboratory) collection of samples is a function of the type of sampling (e.g. manual sample, time/flow weighted auto-sample), and the time interval at which a sample is taken (Grizzard et al., 1976; McDonald et al., 2016). Uncertainty can also be introduced by poor sampling location, though this was not included, as the size of conveyances sampled are generally small, and because it was not possible to ascribe a value from the information in most studies. As the information from source studies regarding the specific details of sampling was not consistently available, a central value from the meta-analysis in (Harmel et al., 2006) was used, such that the u_{obs} associated with manual sampling was 27.5%, time-weighted auto-sampling was 21%, and a flow-weighted auto-sampling was 11%.

After a sample is collected in the field, various physical and bio-chemical processes can alter the measured concentration of nutrients; an additional source of uncertainty (Jarvie et al., 2002). Standard analytical methods generally specify that samples be analyzed as soon as possible after collection, or that they be acidified, refrigerated, or frozen (see References in Table 5.3). As the source studies reviewed do not consistently report sample handling, it was assumed that all samples were iced and analyzed within six hours of collection, and uncertainty values of 7% and 3% for TP and TN respectively were used based on central tendencies from Harmel et al. (2006) and Kotlash and Chessman (1998).

Table 5.3 - Analytical uncertainty associated with laboratory characterization of phosphorus, nitrogen, and total suspended solids using several methods. Observation uncertainty, u_{obs} is estimated as the probable error range of bias and precision uncertainty

Method	Method Description	Uncertainty Estimation	u_{bias} (±)	$u_{precision}$ (±)	$u_{obs,lab}$ (±)	Reference(s)
Phosphorus						
S.M. 4500-P B, E, F.	Digestion and Ascorbic Acid Reduction	Digestion by Persulfate or Sulfuric-Nitric Acid, Colorimetry by manual or automated ascorbic acid. Bias and precision as median of values reported in S.M. 4500 P B.4,5; E; F	4.0%	7.5%	8.5%	(APHA et al., 2012)
EPA 200.7	Inductively Coupled Plasma Spectrometry	Bias: average recovery of low spike, Precision: relative percent different in duplicate spikes	2%	9.4%	9.6%	(USEPA, 1994)
EPA 365.1	Semi-Automated Colorimetry	Bias: Median absolute residual from true value, Precision: Median relative std. dev.	2.3%	7.4%	7.8%	(USEPA, 1993a)
EPA 365.2	Colorimetric, Ascorbic Acid, Single Reagent	Used the PER of bias and precision for organic P and ortho-P. Bias: median absolute bias, Precision: median relative standard deviation	4.9%	27%	27%	(USEPA, 1971a)
EPA 365.3	Colorimetric, Ascorbic Acid, Two Reagent	Bias: Single laboratory recoveries of 7.6 and 0.55 mg P/L were 99 and 100%	1%	N/A	1%	(USEPA, 1978a)
EPA 365.4	Colorimetric, Automated, Block Digester AA II	Single laboratory. Bias: absolute value of recoveries (2 samples), Precision: median given precision (3 samples)	4%	3%	5%	(USEPA, 1974a)
Hach Method 8190	Acid Persulfate Digestion	Precision: stated as 95% confidence intervals on 3.00 mg P/L standard. Bias not given	N/A	2.3%	2.3%	(Hach, 2009)
Nitrogen – Total Nitrogen is frequently estimated as the sum of Total Kjeldahl Nitrogen (TKN), Nitrate and Nitrite-N						
TKN – S.M. 4500 N _{org} B.	Macro-Kjeldahl	Single laboratory, 18 samples of nicotinic acid. Bias: median recovery. Precision: median relative std. dev.	4.5%	3.3%	5.6%	(APHA et al., 2012)
S.M. 4500-NO ₃ ⁻ E,F,H,I	Reduction – Cadmium or Hydrazine, Auto or Manual	Used median bias and precision data from four reported reduction methods	4.0%	1.1%	4.1%	(APHA et al., 2012)
S.M. 4500-NO ₂ ⁻ B.	Colorimetric Method	Single laboratory, 4 sample. Bias: median of recoveries, Precision: median relative std. dev.	0%	1.3%	1.3%	(APHA et al., 2012)
TKN – EPA 351.1	Colorimetric, Automated Phenate	Six laboratory, 4 sample. Bias: median absolute bias, Precision: median std. dev as percent of lab standard	24%	28%	37%	(USEPA, 1978b)
TKN – EPA 351.2	Semi-Automated Colorimetry	Inter-laboratory, 12 sample. Bias: median absolute bias, Precision: median std. dev as percent of lab standard	1.1%	9.3%	9.3%	(USEPA, 1993b)
TKN – EPA 351.3	Colorimetric, Titrimetric, Potentiometric	31 analysts, 20 labs, 4 samples. Bias: median absolute bias, Precision: median std. dev as percent of lab standard	3.7%	53%	53%	(USEPA, 1978c)
NO ₃ ⁻ - EPA 300.0	Ion Chromatography	19 laboratories. Bias: median absolute bias. Precision: median std. dev as percent of lab standard	2.8%	7.2%	7.7%	(USEPA, 1993c)
NO ₂ ⁻ - EPA 300.0	Ion Chromatography	19 laboratories. Bias: median absolute bias. Precision: median std. dev as percent of lab standard	3.1%	5.1%	5.9%	(USEPA, 1993c)
NO _{2,3} ⁻ - EPA 353.2	Automated Colorimetry	3 laboratories, 4 samples. Bias: median absolute bias. Precision: median std. dev as percent of lab standard	5.1%	10%	12%	(USEPA, 1993d)
NO _{2,3} ⁻ - EPA 353.3	Spectrophotometric, Cadmium Reduction	Single laboratory, 4 samples. Bias: median of recoveries, Precision: median of relative std. dev	0%	1.3%	1.3%	(USEPA, 1974b)
Hach Method 10071	Persulfate Digestion	Precision: stated as 95% confidence intervals on 10.0 mg N/L standard. Bias not given	N/A	4%	4%	(Hach, 2014)
Total Suspended Solids						
S.M. 2540 D	Total Suspended Solids Dried at 103 – 105° C	Two analysts, four replicates. Bias: Slope of SiO ₂ vs. standard taken at mid-depth (Clark and Siu, 2008a). Precision: median std. dev as percent of lab standard from Standard Methods	7.0%	10%	12%	(APHA et al., 2012; Clark and Siu, 2008a)
EPA 160.2	Residue, Non-Filterable	Precision and accuracy not available in method. See (Clark and Siu, 2008a) Bias: Slope of SiO ₂ vs. standard for 100 mL sample. Precision: ½ width of 95% confidence interval on slope	34%	3.2%	34%	(Clark and Siu, 2008a; USEPA, 1971b)
ASTM D3977-97(B)	Filtration	9 laboratory, 3 sample. Bias: median absolute bias, Precision: median std. dev as percent of lab standard.	9%	5.3%	10%	(ASTM, 2013)

Finally, the laboratory methods used to estimate constituent concentrations are subject to uncertainty caused by instrument and human error. This is generally reported in the method documentation as “Bias” and “Precision”, which is denoted in Table 5.3 as u_{bias} and $u_{precision}$ respectively, and was propagated using the PER (Equation 4.3) to estimate an overall laboratory observation uncertainty $u_{obs,lab}$ for the method.

In order to estimate an overall u_{obs} for an individual SCM studied, all components of flow and EMC monitoring at the influent and effluent were noted, and a value of uncertainty was ascribed to each individual component using Table 5.2 and Table 5.3. If a source study did not provide monitoring information, the median value from each monitoring component was ascribed to the SCM studied. The PER (Equation 4.3) was then used to propagate the component uncertainties to arrive at the overall u_{obs} for the SCM studied. An example of the estimation of overall u_{obs} for nutrients in a study that uses depth sensors and a primary device to measure flow is shown in Equation 5.5:

$$u_{obs} = \sqrt{[u_{depth}^2 + u_{primary}^2 + u_{sampling}^2 + u_{handling}^2 + u_{bias}^2 + u_{precision}^2]_{in,out}} \quad \text{Equation 5.5}$$

where uncertainty is designated as u and the subscripts represent uncertainty due to depth measurement, the conversion of depth to discharge using a primary device, sampling method (e.g. manual, time/flow weighted interval), sample handling, laboratory bias, and laboratory precision respectively. The subscripts *in,out* designate that the sum of squares must include measurement components at both the influent (or reference) and effluent of the SCM, which may be different, depending on the structure’s configuration.

Several assumptions were necessary in order to estimate u_{obs} for the source studies. First, this method of ascribing uncertainty does not account for varying levels of quality control and other random factors that contribute to u_{obs} in a study, however this was the extent of what was possible given the details provided in source studies, and is sufficient for the goals of this study. Second, the observation uncertainties shown in Table 5.2 and Table 5.3 are associated with a single measurement of discharge or concentration, integrated over a given time duration as shown in Equation 5.1 and Equation 5.2, in order to generate a single value of RR or PR. Applying these uncertainties to individual observations was not possible, as individual observations are rarely reported in source studies. Instead, the uncertainties in Table 5.2 and Table 5.3 were added bi-directionally to the RR and PR values, which assumes that the

uncertainty associated with a single discharge or EMC observation propagated over the effluent hydrograph or pollutograph, and divided by the influent hydrograph or pollutograph, will be at least as large as the uncertainty of an individual measurement.

5.4.3 Category Uncertainty

Once u_{obs} and u_{env} had been estimated for each individual SCM studied (j), they were combined by adding and subtracting u_{obs} from the upper and lower bounds of u_{env} , as shown in Figure 5.1. The PER was not used to propagate u_{obs} and u_{env} , as it was determined that the effect of observation uncertainty was a bi-directional shift of each individual RR or PR value, which leads to a proportional bi-directional shift of u_{env} . The overall uncertainty associated with an SCM studied was then used to estimate a category uncertainty for the k individual SCM types described in Table 5.1. To do this, the upper and lower uncertainty bounds of each SCM studied were weighted according to the number of individual RR or PR data points available in the study (n_j), and the type of statistic reported (w_j), as shown in Equation 5.6.

$$u_{cat_k} = \frac{\sum_{j=1}^n (n_j \times w_j \times [u_{env,lower} - u_{obs}, u_{env,upper} + u_{obs}]_j)}{\sum_{j=1}^n (n_j \times w_j)} \quad \text{Equation 5.6}$$

A variety of different statistics were reported (discussed previously), and as these statistics are not directly comparable, each study was weighted based on the proportion of the overall population that the statistic is meant to represent, w_j . For example, for a study that reported standard deviation, w_j was set to 0.683, the bi-directional width of one standard deviation of a standard normal distribution. In these instances, it had to be assumed that the underlying distributional assumption was faithful to the actual distribution of the results in that study, as there was no other possible method of extracting this information.

This estimation of category uncertainty amounts to a weighted average of the uncertainty bounds of all studies in category k based on environmental stochasticity and hydrometric error, the number of data points in each study, and the proportion of the overall population that each study's reported statistic is meant to represent. Once category uncertainty for RR and PR had been estimated for each SCM type, the upper and lower u_{cat} bounds were individually input into Equation 5.3 in order to estimate the uncertainty bounds for TR: mass load removal percent removal. These uncertainty bounds for TR for each category of SCM and POC represent the

final results of this study, as TR is the metric used for the estimation of SCM load reduction, and subsequently SCM cost-effectiveness.

5.5 RESULTS AND DISCUSSION

The results of this study are presented as follows: first, some general observations on the meta-analysis are given, then the effects of observation and environmental uncertainty are shown for RR and PR (TP, TN, and TSS) for a single category of SCM – bioretention – as an example of the 15 SCM categories. Results of the remaining categories can be found in Appendix C. Finally, the u_{cat} for each SCM category and POC is presented, and the importance of the uncertainty bounds is demonstrated using an existing SCM as an example.

SCM data were collected from 164 different studies of 308 unique SCMs. These studies are catalogued in Appendix D according to the SCM type and POCs that were evaluated in the study, and a summary of the number of SCMs studied, and data points within each of these studies is shown in Table 5.4. Data was least available for TN, as studies frequently reported components of TN (TKN and $\text{NO}_{2,3}^-$), but did not present them in a manner that allowed for the estimation of TN PR. The number of studies demonstrating volume reduction capabilities (RR) of SCMs was also less than the other POCs, identifying a need for further hydrologic monitoring of SCMs. Wet Ponds and Constructed Wetlands had the most SCMs studied, but Wet Ponds and Bioretention facilities had the most individual data points.

Table 5.4 - Number of stormwater control measures (SCMs) for which data was collected for each pollutant of concern (POC), and the number of individual data points in these studies. Many studies reported results for multiple SCMs and/or multiple POCs.

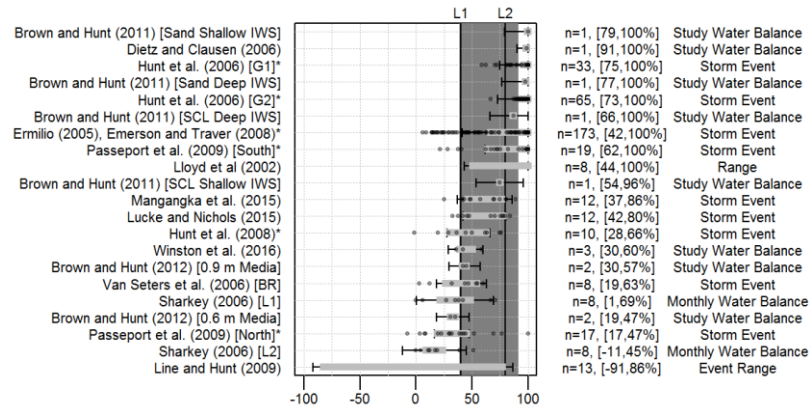
Stormwater Control Measure Category	No. of SCMs Studied, <i>j</i> (No. of Individual Data Points, <i>i</i>)			
	RR	TN	TP	TSS
Bioretention	21 (403)	14 (167)	24 (242)	10 (177)
Compost Soil Amendment	3 (57)			4 (8)
Constructed Wetland		10 (122)	20 (152)	19 (161)
Dry Swale	2 (3)			
Extended Detention		7 (21)	19 (183)	19 (182)
Filtering Practices		8 (121)	24 (291)	25 (312)
Grass Channel	10 (56)	11 (128)	28 (237)	25 (222)
Infiltration	3 (51)	3 (18)	4 (22)	3 (18)
Permeable Pavement	20 (168)	13 (40)	17 (148)	18 (130)
Rainwater Harvesting	2 (2)	11 (11)	5 (5)	5 (5)
Rooftop Disconnect	2 (16)			
Vegetated Filter Strip	21 (269)	13 (106)	24 (241)	25 (261)
Vegetated Roof	7 (120)	7 (14)	8 (15)	
Wet Pond		21 (219)	42 (492)	40 (539)
Wet Swale		7 (59)	12 (117)	11 (130)
Totals:	91 (1145)	125 (1026)	227 (2145)	204 (2145)

The datasets used in this study were identified starting with the sources in the RRM document (CWP, 2008), and following a forward and reverse citation search. Other datasets were extracted from the 12/20/2014 version of the International Stormwater BMP Database (ISWBMPDB) (WERF et al., 2014), and documentation for each of the SCMs in this database was reviewed. Runoff volume reduction data from the ISWBMPDB was extracted from the dataset published along with the report by Geosyntec Consultants and Wright Water Engineers (2011), as these data were quality-reviewed. As many sources were included in this study as possible, though the authors acknowledge that there are likely to be further datasets that were not included in this study.

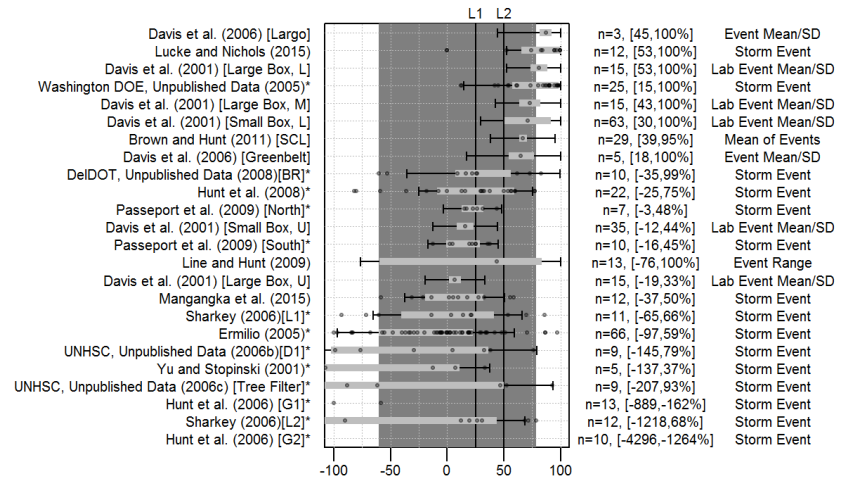
5.5.1 Environmental Uncertainty

The uncertainty associated with bioretention SCMs, based on the literature is shown in Figure 5.2 (A) – (D); the details of this figure are described below, and it is also recommended that the reader refer to Figure 5.1 for interpretation of this graphic. Each row in Figure 5.2 represents a unique SCM, modification to an SCM, or laboratory simulation of an SCM. A reference to each study is shown to the left of the bar chart, and text is bracketed if a single study includes multiple sites or treatment types. For example, R. A. Brown and Hunt (2011) monitor the hydrologic and water quality effects of two SCMs – one with sand and another with sandy

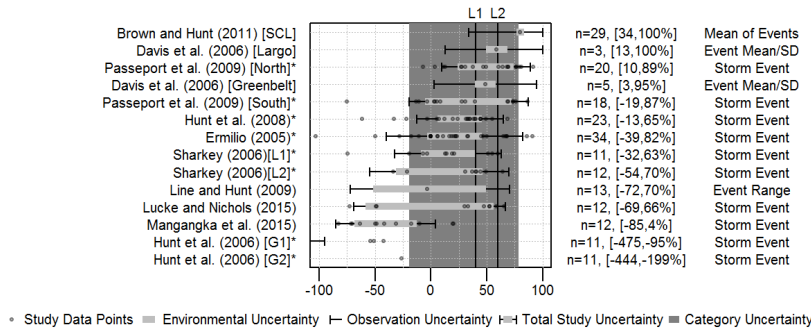
clay loam (SCL) – with two different internal water storage (IWS) zone configurations – deep and shallow. The text shown to the right of the graph indicates from left to right: the number of data points extracted from the study, the total study uncertainty (including both environmental and observation uncertainty) in brackets, and the type of information reported in the study. The variety in the form of the reported information is notable, as laboratory controlled studies reporting the mean and standard deviation of PR values are compared with individual storm events monitored in a field setting.



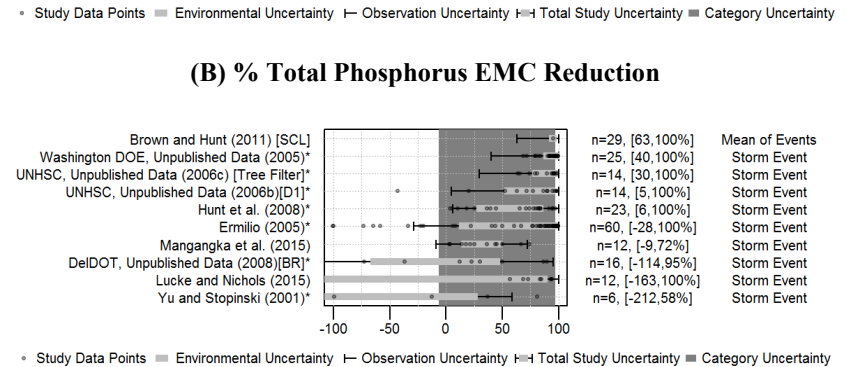
(A) % Runoff Volume Reduction



(B) % Total Phosphorus EMC Reduction



(C) % Total Nitrogen EMC Reduction



(D) % Total Suspended Solids Reduction

• Study Data Points ■ Environmental Uncertainty — Observation Uncertainty — Total Study Uncertainty ■ Category Uncertainty

Figure 5.2 - Uncertainty associated with removal capabilities for bioretention stormwater control measures for: (A) Runoff Volume Reduction (RR), (B) Total Phosphorus EMC Reduction (PR), (C) Total Nitrogen EMC Reduction (PR), and (D) Total Suspended Solids EMC Reduction (PR). Information on the right of each Figure is as follows: “n” represents the number of data points in each study, the bracketed number represents upper and lower uncertainty bounds of the Total Study Uncertainty ($u_{env} + u_{obs}$), and the text describes the type of information reported. Citations noted with an * use RR data from Geosyntec Consultants and Wright Water Engineers (2011) and PR data from the 12/20/2014 version of the International Stormwater BMP Database (WERF et al., 2014).

The data points from each study are plotted horizontally (if available) as semi-transparent circles, to demonstrate the distribution of the results of a single study. The horizontal axes are bounded by -100% and +100%, where values less than zero represent a higher magnitude of volume or EMC at the effluent than the influent; it is not possible to have an RR or PR value greater than 100%. Values of RR less than zero occur when the invert of the SCM is below the groundwater table, causing infiltration into (rather than out of) the system (see, e.g. Line and Hunt, 2009). Values of PR less than zero are common, and occur when there are latent nutrients or sediment in the soil matrix either because of the selected soil, or from previous storm events.

The light gray horizontal bars behind the data points in Figure 5.2 represent the environmental uncertainty, estimated as the IQR if individual storm event data was available, or the measure of dispersion stated in the source study, otherwise. The width of the bioretention environmental uncertainty bounds for studies with some measure of dispersion (i.e. not only a mean or median), increase as: $RR < TN < TSS < TP$, although this was not true for all SCM types. For all SCMs (including bioretention), the median u_{env} was 26%, 28%, 38%, and 48% for RR, TSS, TN, and TP respectively. This suggests that for all SCM categories, the effects of environmental stochasticity on runoff volume reduction is less than the effects on TSS reduction, which is less than the effects on nutrient reduction; likely a function of the magnitude of the influent, as discussed previously.

5.5.2 Observation Uncertainty

The black dimension lines (i.e. the whiskers) in Figure 5.2 demonstrate the total source study uncertainty when observation uncertainty is included. These uncertainty bounds were estimated based on literature values of uncertainty for flow and concentration measurement. In general, observation uncertainty is larger for the estimation of pollutant removal than for runoff reduction, as RR and PR both require flow measurement, but the estimation of influent and effluent EMCs also requires the collection of sample aliquots, sample handling and laboratory analysis – introducing further potential for uncertainty.

Figure 5.3 demonstrates the magnitude of observation uncertainty as compared to environmental uncertainty for each POC type, and in the case of TSS, u_{obs} accounted for a substantial proportion of overall uncertainty. This is because many studies reported using EPA Method 160.2 for TSS, introducing 34% uncertainty. This significant uncertainty attributed to

EPA 160.2 is a result of the inability to pour the sample volume from the sample bottle into a graduated cylinder sufficiently quickly to transfer larger diameter solids (Clark and Siu, 2008a).

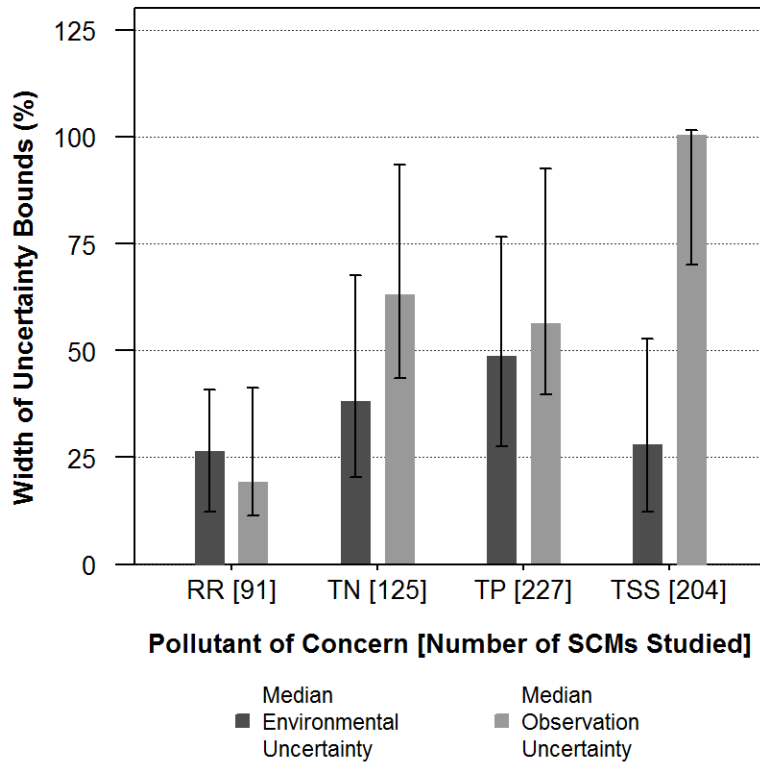


Figure 5.3 - Summary of environmental and observation uncertainty for all stormwater control measures (SCMs) studied in each of the four pollutants of concern: runoff volume (RR), total nitrogen (TN), total phosphorus (TP), and total suspended solids (TSS). Bar heights represent the median uncertainties for j SCMs studied in each group (shown in brackets), and whiskers represent 25th and 75th percentile values.

A large number of studies ($n = 70$) used an inflow-outflow monitoring technique, with a depth measurement device, paired with a primary device for discharge measurement, and a flow-weighted autosampler. Studies also generally reported using Standard Methods or EPA Methods for laboratory characterization, but the specific method varied. Many of the studies reviewed did not provide sufficient information to determine how monitoring was performed, and as a result, the median uncertainty from each component monitoring method was ascribed. In addition to reporting the recommended SCM parameters (Urbonas, 1995), it would also be of value to systematize the reporting of monitoring methods used for SCM studies, as this would provide a better understanding of the study’s data quality, and would also provide guidance for sampling design and sensor purchase for future studies.

The large contribution of u_{obs} to the overall uncertainty associated with an SCM study suggests that further improvements for SCM monitoring are needed – especially with respect to

laboratory characterization. This also suggests that the value added by monitoring influent and effluent concentrations in addition to discharge monitoring may be marginally small, as the additional monitoring steps required to characterize EMC reduction have the potential to introduce uncertainty that is greater than the range of performance caused by environmental variability. Discharge monitoring alone, though it cannot demonstrate biogeochemical processes in an SCM, may be a more efficient means of monitoring SCM function, as there are fewer sources of observation uncertainty.

5.5.3 *Category Uncertainty*

Once the overall study uncertainty had been estimated as the sum of environmental and observation uncertainty, the upper and lower bounds of each study uncertainty were then used to calculate category uncertainty for each DEQ defined SCM category. This was done by averaging uncertainty bounds for all studies in each SCM category, weighted by the number of observations reported in each source study, and the statistic reported (Equation 5.6). This is shown in the dark gray region in the background of Figure 5.2, where the horizontal extent represents the probable range of RR or PR for a given SCM category based on the existing data.

The region demonstrating SCM category uncertainty can be compared to the black vertical lines, which represent credit given by the Virginia DEQ for bioretention cells constructed according to their BMP specifications (VDEQ, 2013) for Level 1 (L1) and Level 2 (L2) design (see also Table 5.1). A L2 design is larger, has an improved internal water storage zone, requires sub-soil testing for infiltration rate, and additional pre-treatment and planting requirements. It is unlikely that any of the bioretention SCMs shown were built according to VDEQ specifications, as only the study by (Yu and Stopinski, 2001) was performed in Virginia, and even this study was performed before these design standards were adopted. However, it should be noted that the L1 and L2 credit was determined based on the range of percent removal values from CWP (2008), which used some of the same studies as those shown in Figure 5.2 – in fact, the CWP review was the starting point for the literature review performed for this research.

Some of the studies shown in Figure 5.2 were not included in the estimation of category uncertainty, as a unique feature of their design led to erratic results. An example of this, is the drastic increase in TP concentrations in the effluent, due to the use of soil media with a high Mehlich-3 P-Index (Hunt et al., 2006). This particular study was marked as an outlier study for

nutrients, because of the dramatic increases in concentration across the SCM due to a design feature, but nearly every study had some nuance that could have excepted it from analysis. This included abnormal precipitation amounts during the study (Greb et al., 2000), erratic behavior in sub-surface media (Collins et al., 2008), and even disruptions of monitoring devices by vandalism and wildlife (Foad Hussain et al., 2006; Rushton, 2002). Studies were excluded from u_{cat} calculations if there was a significant discrepancy in SCM design that had an obvious outcome for RR or PR performance. Otherwise, it was determined that studies with other issues would not be excluded, as these issues are representative of this field of research.

The bounds of category uncertainty estimated for runoff volume reduction (RR); and EMC reduction (TP, TN, and TSS) were then used as input into the RRM's equation to estimate total mass load reduction (TR) as a function of these four inputs (Equation 5.3). This generated upper and lower mass load reduction efficiencies for each of the SCM categories for TP, TN, and TSS, as shown in Figure 5.4. The columns to the right of the bar graph show (from left to right) the number of SCMs studied in each category for each POC, category uncertainty bounds in brackets, and the total width of the uncertainty bounds for TR. The VDEQ's L1 and L2 credit for each of the 15 SCM categories (VDEQ, 2013) and POCs are shown as points (TSS credit is not given in the BMP specifications), overlaid on the respective TR uncertainty bounds. Again, not all of the SCM's reviewed in this research would conform to the DEQ's specifications for L1 and L2 design, but the values are shown because the credit system is based on a review of SCM studies that used central tendencies of RR and PR values from the SCM monitoring literature.

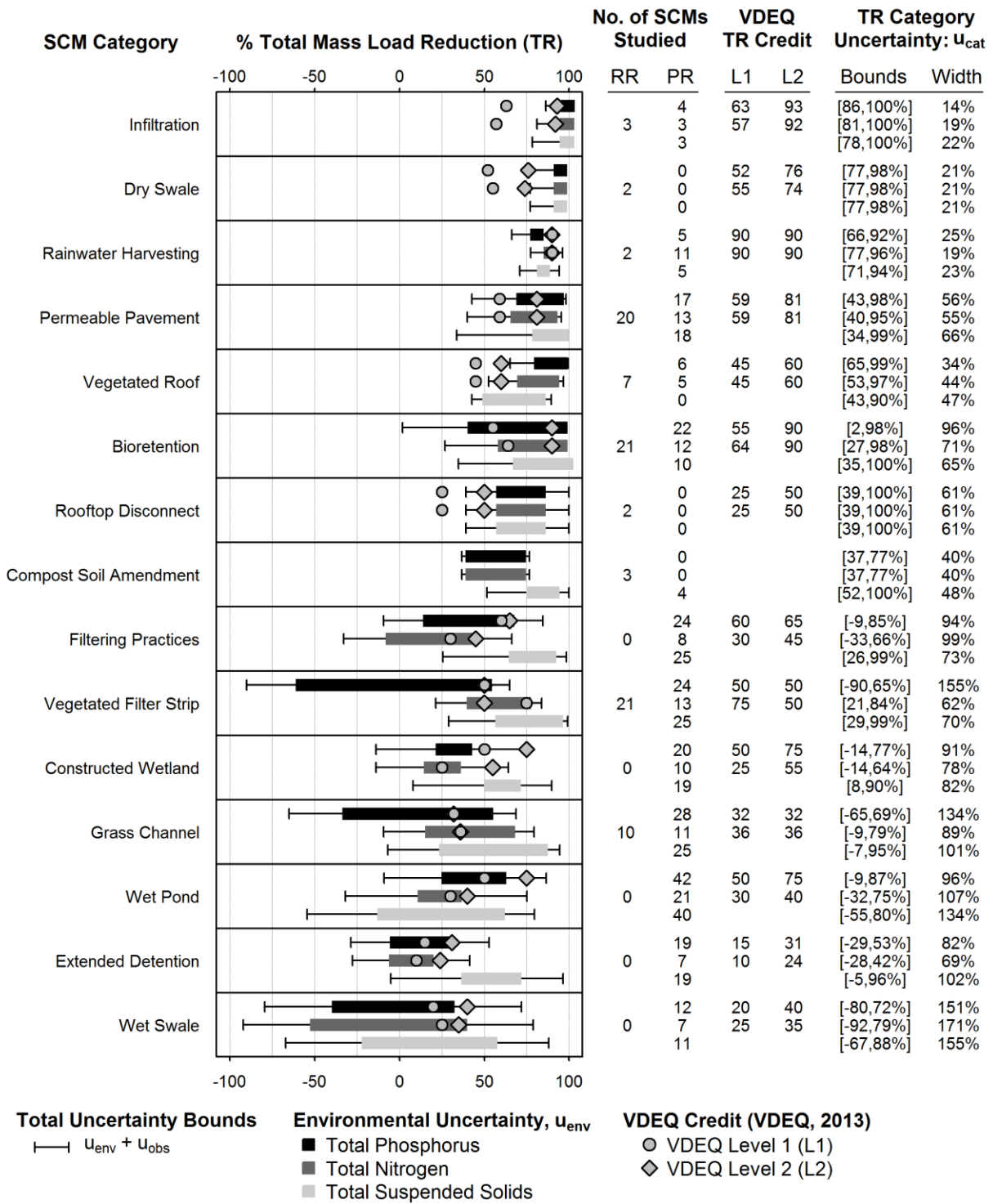


Figure 5.4 - Uncertainty in total mass load reduction (TR): $1 - (\text{effluent mass}/\text{influent mass})$, estimated based on the Runoff Reduction Method (Equation 5.3). VDEQ Level 1 and 2 designate the credit given by the VDEQ for the two design levels of the SCM (VDEQ, 2013); TSS credit is not specified by the VDEQ.

The width of the uncertainty bounds shown in Figure 5.4 is the result of an aggregation (by SCM category) of uncertainty bounds for each individual SCM studied within each category. Category uncertainty is not an additional source of uncertainty, per se, but an artifact of the distillation of results of different SCM designs and study designs into a single SCM category, and their input into the RRM. The overall width of u_{cat} bounds range from 14% ($\pm 7\%$) for TP removal in infiltration practices, to 171% ($\pm 85.5\%$) for TN removal in wet swales. This means that given the existing data and methods for distilling datasets into a single u_{cat} , a wet swale may affect TN loading such that the effluent is between 21% and 192% of the influent loading. The median width of the uncertainty bounds for all SCMs were 82%, 62%, and 71% for TP, TN, and TSS respectively.

Infiltration practices appear to outperform all other SCMs, although this is likely caused by a lack of data in this category. The load reduction performance in the infiltration category was heavily weighted by the large amount of runoff volume reduction demonstrated in (Bright, 2007), but the SCMs in this study were sand dune infiltration systems, and the results may not be readily transposable to other infiltration SCMs. Similarly, vegetated roofs appear to have uniformly positive nutrient and solids removal benefits, but only one of the six studies presenting TP data provided individual storm events, and this study was excluded because of the use of cow manure as a soil amendment (Moran, 2005). This is the sort of variation in SCM design that u_{cat} was designed to demonstrate. The smaller width of uncertainty bounds for the dry swale, rainwater harvesting, and compost soil amendment categories were also due to a lack of available percent removal data for these studies, or because the available data did not include any measures of dispersion.

It was not clear if additional studies reporting the percent removal metric led to increased certainty about the long term pollutant removal capacity of the 15 SCM categories, as there was no discernable relationship between the number of studies or data points in each SCM category, and the width of the uncertainty bounds for that category. Vegetated Filter Strips, for example, have a relatively large amount of RR and TP literature ($n = 21$ and 24 SCMs studied), but the uncertainty bounds associated with the capacity of these SCMs to reduce TP loading is still between -90 and $+65\%$ - a width of 155% . Other categories, such as constructed wetlands and extended detention ponds, had a relatively large amount of data, and relatively smaller uncertainty widths, though more data did not systematically lead to less uncertainty. This lack of

convergence of percent removal uncertainty bounds with additional data suggests that a framework that attempts to explain the effectiveness of SCMs as a static function of a regulator-defined category may not lead to optimal, or even near-optimal SCM implementation.

Even so, there were some general observations that could be made about the results shown in Figure 5.4. Permeable pavement and bioretention are the only two categories whose uncertainty bounds for all three POCs are entirely in the positive domain, and have a substantial amount of studies for all POCs (including RR). This means that using the percent removal metric, the existing literature suggests that both of these SCMs uniformly reduce TP, TN, and TSS loading, but the extent of removal is still subject to large uncertainty bounds. Several other SCMs had TSS (but not TP or TN) uncertainty bounds entirely in the positive domain, including filtering practices, vegetated filter strips, grass channels, and constructed wetlands. It was difficult to draw conclusions about the remaining SCMs, either because of insufficient data, or because of excessively large uncertainty bounds.

The percent removal values attributed to the various SCMs by the VDEQ are shown as gray circles and diamonds in Figure 5.4. These values are based on a review of many source studies similar to this research, but instead used expert judgement of central tendencies (CWP, 2008). This review by the CWP did not attempt to calculate the uncertainty in their values, but clearly acknowledged that a large amount of uncertainty existed, due to the lack of data at the time. The CWP's percent removal values were adopted in the 2013 Virginia stormwater regulations [9 VAC 25-870-65 and VDEQ (2013)], and as such, the uncertainty bounds shown in Figure 5.4 can be applied to the VDEQ values shown. For example, a bioretention cell constructed according to VDEQ Level 1 (L1) design specifications receives credit for removing 55% of the annual influent TP loading, but the literature on which this credit is based can only conclude that bioretention cells (not specifically in accordance with L1 specifications) remove between 2 and 98% of influent loads based on the RRM.

The importance of this is demonstrated with an example bioretention cell constructed in 2013 in Fairfax County, Virginia (Fairfax County Virginia, 2016), designed to treat a 0.62 acre, 39% impervious drainage area. The total design and construction (D&C) cost of this SCM was \$40,350 and the annual maintenance cost was estimated as \$368/yr. The D&C cost, annuitized over a 30 year design life at an interest rate of 4%, was \$2,333/yr, leading to a total annual cost of \$2,701/yr. The estimated annual yield of TP from this SCM's CDA was 0.54 lbs/yr, based on

the loading rates in the Chesapeake Bay Model Phase 5.3.2 (USEPA, 2010c). The cost-effectiveness of this bioretention cell over its design life, is the annuitized cost of the practice divided by the annual load reduction, as shown in Equation 5.7:

$$\text{Cost Effectiveness} = \frac{\text{Annuitized Cost (\$)}}{\text{Annual Load}_{in} \times TR (\%)} \quad \text{Equation 5.7}$$

where $Load_{in}$ is the influent mass load and $TR (\%)$ is the load removal percent efficiency. Credit for $TR (\%)$ given by the DEQ's specification for TP load reduction for a L1 bioretention cell is 55% (VDEQ, 2013), meaning that this cell will remove 0.30 lbs TP/yr (55% of the influent 0.54 lbs/yr). Based on the annuitized cost of \$2,701 estimated above, the load reduction cost-effectiveness of this SCM is \$9,094/lb TP removed, using Equation 5.7. However, if the TP load removal efficiency, $TR (\%)$, of bioretention cells is between 2 and 98% as shown in Figure 5.4, the cost-effectiveness of this particular cell is between \$5,103 and \$250,093 per pound of TP removed. By comparison, estimates for the cost to remove one pound of TP at WWTPs ranges from \$0.96 – \$20 depending on the level of treatment (Barr Engineering Company, 2004), and the addition of nutrient reduction technologies to wastewater treatment plants (WWTPs) has been estimated at \$74/lb of additional TP removal (Chesapeake Bay Commission, 2004).

This simple example is not intended to present a method for cost estimation, but to demonstrate the consequence of the pollutant removal uncertainty associated with widely accepted regulator-defined percent removal values on a single SCM facility. If the current input-output percent removal paradigm of SCM function remains codified in state and federal regulations without consideration to the inherent uncertainty of these metrics, the use of resources on SCMs may result in water quality outcomes that are vastly different than the expected outcome, based on the regulator-defined values. Cost-effectiveness uncertainty is further amplified by the variability in whole-life costs of SCMs due to geographical variation in material and labor costs, and a lack of operation and maintenance cost data (Hodges et al., 2016; S. Taylor et al., 2014).

The final consequences of SCM performance uncertainty are manifest at the local government scale, as MS4 program managers attempt to optimize return-on-investment, and demonstrate progress towards TMDL WLAs by constructing SCMs in their watersheds. As the acknowledgement of the substantial uncertainty bounds shown in Figure 5.4 would render these actions very difficult, program managers are faced with the decision of ignoring the uncertainty

and spending revenue in a potentially inefficient manner, or seeking a method that more faithfully represents the removal dynamics of SCMs.

5.6 CONCLUSION

The category uncertainty shown in Figure 5.4 is the unexplained variance in a model that attempts to predict total mass load removal as a function of a single categorical variable: SCM type. The alternative hypotheses that would be tested using this model are: given the amount of available data, and the variability in these data, the mass load removal for each SCM – POC category are (1) significantly greater than zero, and (2) significantly different from each other. The widespread use of the percent removal metric in engineering practice implies that there is sufficient evidence to reject the null hypothesis in both cases, and that the metric itself is representative of the pollutant removal processes in an SCM.

These statistical tests were not performed in this research, because the results of this study confirm the nearly unanimous consent in the SCM monitoring literature, that the percent removal metric does not sufficiently characterize the pollutant removal dynamics of SCMs (e.g. Barrett, 2005, 2008; City of Austin, 2013; Jones et al., 2008; Strecker et al., 2001). It is possible, for example, that an infiltration SCM has TP removal capabilities significantly greater than null, and also greater than other SCMs, but it is not possible to draw this conclusion using the percent removal metric alone. A move to a more process-based approach to SCM modeling has begun to appear in the literature (e.g. Hoss et al., 2016), but has not been adopted in the general practice of SCM selection and design, nor stormwater regulations. Instead, the ease of application of the percent removal metric, in spite of its known shortcomings, seem to preserve it as the default method for stormwater treatment computations, and it is unlikely to be supplanted without a straightforward replacement.

The results of this study point to a method that accounts for the episodic, stochastic nature of stormwater runoff, that also minimizes the number of measurement methods needed to characterize effectiveness. Two examples of this are the effluent probability method, which relates the effluent load from an SCM to a probability of non-exceedance (City of Austin, 2013), or the process-based methods used in S. Taylor et al. (2014). Both of these methods would need to be further developed to be suitable for SCM selection and design in practice.

Another alternative that has received recent attention is a move from passive to active treatment of stormwater, by the use of real-time control systems (Kerkez et al., 2016). If active treatment is adopted at a large scale, the search for the critical variables for passive stormwater treatment could diminish in importance, as real-time control would enable the active modification of storage and release mechanisms in SCMs. Further research is needed in this field to assure the function of these systems, and to elucidate questions about engineering liability.

Finally, the width of the POC removal uncertainty, and therefore cost-effectiveness uncertainty, demonstrated in this research leads to three questions about the effects of uncertainty on stormwater treatment decision making, as framed by Fischhoff and Davis (2014). (1) Action thresholds: how does the uncertainty about the cost-to-treat stormwater runoff from a development site affect the decision to develop that site? (2) Fixed options: which type of SCM is the most cost-effective? (3) Potential options: what is the range of expected outcomes from a complete implementation of SCMs in a watershed? A new method for evaluating SCM performance – whatever it may be - would need to demonstrate a level of certainty that allows for a discrete answer to each of these questions that would not be subsumed by the uncertainty of the method. As it is, the uncertainty associated with the input-output paradigm of SCM performance that is currently used in practice does not allow for sufficient answers to these questions, and if perpetuated, will not lead to increased information about SCM function – a knowledge gap of critical importance in the field of urban stormwater management.

6. CONCLUSION

The results of each Chapter of this dissertation are discussed specifically within the Chapter, with associated conclusions. The purpose of this final Chapter is to review the results and conclusions of this dissertation at a synoptic scale, and discuss their importance in the body of stormwater management literature. First, the outcomes of this dissertation are revisited briefly, then the broader impacts of these outcomes are discussed. Finally, the knowledge gaps that this dissertation identifies are discussed in the Future Work section.

6.1 OUTCOMES

The first objective of this work – to enumerate and explain the use of non-structural SCMs in Virginia demonstrates a wide range of compliance across the State, with SCM use varying based on the type of regulated entity (e.g. City, County, Military Institution). Other socio-economic and environmental variables did not appear to affect SCM selection, and there was a notable lack of inter-agency cooperation in the reviewed stormwater programs. Inefficiencies within stormwater programs were also noted, such as frequently reported, but piecemeal use of GIS, and the widespread use of stormwater inspection programs with minimal training for stormwater inspectors. The outcomes of the National Pollutant Discharge Elimination System (NPDES) Stormwater Program’s adaptive goal of reducing the discharge of pollutants to the maximum extent practicable (MEP) was manifest in this research, as the stormwater programs in compliance with the NPDES program in Virginia varied widely in the character and extent of non-structural SCMs that comprised a “compliant” program.

The MEP paradigm implies a scaling of compliance as a function of economics and institutional capacity, and though the results of this study showed dramatic variety in Virginia stormwater programs, this variety did not appear to be a function of socio-economic or geographic indicators. Instead, the choice of non-structural SCMs was based largely on the form of governance of the regulated entity, and was otherwise apparently random. Conversations with stormwater authorities suggest that these choices are not likely to fit into a statistical model, but are based on the experience, expertise, and preference of the authorities and their local collaborators. This leads to an important question about the MEP paradigm: is “practical” defined based on the state of the science, local capacity, or somewhere in between? This is discussed in the Broader Impacts section.

The second objective of the dissertation was to quantify the magnitude of uncertainty associated with several commonly used storm sewer measurement devices, as accurate flow measurement is fundamental to, and implicit in the development of TMDL waste load allocations, and the prescription of SCMs at a given pollutant removal efficiency. The outcomes of this study show that small uncertainties in the measurement of depth and velocity propagate to result in large uncertainties in the estimation of storm event peak discharge and runoff volume. The storm hydrographs generated based on the measurements of two separate sensors in a field culvert did not match each other well enough in 8 of 9 storm events recorded, to use the hydrographs for model calibration for those events – even after rigorous laboratory testing and calibration. The outcomes of this study do not imply that all flow measurements are subject to this magnitude of uncertainty – though all measurement methods are uncertain to some degree.

Observation uncertainty limits the extent to which urban watershed processes can be understood, and therefore the extent to which urban watershed issues can be treated. The TMDL framework assumes that a reasonable amount is known about both watershed processes and issues, and that an SCM accounting and enforcement system will eventually lead to enough treatment by these SCMs to de-list receiving bodies. However, if the results of the second objective are characteristic of discharge measurement uncertainty, certainty about watershed function is diminished, and so is the validity of the TMDL framework.

The outcomes of research question 2 led to research question 3, which focused on the overall uncertainty in the effectiveness of the current methods for stormwater treatment, presented as uncertainty bounds on the regulator-defined metric: percent removal. The outcomes of this study show that in many cases it is not possible to determine if an SCM has a positive treatment benefit, and in the cases where it was, the width of uncertainty bounds would still prevent effective SCM selection. SCM categories with more existing performance data did not necessarily have smaller uncertainty bounds, suggesting that use of the percent removal metric may not lead to convergence on the most effective SCM, and that the categorization of SCMs may need to be revised.

6.2 BROADER IMPACTS

The outcomes of this dissertation point to some important questions about the urban ecosystem. If the cost-effectiveness of the existing technology used to prevent flooding and improve water quality is as uncertain as shown in Chapter 5, are stormwater control measures (SCMs) a sustainable solution to the urban stormwater problem? This question will become more salient as the global population increases, and continues to shift to urban areas. These additional urban dwellers will require further conversion of the soil cover to impervious surfaces, and will continue to redistribute biogeochemical balances, further exacerbating urban stormwater issues. These impacts may also be compounded by global changes in precipitation characteristics (Karamouz et al., 2011; Olsson et al., 2012).

The infiltrative principles adopted by SCMs attempt to restore a pre-development water balance, however it is unlikely that the loss of infiltration due to widespread compaction and paving required to support urban development can be fully restored by the use of these SCMs, as the cost to do so would likely outweigh the economic value of the development. Nevertheless, some treatment is needed to prevent the ecosystem from exceeding its resiliency bounds - the point beyond which the system cannot return to equilibrium (e.g. Aral, 2011) – though it is difficult to say exactly where these bounds lie. The maximum extent practicable (MEP) paradigm acknowledges that both water quality objectives, and treatment effectiveness are moving targets, allowing for a flexible definition of compliance, as seen in Chapter 3.

Proponents of the MEP paradigm argue that its use has led to significant improvements in water quality, and point to the pseudo-numeric effluent limitations in the TMDL program as a failed attempt at science-based environmental policy (Houck, 2003). However, the shortcomings of the TMDL program may not be a result of the program itself, but rather, the enforcement mechanism through the NPDES Stormwater program's MS4 permits. Implementation of the program has brought with it all the uncertainties surrounding the development of pseudo-numeric effluent limitations of TMDLs and their prescribed SCMs, and also incorporates the NPDES Stormwater program's MEP principle, such that no monitoring of SCMs or watersheds is required, and as a result, the adaptive management feedback loop is broken (NRC, 2004).

But if adaptive management is adopted into the NPDES/TMDL approach to urban stormwater management, it implies that measurements and analysis will lead to improved knowledge about the watershed system, and subsequently, the most effective means to treat

watershed issues. The outcomes of Chapter 4 and Chapter 5 both question the validity of this claim, as Chapter 4 demonstrates the uncertainty in the measurements themselves, and Chapter 5 shows that additional measurements do not always converge on certainty. However, these two Chapters suggest that (1) the methods used to take the measurements and (2) the manipulation of the measurements into useful management metrics are both of great importance. Adaptive management will only lead to improved management if the measurements and derived metrics are suitable indicators for evaluating past decisions, and making new ones.

6.3 FUTURE WORK

The outcomes of Chapter 5 demonstrate that the static percent removal metric is not sufficient for describing effectiveness of SCMs that varies with episodic and stochastic watershed processes. However, some metric is needed for SCM performance that more faithfully represents the removal dynamics, while still being straightforward enough to use for SCM selection, design, and regulatory accounting. This presents a challenging problem, as several methods have been proposed that are more true to the dynamics, but would require a shift to a probabilistic paradigm (City of Austin, 2013; Granato, 2014; Hoss et al., 2016; Park and Roesner, 2012).

In addition to a metric that varies with stochastic watershed processes, the existing SCM categories (e.g. bioretention) used to explain pollutant removal performance could also be improved upon by incorporating the most important physical characteristics of the SCM. Two examples of this are: (1) the incorporation of SCM size into the performance calculation (Bahr et al., 2012; Koch et al., 2014), and (2) the re-categorization of SCMs according to the removal processes employed Scholes et al. (2008). These studies find that there is insufficient data to build robust statistical models of long-term performance, but short-term results suggest that this could be a promising path forward.

Further research is also needed to quantify the effects of SCM performance uncertainty across a watershed, to demonstrate the range of potential outcomes of SCM implementation plans. This work could also incorporate alternative methods of characterizing SCM performance proposed above, in order to demonstrate how optimal SCM selection is confounded by the uncertainty in the existing paradigm, but how an improved paradigm might provide optimal or near-optimal solutions.

Finally, the outcomes of this dissertation suggest that the magnitude of uncertainty in urban stormwater impedes management decisions based on traditional ratio scale data. For example, two continuous, ratio scale variables - pollutant removal effectiveness and life cycle cost – are used to make decisions about the most cost-effective stormwater treatment options, but this decision requires a level of certainty about both of these variables that may not actually exist (see Chapter 5). A potential alternative that would allow for data-driven management decisions, while acknowledging underlying uncertainty, could be an ordinal scale paradigm. In this framework, rank scale information with uncertainty bounds could be assigned based on the extent of available information. This type of analysis has appeared as illicit discharge risk mapping (Bender et al., 2016; E. Brown et al., 2004), but could present opportunities for other research in the stormwater management field.

REFERENCES

- 33 USC 26 - Federal Water Pollution Control Act (1972). Retrieved from <http://www.law.cornell.edu/uscode/text/33/chapter-26>.
- Aguilar, M. F., and Dymond, R. L. (2014). Innovative Technologies for Stormwater Management Programs in Small Urbanized Areas. *Journal of Water Resources Planning and Management*, 140(11), 4014029.
- Aguilar, M. F., and Dymond, R. L. (2015). Evaluation of Variability in Response to the NPDES Phase II Stormwater Program in Virginia. *Journal of Sustainable Water in the Built Environment*, 4015006(13), 1–13. doi:10.1061/JSWBAY.0000802.
- Aguilar, M. F., McDonald, W. M., and Dymond, R. L. (2016). Benchmarking Laboratory Observation Uncertainty for In-Pipe Storm Sewer Discharge Measurements. *Journal of Hydrology*, 534, 73–86. doi:10.1016/j.jhydrol.2015.12.052.
- Ahlman, S., Malm, A., Kant, H., Svensson, G., and Karlsson, P. (2005). Modelling non-structural Best Management Practices- focus on reductions in stormwater pollution. *Water Science & Technology*, 52(5), 9–16.
- Alig, R. J., Kline, J. D., and Lichtenstein, M. (2004). Urbanization on the US landscape: Looking ahead in the 21st century. *Landscape and Urban Planning*, 69(2), 219–234.
- AMEC. (2012). [Draft] *Fairfax County Chesapeake Bay TMDL Compliance – Planning Level Analysis*.
- Angrisani, L., Baccigalupi, A., Schiano, R., and Moriello, L. (2009). Ultrasonic-Based Distance Measurement Through Discrete Extended Kalman Filter. In V. M. Moreno & A. Pigazo (Eds.), *Kalman Filter Recent Advances and Applications* (pp. 269–296). InTech. doi:10.5772/56411.
- APHA, AWWA, and WEF. (2012). *Standard Methods for the Examination of Water and Wastewater* (22nd ed.). American Water Works Association, American Public Health Association.
- Aral, M. M. (2011). Surgery Was Successful but the Patient Died. *Journal of Hydrologic Engineering*, 16(2), 91–92. doi:10.1061/(ASCE)HE.1943-5584.0000313, 91-92.
- ARC, AECOM, Center for Watershed Protection, Center Forward, Georgia Environmental Protection Division, and Mandel Design. (2016). *Georgia Stormwater Management Manual*. Atlanta, GA: Atlanta Regional Commission. Retrieved from

- <http://atlantaregional.com/environment/georgia-stormwater-manual>.
- ASTM. (2013). *ASTM D3977-97 - Standard Test Methods for Determining Sediment Concentration in Water Samples*. Conshocken, PA: American Society for Testing and Materials. doi:10.1520/D3977-97R13E01.2.
- AWWA. (2012). Testing of Meters- Test Procedures and Equipment. In *Water Meters - Selection, Installation, Testing , and Maintenance - Manual of Water Supply Practices, M6* (pp. 53–85). American Water Works Association.
- Bahr, R., Brown, T., Hansen, L. J., Kelly, J., Papacosma, J., Snead, V., Stack, B., Stack, R., and Stewart, S. (2012). *Recommendations of the Expert Panel to Define Removal Rates for Urban Stormwater Retrofit Projects*. (T. Schueler & C. Lane, Eds.). Chesapeake Stormwater Network. Retrieved from <http://chesapeakestormwater.net>.
- Barr Engineering Company. (2004). Effluent Total Phosphorus Reduction Efforts by Wastewater Treatment Plants. In *Detailed Assessments of Phosphorus Sources to Minnesota Watersheds* (pp. 217–223). St. Paul, MN, MN: Minnesota Pollution Control Agency. Retrieved from <https://www.pca.state.mn.us/water/detailed-assessments-phosphorus-sources-minnesota-watersheds>.
- Barrett, M. E. (2005). Performance Comparison of Structural Stormwater Best Management Practices. *Water Environment Research*, 77(1), 78–86.
- Barrett, M. E. (2008). Comparison of BMP performance using the International BMP Database. *Journal of Irrigation and Drainage Engineering*, 134(5), 556–561.
- Battiata, J., Collins, K., Hirschman, D., and Hoffmann, G. (2010). The Runoff Reduction Method. *Journal of Contemporary Water Research & Education*, 146(1), 11–21. doi:10.1111/j.1936-704X.2010.00388.x.
- Bende-Michl, U., and Hairsine, P. B. (2010). A systematic approach to choosing an automated nutrient analyser for river monitoring. *Journal of Environmental Monitoring*, 12(1), 127–34.
- Bender, P. R., Dymond, R. L., and Aguilar, M. F. (2016). Development of Effective Procedures for Illicit Discharge Risk Mapping. In Press. *Journal of Water Resources Planning and Management*.
- Bertrand-Krawjewski, J. L., and Muste, M. (2008a). Understanding and managing uncertainty. In T. D. Fletcher & A. Deletic (Eds.), *Data Requirements for Integrated Urban Water Management* (pp. 65–90). The Netherlands: Taylor & Francis.

- Bertrand-Krawjewski, J. L., and Muste, M. (2008b). Data validation: principles and implementation. In T. D. Fletcher & A. Deletic (Eds.), *Data Requirements for Integrated Urban Water Management* (pp. 103–124). The Netherlands: Taylor & Francis.
- Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., and West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(17), 7301–7306.
- Bettencourt, L. M. A., and West, G. B. (2010). A unified theory of urban living. *Nature*, *467*(7318), 912–3. doi:10.1038/467912a.
- Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41–51. doi:10.1016/0309-1708(93)90028-E.
- 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), 11–29. doi:10.1016/S0022-1694(01)00421-8.
- Bogardi, I., and Szidarovsky, F. (1974). Induced Safety Algorithm for Hydrologic Design Under Uncertainty. *Water Resources Research*, *10*(2).
- Bonakdari, H., and Zinatizadeh, A. A. (2011). Influence of position and type of Doppler flow meters on flow-rate measurement in sewers using computational fluid dynamic. *Flow Measurement and Instrumentation*, *22*(3), 225–234. doi:10.1016/j.flowmeasinst.2011.03.001.
- Boughton, W. C. (1989). A Review of the USDA SCS Curve Number Method. *Australian Journal of Soil Research*, *27*, 511–523. doi:10.1061/9780784410042.ch02.
- Brabec, E. A. (2009). Imperviousness and land-use policy: Toward an effective approach to watershed planning. *Journal of Hydrologic Engineering*, *14*(4), 425–433.
- Brandeis, L. D. (1932). *New State Ice Co. v. Liebmann, Dissenting Opinion*, (463). Dissenting Opinion.
- Bright, T. M. (2007). An Examination of a Dune Infiltration System’s Impact on Coastal Hydrology and Bacteria Removal. M.S. Thesis, North Carolina State University. Dept. of Biological and Agricultural Engineering. Raleigh, NC. Retrieved from <http://www.lib.ncsu.edu/resolver/1840.16/575>.
- Brown, E., Caraco, D., Pitt, R., and Rittenhouse, B. (2004). *Illicit Discharge Detection and Elimination: A Guidance Manual for Program Development and Technical Assessments*.

Retrieved from http://www.cwp.org/online-watershed-library/doc_download/72-illicit-discharge-detection-and-elimination-a-guidance-manual-for-program-development-and-technical-assessments.

- Brown, R. A., and Hunt, W. F. (2011). Underdrain Configuration to Enhance Bioretention Exfiltration to Reduce Pollutant Loads. *Journal of Environmental Engineering*, 137(11), 1082–1091. doi:10.1061/(ASCE)EE.1943-7870.0000437.
- Burns, M. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., and Hatt, B. E. (2012). Hydrologic shortcomings of conventional urban stormwater management and opportunities for reform. *Landscape and Urban Planning*, 105(3), 230–240. doi:10.1016/j.landurbplan.2011.12.012.
- Campbell, C. W., Dymond, R. L., Kea, K., and Dritschel, A. (2014). Western Kentucky University Stormwater Utility Survey 2014. Report, Bowling Green, Kentucky: Western Kentucky University.
- Carmen, N. B., Hunt, W. F., and Anderson, A. R. (2016). Volume Reduction Provided by Eight Residential Disconnected Downspouts in Durham, North Carolina. *Journal of Environmental Engineering*, 1–9. doi:10.1061/(ASCE)EE.1943-7870.0001107.
- Chesapeake Bay Commission. (2004). Cost Effective Strategies for the Bay. Report, Chesapeake Bay Commission. Retrieved from http://www.chesbay.us/Publications/cost_effective.pdf.
- City of Austin. (2009). *Stormwater Runoff Quality and Quantity from Small Watersheds in Austin, TX: Updated through 2008 - Report Number CM-09-03*. Report, Austin, TX: Watershed Protection Department, City of Austin.
- City of Austin. (2013). *Impacts of Stormwater Control Measures on Water Quality in Austin, TX - Report Number CM 13-02*. Report, Austin, TX: Watershed Protection Department, City of Austin. Retrieved from http://www.austintexas.gov/watershed_protection/publications/default.cfm.
- Clark, S. E., and Siu, C. Y. S. (2008a). Measuring solids concentration in stormwater runoff: Comparison of analytical methods. *Environmental Science and Technology*, 42(2), 511–516. doi:10.1021/es070371g.
- Clark, S. E., and Siu, C. Y. S. (2008b). [Supporting Information] Measuring solids concentration in stormwater runoff: Comparison of analytical methods. *Environmental Science and Technology*, 42(2), 511–516. doi:10.1021/es070371g.
- Coleman, H. W., and Steele, W. G. (1995). Engineering application of experimental uncertainty

- analysis. *AIAA Journal*, 33(10), 1888–1896. doi:10.2514/3.12742.
- Collins, K. A., Hunt, W. F., and Hathaway, J. M. (2008). Hydrologic Comparison of Four Types of Permeable Pavement and Standard Asphalt in Eastern North Carolina. *Journal of Hydrologic Engineering*, 15(6), 512–521. doi:10.1061/(ASCE)HE.1943-5584.0000139.
- Cooter, W. S. (2004). Clean Water Act assessment processes in relation to changing U.S. Environmental Protection Agency management strategies. *Environmental science & technology*, 38(20), 5265–5273. doi:10.1021/es030690h.
- Copeland, C. (2012). Clean Water Act and Pollutant Total Maximum Daily Loads (TMDLs). Report for Congress, Washington, D.C.: Congressional Research Service. Retrieved from <http://fas.org/sgp/crs/misc/R42752.pdf>.
- Costanza, R., and Daly, H. E. (1992). Natural Capital and Sustainable Development. *Conservation Biology*, 6(1), 37–46. doi:10.1046/j.1523-1739.1992.610037.x.
- Coz, J. Le, Ph, D., Blanquart, B., Ph, D., Pobanz, K., Dramais, G., Pierrefeu, G., Hauet, A., Ph, D., and Despax, A. (2012). Estimating the Uncertainty of Streamgauging Techniques Using In Situ Collaborative Interlaboratory Experiments. *Journal of Hydraulic Engineering*, 4016011, 1–15. doi:10.1061/(ASCE)HY.1943-7900.0001109.
- CWP. (2008). Technical Memorandum: The Runoff Reduction Method. Report, Ellicott City, Maryland: Center for Watershed Protection. Retrieved from [http://www.vwrrc.vt.edu/swc/documents/pdf/CWP Technical Memo RRMetho_041808w_Apps.pdf](http://www.vwrrc.vt.edu/swc/documents/pdf/CWP%20Technical%20Memo%20RRMethod_041808w_Apps.pdf).
- D'Asaro, F., and Grillone, G. (2012). Empirical Investigation of Curve Number Method Parameters in the Mediterranean Area. *Journal of Hydrologic Engineering*, 17(10), 1141–1152. doi:10.1061/(ASCE)HE.1943-5584.0000570.
- Davis, A. P. (2008). Field Performance of Bioretention: Hydrology Impacts. *Journal of Hydrologic Engineering*, 13(2), 90–95. doi:10.1061/(ASCE)1084-0699(2008)13:2(90).
- Dedrick, A. R., and Clemmens, A. J. (1984). Double-Bubblers Coupled with Pressure Transducers for Water Level Sensing. *Transactions of the ASAE*, 27(3), 779–783.
- Dietz, M. E., Clausen, J. C., and Filchak, K. K. (2004). Education and changes in residential nonpoint source pollution. *Environmental Management*, 34(5), 684–690.
- 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*, 130(6),

- 690–694. doi:10.1061/(ASCE)0733-9372(2004)130:6(690).
- Dolowitz, D., Keeley, M., and Medearis, D. (2012). Stormwater management: can we learn from others? *Policy Studies*, 33(6), 501–521.
- Dotto, C. B. S., Kleidorfer, M., Deletic, A., Rauch, W., and McCarthy, D. T. (2014). Impacts of measured data uncertainty on urban stormwater models. *Journal of Hydrology*, 508, 28–42. doi:10.1016/j.jhydrol.2013.10.025.
- 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), 2545–2558. doi:10.1016/j.watres.2012.02.009.
- Doyle, M. W. (2012). America's Rivers and the American Experiment. *JAWRA Journal of the American Water Resources Association*, 48(4), 820–837.
- Doyle, M. W., Stanley, E. H., Strayer, D. L., Jacobson, R. B., and Schmidt, J. C. (2005). Effective discharge analysis of ecological processes in streams. *Water Resources Research*, 41(11), 1–16. doi:10.1029/2005WR004222.
- EDC and NRDC v. USEPA. (2003). Environmental Defense Center and Natural Resources Defense Council v. USEPA. *United States Court of Appeals for the Ninth Circuit Court - 344 F.3d 832*.
- EDC and NRDC v. USEPA. (2014). Environmental Defense Center and Natural Resources Defense Council v. USEPA. Petition for a Writ of Mandamus. *United States Court of Appeals for the Ninth Circuit Court*. Retrieved July 1, 2014, from http://docs.nrdc.org/water/files/wat_14121901a.pdf.
- Elmore, A. J., and Kaushal, S. S. (2008). Disappearing headwaters: patterns of stream burial due to urbanization. *Frontiers in Ecology and the Environment*, 6(6), 308–312. doi:10.1890/070101.
- Emerson, C. H., and Traver, R. G. (2008). Multiyear and Seasonal Variation of Infiltration from Storm-Water Best Management Practices. *Journal of Irrigation and Drainage Engineering*, 134(5), 598–605. doi:10.1061/(ASCE)0733-9437(2008)134:5(598).
- Emerson, C. H., Welty, C., and Traver, R. G. (2005). Watershed-Scale Evaluation of a System of Storm Water Detention Basins. *Journal of Hydrologic Engineering*, 10(3), 237–242.
- Engineering Laboratory Design. (1999). 20'0" Sediment Demonstration Channel - Installation,

- Operation, and Maintenance Instructions. Manual, Lake City, MN.
- Fairfax County Virginia. (2016). BMP Database. Access Database, Fairfax County, VA: Department of Public Works and Environmental Services.
- Fischhoff, B., and Davis, A. L. (2014). Communicating scientific uncertainty. *Proceedings of the National Academy of Sciences*, 111(Supplement 4), 13664–13671. doi:10.1073/pnas.1317504111.
- Fisher, R. A. (1970). *The Exact Treatment of 2X2 Tables* (14th Ed.). Darien, Connecticut: Hafner Publishing Company.
- FlowWav. (2013). Pipeline Model PSA-AV. *Product Datasheet*. Retrieved from <http://thinkhd.com/dev/flowav/wp-content/uploads/2013/06/Pipeline-Model-PSA-AV-v19.pdf>.
- Foad Hussain, C., Brand, J., Gulliver, J., and Weiss, P. (2006). Water Quality Performance of Dry Detention Ponds with Under-Drains. Report, St. Paul, MN: University Department of Transportation Research Services Section. Retrieved from <http://www.lrrb.org/PDF/200643.pdf>.
- Folke, C., Jansson, Å., Larsson, J., and Costanza, R. (1997). Ecosystem Appropriation by Cities. *Ambio*, 26(3), 167–172.
- Fowler et al. v. USEPA. (2010). C. Bernard Fowler, Chesapeake Bay Foundation et al. v. USEPA. Settlement Agreement. *United States District Court for the District of Columbia*. Settlement Agreement. Retrieved November 12, 2014, from <http://www.cbf.org/Document.Doc?id=512>.
- Gain, W. S. (1996). *The Effects of Flow-Path Modification on Water-Quality Constituent Retention in an Urban Stormwater Detention Pond and Wetland System, Orlando, Florida*. *Water Resources Investigations Report 95-4297*. Tallahassee, FL: U.S. Geological Survey.
- Galavotti, H., Herbert, R., Pittman, J., Montague-Breakwell, C., Kosco, J., and Frey, M. (2012). National Overview of the Municipal Separate Storm Sewer System (MS4) Program: Status and Progress. In *Proceedings of the Water Environment Federation* (Vol. 402, pp. 411–428).
- Gauron, T. (2015). State promulgates TMDLs based on sporadic grab sampling. *World Water: Stormwater Management*, 3(3), 19–21.
- Geosyntec Consultants, and Wright Water Engineers. (2011). International Stormwater Best

- Management Practices (BMP) Database. Technical Summary: Volume Reductions. Report, WERF, FHWA, ASCE-EWRI.
- Geosyntec Consultants, and Wright Water Engineers. (2014). International Stormwater Best Management Practices (BMP) Database Pollutant Category Statistical Summary Report: Solids, Bacteria, Nutrients, and Metals. Report, WERF, FHWA, ASCE-EWRI.
- Gippel, C. J., O'Neill, I. C., Finlayson, B. L., and Schnatz, I. (1996). Hydraulic guidelines for the re-introduction and management of large woody debris in lowland rivers. *Regulated Rivers: Research & Management*, 12, 223–236.
- Glaeser, E. (2012). *Triumph of the City: How our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*. New York: Penguin Books.
- Global Water. (2014). WL705 Ultrasonic Water Level Sensor. *Product Datasheet*. Retrieved from <http://www.globalw.com/downloads/WL700/WL705B.pdf>.
- Granato, G. E. (2014). Statistics for Stochastic Modeling of Volume Reduction, Hydrograph Extension, and Water- Quality Treatment by Structural Stormwater Runoff Best Management Practices (BMPs). *U.S. Geological Survey Scientific Investigations Report*, 37. Retrieved from <http://pubs.usgs.gov/sir/2014/5037/>.
- Greb, S. R., Bannerman, R. T., Corsi, S. R., and Pitt, R. E. (2000). Evaluation of the Multichambered Treatment Train, a Retrofit Water-Quality Management Practice. *Water Environment Research*, 72(2), 207–216. Retrieved from <http://www.jstor.org/stable/25045360>.
- Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., and Briggs, J. M. (2008). Global change and the ecology of cities. *Science*, 319(5864), 756–60. doi:10.1126/science.1150195.
- Grizzard, T. J., Randall, C. W., and Hoehn, R. C. (1976). Data Collection for Water Quality Modeling in the Occoquan Watershed of Virginia. In *Conference on Environmental Modeling and Simulation* (pp. 819–823). Cincinnati, OH: U.S. EPA.
- Guo, Y., and Gao, T. (2015). Analytical Equations for Estimating the Total Runoff Reduction Efficiency of Infiltration Trenches. *Journal of Sustainable Water in the Built Environment*, 1–9. doi:10.1061/JSWBAY.0000809.
- Hach. (2009). Method 8190. Phosphorus, Total. USEPA PhosVer3 with Acid Persulfate Digestion. Laboratory Method, Hach.

- Hach. (2014). Method 10071. Nitrogen, Total. Persulfate Digestion Method. Laboratory Method, Hach.
- Hardin, G. (1968). The Tragedy of the Commons. *Science*, 162(3859), 1243–8. doi:10.1126/science.162.3859.1243.
- Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., and Arnold, J. G. (2006). Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Transactions of the ASABE*, 49(3), 689–701.
- Harmel, R. D., and Smith, P. K. (2007). Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *Journal of Hydrology*, 337(3–4), 326–336. doi:10.1016/j.jhydrol.2007.01.043.
- Harmel, R. D., Smith, P. K., and Migliaccio, K. W. (2010). Modifying Goodness-of-Fit Indicators to Incorporate Both Measurement and Model Uncertainty in Model Calibration and Validation. *Transactions of the ASABE*, 53(1), 55–63.
- Hartigan, J. A. (1975). The K-Means Algorithm. In *Clustering Algorithms* (pp. 84–112). New York: John Wiley & Sons, Inc.
- Hartigan, J. A., and Wong, M. A. (1979). Algorithm AS 136 A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100–108.
- Heiner, B. J., and Vermeyen, T. B. (2013). Laboratory Evaluation of Open Channel Area-Velocity Flow Meters. Hydraulic Laboratory Report HL-2012-03. (R. F. Einhellig & K. W. Frizell, Eds.). Report, Denver, CO: U.S. Bureau of Reclamation.
- Helsel, D. R., and Hirsch, R. M. (1994). Statistical Methods in Water Resources. In *Techniques of Water-Resources Investigations of the United States Geological Survey* (Vol. 36, p. 323). USGS.
- Herschy, R. W. (1995). General purpose flow measurement equations for flumes and thin plate weirs. *Flow Measurement and Instrumentation*, 6(4), 283–293. doi:10.1016/0955-5986(95)00016-X.
- Hodges, C. C., Dymond, R. L., and Young, K. D. (2016). Generation of Cost Functions for use in Stormwater BMP Selection Optimization. In Review. *Journal of Sustainable Water in the Built Environment*.
- Hoogstraal, G. K. (2015). *Water-Quality Characteristics of Stormwater Runoff in Rapid City , South Dakota, 2008 – 14: Scientific Investigations Report 2015 – 5069*. Reston, VA: U.S.

Geological Survey.

- Horner, R., May, C., Livingston, E., Blaha, D., Scoggins, M., Tims, J., and Maxted, J. (2002). Structural and Non-Structural BMPs for Protecting Streams. In *Linking Stormwater BMP Designs and Performance to Receiving Water Impact Mitigation* (pp. 60–77). Reston, VA: American Society of Civil Engineers.
- Hoss, F., Fischbach, J., and Molina-Perez, E. (2016). Effectiveness of Best Management Practices for Stormwater Treatment as a Function of Runoff Volume. *Journal of Water Resources Planning and Management*, 5016009, 1–12. doi:10.1061/(ASCE)WR.1943-5452.0000684.
- Houck, O. (2003). Tales from a troubled marriage: science and law in environmental policy. *Science*, 302(5652), 1926–9. doi:10.1126/science.1093758.
- Houle, J. J., Roseen, R. M., Ballestero, T. P., Puls, T. A., and Sherrard, J. (2013). Comparison of maintenance cost, labor demands, and system performance for LID and conventional stormwater management. *Journal of Environmental Engineering*, 139(7), 932–938. doi:10.1061/(ASCE)EE.1943-7870.0000698.
- 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*, 132(6), 600–608.
- ISCO. (2011). 2150 Area Velocity Flow Module and Sensor - Installation and Operation Guide. User's Manual, Teledyne ISCO.
- Jarvie, H. P., Withers, J. A., and Neal, C. (2002). Review of robust measurement of phosphorus in river water: sampling, storage, fractionation and sensitivity. *Hydrology and Earth System Sciences*, 6(1), 113–131. doi:10.5194/hess-6-113-2002.
- Jastram, J. D. (2014). *Streamflow, Water Quality, and Aquatic Macroinvertebrates of Selected Streams in Fairfax County, Virginia, 2007 – 12: Scientific Investigations Report 2014 – 5073*. Reston, VA: U.S. Geological Survey.
- Jones, J. E., Clary, J. K., Strecker, E. W., and Quigley, M. M. (2008). 15 Reasons You Should Think Twice Before Using Percent Removal to Assess BMP Performance. *Stormwater*, 10–14.
- Karamouz, M., Hosseinpour, A., and Nazif, S. (2011). Improvement of Urban Drainage System Performance under Climate Change Impact : Case Study. *Journal of Hydrologic*

- Engineering*, 16(5), 395–412. doi:10.1061/(ASCE)HE.1943-5584.0000317.
- Kaufman, M. M. (1995). Community Response To Stormwater Pollution In An Urbanized Watershed. *JAWRA Journal of the American Water Resources Association*, 31(3), 491–504.
- Kaushal, S. S., and Belt, K. T. (2012). The urban watershed continuum: Evolving spatial and temporal dimensions. *Urban Ecosystems*, 15(2), 409–435. doi:10.1007/s11252-012-0226-7.
- Kea, K., and Dymond, R. L. (2016). An Analysis of Patterns and Trends in U.S. Stormwater Utility Systems. In Press. *JAWRA Journal of the American Water Resources Association*.
- Kerkez, B., Gruden, C., Lewis, M. J., Montestruque, L., Quigley, M., Wong, B. P., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O., Poresky, A., and Pak, C. (2016). Smarter Stormwater Systems. *Environmental Science & Technology*, 50, 7267–7273. doi:10.1021/acs.est.5b05870.
- Kirby, W. H. (1991). *Temperature Sensitivity of Mercury-Manometer Bubble Gages*. *Water Resources Investigations Report 91 - 4038*. Reston, VA: U.S. Geological Survey.
- Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1), 13–24.
- Koch, B. J., Febria, C. M., Gevrey, M., Wainger, L. A., and Palmer, M. A. (2014). Nitrogen Removal by Stormwater Management Structures: A Data Synthesis. *Journal of the American Water Resources Association*, 50(6), 1594–1607. doi:10.1111/jawr.12223.
- Komlos, J., and Traver, R. G. (2012). Long-Term Orthophosphate Removal in a Field-Scale Storm-Water Bioinfiltration Rain Garden. *Journal of Environmental Engineering*, 138(10), 991–998. doi:10.1061/(ASCE)EE.1943-7870.0000566.
- Kotlash, A. R., and Chessman, B. C. (1998). Effects of water sample preservation and storage on nitrogen and phosphorus determinations: Implications for the use of automated sampling equipment. *Water Research*, 32(12), 3731–3737. doi:10.1016/S0043-1354(98)00145-6.
- Kotula, A., and Capiella, K. (2013). Cost-Effectiveness Study of Urban Stormwater BMPs in the James River Basin. Report, Ellicott City, Maryland: Center for Watershed Protection.
- Krueger, T., Freer, J. E., Quinton, J. N., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Butler, P., and Haygarth, P. M. (2010). Ensemble evaluation of hydrological model hypotheses. *Water Resources Research*, 46(W07516), 1–17. doi:10.1029/2009WR007845.
- Kunstler, J. H. (2003). *The City in Mind: Notes on the urban condition*. Simon and Schuster.
- Larrarte, F., Bernard Bardiaux, J., Battaglia, P., and Joannis, C. (2008). Acoustic Doppler flow-

- meters: A proposal to characterize their technical parameters. *Flow Measurement and Instrumentation*, 19(5), 261–267. doi:10.1016/j.flowmeasinst.2008.01.001.
- Lee, H., Lau, S.-L., Kayhanian, M., and Stenstrom, M. K. (2004). Seasonal first flush phenomenon of urban stormwater discharges. *Water Research*, 38(19), 4153–63. doi:10.1016/j.watres.2004.07.012.
- Lee, K., Ho, H.-C., Marian, M., and Wu, C.-H. (2014). Uncertainty in open channel discharge measurements acquired with StreamPro ADCP. *Journal of Hydrology*, 509, 101–114. doi:10.1016/j.jhydrol.2013.11.031.
- Leopold, L. B. (1968). *Hydrology for Urban Land Planning - A Guidebook on the Hydrologic Effects of Urban Land Use*. Washington D.C.: United States Department of the Interior.
- Line, D. E., and Hunt, W. F. (2009). Performance of a Bioretention Area and a Level Spreader-Grass Filter Strip at Two Highway Sites in North Carolina. *Journal of Irrigation and Drainage Engineering*, 135(2), 217–224. doi:10.1061/(ASCE)0733-9437(2009)135:2(217).
- Maheepala, U. K., Takyi, A. K., and Perera, B. J. . (2001). Hydrological data monitoring for urban stormwater drainage systems. *Journal of Hydrology*, 245, 32–47. doi:10.1016/S0022-1694(01)00342-0.
- Makepeace, D. K., Smith, D. W., and Stanley, S. J. (1995). Urban stormwater quality: Summary of contaminant data. *Critical Reviews in Environmental Science and Technology*, 25(2), 93–139.
- Mangangka, I. R., Liu, A., Egodawatta, P., and Goonetilleke, A. (2015). Performance characterisation of a stormwater treatment bioretention basin. *Journal of Environmental Management*, 150, 173–178. doi:10.1016/j.jenvman.2014.11.007.
- Massa. (2014). Model M-300/95 Low Cost Sensor. *Product Datasheet*. Retrieved from <http://www.massa.com/wp-content/uploads/2014/11/M-300-95-Datasheet-141107.pdf>.
- McCuen, R. H. (1974). A Regional Approach to Urban Storm Water Detention. *Geophysical Research Letters*, 1(7), 7–8.
- McCuen, R. H., and Moglen, G. E. (1988). Multicriterion Stormwater Management Methods. *Journal of Water Resources Planning and Management*, 114(4), 414–431.
- McCuen, R. H., Sharkey, M., and Walker, A. (2014). Trash Deposition in Rivers and Streams : Guidelines for Prediction of Counts , Volumes , and Masses. *Journal of Hydrologic Engineering*, 4014001, 1–10. doi:10.1061/(ASCE)HE.1943-5584.0000911.

- McDonald, W. M., Dymond, R. L., and Lohani, V. K. (2016). A Methodology to Select Stormflow Sampling Frequencies using Regional Regression Techniques. In Review. *Journal of Hydrology*.
- McIntyre, N., and Marshall, M. (2008). Field verification of bed-mounted ADV meters. *Proceedings of the ICE - Water Management*, 161(4), 199–206.
- McMillan, H., Krueger, T., and Freer, J. (2012). Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26(26), 4078–4111. doi:10.1002/hyp.9384.
- McNett, J. K., Hunt, W. F., and Davis, A. P. (2011). Influent Pollutant Concentrations as Predictors of Effluent Pollutant Concentrations for Mid-Atlantic Bioretention. *Journal of Environmental Engineering*, 137(9), 790–799. doi:10.1061/(ASCE)EE.1943-7870.0000373.
- Meals, D. W., Dressing, S. A., and Davenport, T. E. (2010). Lag time in water quality response to best management practices: a review. *Journal of Environmental Quality*, 39(1), 85–96.
- Mehta, C. R., and Patel, N. R. (1986). Algorithm 643: FEXACT: a FORTRAN subroutine for Fisher's exact test on unordered $r \times c$ contingency tables. *ACM Transactions on Mathematical Software*, 12(2), 154–161.
- Merz, B., and Thielen, A. H. (2005). Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, 309, 114–132. doi:10.1016/j.jhydrol.2004.11.015.
- Moffat, R. J. (1988). Describing the uncertainties in experimental results. *Experimental Thermal and Fluid Science*, 1(1), 3–17. doi:10.1016/0894-1777(88)90043-X.
- Moran, A. C. (2005). A North Carolina Field Study to Evaluate Greenroof Runoff Quantity, Runoff Quality, and Plant Growth. M.S. Thesis, Raleigh, NC: North Carolina State University. Dept. of Biological and Agricultural Engineering. Retrieved from <http://repository.lib.ncsu.edu/ir/bitstream/1840.16/803/1/etd.pdf>.
- Moriasi, D. N., Arnold, J. G., Liew, M. W. Van, Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50(3), 885–900.
- Morison, P. J., and Brown, R. R. (2011). Understanding the nature of publics and local policy commitment to Water Sensitive Urban Design. *Landscape and Urban Planning*, 99(2), 83–92.

- NAFSMA. (2006). Guidance for Municipal Stormwater Funding. Report. Prepared for U.S. Environmental Protection Agency, National Association of Flood and Stormwater Management Agencies.
- Nash, J., and Sutcliffe, J. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
- Neff, E. L. (1977). How much rain does a rain gage gage? *Journal of Hydrology*, 35, 213–220.
- Nesbitt, J. B. (1969). Phosphorus Removal: The State of the Art. *Journal (Water Pollution Control Federation)*, 41(5), 701–713.
- NJDEP. (2004). Stormwater Pollutant Removal Criteria. In *New Jersey Stormwater Best Management Practices Manual*. New Jersey Department of Environmental Protection. Retrieved from http://www.njstormwater.org/bmp_manual2.htm.
- NOAA. (2001). Long-Term Average Annual Precipitation Totals (1981-2001). Dataset, National Oceanic and Atmospheric Administration - National Climatic Data Center. Retrieved from <http://www.ncdc.noaa.gov/land-based-station-data/climate-normals/1981-2010-normals-data>.
- Nobles, A. L., Goldstein, H. D., Goodall, J. L., and Fitch, G. M. (2014). Investigating the Cost-Effectiveness of Nutrient Credit Use As an Option for VDOT Stormwater Permitting Requirements. Report, Charlottesville, Virginia: Virginia Center for Transportation Innovation & Research.
- Nolon, J. R. (1996). The National Land Use Policy Act. *Pace Environmental Law Review*, 13(2), 519–523.
- Nord, G., Gallart, F., Gratiot, N., Soler, M., Reid, I., Vachtman, D., Latron, J., Martín-Vide, J. P., and Laronne, J. B. (2014). Applicability of acoustic Doppler devices for flow velocity measurements and discharge estimation in flows with sediment transport. *Journal of Hydrology*, 509, 504–518. doi:10.1016/j.jhydrol.2013.11.020.
- Nortek. (2013). Vectrino II. *Product Datasheet*. Retrieved December 11, 2013, from <http://www.nortek-as.com/lib/brochures/vectrino-ii>.
- NRC. (2004). *Adaptive Management for Water Resources Project Planning*. Washington D.C.: National Research Council.
- NRC. (2009). *Urban Stormwater Management in the United States*. Washington, D.C.: National Research Council.

- Olsson, J., Amaguchi, H., Alsterhag, E., Däverhög, M., Adrian, P.-E., and Kawamura, A. (2012). Adaptation to climate change impacts on urban storm water: a case study in Arvika, Sweden. *Climatic Change*, 116(2), 231–247. doi:10.1007/s10584-012-0480-y.
- Owen, D. (2011). Urbanization, Water Quality, and the Regulated Landscape. *University of Colorado Law Review*, 82, 431–504.
- Owen, H. J. (1979). *Information for Local Officials on Flood Warning Systems*. Palo Alto, California: National Weather Service. Retrieved from <http://hdl.handle.net/2027/uc1.31822017076480>.
- Pampel, F. C. (2000). *Logistic Regression: a primer*. Thousand Oaks, CA: SAGE Publications, Inc.
- 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), 6881–6890. doi:10.1016/j.watres.2012.04.023.
- Patterson, L. A., Hughes, J., Barnes, G., and Berahzer, S. I. (2012). A Question of Boundaries: The Importance of “Revenuesheds” for Watershed Protection. *JAWRA Journal of the American Water Resources Association*, 48(4), 838–848.
- Paul, M. J., and Meyer, J. L. (2001). Streams in the Urban Landscape. *Annual Review of Ecology and Systematics*, 32(2001), 333–365. doi:doi:10.1146/annurev.ecolsys.32.081501.114040.
- PDEP. (2006). Chapter 6 Structural BMPs. In *Stormwater Best Management Practices Manual* (pp. 1–259). Pennsylvania Dept. of Environmental Protection.
- Pitt, R., Maestre, A., and Morquecho, R. (2004). The National Stormwater Quality Database (NSQD, version 1.1). In *1st Annual Stormwater Management Research Symposium* (pp. 1–9). Tuscaloosa, AL: University of Alabama and Center for Wateshed Protection. Retrieved from <http://unix.eng.ua.edu/~rpitt/Research/ms4/Paper/MS4 Feb 16 2004 paper.pdf>.
- R Core Development Team. (2014). R Language and Environment for Statistical Computing Version 3.0.2. Vienna Austria: R Foundation for Statistical Computing. Retrieved from www.r-project.org.
- Regional Cooperation Act. (1968). *Code of Virginia, Title 15.2*(Chapter 4207).
- Rehmel, M. (2008). Application of Acoustic Doppler Velocimeters for Streamflow Measurements. *Journal of Hydraulic Engineering*, 133(12), 1433–1438.
- Routledge, R. (2005). Fisher’s Exact Test. In *Encyclopedia of Biostatistics*. Wiley.

- Roy, A. H., Wenger, S. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., Shuster, W. D., Thurston, H. W., and Brown, R. R. (2008). Impediments and solutions to sustainable, watershed-scale urban stormwater management: lessons from Australia and the United States. *Environmental Management*, 42(2), 344–59.
- Ruhl, J. B. (1999). The (Political) Science of Watershed Management in the Ecosystem Age. *JAWRA Journal of the American Water Resources Association*, 35(3), 519–526.
- Rushton, B. T. (2002). Treatment of Stormwater Runoff from an Agricultural Basin by a Wet-Detention Pond in Ruskin, Florida. Final Report, Brooksville, Florida: Southwest Florida Water Management District. Retrieved from <http://www15.swfwmd.state.fl.us/LibraryImages/04645.pdf>.
- Sample, D. J., Grizzard, T. J., Sansalone, J., Davis, A. P., Roseen, R. M., and Walker, J. (2012). Assessing performance of manufactured treatment devices for the removal of phosphorus from urban stormwater. *Journal of Environmental Management*, 113, 279–91.
- Sauer, V. B., and Meyer, R. W. (1992). *Determination of Error in Individual Discharge Measurements. Open-File Report 92-144*. Norcross, GA: U.S. Geological Survey.
- Scholes, L., Revitt, D. M., and Ellis, J. B. (2008). A systematic approach for the comparative assessment of stormwater pollutant removal potentials. *Journal of Environmental Management*, 88(3), 467–478. doi:10.1016/j.jenvman.2007.03.003.
- Schueler, T., and Holland, H. K. (2000). Article 65: Irreducible pollutant concentrations discharged from stormwater practices. *Watershed Protection Techniques*, 377–380.
- Science Applications International Corporation. (1999a). Executive Summary. In *Economic Analysis of the Final Phase II Stormwater Rule* (pp. 1–15). Reston, VA: USEPA.
- Science Applications International Corporation. (1999b). Potential Costs, Pollutant Load Reductions, and Cost Effectiveness. In *Economic Analysis of the Final Phase II Stormwater Rule* (pp. 4-1-4–36). Reston, VA.
- SDI-12 Support Group. (2013). SDI-12: A Serial-Digital Interface Standard for Microprocessor-Based Sensors. *Technical Committee Specification*. Retrieved from <http://www.sdi-12.org>.
- Sieck, L. C., Burges, S. J., and Steiner, M. (2007). Challenges in obtaining reliable measurements of point rainfall. *Water Resources Research*, 43(1), 1–23. doi:10.1029/2005WR004519.
- Simpson, T., and Weammert, S. (2009). Developing Nitrogen, Phosphorus, and Sediment

- Reduction Efficiencies for Tributary Strategy Practices. BMP Assessment : Final Report. Report, University of Maryland Mid-Atlantic Water Program. Retrieved from http://archive.chesapeakebay.net/pubs/bmp/BMP_ASSESSMENT_FINAL_REPORT.pdf.
- Smith, E. P., Ye, K., Hughes, C., and Shabman, L. (2001). Statistical Assessment of Violations of Water Quality Standards under Section 303 (d) of the Clean Water Act. *Environmental Science & Technology*, 35(3), 606–612.
- Storms, E. F., Oelsner, G. P., Locke, E. A., Stevens, M. R., and Romero, O. C. (2015). *Summary of Urban Stormwater Quality in Albuquerque , New Mexico , 2003 – 12 Scientific Investigations Report 2015 – 5006*. Reston, VA: U.S. Geological Survey.
- 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*, 127(3), 144–149.
- Taylor, A. C., and Fletcher, T. D. (2007). Nonstructural urban stormwater quality measures: building a knowledge base to improve their use. *Environmental Management*, 39(5), 663–77.
- Taylor, A., and Wong, T. (2002). Non-Structural Stormwater Quality Best Management Practices - An Overview of Their Use, Value , Cost and Evaluation. Technical Report 02/11, Victoria, Australia: Cooperative Research Centre for Catchment Hydrology. Retrieved from <http://www.catchment.crc.org.au/pdfs/technical200211.pdf>.
- Taylor, S., Barrett, M., Leisenring, M., Sahu, S., Pankani, D., Poresky, A., Questad, A., Strecker, E., Weinstein, N., and Venner, M. (2014). *Long-Term Performance and Life-Cycle Costs of Stormwater Best Management Practices*. Washington, D.C.: Transportation Research Board.
- Turnipseed, D. P., and Sauer, V. B. (2010). Discharge Measurements at Gaging Stations. Book 3, Chapter 8, Section A. In *U.S. Geological Survey Techniques and Methods* (p. 87). Reston, VA: U.S. Geological Survey. Retrieved from <http://pubs.usgs.gov/tm/tm3-a8/>.
- U.S. Census Bureau. (2009). 2009 TIGER/Line* Shapefiles. United States Census Bureau. Retrieved from <http://www.census.gov/geo/maps-data/data/tiger-line.html>.
- U.S. Census Bureau. (2010). 2010 Census. Retrieved August 8, 2014, from <http://factfinder2.census.gov/>.
- U.S. Census Bureau. (2013). American Community Survey 5-year Estimates. Retrieved October

- 4, 2014, from <http://factfinder2.census.gov>.
- United Nations. (2014). *World Urbanization Prospects: The 2014 Revision, Highlights*. New York: United Nations Department of Economic and Social Affairs, Population Division.
- University of Virginia (Ed.). (2008). Real Property Tax. In *Virginia Local Tax Rates 2008* (pp. 5–28). Charlottesville, Virginia: Weldon Cooper Center for Public Service.
- Urbonas, B. R. (1995). Recommended Parameters to Report with BMP Monitoring Data. *Journal of Water Resources Planning and Management*, 121(1), 23–34.
- USBR. (2001). *Water Measurement Manual*. U.S. Bureau of Reclamation.
- USEPA. (1971a). *Method 365.2. Phosphorus, All Forms (Colorimetric, Ascorbic Acid, Single Reagent)*. Cincinnati, OH: U.S. EPA National Exposure Research Laboratory. Retrieved from <http://water.epa.gov/type/rsl/monitoring/vms56.cfm>.
- USEPA. (1971b). *Method 160.2 Residue, Non-Filterable (Gravimetric, Dried at 103-105 Degrees C)*. U.S. Environmental Protection Agency; Office of Water. Retrieved from https://www.nemi.gov/methods/method_summary/5213/.
- USEPA. (1974a). *Method 365.4 Total Phosphorus (Colorimetric, Automated, Block Digester AA II)*.
- USEPA. (1974b). *Method 353.3. Nitrogen, Nitrate-Nitrite (Spectrophotometric, Cadmium Reduction)*. U.S. EPA National Exposure Research Laboratory.
- USEPA. (1978a). *Method 365.3. Phosphorus, All Forms (Colorimetric, Ascorbic Acid, Two Reagent)*. U.S. Environmental Protection Agency; Office of Water. Retrieved from https://www.nemi.gov/methods/method_summary/5255/.
- USEPA. (1978b). *Method 351.1. Nitrogen, Kjeldahl, Total (Colorimetric, Automated Phenate)*.
- USEPA. (1978c). *Method 351.3. Nitrogen, Kjeldahl, Total (Colorimetric; Titrimetric; Potentiometric)*. U.S. Environmental Protection Agency; Office of Water. Retrieved from https://www.nemi.gov/methods/method_summary/5400/.
- USEPA. (1983). *Results of the Nationwide Urban Runoff Program. Volume 1 - Final Report*. Washington, D.C. Retrieved from https://www3.epa.gov/npdes/pubs/sw_nurp_vol_1_finalreport.pdf.
- USEPA. (1990). *National Pollutant Discharge Elimination System Regulations for Storm Water Discharges; Phase I Final Rule*. Environmental Protection Agency.
- USEPA. (1993a). *Method 365.1. Determination of Phosphorus by Semi-Automated Colorimetry*.

- Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development.
- USEPA. (1993b). *Method 351.2. Determination of Total Kjeldahl Nitrogen by Semi-Automated Colorimetry. Revision 2.0*. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development.
- USEPA. (1993c). *Method 300.0 Determination of Inorganic Anions By Ion Chromatography. Revision 2.1*. U.S. Environmental Protection Agency, Office of Research and Development. Retrieved from https://www.epa.gov/sites/production/files/2015-08/documents/method_300-0_rev_2-1_1993.pdf.
- USEPA. (1993d). *Method 353.2: Determination of Nitrate-Nitrite Nitrogen By Automated Colorimetry*. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development.
- USEPA. (1994). *Method 200.7: Determination of Metals and Trace Elements in Water and Wastes By Inductively Coupled Plasma-Atomic Emission Spectrometry (4.4.)*. Cincinnati, OH: U.S. Environmental Protection Agency, Office of Research and Development.
- USEPA. (1999). *National Pollutant Discharge Elimination System - Regulations for Revision of the Water Pollution Control Program Addressing Storm Water Discharges. Phase II Final Rule* (Vol. 64). Washington D.C.: Environmental Protection Agency.
- USEPA. (2010a). *MS4 Permit Improvement Guide*. Report, Washington D.C.: U.S. EPA Office of Water.
- USEPA. (2010b). Section 6. Best Management Practices for Nutrients and Sediment. In *Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10* (pp. 1–77). Annapolis, MD: U.S. Environmental Protection Agency, Chesapeake Bay Program Office. Retrieved from ftp://ftp.chesapeakebay.net/modeling/P5Documentation/SECTION_6.pdf.
- USEPA. (2010c). *Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10*. Annapolis, MD: U.S. Environmental Protection Agency, Chesapeake Bay Program Office.
- USEPA. (2013). NPDES Electronic Reporting Rule; Proposed Rule. *Federal Register*, 78(146), EPA-HQ-OECA-2009-0274-0097. Retrieved from <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OECA-2009-0274-0097>.

- USEPA. (2014). *Municipal Separate Storm Sewer System Permits - Post-Construction Performance Standards & Water Quality-Based Requirements - A Compendium of Permitting Approaches*. Washington D.C.: Office of Water - Water Permits Division. Retrieved from http://water.epa.gov/polwaste/npdes/stormwater/upload/sw_ms4_compendium.pdf.
- USGS. (2008). National Hydrography Dataset. *The National Map*. Retrieved October 12, 2013, from <http://viewer.nationalmap.gov/viewer/nhd.html?p=nhd>.
- VDEQ. (2013). *BMP Standards and Specifications [DRAFT]*. Virginia Department of Environmental Quality. Retrieved from http://www.deq.virginia.gov/fileshare/wps/2013_DRAFT_BMP_Specs/.
- VDEQ. (2015). *Guidance Memo No. 15-2005 - Chesapeake Bay TMDL Action Plan Guidance*. Virginia Department of Environmental Quality. Retrieved from <http://www.deq.virginia.gov/Programs/Water/Laws,Regulations,Guidance/Guidance/WaterPermitGuidance.aspx>.
- VDOT v. USEPA. (2013). Virginia Department of Transportation, et al., versus USEPA, et al. *United States District Court for the Eastern District of Virginia - 1:12-CV-775*.
- Venables, W. N., and Ripley, B. D. (2002). *Modern applied statistics with S*. book, Springer Science & Business Media.
- Virginia Department of Environmental Quality. (2013). *9 VAC 25-890 - General Permit for Discharges of Stormwater from Small Municipal Separate Storm Sewer Systems*. Permit, Richmond: Virginia Department of Environmental Quality. Retrieved from <http://leg1.state.va.us/cgi-bin/legp504.exe?000+reg+9VAC25-890-40>.
- Virginia Senate Finance Committee. (2011). Chesapeake Bay TMDL Watershed Implementation Plan : What Will it Cost to Meet Virginia's Goals. Presentation. Presentation, Richmond, Virginia: Senate of Virginia. Retrieved from http://sfc.virginia.gov/pdf/retreat/2011Retreat/Presentation_Final_PDF_for_Website/5.Chesapeake_Bay_TMDL_FINAL.pdf.
- Wagner, W. E. (2005). Stormy Regulation: The Problems that Result when Stormwater (and Other) Regulatory Programs Neglect to Account for Limitations in Scientific and Technical Information. *Chapman Law Review*, 9(191).
- Walsh, C. J., Fletcher, T. D., and Ladson, A. R. (2005). Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream.

- Journal of the North American Benthological Society*, 24(3), 690–705. doi:10.1899/0887-3593(2005)024\{0690:SRIUCT\}2.0.CO;2.
- Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., and Morgan, R. P. (2005). The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society*, 24(3), 706–723. doi:10.1899/04-028.1.
- WERF, EWRI, USEPA, U.S. DOT, APWA, Wright Water Engineers, and Geosyntec Consultants. (2014). International Stormwater BMP Database. Retrieved January 1, 2016, from <http://www.bmpdatabase.org/index.htm>.
- White, S. S., and Boswell, M. R. (2006). Planning for Water Quality: Implementation of the NPDES Phase II Stormwater Program in California and Kansas. *Journal of Environmental Planning and Management*, 49(1), 141–160.
- Wild, T., and Davis, A. (2009). Simulation of the Performance of a Storm-Water BMP. *Journal of Environmental Engineering*, 135(12), 1257–1267. doi:10.1061/(asce)ee.1943-7870.0000106.
- Willink, R., and White, R. (2012). *Disentangling Classical and Bayesian Approaches to Uncertainty Analysis. Technical Report No. CCT/12-08*. Bureau International des Poids et Mesures.
- Winston, R. J., Dorsey, J. D., and Hunt, W. F. (2016). Quantifying volume reduction and peak flow mitigation for three bioretention cells in clay soils in northeast Ohio. *Science of the Total Environment*, 553, 83–95. doi:10.1016/j.scitotenv.2016.02.081.
- WMO. (2010). *Manual on Stream Gauging - Vol I - Fieldwork. WMO-No. 1044* (Vol. I). Geneva, Switzerland: World Meteorological Organization.
- WVDEP. (2012). Chapter 3. Best Management Practice Selection and Design Methodology. In *Stormwater Management and Design Guidance Manual* (pp. 1–36). Charleston, WV: West Virginia Dept. of Environmental Protection. Retrieved from <http://www.dep.wv.gov/WWE/Programs/stormwater/MS4/Pages/StormwaterManagementDesignandGuidanceManual.aspx>.
- Xylem. (2009). SonTek Argonaut-SW Shallow Water. *Product Datasheet*. Retrieved from <http://www.water-monitor.com/documents/ArgSWbrochurerev1.pdf>.
- Yu, S. L., Kuo, J.-T., Fassman, E. A., and Pan, H. (2001). Field Test of Grassed-Swale

Performance in Removing Runoff Pollution. *Journal of Water Resources Planning and Management*, 127(3), 168–171.

Yu, S. L., and Stopinski, M. D. (2001). *Testing of Ultra-Urban Stormwater Best Management Practices*. Charlottesville, Virginia: Virginia Transportation Research Council.

APPENDIX A - FLOW SENSOR INFORMATION AND FLUME RATING

This Appendix includes information about the various flow measurement sensors used in Chapter 4, as well as documentation of the rating of the flume using a venturi meter-manometer configuration. References for this Appendix are in the main References section.

A.1 ULTRASONICS

Ultrasonic (US) sensors operate based on the time-of-flight principle by emitting a high frequency sinusoidal sound wave, and measuring the elapsed time before this burst returns to the sensor. Based on the known velocity of the sound wave and the time measurement, distance from the sensor to the water surface can be estimated. For fluid measurement applications, these sensors are generally mounted above the control section (Figure A.1), such that the sound wave is normal to the fluid surface. For a review of ultrasonic theory, see Angrisani et al. (2009).



Figure A.1 – The Massa M-300/95 mounted on the soffit of a concrete box culvert

The instruments used for testing were the Massa M-300/95 (relabelled as the Telog UT-33u/95) and Global Water WL705, known henceforth as the Massa and GW respectively. These sensors are similar in make, and their manufacturer reported specifications and costs are shown in Table 4.1. The primary difference between these two sensors is that the GW includes a data logger that contains the battery power source, while the Massa must be connected to a separate logger for data collection and power.

Table A.1 – Manufacturer Specifications and Cost of Ultrasonic Sensors

Sensor	Min. Distance (in.)	Max Distance (ft.)	Measurement Resolution (in)	Measurement Accuracy ($\pm\%$)	Source	Cost (USD)
Massa M-300/95	12	13	0.01	0.1	Massa (2014)	900
Global Water WL 705	4	12	0.035	<0.5	Global Water (2014)	800

The shortcomings of US sensors is that they require a distance between the sensor and water surface – known as a dead zone or blanking distance - above which the sensor is not able to take measurements. This varies based on the sensor itself, but can range from inches to feet depending on the scale of distance the sensor is designed to measure. Installation locations will be limited based on this dead zone if a sensor has already been purchased – for example a sensor with a twelve inch dead zone will not be useful in a fifteen inch pipe during a rainstorm event.

A.2 ACOUSTIC DOPPLER VELOCIMETERS (ADVS)

Acoustic Doppler Velocimeters (ADVs) measure velocity by continuously transmitting sound waves into a control volume, then measuring the Doppler frequency shift caused by reflection off particulates in the water, and finally integrating these samples across the volume to provide a single measure of water velocity (Bonakdari and Zinatizadeh, 2011). Generally, these sensors also measure depth in order to estimate a discharge using $Q = VA$ given a known depth-area relationship. Depth is estimated using a submerged piezo-resistive chip that is exposed to water pressure and open to the atmosphere through a hose in the communication cable such that the electrical signal from the chip can be calibrated to water depth. Depending on the sensor design, these electrical signals are processed within the device, or relayed to a recording and telemetry unit (RTU, also known as a data logger) for processing.

The instruments used for these tests were the ISCO 2150, FloWav PSA-AV, SonTek Argonaut SW ADCP, and Nortek Vectrino II, and will henceforth be known as the ISCO, FloWav, SonTek, and Nortek. Their manufacturer reported specifications are shown in Table A.2.

Table A.2 - Manufacturer Specifications and Cost of Acoustic Doppler Sensors

Sensor	Depth			Velocity			Source	Cost (USD)
	Max (ft)	Min (in)	Accuracy	Max (fps)	Min (fps)	Accuracy		
Teledyne ISCO 2150	10	0.4	±0.01 ft	20	-5	±0.1 fps ¹	(ISCO, 2011)	3,000
FloWav PSA-AV	15	0.90	±0.25%	20	-5	±2%	(FloWav, 2013)	1,800
SonTek Argonaut SW	16	7.2	±0.1%	16	-16	±1%	(Xylem, 2009)	5,000 ²
Nortek Vectrino II	N/A	N/A	N/A	9.8	0.3	±0.5%	(Nortek, 2013)	10,000

¹The reported accuracy of the ISCO 2150 is ±0.1 fps up to 5 fps, than ±2% between 5 and 20 fps

²The SonTek Argonaut SW has since been discontinued. The price shown is a the cost of a SonTek IQ, a comparable unit

The ISCO and FloWav are designed in a similar fashion; they provide a single, one-dimensional velocity measurement for a conical control volume and a single depth at each time interval. They are both bottom mounted, and both require an external data logger and power source. The SonTek is installed in a similar fashion, though this sensor collects two velocity measurements – along the channel and vertical velocity - and is larger than the first two. The Nortek is top or side mounted, does not provide a depth measurement, and is the most dissimilar of the four velocimeters used in this study.

These four sensors estimate discharge based on the equation adapted from Marshall and McIntyre (2008):

$$Q = \alpha \frac{f_s}{2f_e} \frac{c}{\cos \theta} A + \beta \quad \text{Equation A.1}$$

where f_s represents the observed frequency shift from the emitted signal (f_e), as a proportion of the velocity of the sound wave in water (c) with respect to the angle of the emitted wave (θ) times the flow cross sectional area (A), and adjusted for the velocity cross section and physical limitations of the sensor using α and β .

The ability of these sensors to accurately characterize discharge in a storm sewer pipe is contingent on several factors. First, the position of the sensor on the pipe bottom subjects them to sedimentation at low velocities, and debris accumulation during storm events, potentially leading to data loss or inaccuracy if the velocity transducer is blocked, the pressure transducer clogged, debris is caught on the pressure hose, or excessive turbulence is created. Second, the velocity is estimated based on the magnitude of the frequency shift from suspended particles within the conical control volume; measurement error occurs if this volume is not representative of the cross sectional velocity (Bonakdari and Zinatizadeh, 2011), or if there is an uneven vertical distribution of sediment within the volume (McIntyre and Marshall, 2008).

The parameter α corrects for velocity distribution in a pipe, and is site-specific. β corrects for depth limitations of the sensor, is sensor-specific, and should be accounted for in the sensor's internal processor, but may also need to be post-processed if warranted by flume testing. Further sources of error and considerations for field implementation are outside the scope of this work, but are enumerated in Aguilar and Dymond (2014) and Marshall and McIntyre (2008).

The Nortek operates in a slightly different way than the other three devices in that it uses bistatic sonar. It uses separate transmit and receive beams by transmitting through a central beam

and receiving through four beams displaced off to the side at a 30 degree angle (Figure A.2). In this way, it measures a three dimensional velocity through all four receivers focusing on the same volume. This device is primarily designed for laboratory use, but has found use in field applications as well. Due to its sensitivity to damage from external forces commonly found in storm flow situations, it would not be an ideal sensor for continuous storm flow measurement. However, in the case of the laboratory testing, it provides a good velocity measurement for performance evaluations between sensors.

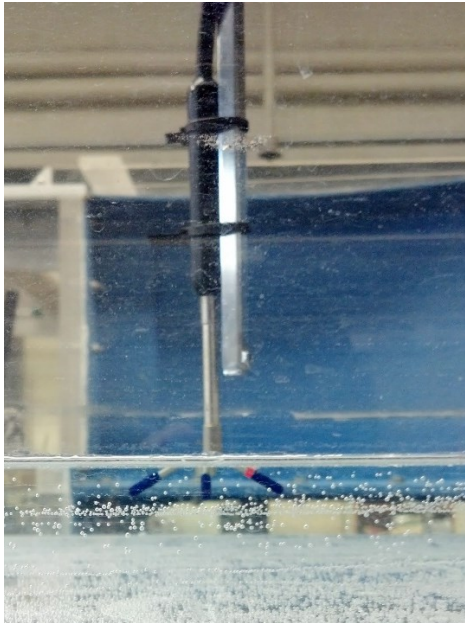


Figure A.2 – Vectrino II Velocimeter mounted in the flume

A.3 SENSOR COMMUNICATION AND POWER SPECIFICATIONS

As the numerical hydraulic data that the sensor provides is its end-product, it is important to consider the data storage and retrieval methods employed. The sensors used for this study store data for download in one of two ways – storage can be built into the sensor itself, or the sensor can send data to an external logger of some variety (Table A.3). When purchasing sensors, the ability of a sensor or group of sensors to interact with a logger is an important consideration. Most of the sensors in this study use an electronics standard known as RS-485 to communicate with a data recorder (logger). Depending on the construction of these loggers, measurement data can either be directly accessed through a connection to a PC, or sent via cellular modem to a remote server. Since these tests were performed in a laboratory setting, data

were retrieved by direct connection with a PC; generally using an RS-232 (communication protocol) to USB converter. The specifications for each sensor are shown in Table A.3.

Table A.3 - Sensor Communication and Power Specifications

Sensor	Connection to Logger	Logger Make	Data Retrieval	Power Supply
Global Water	Analog 4-20mA	Global Water	USB	2 - 9V batteries
Massa	RS-485	Telog	RS-232/USB or Cell Phone	12V battery
FloWav	RS-485	Telog	RS-232/USB or Cell Phone	12V battery
ISCO	RS-485	Teledyne ISCO	RS-232/USB	2 – 6V lantern batteries
SonTek	RS-232/SDI-12 ¹	N/A	RS-232/USB	12V battery
Nortek	RS-485	Nortek	USB	USB

¹The SonTek has an internal data logger, and can communicate directly with a desktop computer. It is also capable of communication using the serial digital interface at 1200 baud (SDI-12) standard (SDI-12 Support Group, 2013).

The sensors are designed for field use, and are powered by batteries (DC), which in some cases can be recharged using solar panels (the Massa and FloWav). However solar recharge was not available in the lab setting, and the testing was power intensive as samples were needed at small time intervals (one minute). The battery power supply was impractical for long term testing of the Massa and FloWav, so the data logger for these sensors was wired to a laboratory DC power supply set at a constant 12V.

A.4 MANOMETER AND VENTURI MEASUREMENTS

To provide a direct benchmark with which to compare sensor observations, the ELD flume provides a venturi – differential manometer configuration with an established rating curve that can be derived from Bernoulli’s equation, flow continuity, and hydrostatic pressure. The venturi reduces flow from a diameter of 6.068” to 3.040”, and has a discharge efficiency coefficient, $C_d = 0.987$. The manometer measures the pressure drop in the venturi, and reports it as a change in height (Δh) to the thousandths of a foot. The equation relating discharge, Q in cfs with Δh in ft. is:

$$Q = C_d A_2 \sqrt{\frac{2g\Delta h}{1 - r^2}} \quad \text{Equation A.2}$$

where A_2 is the cross sectional area of the contraction in the venturi in ft², g is acceleration due to gravity (32.2 ft/s²), r is the contraction cross section area divided by the entrance cross section area (0.251), and C_d is the discharge efficiency coefficient (0.987). Information regarding the calibration of this venturi meter is found in Engineering Laboratory Design (1999).

The ELD flume was equipped with two pumps – one variable speed pump that could be set to a range of pump frequencies between 0 and 60 Hz at 0.1 Hz increments, and one that only operated at 0 or 60 Hz frequencies. The pump frequency-initial channel depth-discharge relationship was created by filling the flume to zero depth in the channel, running the variable speed pump from 0 – 60 Hz, and estimating discharge based on manometer readings. The initial depth in the flume was then increased by one inch, and the process was repeated.

APPENDIX B - LABORATORY METHODS FOR PHOSPHORUS, NITROGEN, AND TOTAL SUSPENDED SOLIDS

This appendix describes the methods used to develop several of the values of laboratory observation uncertainty used in Chapter 5. This appendix is included, as many of the stormwater control measure (SCM) monitoring studies report the use of “Standard Methods” (APHA et al., 2012), but do not specify which particular method was used. As a result, central tendencies were developed based on the reported bias and precision error in the various methods as described below for Phosphorus, Nitrogen, and Total Suspended Solids.

B.1 PHOSPHORUS

In general, total phosphorus is determined in a stormwater sample by first digesting the sample using perchloric, sulfuric-nitric acid, or persulfate (S.M. 4500 P. B.3-5). Perchloric acid is rarely used, persulfate is the simplest, and sulfuric-nitric provides the highest level of accuracy. The sample is then determined by colorimetry, using:

- S.M. 4500 P. C. Vanadomolybdophosphoric Acid if the concentration is 1 – 20 mg/L
- S.M. 4500 P. D, E. Stannous Chloride or Ascorbic Acid if the concentration is 0.01 – 6 mg/L

Based on the stated recoveries and relative standard deviations reported in the corresponding methods, bias and precision for the TP Standard Methods are shown in Figure B.1. Total reactive phosphorus can be determined by direct colorimetry (i.e. without digestion), and is also included in Figure B.1.

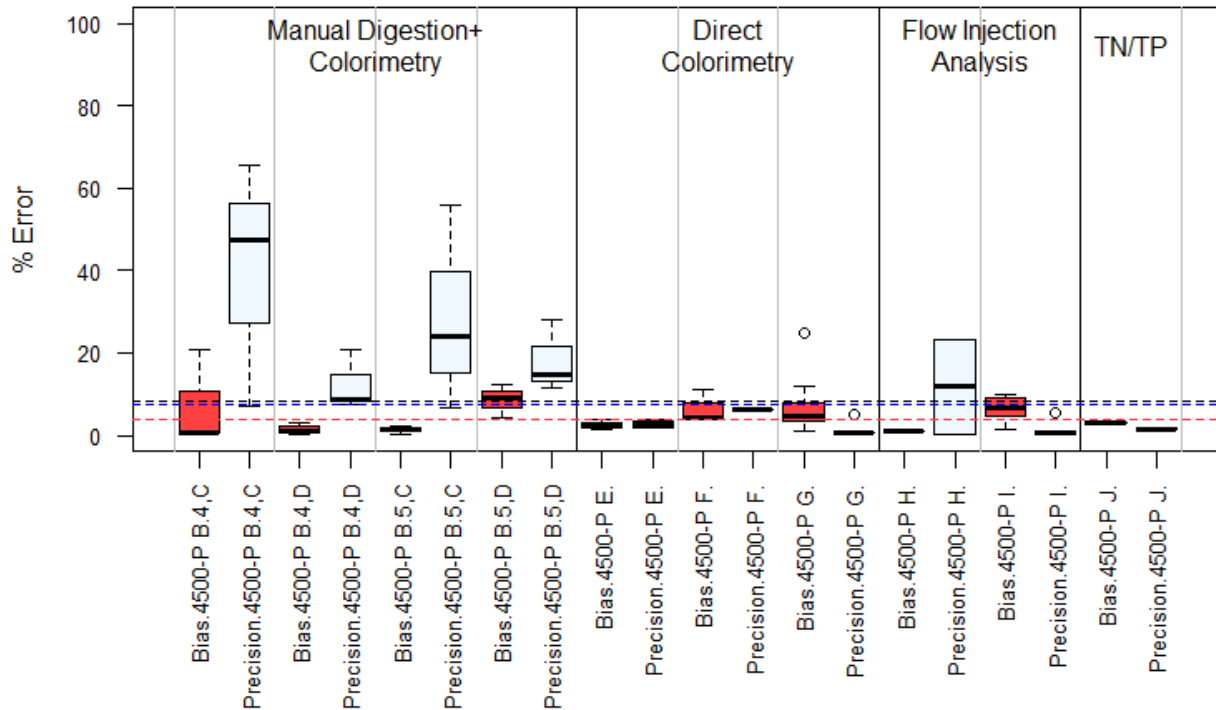


Figure B.1 - Bias and Precision associated with various methods for determining total phosphorus in the laboratory. The median bias, precision, and overall uncertainty used in Chapter 5 are shown as red, blue, and black dashed lines respectively.

The methods that perform manual digestion, followed by a colorimetric method are as follows:

- 4500-P B.4, C -Sulfuric-Nitric Acid Digestion + Vanadomolybdophosphoric Acid Colorimetry
- 4500-P B.4, D - Sulfuric-Nitric Acid Digestion + Stannous Chloride Colorimetry
- 4500-P B.5, C - Persulfate Digestion + Vanadomolybdophosphoric Acid Colorimetry
- 4500-P B.5,D - Persulfate Digestion + Stannous Chloride Colorimetry

The following methods perform direct colorimetry on the sample:

- 4500-P E. – Ascorbic Acid Method
- 4500-P F. – Automated Ascorbic Acid Method
- 4500-P G. – Flow Injection Analysis for Orthophosphate

These methods perform digestion, prior to some type of flow injection analysis

- 4500-P H. - Manual Digestion and Flow Injection Analysis for Total Phosphorus
- 4500-P I. - In-line UV/Persulfate Digestion and Flow Injection Analysis for Total Phosphorus

The final method can characterize both TN and TP

- 4500-P J. - Persulfate Method for Simultaneous Determination of Total Nitrogen and Total Phosphorus.

The overall uncertainty associated with Standard Method 4500 – P was estimated using the bias and precision estimates of S.M. 4500-P B, E, F, as the Ascorbic Acid methods were the most commonly cited methods.

B.2 NITROGEN

The bias and precision for the Nitrogen Standard Methods are shown in Figure B.2. Total Nitrogen is typically determined as the sum of Total Oxidized N, which is comprised of Nitrite (NO_2^-) and Nitrate (NO_3^-); and Total Kjeldahl Nitrogen (TKN), which is comprised of organic Nitrogen species and Ammonia (NH_3). The determination of TKN is generally performed by digestion with sulfuric acid, potassium sulfate, and cupric sulfate, followed by distillation, and colorimetric or titrimetric determination. The values used for observation uncertainty in for nitrogen species are the median values of the S.M. 4500 N_{org} B. – Macro Kjeldahl Method, as this is the only TKN method for which accuracy and precision data were provided. The values used for observation uncertainty for NO_2^- and NO_3^- are also median values of the various methods used to determine these particular species.

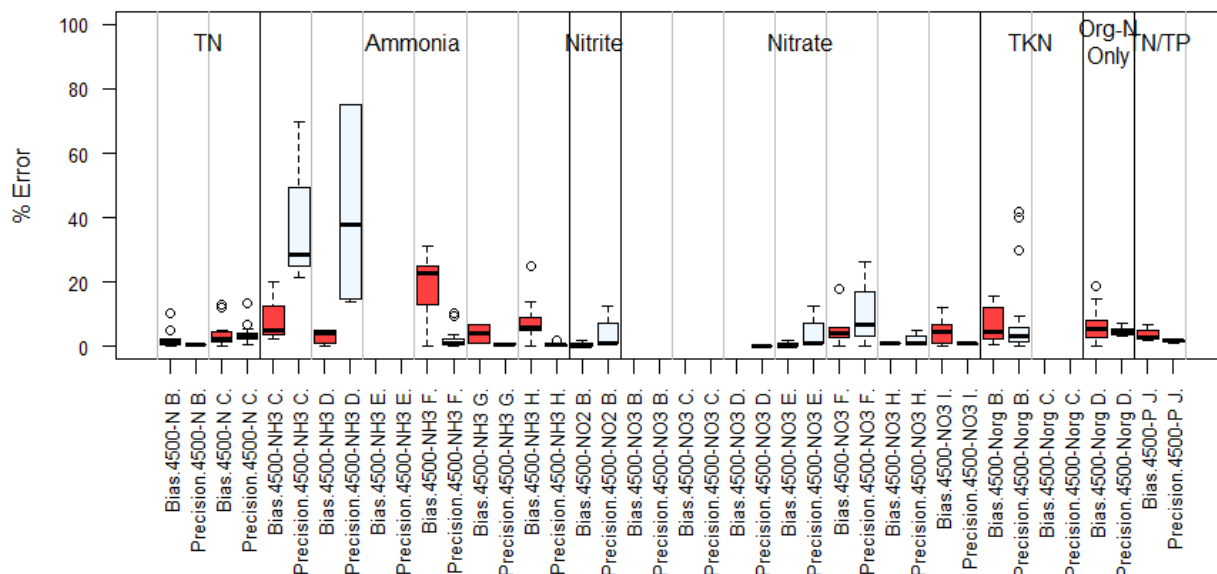


Figure B.2 - Bias and Precision associated with various methods for determining total nitrogen in the laboratory.

B.3 TOTAL SUSPENDED SOLIDS

Precision and Bias information was not provided in EPA Method 160.2 for the determination of total suspended solids (TSS), and as such, data from another study by (Clark and Siu, 2008a) was used to estimate uncertainty. In this study, a known amount of suspended solids was added to 11 samples, and three methods were used to determine concentrations:

Standard Method 2540D, EPA 160.2, and ASTM SSC – D3977-97(B). The results are given in a supplemental information document (Clark and Siu, 2008b), that provides the results of the ordinary least squares regression model of determined concentration vs. known concentration. The slope of this line was used to ascribe bias error, and the 95% confidence intervals of the slope, as a proportion of the slope were used to estimate precision error. Data from the SiO₂ analyses were used, as this appeared to better emulate stormwater characteristics.

APPENDIX C - STORMWATER CONTROL MEASURE PERFORMANCE UNCERTAINTY

The following Appendix contains the graphical results of the estimation of uncertainty for 15 stormwater control measure (SCM) categories, and four pollutants of concern: runoff volume reduction (RR); and event mean concentration (EMC) reduction for total phosphorus (TP), total nitrogen (TN) and total suspended solids (TSS).

C.1 RUNOFF VOLUME REDUCTION

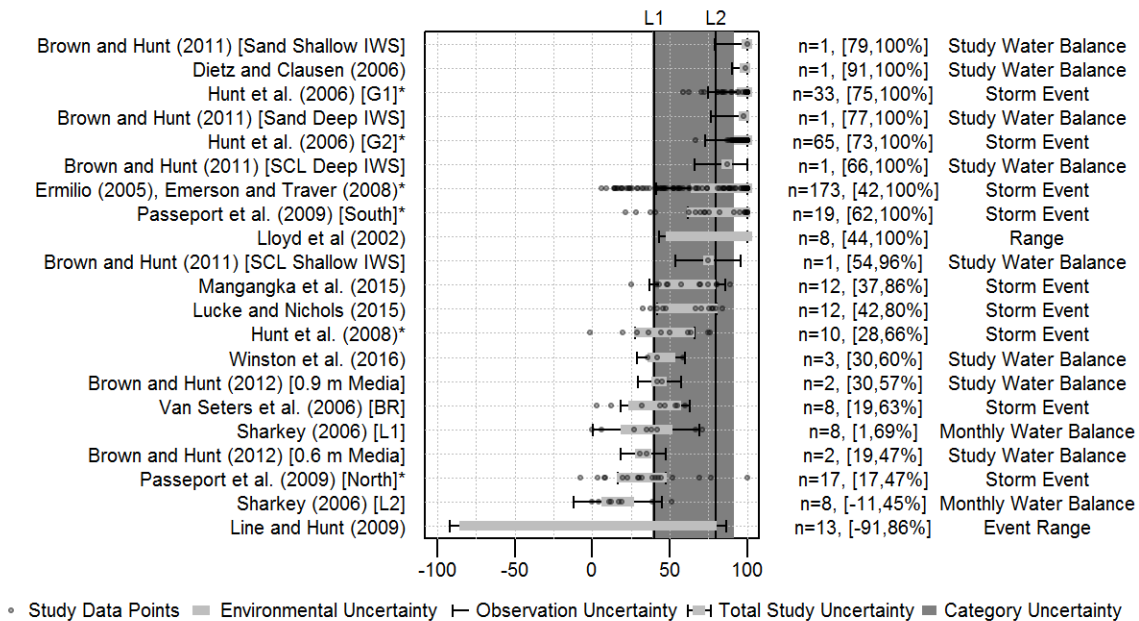


Figure C.1 – Runoff Volume Reduction capabilities of Bioretention Cells

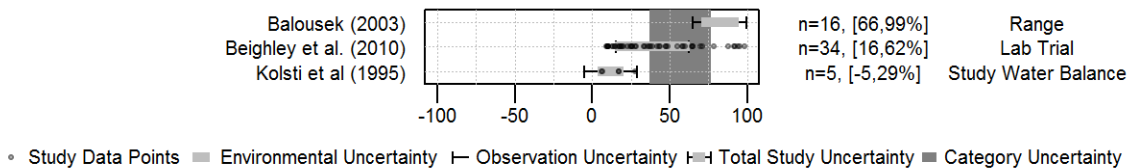


Figure C.2 – Runoff Volume Reduction capabilities of Compost Soil Amendment

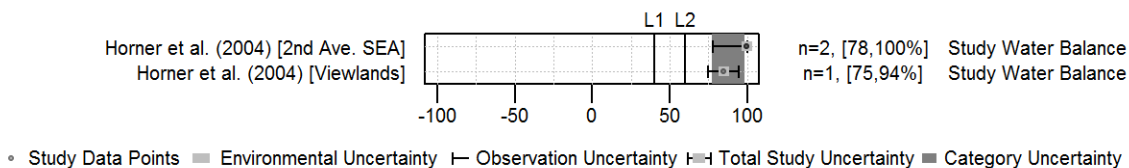


Figure C.3 – Runoff Volume Reduction capabilities of Dry Swales

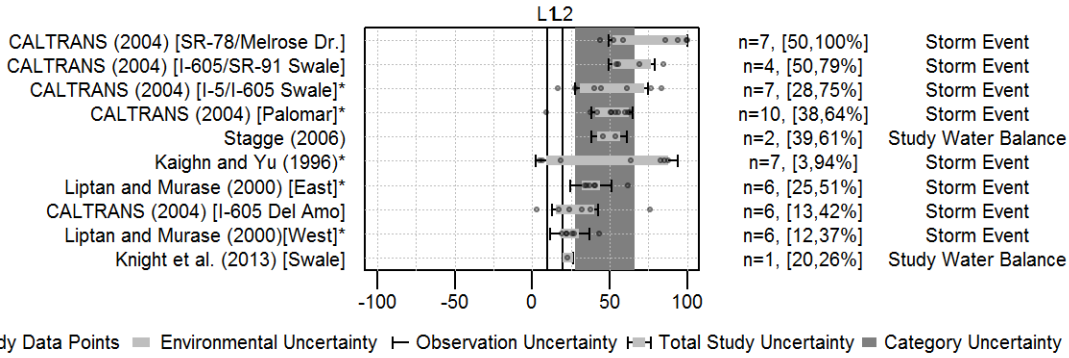


Figure C.4 – Runoff Volume Reduction capabilities of Grass Channels

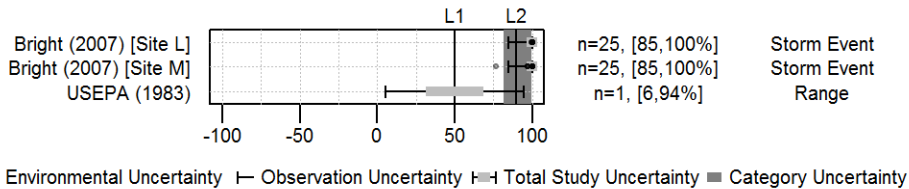


Figure C.5 – Runoff Volume Reduction capabilities of Infiltration

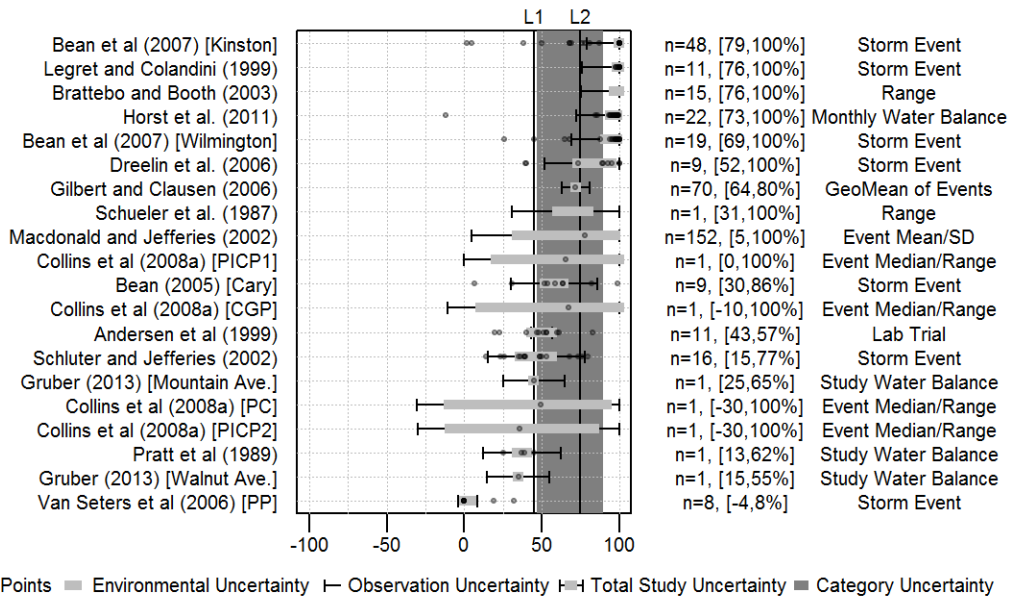


Figure C.6 – Runoff Volume Reduction capabilities of Permeable Pavement

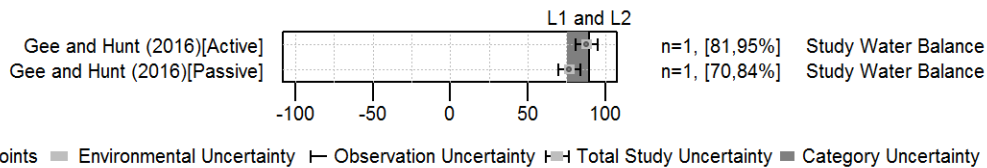


Figure C.7 – Runoff Volume Reduction capabilities of Rainwater Harvesting

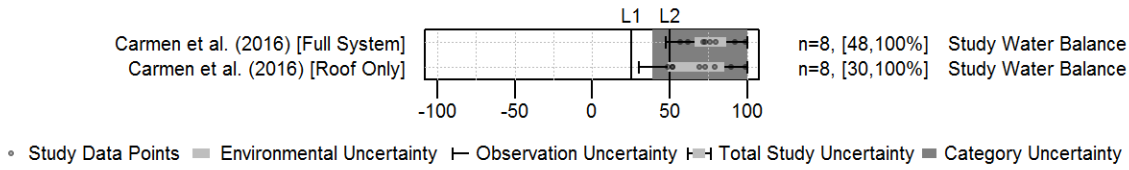


Figure C.8 – Runoff Volume Reduction capabilities of Rooftop Disconnect

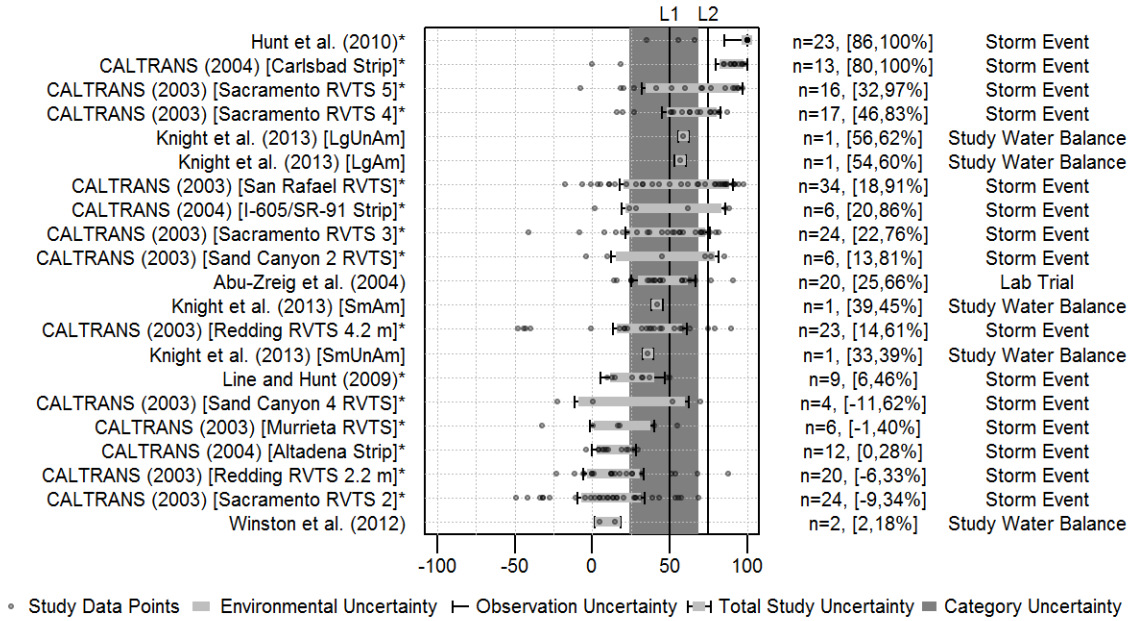


Figure C.9 – Runoff Volume Reduction capabilities of Vegetated Filter Strip

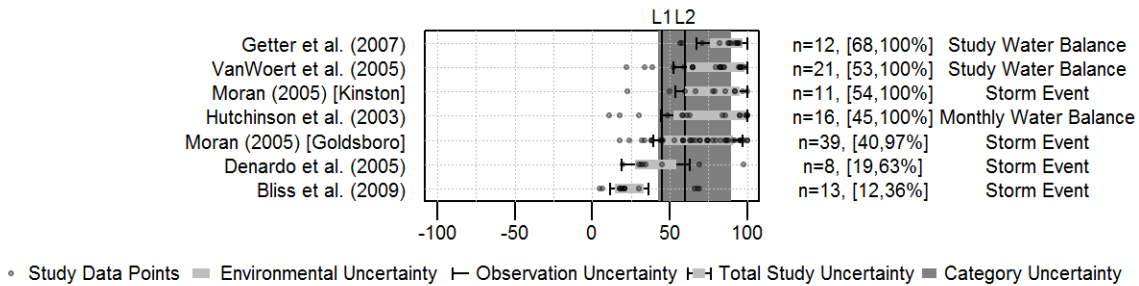


Figure C.10 – Runoff Volume Reduction capabilities of Vegetated Roof

C.2 TOTAL PHOSPHORUS POLLUTANT REMOVAL

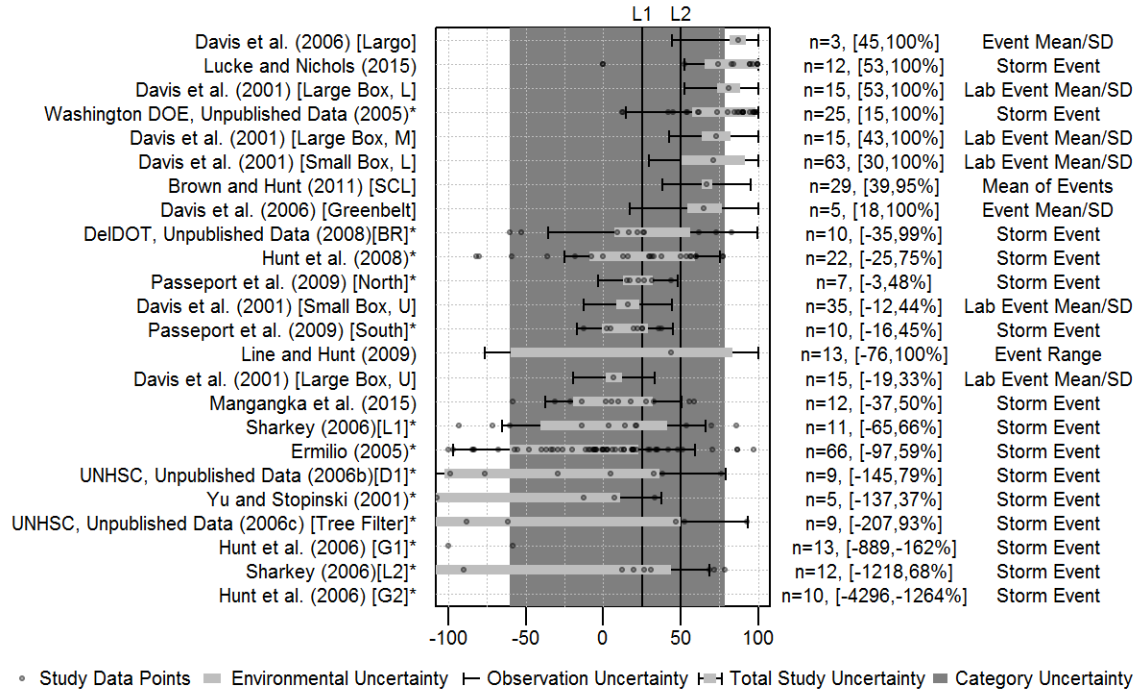


Figure C.11 – Total Phosphorus pollutant removal capabilities of Bioretention cells

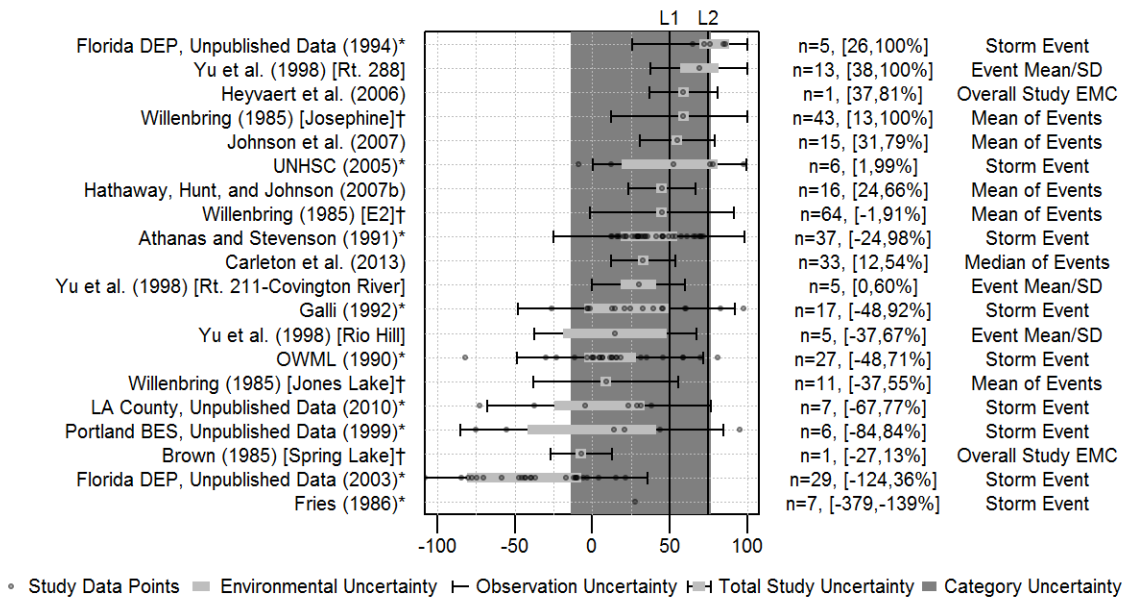


Figure C.12 – Total Phosphorus pollutant removal capabilities of Constructed Wetlands

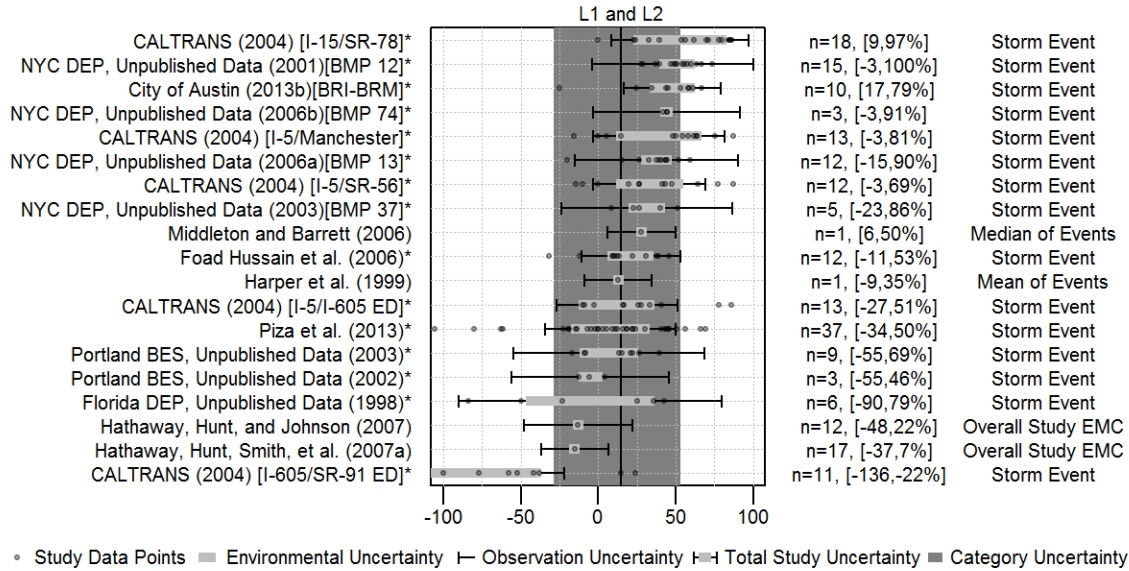


Figure C.13 – Total Phosphorus pollutant removal capabilities of Extended Detention

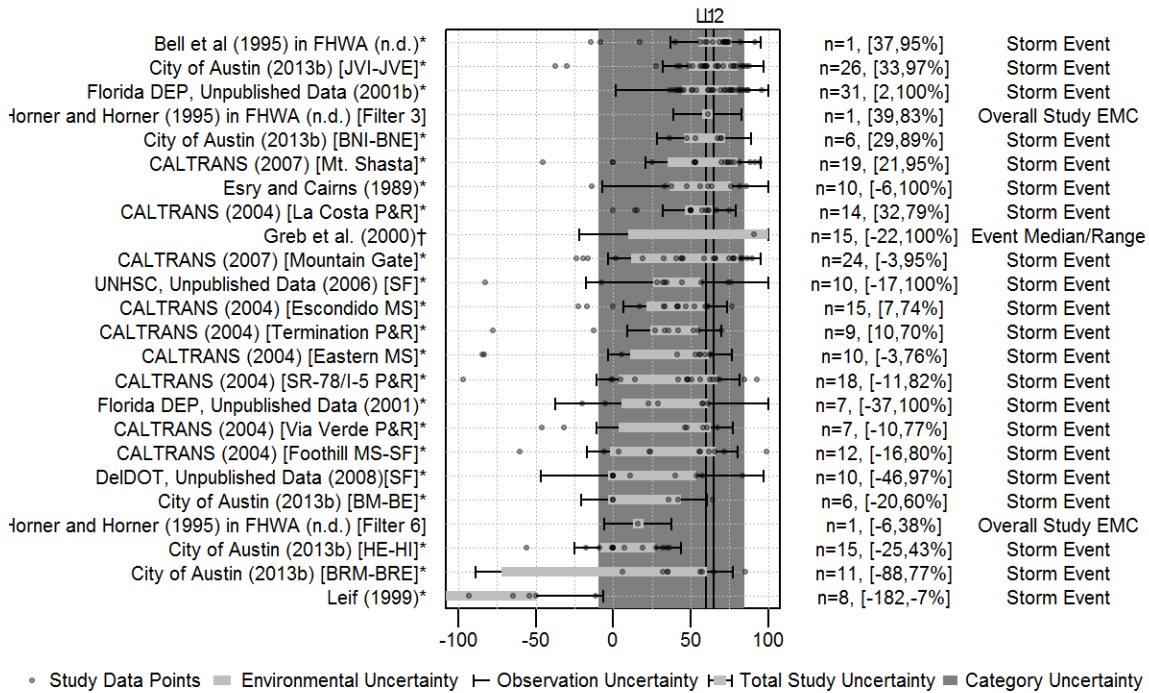


Figure C.14 – Total Phosphorus pollutant removal capabilities of Filtering Practices

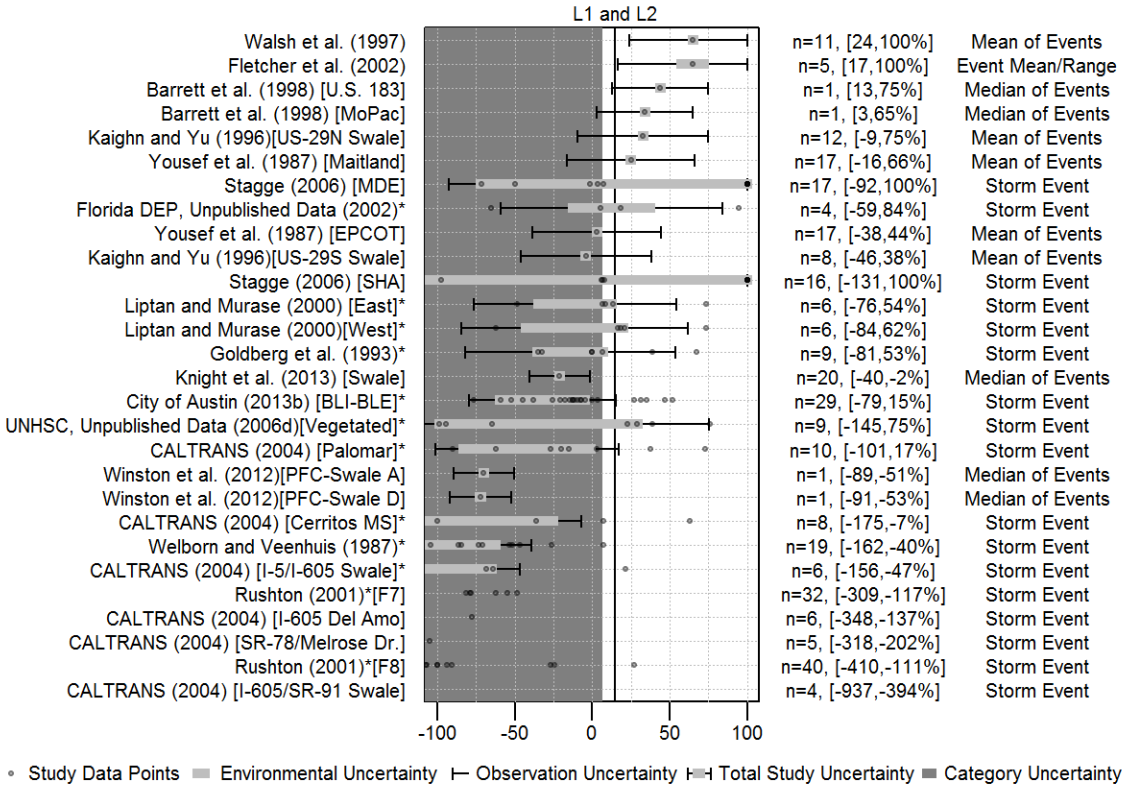


Figure C.15 – Total Phosphorus pollutant removal capabilities of Grass Channels

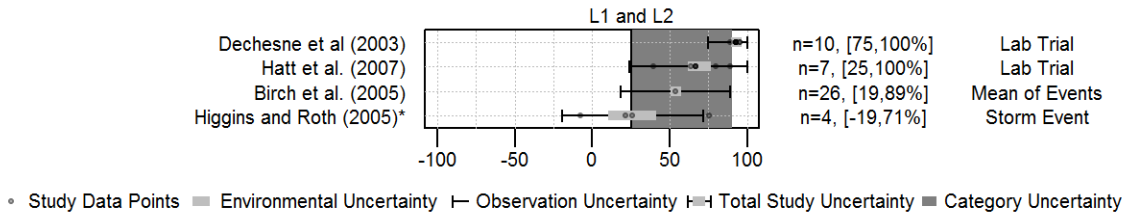


Figure C.16 – Total Phosphorus pollutant removal capabilities of Infiltration

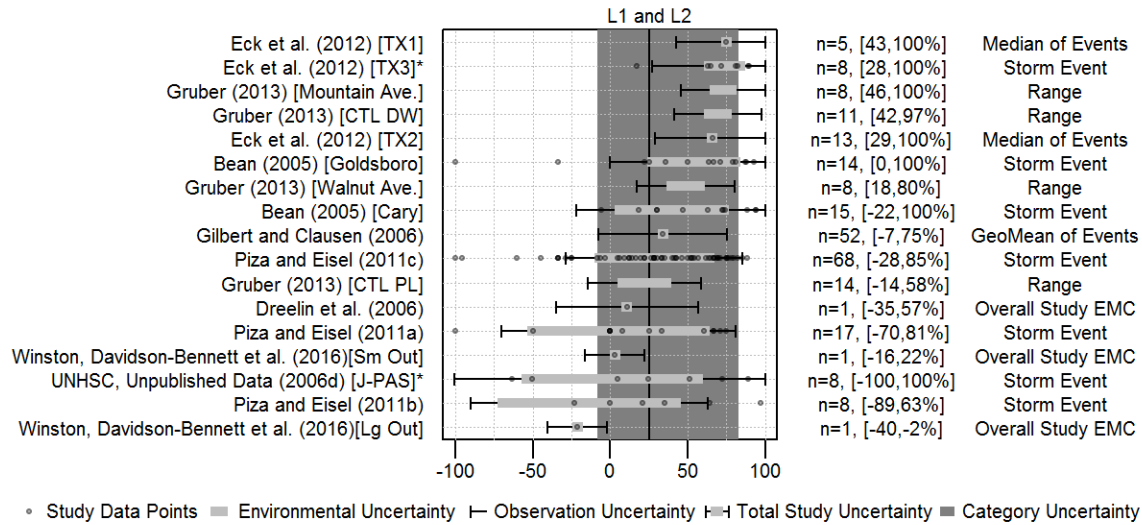


Figure C.17 – Total Phosphorus pollutant removal capabilities of Permeable Pavement

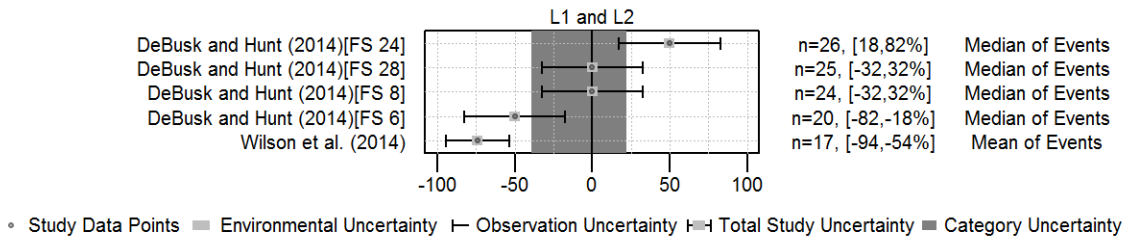


Figure C.18 – Total Phosphorus pollutant removal capabilities of Rainwater Harvesting

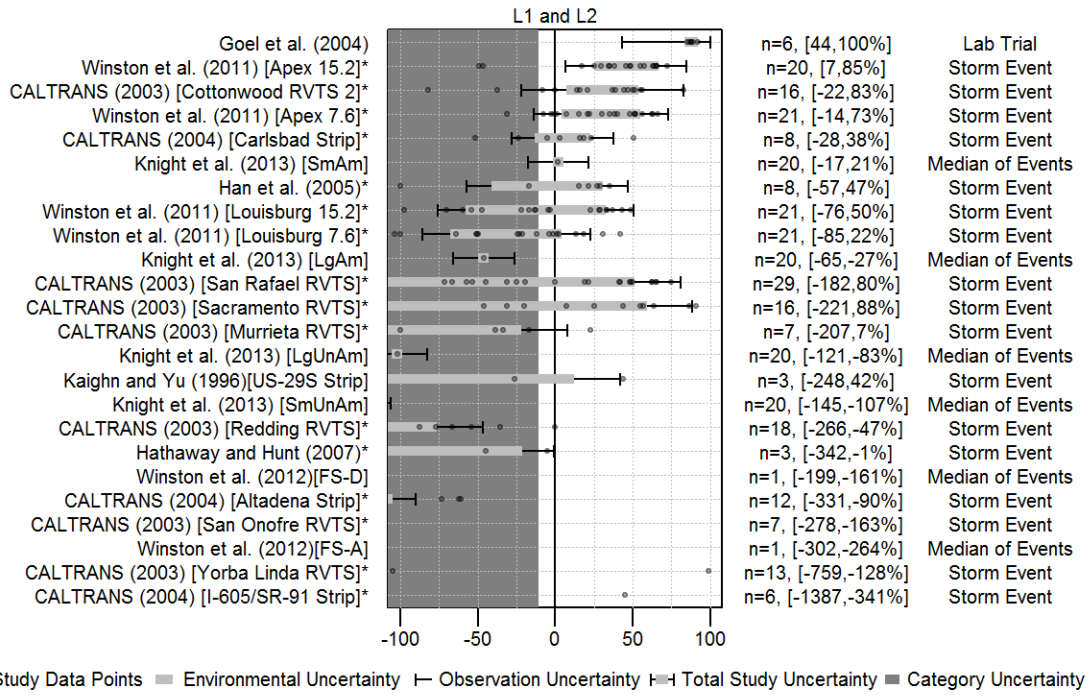


Figure C.19 – Total Phosphorus pollutant removal capabilities of Vegetated Filter Strip

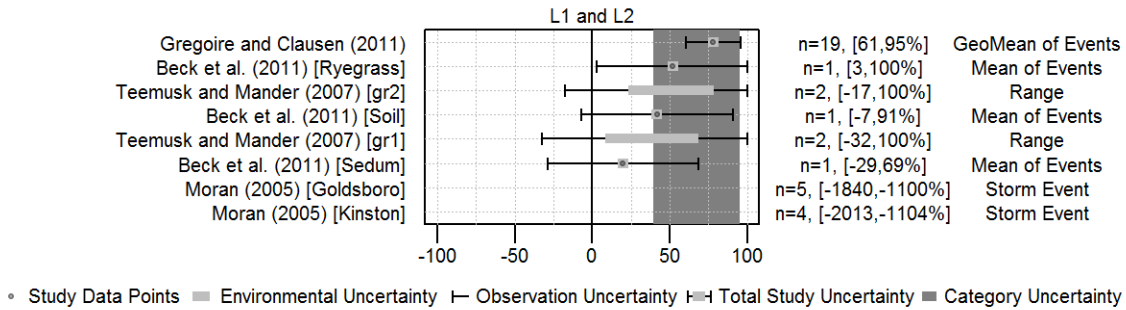


Figure C.20 – Total Phosphorus pollutant removal capabilities of Vegetated Roof. Results from Moran (2005) were excluded, as the soil media used composted cow manure to improve vegetation growth, leading to significant export of nutrients.

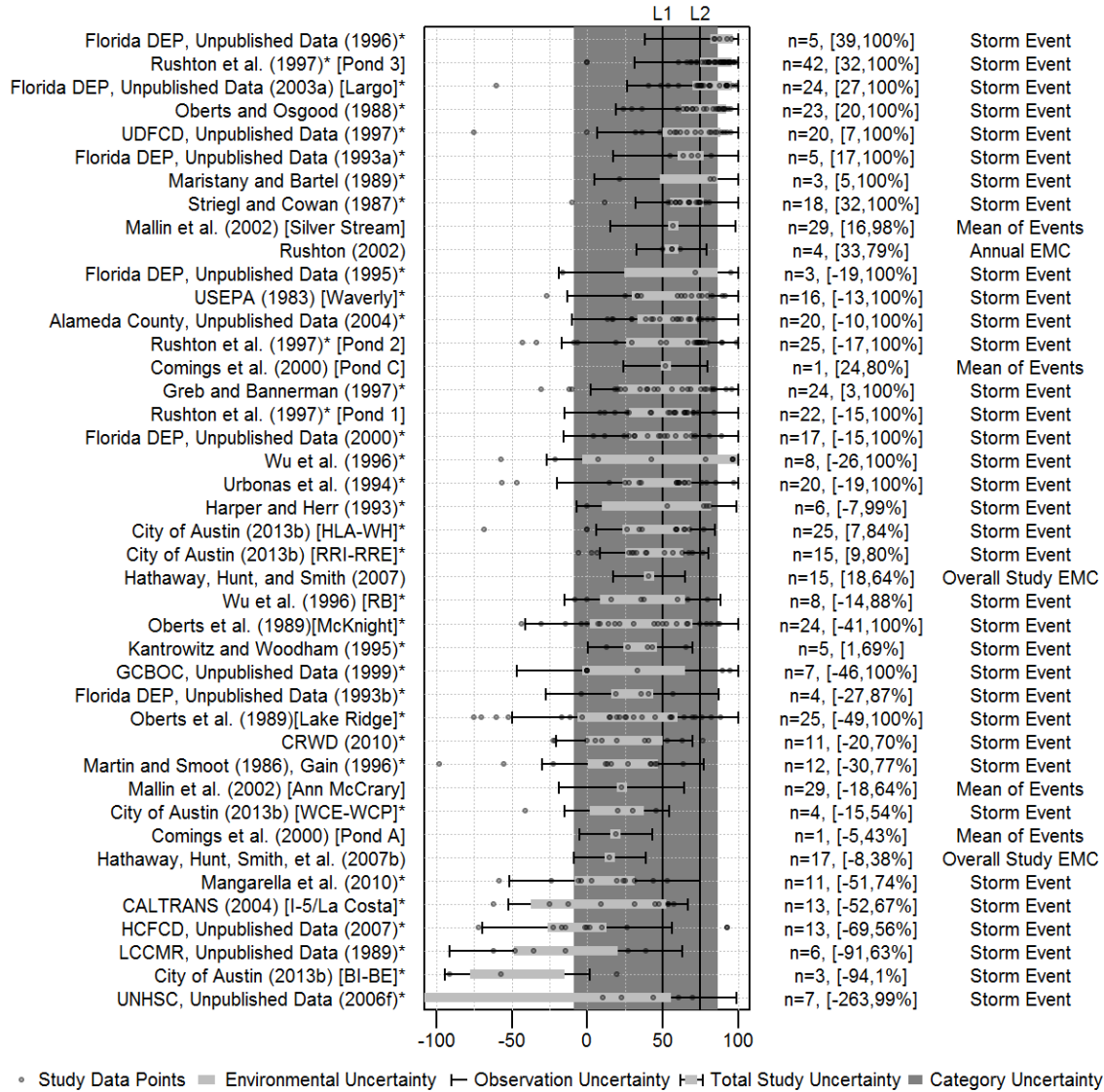


Figure C.21 – Total Phosphorus pollutant removal capabilities of Wet Ponds

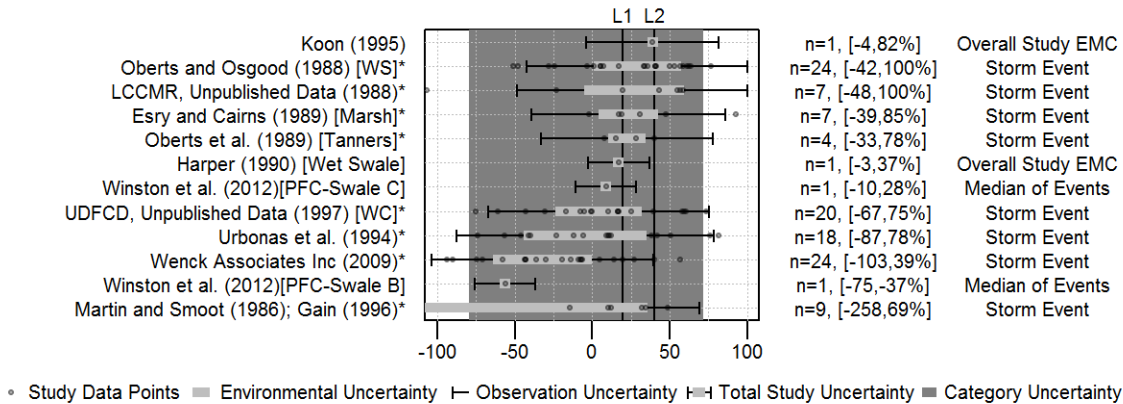


Figure C.22 – Total Phosphorus pollutant removal capabilities of Wet Swales

C.3 TOTAL NITROGEN POLLUTANT REMOVAL

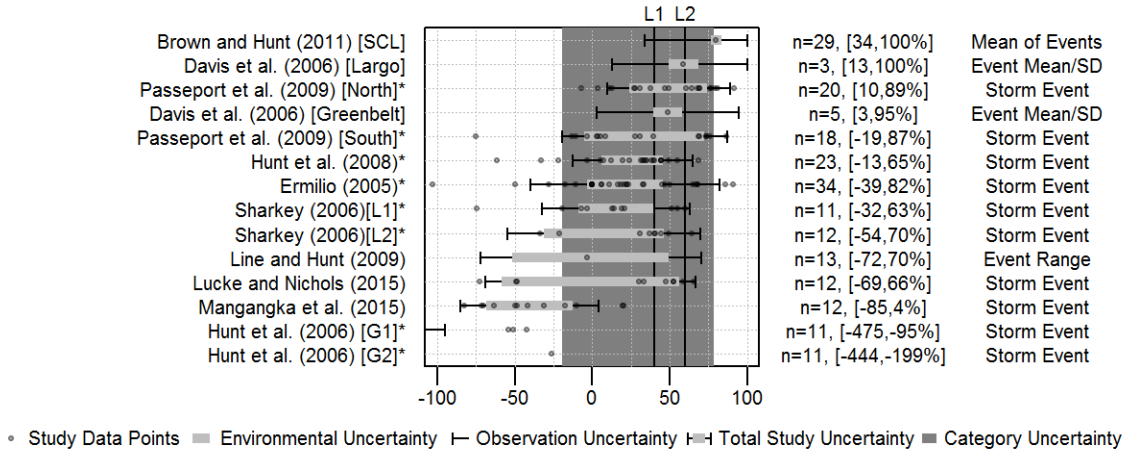


Figure C.23 – Total Nitrogen removal capabilities of Bioretention

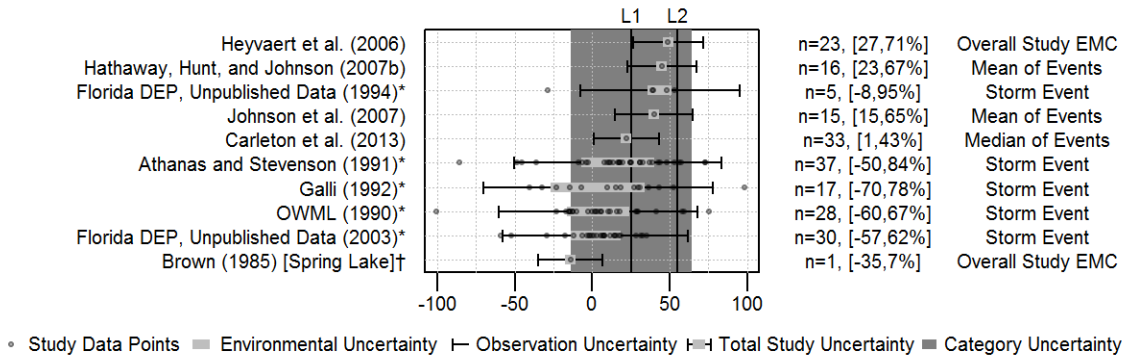


Figure C.24 – Total Nitrogen removal capabilities of Constructed Wetlands

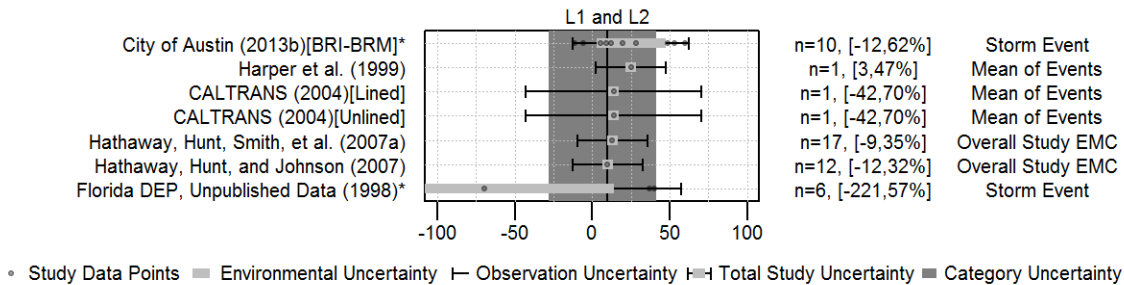


Figure C.25 – Total Nitrogen removal capabilities of Extended Detention

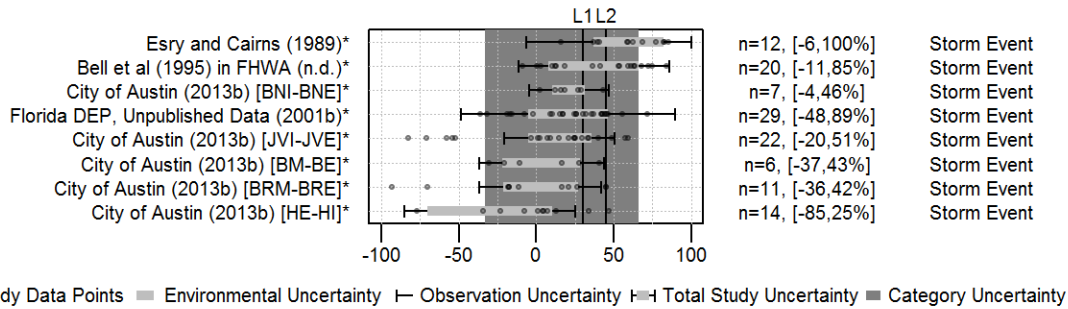


Figure C.26 – Total Nitrogen removal capabilities of Filtering Practices

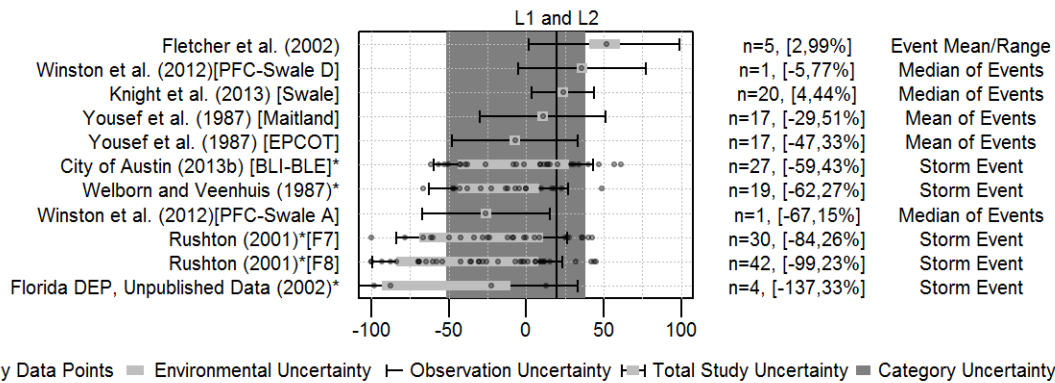


Figure C.27 – Total Nitrogen removal capabilities of Grass Channels

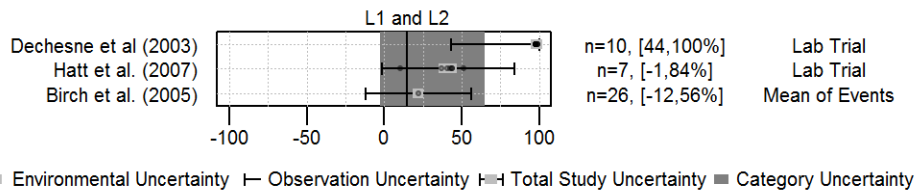


Figure C.28 – Total Nitrogen removal capabilities of Infiltration

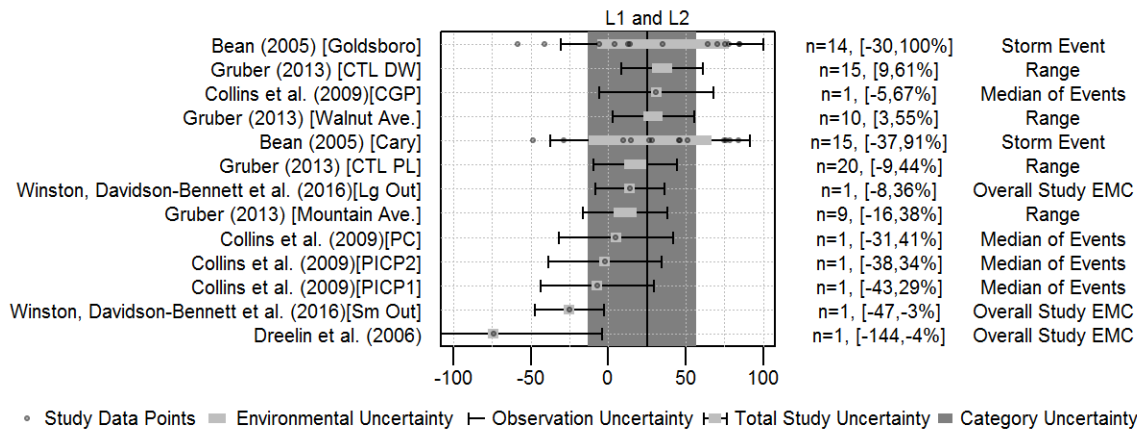


Figure C.29 – Total Nitrogen removal capabilities of Permeable Pavement

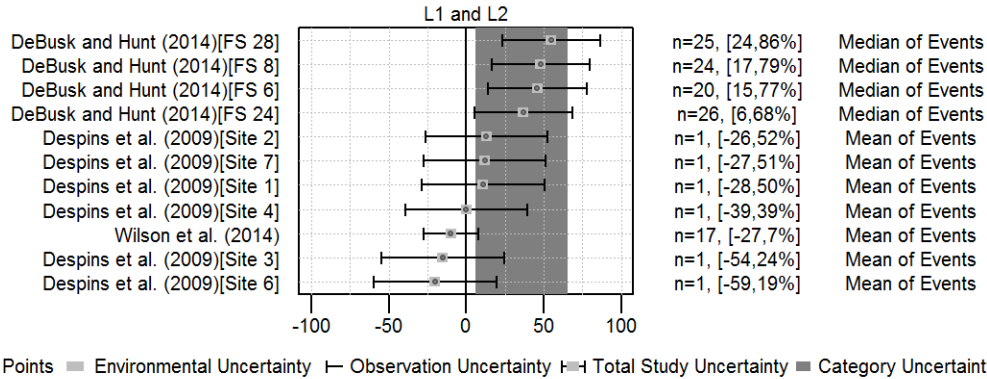


Figure C.30 – Total Nitrogen removal capabilities of Rainwater Harvesting

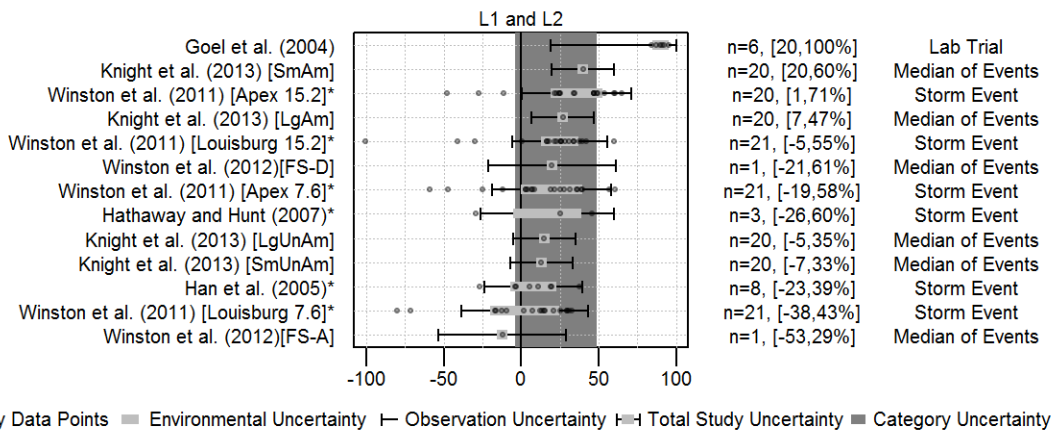


Figure C.31 – Total Nitrogen removal capabilities of Vegetated Filter Strips

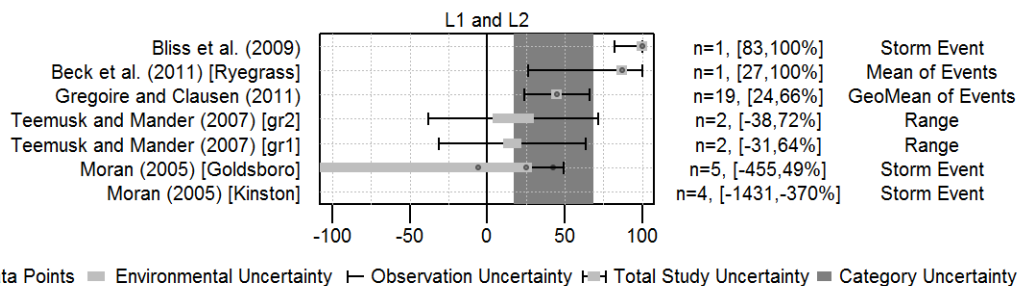


Figure C.32 – Total Nitrogen removal capabilities of Vegetated Roof. Results from Moran (2005) were excluded, as the soil media used composted cow manure to improve vegetation growth, leading to significant export of nutrients.

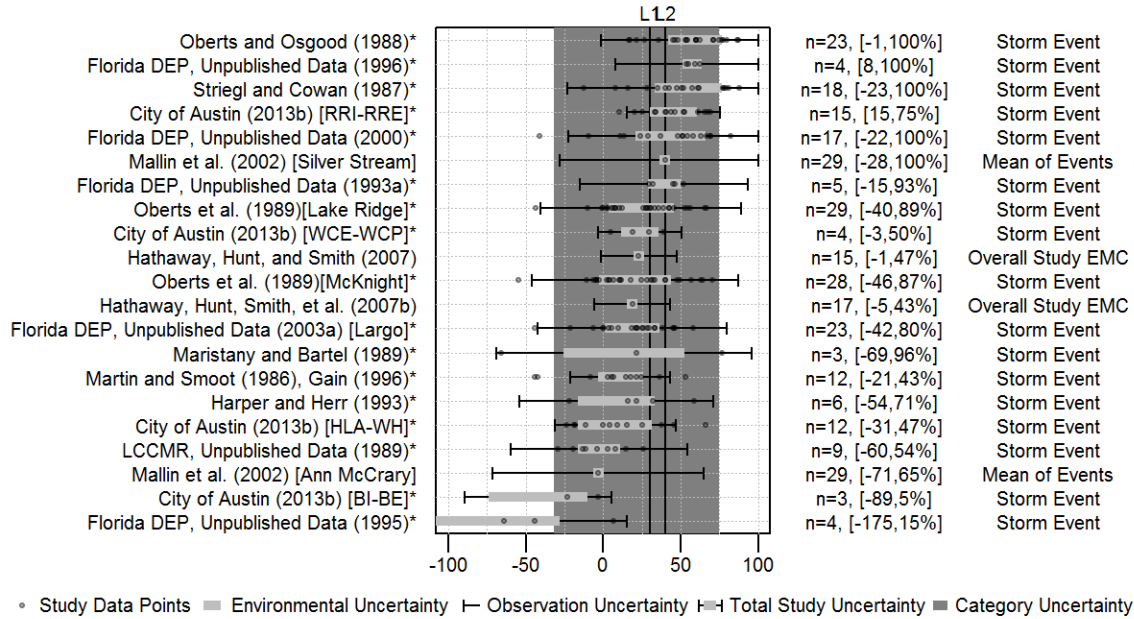


Figure C.33 – Total Nitrogen removal capabilities of Wet Pond

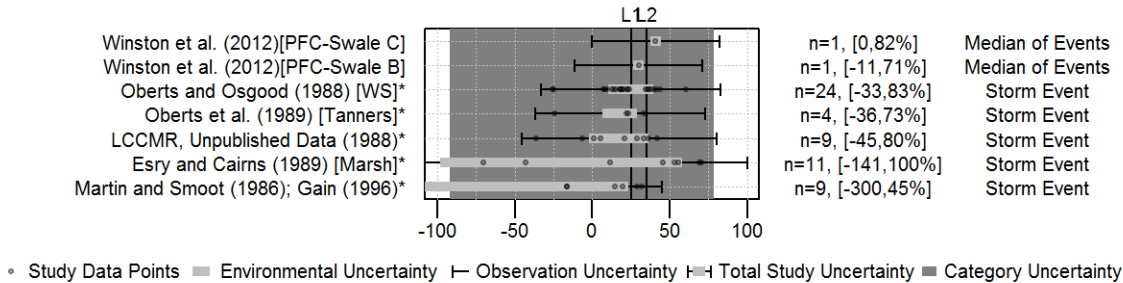


Figure C.34 – Total Nitrogen removal capabilities of Wet Swale

C.4 TOTAL SUSPENDED SOLIDS POLLUTANT REMOVAL

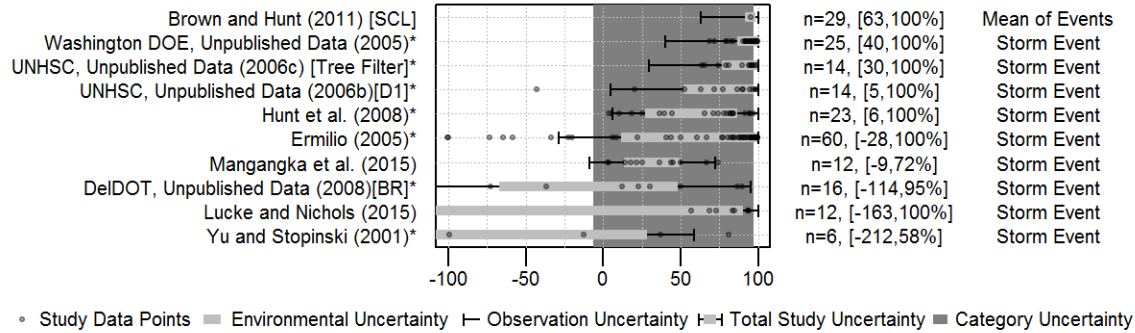


Figure C.35 – Total Suspended Solids removal capabilities of Bioretention

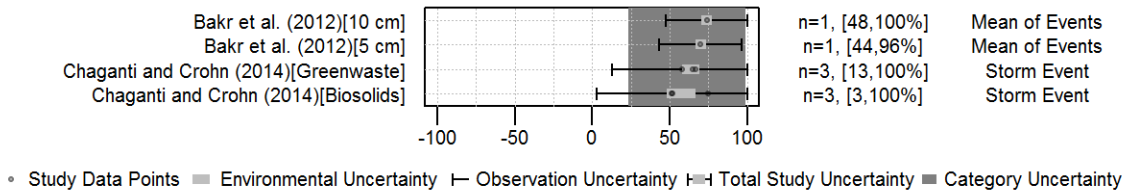


Figure C.36 – Total Suspended Solids removal capabilities of Compost Soil Amendment

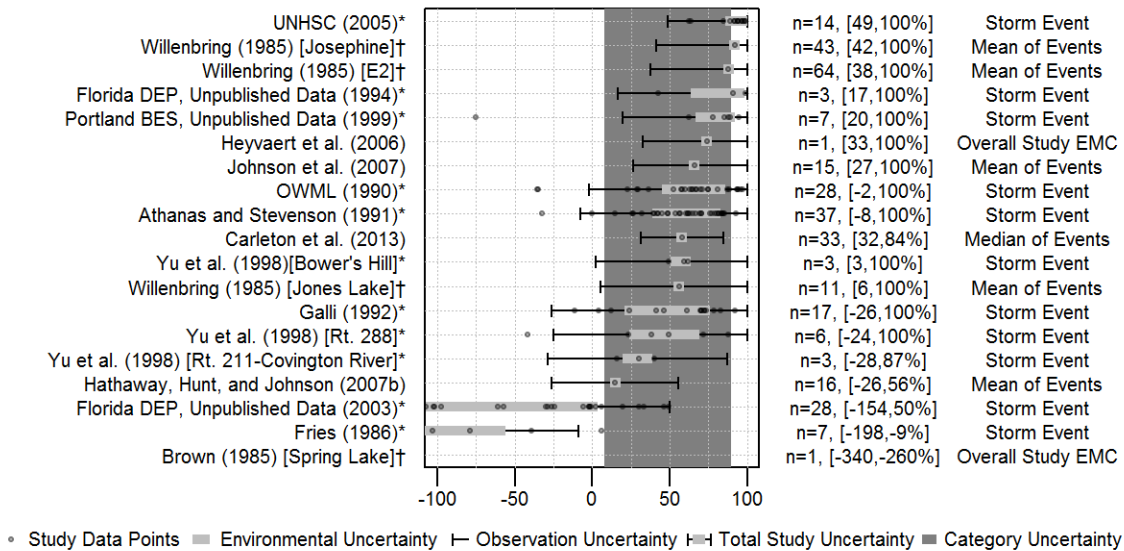


Figure C.37 – Total Suspended Solids removal capabilities of Constructed Wetlands

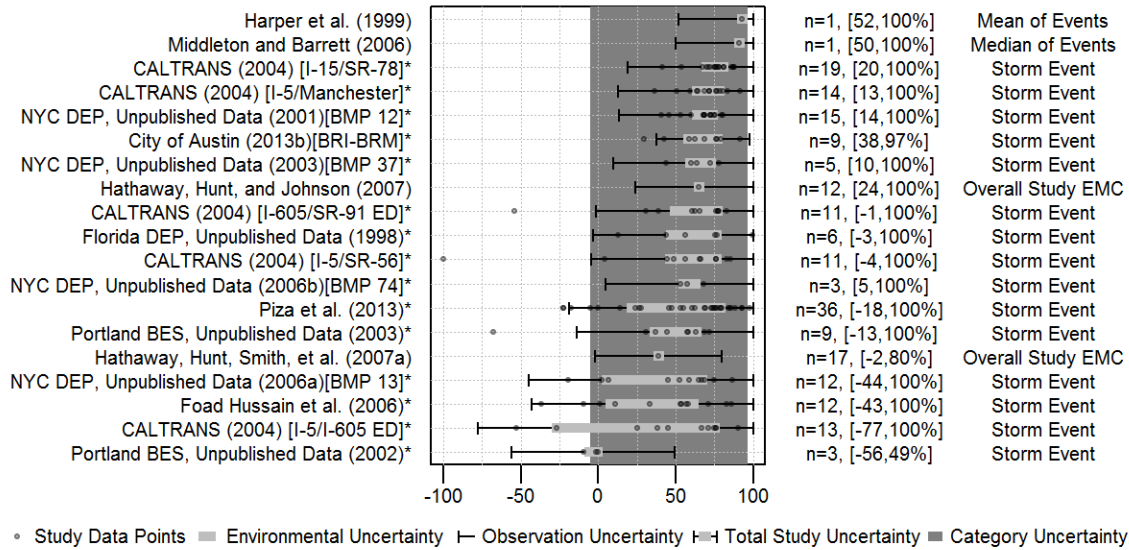


Figure C.38 – Total Suspended Solids removal capabilities of Extended Detention

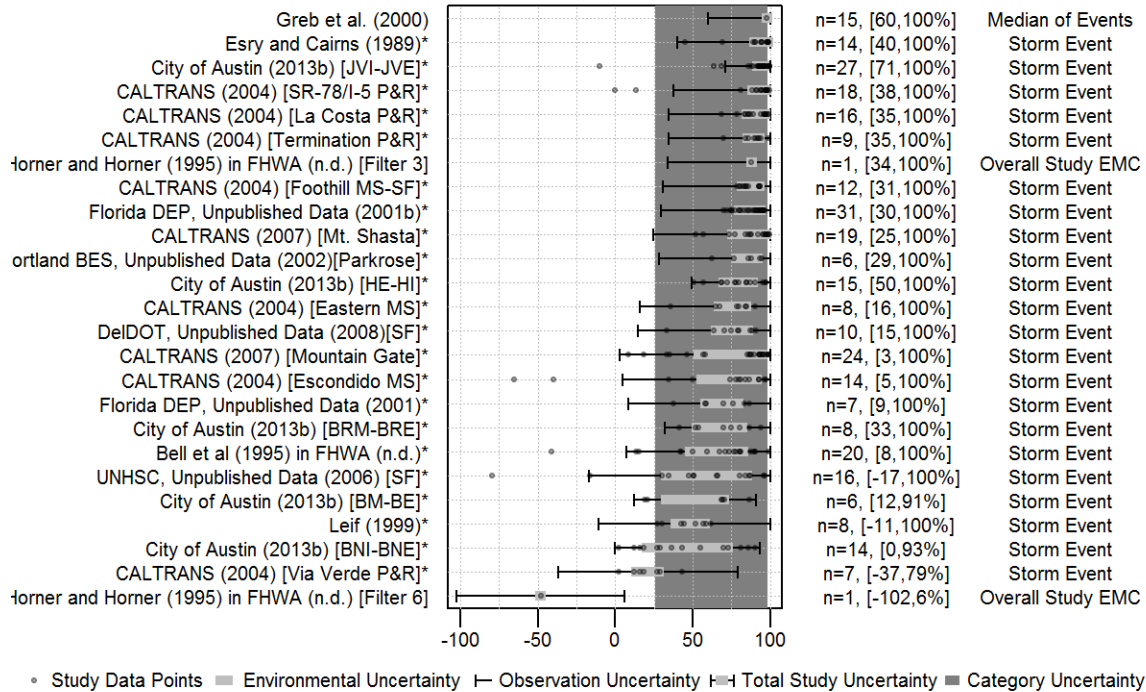


Figure C.39 – Total Suspended Solids removal capabilities of Filtering Practices

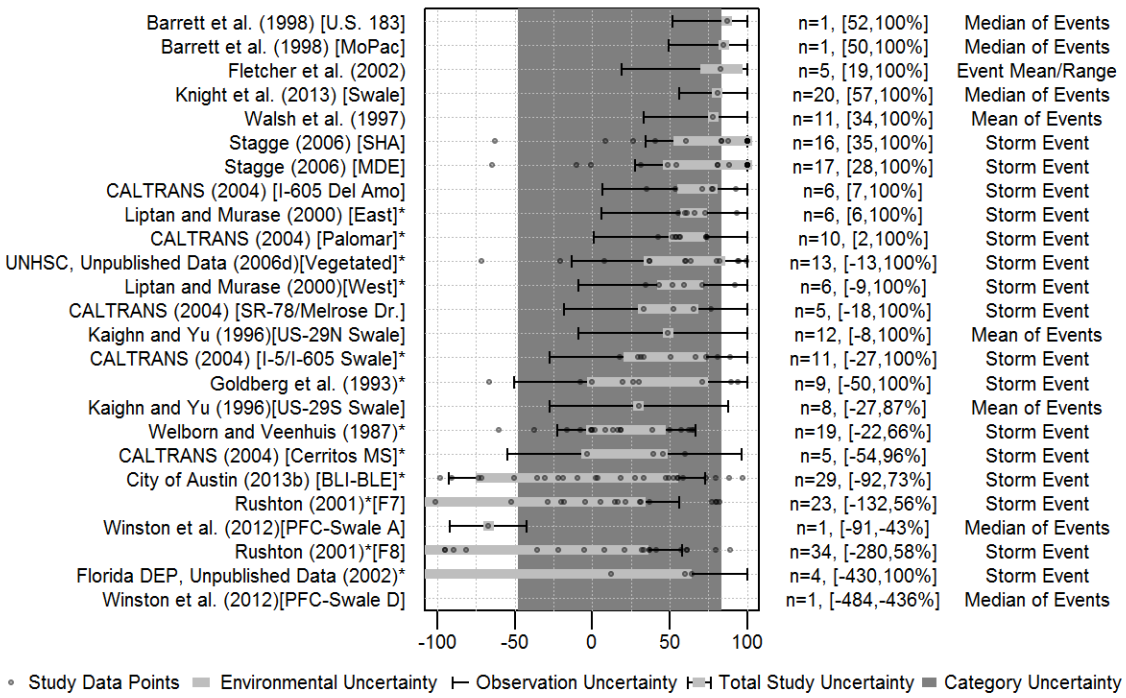


Figure C.40 – Total Suspended Solids removal capabilities of Grass Channels

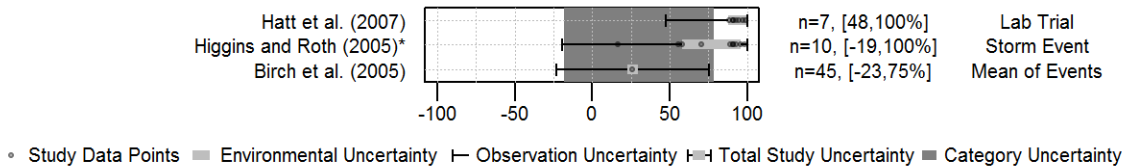


Figure C.41 – Total Suspended Solids removal capabilities of Infiltration

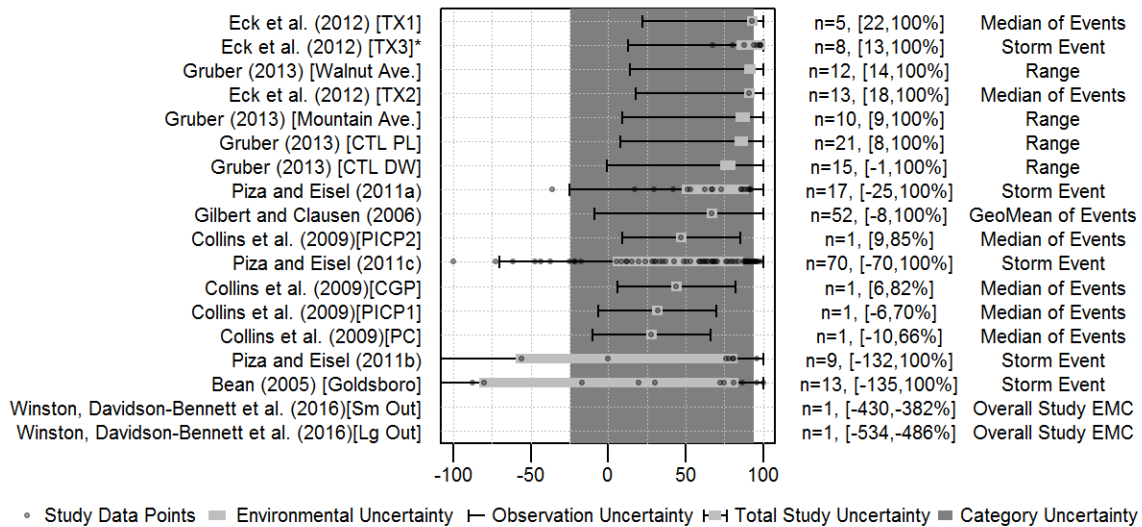


Figure C.42 – Total Suspended Solids removal capabilities of Permeable Pavement

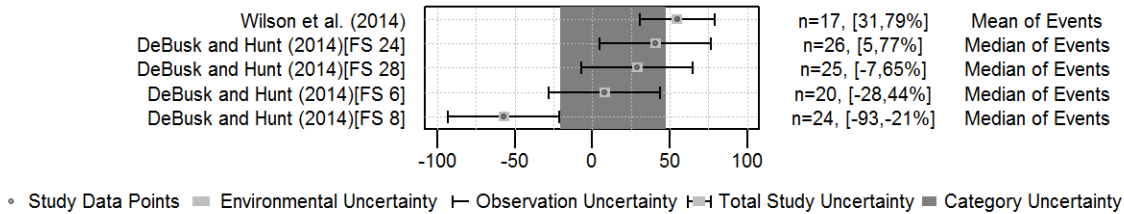


Figure C.43 – Total Suspended Solids removal capabilities of Rainwater Harvesting

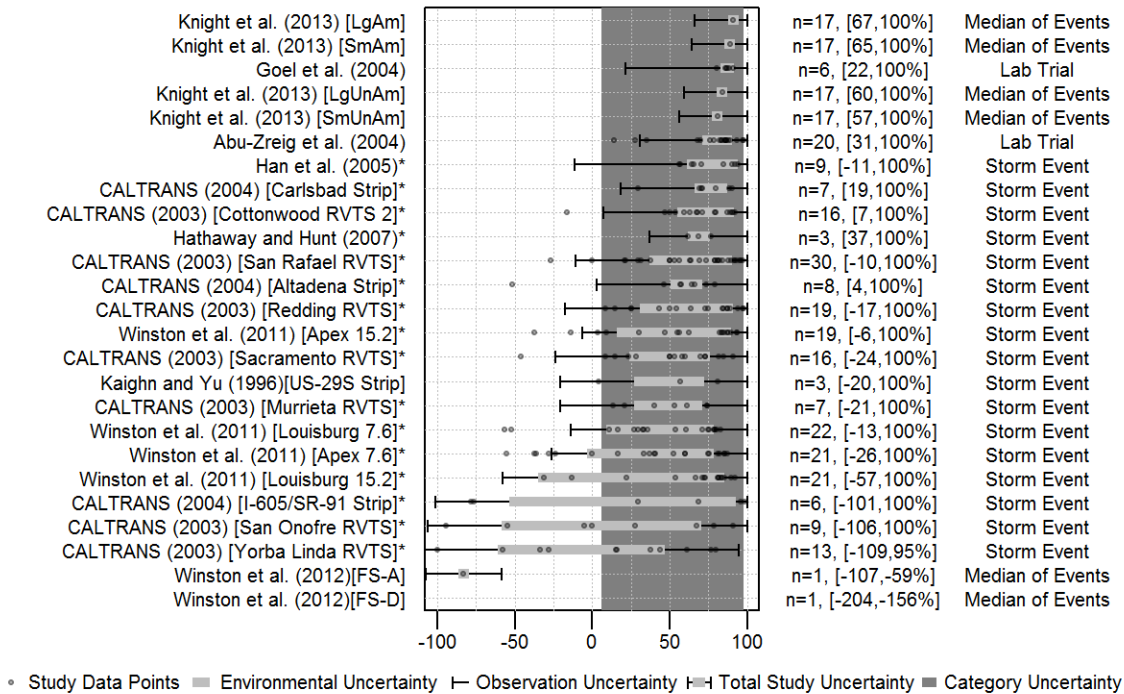


Figure C.44 – Total Suspended Solids removal capabilities of Vegetated Filter Strips

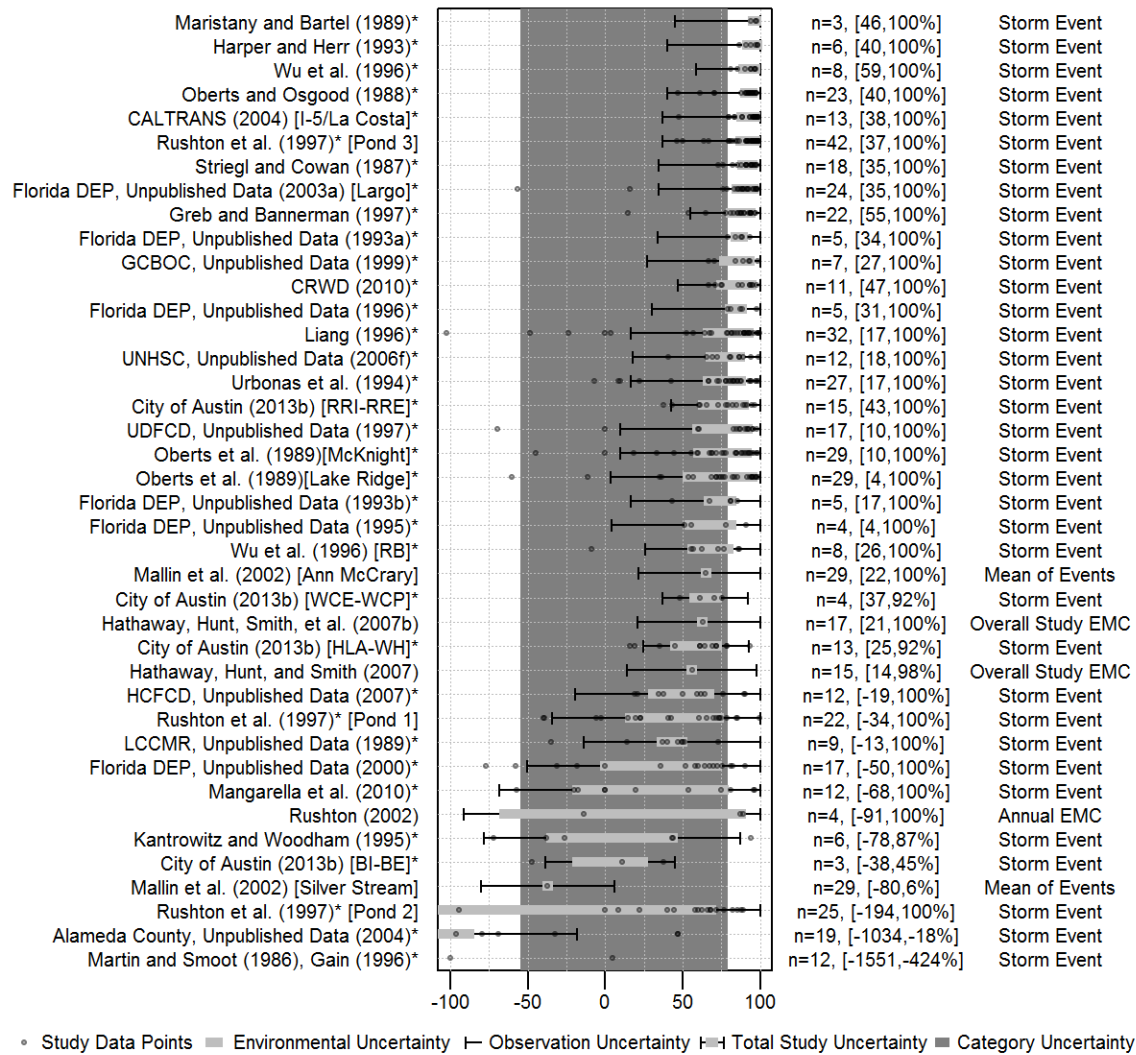


Figure C.45 – Total Suspended Solids removal capabilities of Wet Ponds

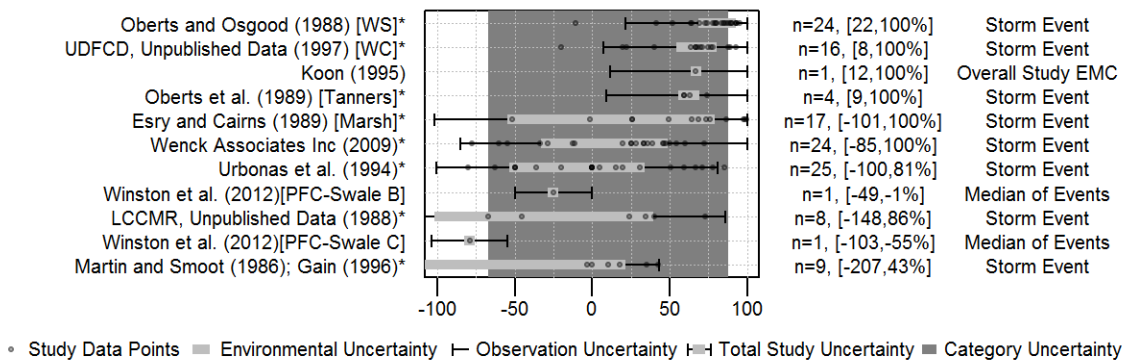


Figure C.46 – Total Suspended Solids removal capabilities of Wet Swales

APPENDIX D - REFERENCES USED TO DEVELOP STORMWATER CONTROL MEASURE PERFORMANCE UNCERTAINTY

D.1 TABLE OF REFERENCES

This Appendix provides a table of all the references used to develop environmental and observation uncertainty bounds presented in Chapter 5. The table is organized into the 15 stormwater control measure (SCM) categories, and four pollutants of concern: runoff volume reduction (RR), and pollutant removal (PR) for total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS). References are shown in the table as in-line citations, and the full reference is available in the References section. Sources noted as “Unpublished Data” were extracted from the 12/20/14 version of the from the International Stormwater BMP Database (WERF et al., 2014), but no documentation could be found describing how the data were collected.

Table D.1 - Studies reviewed for percent removal values for runoff volume reduction (RR), and reduction in EMC concentration as total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS).

Stormwater Control Measure Category	Runoff Reduction (RR)	Total Phosphorus (TP)	Total Nitrogen (TN)	Total Suspended Solids (TSS)
Bioretention	(R. A. Brown and Hunt, 2011, 2012; Dietz and Clausen, 2006; Emerson and Traver, 2008; Ermilio, 2005; William F. Hunt et al., 2006, 2008; Line and Hunt, 2009; Lloyd et al., 2002; Lucke and Nichols, 2015; Mangangka et al., 2015; Passeport et al., 2009; Sharkey, 2006; Smith and Hunt, 2007; Van Seters et al., 2006; Winston, Dorsey, et al., 2016)	(Davis et al., 2001, 2006; DelDOT Unpublished Data, 2008a; Ermilio, 2005; William F. Hunt et al., 2006, 2008; Lucke and Nichols, 2015; Mangangka et al., 2015; Passeport et al., 2009; Sharkey, 2006; UNHSC Unpublished Data, 2006a, 2006b; Washington DOE Unpublished Data, 2005; Yu and Stopinski, 2001)	(Davis et al., 2006; Ermilio, 2005; William F. Hunt et al., 2006, 2008; Lucke and Nichols, 2015; Mangangka et al., 2015; Passeport et al., 2009; Sharkey, 2006)	(DelDOT Unpublished Data, 2008a; Ermilio, 2005; William F. Hunt et al., 2008; Lucke and Nichols, 2015; Mangangka et al., 2015; UNHSC Unpublished Data, 2006a, 2006b; Washington DOE Unpublished Data, 2005; Yu and Stopinski, 2001)
Compost Soil Amendment	(Balousek, 2003; Beighley et al., 2010; Kolsti et al., 1995)	No EMC PR values found, but see (Faucette et al., 2005) for TR values	No EMC values found, but see (Faucette et al., 2005) for TR values	(Bakr et al., 2012)
Dry Swale	(Richard R Horner et al., 2004)	No studies found	No studies found	No studies found
Extended Detention	No RR expected	(CALTRANS, 2004; City of Austin, 2013; Florida DEP Unpublished Data, 1998; Foad Hussain et al., 2006; Harper et al., 1999; Hathaway, Hunt, and Johnson, 2007a; Hathaway, Hunt, Smith, et al., 2007a; Middleton and Barrett, 2006; NYC DEP Unpublished Data, 2006a, 2006b, 2001, 2003; Piza et al., 2013; Portland BES Unpublished Data, 2002a, 2003)	(CALTRANS, 2004; City of Austin, 2013; Florida DEP Unpublished Data, 1998; Harper et al., 1999; Hathaway, Hunt, and Johnson, 2007a; Hathaway, Hunt, Smith, et al., 2007a)	(CALTRANS, 2004; City of Austin, 2013; Florida DEP Unpublished Data, 1998; Foad Hussain et al., 2006; Harper et al., 1999; Hathaway, Hunt, Smith, et al., 2007b; Middleton and Barrett, 2006; NYC DEP Unpublished Data, 2006b, 2001, 2003, 2006a; Piza et al., 2013; Portland BES Unpublished Data, 2003)

Stormwater Control Measure Category	Runoff Reduction (RR)	Total Phosphorus (TP)	Total Nitrogen (TN)	Total Suspended Solids (TSS)
Filtering Practices	No RR expected	(Barrett, 2003; Bell et al., 1995; CALTRANS, 2007; City of Austin, 2013; Esry and Cairns, 1989; FHWA, n.d.-a, n.d.-b; Florida DEP Unpublished Data, 2001a, 2001b; Greb et al., 2000; R. R. Horner and Horner, 1995; Leif, 1999; New Castle County Delaware, 2008; UNHSC Unpublished Data, 2006c)	(Bell et al., 1995; City of Austin, 2013; Esry and Cairns, 1989; FHWA, n.d.-a; Florida DEP Unpublished Data, 2001b)	(Bell et al., 1995; CALTRANS, 2004, 2007; City of Austin, 2013; DeIDOT Unpublished Data, 2008b; Esry and Cairns, 1989; FHWA, n.d.-a; Florida DEP Unpublished Data, 2001a, 2001b; R. R. Horner and Horner, 1995; Leif, 1999; Portland BES Unpublished Data, 2002b; UNHSC Unpublished Data, 2006c)
Grass Channel (“Biofilter – Grass Swale”)	(CALTRANS, 2004; Kaighn and Yu, 1996; Knight et al., 2013; Liptan and Murase, 2000; Stagge, 2006)	(Barrett et al., 1998; CALTRANS, 2004; City of Austin, 2013; Fletcher et al., 2002; Florida DEP Unpublished Data, 2002; Kaighn and Yu, 1996; Knight et al., 2013; Liptan and Murase, 2000; Rushton, 2001; Stagge, 2006; UNHSC Unpublished Data, 2006d; Walsh et al., 1997; Welborn and Veenhuis, 1987; Winston et al., 2012; Yousef et al., 1987)	(City of Austin, 2013; Fletcher et al., 2002; Florida DEP Unpublished Data, 2002; Knight et al., 2013; Rushton, 2001; Welborn and Veenhuis, 1987; Winston et al., 2012; Yousef et al., 1987)	(Barrett et al., 1998; CALTRANS, 2004; City of Austin, 2013; Fletcher et al., 2002; Florida DEP Unpublished Data, 2002; Goldberg et al., 1993; Kaighn and Yu, 1996; Knight et al., 2013; Liptan and Murase, 2000; Rushton, 2001; Stagge, 2006; Walsh et al., 1997; Welborn and Veenhuis, 1987; Winston et al., 2012)
Green Roof	(DeNardo et al., 2005; Getter et al., 2007; Hutchinson et al., 2003; Moran, 2005; VanWoert et al., 2005)	(Beck et al., 2011; Gregoire and Clausen, 2011; Moran, 2005; Teemusk and Mander, 2007)	(Beck et al., 2011; Gregoire and Clausen, 2011; Moran, 2005; Teemusk and Mander, 2007)	No Studies Found
Infiltration	(Bright, 2007; USEPA, 1983)	(Birch et al., 2005; Dechesne et al., 2004; Hatt et al., 2007; Higgins and Roth, 2005)	(Birch et al., 2005; Dechesne et al., 2004; Hatt et al., 2007)	(Birch et al., 2005; Hatt et al., 2007)

Stormwater Control Measure Category	Runoff Reduction (RR)	Total Phosphorus (TP)	Total Nitrogen (TN)	Total Suspended Solids (TSS)
Permeable Pavement	(Andersen et al., 1999; Bean, 2005; Bean et al., 2007; Brattebo and Booth, 2003; Collins et al., 2008; Dreelin et al., 2006; Gruber, 2013; Horst et al., 2011; Legret and Colandini, 1999; Macdonald and Jefferies, 2002; Pratt et al., 1989; Schlüter and Jefferies, 2002; Schueler, 1987; Van Seters et al., 2006)	(Bean, 2005; Dreelin et al., 2006; Eck et al., 2012; Gilbert and Clausen, 2006; Gruber, 2013; Piza and Eisel, 2011a, 2011b, 2011c; UNHSC Unpublished Data, 2006e; Winston, Davidson-Bennett, et al., 2016)	(Bean, 2005; Collins et al., 2009; Dreelin et al., 2006; Gruber, 2013; Winston, Davidson-Bennett, et al., 2016)	(Bean, 2005; Collins et al., 2009; Eck et al., 2012; Gilbert and Clausen, 2006; Gruber, 2013; Higgins and Roth, 2005; Piza and Eisel, 2011a, 2011b, 2011c; Winston, Davidson-Bennett, et al., 2016)
Rainwater Harvesting	(Gee and Hunt, 2016)	(DeBusk and Hunt, 2014; Wilson et al., 2014)	(DeBusk and Hunt, 2014; Despins et al., 2009; Wilson et al., 2014)	(DeBusk and Hunt, 2014; Wilson et al., 2014)
Rooftop Disconnection	(Carmen et al., 2016)	No Studies Found	No Studies Found	No Studies Found
Sheetflow to Conserved Open Space or Vegetated Filter Strip	(Abu-Zreig et al., 2004; CALTRANS, 2003, 2004; W. F. Hunt et al., 2010; Knight et al., 2013; Line and Hunt, 2009; Winston et al., 2012)	(CALTRANS, 2003, 2004; Goel et al., 2004; Han et al., 2005; Hathaway and Hunt, 2007; W. F. Hunt et al., 2010; Kaighn and Yu, 1996; Winston et al., 2011, 2012)	(Goel et al., 2004; Han et al., 2005; Hathaway and Hunt, 2007; W. F. Hunt et al., 2010; Winston et al., 2011, 2012)	(Abu-Zreig et al., 2004; CALTRANS, 2003, 2004; Goel et al., 2004; Han et al., 2005; Hathaway and Hunt, 2007; W. F. Hunt et al., 2010; Kaighn and Yu, 1996; Winston et al., 2011, 2012)

Stormwater Control Measure Category	Runoff Reduction (RR)	Total Phosphorus (TP)	Total Nitrogen (TN)	Total Suspended Solids (TSS)
Wet Pond	No RR expected	(Alameda County Unpublished Data, 2004; CALTRANS, 2004; City of Austin, 2013; Comings et al., 2000; CRWD, 2010; Florida DEP Unpublished Data, 1993a, 1993b, 1995, 1996, 2000; Gain, 1996; GCBOD Unpublished Data, 1999; Greb and Bannerman, 1997; Harper and Herr, 1993; Hathaway, Hunt, and Smith, 2007; Hathaway, Hunt, Smith, et al., 2007b; HCFCD Unpublished Data, 2007; LCCMR Unpublished Data, 1989; Mallin et al., 2002; Mangarella et al., 2010; Maristany and Bartel, 1989; Martin and Smoot, 1986; Martin, 1988; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; Rushton, 2002; Rushton et al., 1997; Stoker, 1996; Striegl and Cowan, 1987; UDFCD Unpublished Data, 1997; UNHSC Unpublished Data, 2006f; USEPA, 1983; Wu et al., 1996)	(City of Austin, 2013; Florida DEP Unpublished Data, 1993a, 1995, 1996, 2000, 2003a; Gain, 1996; Harper and Herr, 1993; Hathaway, Hunt, and Smith, 2007; Hathaway, Hunt, Smith, et al., 2007b; LCCMR Unpublished Data, 1989; Mallin et al., 2002; Maristany and Bartel, 1989; Martin and Smoot, 1986; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; Striegl and Cowan, 1987)	(Alameda County Unpublished Data, 2004; CALTRANS, 2004; City of Austin, 2013; Florida DEP Unpublished Data, 1993a, 1993b, 1995, 1996, 2000, 2003a; Gain, 1996; GCBOD Unpublished Data, 1999; Greb and Bannerman, 1997; Harper and Herr, 1993; Hathaway, Hunt, and Smith, 2007; Hathaway, Hunt, Smith, et al., 2007b; HCFCD Unpublished Data, 2007; Kantrowitz and Woodham, 1995; LCCMR Unpublished Data, 1989; Liang, 1996; Mallin et al., 2002; Mangarella et al., 2010; Maristany and Bartel, 1989; Martin and Smoot, 1986; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; Rushton, 2002; Rushton et al., 1997; Striegl and Cowan, 1987; UDFCD Unpublished Data, 1997; UNHSC Unpublished Data, 2006f; Urbonas et al., 1994; Wu et al., 1996)

Stormwater Control Measure Category	Runoff Reduction (RR)	Total Phosphorus (TP)	Total Nitrogen (TN)	Total Suspended Solids (TSS)
Wet Swale	No RR expected	(Esry and Cairns, 1989; Gain, 1996; Harper, 1990; Koon, 1995; LCCMR Unpublished Data, 1988; Martin and Smoot, 1986; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; UDFCD Unpublished Data, 1997; Urbonas et al., 1994; Wenck Associates Inc., 2009; Winston et al., 2012)	(Esry and Cairns, 1989; Gain, 1996; LCCMR Unpublished Data, 1988; Martin and Smoot, 1986; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; Winston et al., 2012)	(Esry and Cairns, 1989; Gain, 1996; LCCMR Unpublished Data, 1988; Martin and Smoot, 1986; G. L. Oberts et al., 1989; G. Oberts and Osgood, 1988; UDFCD Unpublished Data, 1997; Urbonas et al., 1994; Wenck Associates Inc., 2009; Winston et al., 2012)
Wetland	No RR expected	(Athanas and Stevenson, 1991; R. G. Brown, 1985; Carleton et al., 2013; Carr and Rushton, 1993; Florida DEP Unpublished Data, 1994, 2003b; Fries, 1986; Gain, 1996; Galli, 1992; Hathaway, Hunt, and Smith, 2007; Heyvaert et al., 2006; Johnson et al., 2007; L.A. County Unpublished Data, 2010; Martin, 1988; OWML, 1990; Portland BES Unpublished Data, 1999; Schiffer, 1989; UNHSC, 2005; Willenbring, 1985; Yu et al., 1998)	(Athanas and Stevenson, 1991; R. G. Brown, 1985; Carleton et al., 2013; Carr and Rushton, 1993; Florida DEP Unpublished Data, 1994, 2003b; Gain, 1996; Galli, 1992; Hathaway, Hunt, and Johnson, 2007b; Heyvaert et al., 2006; Johnson et al., 2007; Martin, 1988; OWML, 1990)	(Athanas and Stevenson, 1991; R. G. Brown, 1985; Carleton et al., 2013; Carr and Rushton, 1993; Florida DEP Unpublished Data, 1994, 2003b; Fries, 1986; Gain, 1996; Galli, 1992; Hathaway, Hunt, and Johnson, 2007b; Heyvaert et al., 2006; Johnson et al., 2007; Martin, 1988; OWML, 1990; Portland BES Unpublished Data, 1999; UNHSC, 2005; Yu et al., 1998)

D.2 LIST OF REFERENCES

- Abu-Zreig, M., Rudra, R. P., Lalonde, M. N., Whiteley, H. R., and Kaushik, N. K. (2004). Experimental investigation of runoff reduction and sediment removal by vegetated filter strips. *Hydrological Processes*, 18(11), 2029–2037. doi:10.1002/hyp.1400.
- Alameda County Unpublished Data. (2004). *Heron Bay Wet Pond*.
- Andersen, C. T., Foster, I. D. L., and Pratt, C. J. (1999). The role of urban surfaces (permeable pavements) in regulating drainage and evaporation : development of a laboratory simulation experiment. *Hydrological Processes*, 13, 597–609.
- Athanas, C., and Stevenson, J. C. (1991). *The use of artificial wetlands in treating stormwater runoff*. Maryland Sediment and Stormwater Administration.
- Bakr, N., Weindorf, D. C., Zhu, Y., Arceneaux, A. E., and Selim, H. M. (2012). Evaluation of compost/mulch as highway embankment erosion control in Louisiana at the plot-scale. *Journal of Hydrology*, 468, 257–267. doi:10.1016/j.jhydrol.2012.08.040.
- Balousek, J. D. (2003). *Quantifying Decreases in Stormwater Runoff from Deep Tilling, Chisel Plowing, and Compost-Amendment*. Dane County, WI: Dane County Land Conservation Department. Retrieved from <http://www.clflwd.org/documents/quantifyingdecreasesinswrunoff.pdf>.
- Barrett, M. E. (2003). Performance, Cost, and Maintenance Requirements of Austin Sand Filters. *Journal of Water Resources Planning and Management*, 129(3), 234–242. doi:10.1061/(ASCE)0733-9496(2003)129:3(234).
- Barrett, M. E., Walsh, P. M., Jr., J. F. M., and Charbeneau, R. J. (1998). Performance of Vegetative Controls for Treating Highway Runoff. *Journal of Environmental Engineering*, 124(November), 1121–1128. doi:10.1061/(ASCE)0733-9372(1998)124:11(1121).
- Bean, E. Z. (2005). *A Field Study to Evaluate Permeable Pavement Surface Infiltration Rates, Runoff Quantity, Runoff Quality, and Exfiltrate Quality*. North Carolina State University.
- Bean, E. Z., Hunt, W. F., and Bidelspach, D. A. (2007). Evaluation of four permeable pavement sites in eastern North Carolina for runoff reduction and water quality impacts. *Journal of Irrigation and Drainage Engineering*, 133(6), 583–592. doi:10.1061/(asce)0733-9437(2007)133:6(583).
- Beck, D. A., Johnson, G. R., and Spolek, G. A. (2011). Amending greenroof soil with biochar to affect runoff water quantity and quality. *Environmental Pollution*, 159(8–9), 2111–2118.

doi:10.1016/j.envpol.2011.01.022.

- Beighley, R. E., Scholl, B., Faucette, L. B., and Governo, J. (2010). Runoff Characteristics for Construction Site Erosion Control Practices. *Journal of Irrigation & Drainage Engineering*, 136(6), 405–413. doi:10.1061/(ASCE)IR.1943-4774.0000202.
- Bell, W., Stokes, L., Gavan, J., and Nguyen, T. N. (1995). *Assessment of the Pollutant Removal Efficiencies of Delaware Sand Filter BMPs*. Alexandria, VA: City of Alexandria, Dept. of Transportation and Environmental Services.
- Birch, G. F., Fazeli, M. S., and Matthai, C. (2005). Efficiency of an infiltration basin in removing contaminants from urban stormwater. *Environmental Monitoring and Assessment*, 101(1–3), 23–38. doi:10.1007/s10661-005-9126-0.
- Brattebo, B. O., and Booth, D. B. (2003). Long-term stormwater quantity and quality performance of permeable pavement systems. *Water Research*, 37(18), 4369–4376. doi:10.1016/S0043-1354(03)00410-X.
- Bright, T. M. (2007). An Examination of a Dune Infiltration System's Impact on Coastal Hydrology and Bacteria Removal. M.S. Thesis, North Carolina State University. Dept. of Biological and Agricultural Engineering. Raleigh, NC. Retrieved from <http://www.lib.ncsu.edu/resolver/1840.16/575>.
- Brown, R. A., and Hunt, W. F. (2011). Underdrain Configuration to Enhance Bioretention Exfiltration to Reduce Pollutant Loads. *Journal of Environmental Engineering*, 137(11), 1082–1091. doi:10.1061/(ASCE)EE.1943-7870.0000437.
- Brown, R. A., and Hunt, W. F. (2012). Improving bioretention/biofiltration performance with restorative maintenance. *Water Science and Technology*, 65(2), 361–367. doi:10.2166/wst.2012.860.
- Brown, R. G. (1985). *Effects of wetlands on quality of runoff entering lakes in the Twin Cities Metropolitan Area, Minnesota*. USGS Water-Resources Investigations Report 85-4170. St. Paul, MN: U.S. Geological Survey. Retrieved from <https://pubs.er.usgs.gov/publication/wri854170>.
- CALTRANS. (2003). *Roadside Vegetated Treatment Sites (RVTS) Study*. CTSW-RT-03-028. Sacramento, CA: CALTRANS. Retrieved from <http://www.dot.ca.gov/hq/env/stormwater/pdf/CTSW-RT-03-028.pdf>.
- CALTRANS. (2004). *BMP Retrofit Pilot Program Final Report*. CTSW-RT-01-050. Sacramento,

- CA: CALTRANS. Retrieved from <http://www.dot.ca.gov/hq/oppd/stormwtr/Studies/BMP-Retro-fit-Report.pdf>.
- CALTRANS. (2007). *Statewide [Austin] Sand Filter Study Final 2006 Stormwater Monitoring Report. CTSW-RT-06-128.01.1*. Sacramento, CA: CALTRANS.
- Carleton, J. N., Grizzard, T. J., Godrej, A. N., Post, H. E., Lampe, L., and Kenel, P. P. (2013). Performance of a Constructed Wetlands in Treating Urban Stormwater Runoff Performance in Treating of a Constructed Stormwater Wetlands Runoff Urban. *Water Environment Research*, 72(3), 295–304.
- Carmen, N. B., Hunt, W. F., and Anderson, A. R. (2016). Volume Reduction Provided by Eight Residential Disconnected Downspouts in Durham, North Carolina. *Journal of Environmental Engineering*, 1–9. doi:10.1061/(ASCE)EE.1943-7870.0001107.
- Carr, D. W., and Rushton, B. (1993). Treatment of stormwater by a native herbaceous marsh. In F. Webb (Ed.), *Proceedings of the twentieth annual conference on wetlands restoration and creation*. Tampa, FL: Hillsborough Community College. Retrieved from <http://images.library.wisc.edu/EcoNatRes/EFacs/Wetlands/Wetlands20/reference/econatres.wetlands20.d carr.pdf>.
- City of Austin. (2013). *Stormwater Control Measures in Austin , TX : Data Report. Report Number: CM-13-01*. Austin, TX: Watershed Protection Department, City of Austin. Retrieved from http://www.austintexas.gov/watershed_protection/publications/document.cfm?id=202219.
- Collins, K. A., Hunt, W. F., and Hathaway, J. M. (2008). Hydrologic Comparison of Four Types of Permeable Pavement and Standard Asphalt in Eastern North Carolina. *Journal of Hydrologic Engineering*, 15(6), 512–521. doi:10.1061/(ASCE)HE.1943-5584.0000139.
- Collins, K. A., Hunt, W. F., and Hathaway, J. M. (2009). Hydrologic and Water Quality Evaluation of Four Permeable Pavements in North Carolina, USA. In *9th International Conference on Concrete Block Paving*. Buenos Aires, Argentina: AABH, ICPA, SEPT.
- 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*, 126(April), 321–330. doi:10.1061/(ASCE)0733-9372(2000)126:4(321).
- CRWD. (2010). *BMP Performance and Cost-Benefit Analysis: Arlington Pascal Project 2007-2010*. St. Paul, MN: Capitol Region Watershed District. Retrieved from

http://www.capitolregionwd.org/wp-content/uploads/2012/09/2007_2010_BMP_Performance_MainBody.pdf.

- Davis, A. P., Shokouhian, M., Sharma, H., and Minami, C. (2001). Laboratory Study of Biological Retention for Urban Management. *Water Environment Research*, 73(1), 5–14.
- Davis, A. P., Shokouhian, M., Sharma, H., and Minami, C. (2006). Water Quality Improvement through Bioretention Media: Nitrogen and Phosphorus Removal. *Water Environment Research*, 78(3), 284–293. doi:10.2175/106143005X94376.
- DeBusk, K. M., and Hunt, W. F. (2014). Impact of rainwater harvesting systems on nutrient and sediment concentrations in roof runoff. *Water Science and Technology: Water Supply*, 14(2), 220–229. doi:10.2166/ws.2013.191.
- Dechesne, M., Barraud, S., and Bardin, J.-P. (2004). Spatial distribution of pollution in an urban stormwater infiltration basin. *Journal of Contaminant Hydrology*, 72(1–4), 189–205. doi:10.1016/j.jconhyd.2003.10.011.
- DelDOT Unpublished Data. (2008a). *I-95 Plaza Bioretention Cell*. International Stormwater BMP Database.
- DelDOT Unpublished Data. (2008b). *I-95 Plaza Delaware Sand Filter*. International Stormwater BMP Database.
- DeNardo, J. C., Jarrett, A. R., Manbeck, H. B., Beattie, D. J., and Berghage, R. D. (2005). Stormwater mitigation and surface temperature reduction by green roofs. *Transactions of the ASAE*, 48(4), 1491–1496. doi:10.13031/2013.19181.
- Despins, C., Farahbakhsh, K., and Leidl, C. (2009). Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. *Journal of Water Supply: Research and Technology - AQUA*, 58(2), 117–134. doi:10.2166/aqua.2009.013.
- Dietz, M. E., and Clausen, J. C. (2006). Saturation to Improve Pollutant Retention in a Rain Garden. *Environmental Science and Technology*, 40(4), 1335–1340. doi:10.1021/es051644f.
- Dreelin, E. A., Fowler, L., and Ronald Carroll, C. (2006). A test of porous pavement effectiveness on clay soils during natural storm events. *Water Research*, 40(4), 799–805. doi:10.1016/j.watres.2005.12.002.
- Eck, B. J., Winston, R. J., Hunt, W. F., and Barrett, M. E. (2012). Water Quality of Drainage from Permeable Friction Course. *Journal of Environmental Engineering*, 138(2), 174–181.

doi:Doi 10.1061/(Asce)Ee.1943-7870.0000476.

- Emerson, C. H., and Traver, R. G. (2008). Multiyear and Seasonal Variation of Infiltration from Storm-Water Best Management Practices. *Journal of Irrigation and Drainage Engineering*, 134(5), 598–605. doi:10.1061/(ASCE)0733-9437(2008)134:5(598).
- Ermilio, J. R. (2005). *Characterization Study of a Bio-Infiltration Stormwater BMP*. Master's Thesis. Villanova University.
- Esry, D. H., and Cairns, D. J. (1989). Overview of the Lake Jackson restoration project with artificially created wetlands for treatment of urban runoff. In *Wetlands: Concerns and Successes*. AWWA. Tampa, Florida: American Water Resources Association.
- Faucette, L. B., Jordan, C. F., Risse, L. M., Cabrera, M., Coleman, D. C., and West, L. T. (2005). Evaluation of stormwater from compost and conventional erosion control practices in construction activities. *Journal of Soil and Water Conservation*, 60(6), 288–297. Retrieved from <Go to ISI>://000233893300009.
- FHWA. (n.d.-a). Monitoring Case Study - Delaware Sand Filter BMPs at Airpark, Alexandria, Virginia. *Stormwater Best Management Practices in an Ultra-Urban Setting: Selection and Monitoring*. Retrieved from <https://www.environment.fhwa.dot.gov/ecosystems/ultraurb/5mcs3.asp>.
- FHWA. (n.d.-b). Monitoring Case Study - Modified Delaware Sand Filter BMPs at Alaska Marine Lines, Seattle, Washington. *Stormwater Best Management Practices in an Ultra-Urban Setting: Selection and Monitoring*. Retrieved from <https://www.environment.fhwa.dot.gov/ecosystems/ultraurb/5mcs4.asp>.
- Fletcher, T. D., Peljo, L., Fielding, J., Wong, T. H. F., and Weber, T. (2002). The Performance of Vegetated Swales for Urban Stormwater Pollution Control. In *Global Solutions for Urban Drainage* (pp. 1–16). Portland, OR: American Society of Civil Engineers.
- Florida DEP Unpublished Data. (1993a). *102nd Avenue Retention Pond - Original*. Largo, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (1993b). *Jungle Lake Original*. St. Petersburg, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (1994). *Greenwood Urban Wetland*. Orlando, Florida: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (1995). *Jungle Lake Re-Configured*. St. Petersburg, FL:

- International Stormwater BMP Database.
- Florida DEP Unpublished Data. (1996). *102nd Avenue Retention Pond - Reconfigured*. Largo, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (1998). *The Reserve at DeBary - Pond D*. DeBary, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2000). *Jan Phyl Village Retrofitted WWTP*. Winter Haven, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2001a). *Appleyard Drive Delaware Sand Filter*. Tallahassee, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2001b). *Florida Aquarium Test Site - Street Pond 3&4*. Tampa, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2002). *Key Colony Detention Pond Swale*. Key Colony Beach, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2003a). *Largo Regional STF*. Largo, FL: International Stormwater BMP Database.
- Florida DEP Unpublished Data. (2003b). *Broadway Outfall Marsh*. International Stormwater BMP Database.
- Foad Hussain, C., Brand, J., Gulliver, J., and Weiss, P. (2006). Water Quality Performance of Dry Detention Ponds with Under-Drains. Report, St. Paul, MN: University Department of Transportation Research Services Section. Retrieved from <http://www.lrrb.org/PDF/200643.pdf>.
- Fries, B. M. (1986). *Fate of Phosphorus from Residential Stormwater Runoff in a Southern Hardwood Wetland*. University of Central Florida. Retrieved from <http://stars.library.ucf.edu/rtd/4945/>
- Gain, W. S. (1996). *The Effects of Flow-Path Modification on Water-Quality Constituent Retention in an Urban Stormwater Detention Pond and Wetland System, Orlando, Florida*. *Water Resources Investigations Report 95-4297*. Tallahassee, FL: U.S. Geological Survey.
- Galli, J. (1992). *Analysis of Urban BMP Performance and Longevity in Prince George's County, Maryland: Final Report*. Department of Environmental Programs, Metropolitan Washington Council of Governments.
- GCBOC Unpublished Data. (1999). *Shawnee Ridge Retention Pond*. International Stormwater

BMP Database.

- Gee, K. D., and Hunt, W. F. (2016). Enhancing Stormwater Management Benefits of Rainwater Harvesting via Innovative Technologies. *Journal of Environmental Engineering*, 142(8), 1–11. doi:10.1061/(ASCE)EE.1943-7870.0001108.
- Getter, K. L., Rowe, D. B., and Andresen, J. A. (2007). Quantifying the effect of slope on extensive green roof stormwater retention. *Ecological Engineering*, 31(4), 225–231. doi:10.1016/j.ecoleng.2007.06.004.
- Gilbert, J. K., and Clausen, J. C. (2006). Stormwater runoff quality and quantity from asphalt, paver, and crushed stone driveways in Connecticut. *Water Research*, 40(4), 826–832. doi:10.1016/j.watres.2005.12.006.
- Goel, P. K., Rudra, R. P., Gharabaghi, B., Das, S., and Gupta, N. (2004). Pollutants Removal by Vegetative Filter Strips Planted with Different Grasses. In *ASAE/CSAE Annual International Meeting* (Vol. 300, pp. 1–15). Retrieved from [http://www.pcwp.tamu.edu/docs/lshs/end-notes/pollutants removal by vegetative filter strips planted with dif-1660940058/pollutants removal by vegetative filter strips planted with different grasses.pdf](http://www.pcwp.tamu.edu/docs/lshs/end-notes/pollutants%20removal%20by%20vegetative%20filter%20strips%20planted%20with%20dif-1660940058/pollutants%20removal%20by%20vegetative%20filter%20strips%20planted%20with%20different%20grasses.pdf).
- Goldberg, J. S., Woelfel, C., Chinn, A., McDaniel, M., Paston, C., Thilbert, N., and She, N. (1993). *Dayton Avenue Swale Biofiltration Study*. Seattle, Washington: International Stormwater BMP Database.
- Greb, S. R., and Bannerman, R. T. (1997). Influence of particle size on wet pond effectiveness. *Water Environment Research*, 69(6), 1134–1138.
- Greb, S. R., Bannerman, R. T., Corsi, S. R., and Pitt, R. E. (2000). Evaluation of the Multichambered Treatment Train, a Retrofit Water-Quality Management Practice. *Water Environment Research*, 72(2), 207–216. Retrieved from <http://www.jstor.org/stable/25045360>.
- Gregoire, B. G., and Clausen, J. C. (2011). Effect of a modular extensive green roof on stormwater runoff and water quality. *Ecological Engineering*, 37(6), 963–969. doi:10.1016/j.ecoleng.2011.02.004.
- Gruber, E. (2013). *Analysis and Evaluation of Stormwater Quantity and Quality Performance for Three Permeable Pavement Systems in Fort Collins, Colorado*. Colorado State University. Retrieved from

http://dspace.library.colostate.edu/webclient/DeliveryManager/digitool_items/csu01_storage/2013/07/08/file_1/208132

- Han, J., Wu, J. S., and Allan, C. (2005). Suspended Sediment Removal by Vegetative Filter Strip Treating Highway Runoff. *Journal of Environmental Science and Health, Part A*, 40(8), 1093–4529. doi:10.1081/ESE-200060683.
- Harper, H. H. (1990). *Effects of Stormwater Management Systems on Groundwater Quality*. Orlando, FL: Environmental Research & Design Inc. Submitted to Florida DEP. Retrieved from <http://www.erd.org/erd-publications.htm>.
- Harper, H. H., and Herr, J. L. (1993). *Treatment Efficiencies of Detention with Filtration Systems*. Orlando, Florida: St. Johns River Watershed Management District. Retrieved from <http://www.sjrwmd.com/technicalreports/pdfs/SP/SJ93-SP12.pdf>.
- Harper, H. H., Herr, J. L., Baker, D., and Livingston, E. H. (1999). Performance Evaluation of Dry Detention Stormwater Management Systems. In *Proceedings of the 6th Biennial Stormwater Research and Watershed Management Conference* (pp. 162–178). Southwest Florida Water Management District. Retrieved from https://www.swfwmd.state.fl.us/documents/publications/6thconference_proceedings.pdf#page=170.
- Hathaway, J., and Hunt, W. F. (2007). *City of Charlotte Pilot BMP Monitoring Program - Little Sugar Creek - Westfield Level Spreader*. Charlotte, NC: North Carolina State University. Retrieved from [http://charmeck.org/stormwater/StormWaterAgencies/Pilot SCMs/Wet Ponds/Shade Valley Pond - Final Report.pdf#search=shade valley pond](http://charmeck.org/stormwater/StormWaterAgencies/Pilot%20SCMs/WetPonds/Shade%20Valley%20Pond-Final%20Report.pdf#search=shade%20valley%20pond).
- Hathaway, J., Hunt, W. F., and Johnson, A. (2007a). *City of Charlotte Pilot BMP Monitoring Program - Morehead Place Dry Detention Basin Final Monitoring Report*. North Carolina State University. Retrieved from <http://charmeck.org/stormwater/StormWaterAgencies/Pages/ListofPilotBMPs.aspx>.
- Hathaway, J., Hunt, W. F., and Johnson, A. (2007b). *City of Charlotte Pilot BMP Monitoring Program - Edwards Branch Wetland Final Monitoring Report*. Raleigh, NC: North Carolina State University. Retrieved from [http://charmeck.org/stormwater/StormWaterAgencies/Pilot SCMs/Wetlands/Edwards Branch Wetland - Final Report.pdf#search=edwards branch wetland](http://charmeck.org/stormwater/StormWaterAgencies/Pilot%20SCMs/Wetlands/Edwards%20Branch%20Wetland-Final%20Report.pdf#search=edwards%20branch%20wetland).
- Hathaway, J., Hunt, W. F., and Smith, J. T. (2007). *City of Charlotte Pilot BMP Monitoring*

- Program - Pierson Pond Final Monitoring Report*. Raleigh, NC: North Carolina State University. Retrieved from [http://charmeck.org/stormwater/StormWaterAgencies/PilotSCMs/WetPonds/Pierson Pond - Final Report.pdf#search=pierson pond](http://charmeck.org/stormwater/StormWaterAgencies/PilotSCMs/WetPonds/PiersonPond-FinalReport.pdf#search=piersonpond).
- Hathaway, J., Hunt, W. F., Smith, J. T., and Johnson, A. (2007a). *City of Charlotte Pilot BMP Monitoring Program - University Executive Park Dry Detention Pond Final Monitoring Report*. Charlotte, NC: North Carolina State University. Retrieved from <http://charmeck.org/stormwater/StormWaterAgencies/Pages/ListofPilotBMPs.aspx>.
- Hathaway, J., Hunt, W. F., Smith, J. T., and Johnson, A. (2007b). *City of Charlotte Pilot BMP Monitoring Program - Shade Valley Pond Final Monitoring Report*. Raleigh, NC: North Carolina State University. Retrieved from [http://charmeck.org/stormwater/StormWaterAgencies/PilotSCMs/WetPonds/Shade Valley Pond - Final Report.pdf#search=shade valley pond](http://charmeck.org/stormwater/StormWaterAgencies/PilotSCMs/WetPonds/ShadeValleyPond-FinalReport.pdf#search=shadevalleypond).
- Hatt, B. E., Fletcher, T. D., and Deletic, A. (2007). Treatment performance of gravel filter media: Implications for design and application of stormwater infiltration systems. *Water Research*, 41(12), 2513–2524. doi:10.1016/j.watres.2007.03.014.
- HCFCU Unpublished Data. (2007). *Greens Bayou Wetland Mitigation System - Surge Basin*. Houston, TX: International Stormwater BMP Database.
- Heyvaert, A. C., Reuter, J. E., and Goldman, C. R. (2006). Subalpine storm water treatment with a constructed surface flow wetland. *Journal of the American Water Resources Association*, 42, 45–54.
- Higgins, K., and Roth, C. (2005). Sun Valley Park Storm Water Infiltration Basin. In *Impacts of Global Climate Change* (pp. 1–10). Anchorage, AK: American Society of Civil Engineers.
- Horner, R. R., and Horner, C. R. (1995). *Design, Construction, and Evaluation of a Sand Filter Stormwater Treatment System. Part II. Performance Monitoring*. Seattle, Washington: Report to Alaska Marine Lines.
- Horner, R. R., Lim, H., and Burges, S. J. (2004). Hydrologic Monitoring of the Seattle Ultra-Urban Stormwater Management Projects: Summary of the 2000-2003 Water Years. In *Water Resources Series Technical Report No. 181*. Seattle, Washington.
- Horst, M., Welker, A. L., and Traver, R. G. (2011). Multiyear Performance of a Pervious Concrete Infiltration Basin BMP. *Journal of Irrigation and Drainage Engineering*, 137(6), 352–358. doi:10.1061/(ASCE)IR.1943-4774.0000302.

- Hunt, W. F., Hathaway, J. M., Winston, R. J., and Jadlocki, S. J. (2010). Runoff Volume Reduction by a Level Spreader–Vegetated Filter Strip System in Suburban Charlotte, N.C. *Journal of Hydrologic Engineering*, 15(June), 499–503. doi:10.1061/(ASCE)HE.1943-5584.0000160.
- 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*, 132(6), 600–608.
- Hunt, W. F., Smith, J. T., Jadlocki, S. J., Hathaway, J. M., and Eubanks, P. R. (2008). Pollutant Removal and Peak Flow Mitigation by a Bioretention Cell in Urban Charlotte, N.C. *Journal of Environmental Engineering*, 134(5), 403–408. doi:10.1061/(ASCE)0733-9372(2008)134:5(403).
- Hutchinson, D., Abrams, P., Retzlaff, R., and Liptan, T. (2003). Stormwater monitoring two ecoroofs in Portland, Oregon, USA. In *Greening Rooftops for Sustainable Communities* (pp. 1–18). Retrieved from <http://www.portlandoregon.gov/bes/article/63098> \n<http://www.portlandonline.com/shared/cfm/image.cfm?id=63098>.
- Johnson, J., Hathaway, J., and Hunt, W. F. (2007). *City of Charlotte Pilot BMP Monitoring Program - Bruns Ave. Elementary School Wetland Final Monitoring Report*. Raleigh, NC: North Carolina State University. Retrieved from [http://charmeck.org/stormwater/StormWaterAgencies/Pilot SCMs/Wetlands/Bruns Ave School Wetland - Final Report.pdf#search=bruns ave wetland](http://charmeck.org/stormwater/StormWaterAgencies/Pilot%20SCMs/Wetlands/Bruns%20Ave%20School%20Wetland-Final%20Report.pdf#search=bruns%20ave%20wetland).
- Kaighn, R., and Yu, S. (1996). Testing of Roadside Vegetation for Highway Runoff Pollutant Removal. *Transportation Research Record*, 1523(1), 116–123. doi:10.3141/1523-14.
- Kantrowitz, I. H., and Woodham, W. M. (1995). Efficiency of a Stormwater Detention Pond in Reducing Loads of Chemical and Physical Constituents in Urban Streamflow, Pinellas County, Florida. *USGS Water-Resources Investigations Report*, 94-4217, 1–18.
- Knight, E. M. P., Hunt, W. F., and Winston, R. J. (2013). Side-by-side evaluation of four level spreader-vegetated filter strips and a swale in eastern North Carolina. *Journal of Soil and Water Conservation*, 68(1), 60–72. doi:10.2489/jswc.68.1.60.
- Kolsti, K. F., Burges, S. J., and Jensen, B. W. (1995). *Hydrologic response of residential-scale laws on till containing various amounts of compost amendment*. University of Washington.

- Koon, J. (1995). *Evaluation of Water Quality Ponds & Swales In The Issaquah/East Lake Sammamish Basins*. Seattle, WA: King County Surface Water Management. Retrieved from <http://www.kingcounty.gov/depts/dnrp/wlr/publications.aspx>.
- L.A. County Unpublished Data. (2010). *Dominguez Gap Wetland*. Los Angeles, CA: International Stormwater BMP Database.
- LCCMR Unpublished Data. (1988). *Carver Ravine Wetland*. Woodbury, MN: International Stormwater BMP Database.
- LCCMR Unpublished Data. (1989). *Carver Ravine Detention Pond*. International Stormwater BMP Database.
- Legret, M., and Colandini, V. (1999). Effects of a porous pavement with reservoir structure on runoff water: Water quality and fate of heavy metals. *Water Science and Technology*. doi:10.1016/S0273-1223(99)00014-1.
- Leif, W. (1999). *Compost Stormwater Filter Evaluation: Final Report*. Everett, WA: Snohomish County Department of Public Works. Surface Water Management Division.
- Liang, W. Y. (1996). *Performance assessment of an off-line stormwater management pond: Heritage Estates Stormwater Management Pond, Richmond Hill, Ontario*. Toronto, ON: Ontario Ministry of Environment and Energy and the Metropolitan Toronto and Region Conservation Authority.
- Line, D. E., and Hunt, W. F. (2009). Performance of a Bioretention Area and a Level Spreader-Grass Filter Strip at Two Highway Sites in North Carolina. *Journal of Irrigation and Drainage Engineering*, 135(2), 217–224. doi:10.1061/(ASCE)0733-9437(2009)135:2(217).
- Liptan, T., and Murase, R. K. (2000). Watergardens as Stormwater Infrastructure in Portland, Oregon. In *Water Sensitive Ecological Planning and Design Symposium* (pp. 125–154). Boston, MA: Harvard Design School.
- Lloyd, S. D., Wong, T. H., and Chesterfield, C. J. (2002). *Water Sensitive Urban Design - A Stormwater Management Perspective*. Melbourne, Australia: Cooperative Research Centre for Catchment Hydrology.
- Lucke, T., and Nichols, P. W. B. (2015). The pollution removal and stormwater reduction performance of street-side bioretention basins after ten years in operation. *Science of the Total Environment*, 536, 784–792. doi:10.1016/j.scitotenv.2015.07.142.
- Macdonald, K. C. B., and Jefferies, C. (2002). Performance of BMPs in Scotland. In *Linking*

- Stormwater BMP Designs and Performance to Receiving Water Impact Mitigation* (pp. 519–523). ASCE. doi:10.1061/40602(263)38.
- Mallin, M. A., Ensign, S. H., Wheeler, T. L., and Mayes, D. B. (2002). Pollutant removal efficacy of three wet detention ponds. *Journal of Environmental Quality*, 31(2), 654–660. doi:10.2134/jeq2002.6540.
- Mangangka, I. R., Liu, A., Egodawatta, P., and Goonetilleke, A. (2015). Performance characterisation of a stormwater treatment bioretention basin. *Journal of Environmental Management*, 150, 173–178. doi:10.1016/j.jenvman.2014.11.007.
- Mangarella, P., Goodman, J., Havens, K., Pankani, D., Lewis, W., and Lin, H. (2010). *Wet Detention Basin Effectiveness Study, Phase II*. Sacramento: City of Sacramento, CA.
- Maristany, A. E., and Bartel, R. L. (1989). Wetlands and stormwater management: a case study of Lake Munson. Part I: long-term treatment efficiencies. In *Wetlands: Concerns and Successes*. AWRA (pp. 215–229). Tampa, Florida: American Water Resources Association.
- Martin, E. H. (1988). Effectiveness of an Urban Runoff Detention Pond-Wetlands System. *Journal of Environmental Engineering*, 114(4), 810–827.
- Martin, E. H., and Smoot, J. L. (1986). *Constituent-Load Changes in Urban Stormwater Runoff Routed Through a Detention Pond-Wetlands System in Central Florida*. Water Resources Investigations Report 85-4310. Tallahassee, FL: U.S. Geological Survey.
- Middleton, J. R., and Barrett, M. E. (2006). Improved Extended Detention Basin Performance through Better Residence Time Control. In *World Environmental and Water Resource Congress 2006* (pp. 1–10). American Society of Civil Engineers. doi:10.1061/40856(200)377.
- Moran, A. C. (2005). A North Carolina Field Study to Evaluate Greenroof Runoff Quantity, Runoff Quality, and Plant Growth. M.S. Thesis, Raleigh, NC: North Carolina State University. Dept. of Biological and Agricultural Engineering. Retrieved from <http://repository.lib.ncsu.edu/ir/bitstream/1840.16/803/1/etd.pdf>.
- New Castle County Delaware. (2008). *NPDES Joint Annual Report*. New Castle County, DE. Retrieved from https://www.deldot.gov/stormwater/permit_docs/Phase_I_Report_2008/DelDOT_Phase_I_Rpt_2008.pdf.
- NYC DEP Unpublished Data. (2001). *BMP 12, Malcolm Brook*. New York: International

- Stormwater BMP Database.
- NYC DEP Unpublished Data. (2003). *BMP 37, West Lake Drive*. New York: International Stormwater BMP Database.
- NYC DEP Unpublished Data. (2006a). *BMP 13, West Lake Drive*. New York: International Stormwater BMP Database.
- NYC DEP Unpublished Data. (2006b). *BMP 74, Route 170*. New York: International Stormwater BMP Database.
- Oberts, G. L., Wotzka, P. J., and Hartsoe, J. A. (1989). *The water quality performance of select urban runoff treatment systems*. St. Paul, MN: Metropolitan Council.
- Oberts, G., and Osgood, R. (1988). The effectiveness of a detention/wetland treatment system and its effect on an urban lake. In *Annual International Symposium on Lake and Watershed Management* (p. 29).
- OWML. (1990). *Final Project Report - The Evaluation of a Created Wetland as an Urban Best Management Practice*. Manassas, VA: Occoquan Watershed Monitoring Laboratory George Mason University Department of Biology.
- 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*, 135(4), 505–510.
doi:10.1061/(ASCE)IR.1943-4774.0000006.
- Piza, H., and Eisel, C. (2011a). *Permeable Interlocking Concrete Pavement (PICP) at Denver Wastewater Management Building Denver, Colorado 2008-2010*. Denver, CO: Urban Drainage and Flood Control District. Retrieved from <http://udfcd.org/technical-papers>.
- Piza, H., and Eisel, C. (2011b). *Pervious Concrete at Lakewood Maintenance Facility Lakewood, Colorado*. Denver, CO: Urban Drainage and Flood Control District. Retrieved from <http://udfcd.org/technical-papers>.
- Piza, H., and Eisel, C. (2011c). *Porous Asphalt at Denver Wastewater Management Building Denver, Colorado 2008-2010*. Denver, CO: Urban Drainage and Flood Control District. Retrieved from <http://udfcd.org/technical-papers>.
- Piza, H., Stawski, J., and Eisel, C. (2013). *Stormwater Quality Monitoring Report Extended Detention Basin at Grant Ranch*. Denver, CO: Urban Drainage and Flood Control District. Retrieved from <http://udfcd.org/technical-papers>.

- Portland BES Unpublished Data. (1999). *Water Garden*. Portland, OR: International Stormwater BMP Database.
- Portland BES Unpublished Data. (2002a). *Whitaker Ponds PRF*. Portland, OR: International Stormwater BMP Database.
- Portland BES Unpublished Data. (2002b). *Parkrose Sand Filter*. Portland, OR: International Stormwater BMP Database.
- Portland BES Unpublished Data. (2003). *Lexington Hills Detention Pond*. Portland, OR: International Stormwater BMP Database.
- Pratt, C. J., Mantle, J. D. G., and Schofield, P. A. (1989). Urban Stormwater Reduction and Quality Improvement Through the Use of Permeable Pavements. *Water Science & Technology*, 21, 769–779.
- Rushton, B. T. (2001). Low-Impact Parking Lot Design Reduces Runoff and Pollutant Loads. *Journal of Water Resources Planning and Management*, 127(3), 172–179.
- Rushton, B. T. (2002). Treatment of Stormwater Runoff from an Agricultural Basin by a Wet-Detention Pond in Ruskin, Florida. Final Report, Brooksville, Florida: Southwest Florida Water Management District. Retrieved from <http://www15.swfwmd.state.fl.us/LibraryImages/04645.pdf>.
- Rushton, B. T., Miller, C., Hull, C., and Cunningham, J. (1997). *Three Design Alternatives for Stormwater Detention Ponds*. Brooksville, Florida: Southwest Florida Water Management District.
- Schiffer, D. M. (1989). *Effects of highway runoff on the quality of water and bed sediments of two wetlands in central Florida: USGS Water-Resource Investigations Report 88-4200*. Tallahassee, FL: U.S. Geological Survey.
- Schlüter, W., and Jefferies, C. (2002). Modelling the outflow from a porous pavement. *Urban Water*, 4(3), 245–253. doi:10.1016/S1462-0758(01)00065-6.
- Schueler, T. R. (1987). *Controlling Urban Runoff: A Practical Manual for Planning and Designing Urban BMPs*. Washington D.C.
- Sharkey, J. L. (2006). *The Performance of Bioretention Areas in North Carolina: A Study of Water Quality, Water Quantity, and Soil Media*. Master's Thesis. North Carolina State University.
- Smith, R. A., and Hunt, W. F. (2007). Pollutant Removal in Bioretention Cells with Grass Cover.

- In *World Environmental and Water Resources Congress*. Environmental and Water Resources Institute, ASCE.
- Stagge, J. H. (2006). *Field Evaluation of Hydrologic and Water Quality Benefits of Grass Swales for Managing Highway Runoff*. University of Maryland, College Park. Retrieved from <http://drum.lib.umd.edu/bitstream/handle/1903/3969/umi-umd-3843.pdf?sequence=1&isAllowed=y>
- Stoker, Y. E. (1996). *Effectiveness of a Stormwater Collection and Detention System for Reducing Constituent Loads from Bridge Runoff in Pinellas County. USGS Open File Report 96-484*. Tampa, Florida: U.S. Geological Survey. Retrieved from <https://pubs.usgs.gov/of/1996/0484/report.pdf>.
- Striegl, R. G., and Cowan, E. A. (1987). *Relations Between Quality of Urban Runoff and Quality of Lake Ellyn at Glen Ellyn, Illinois. Water-Supply Paper 2301*. U.S. Geological Survey. Retrieved from <http://pubs.usgs.gov/wsp/2301/report.pdf>.
- Teemusk, A., and Mander, U. (2007). Rainwater runoff quantity and quality performance from a greenroof: The effects of short-term events. *Ecological Engineering*, 30(3), 271–277. doi:10.1016/j.ecoleng.2007.01.009.
- UDFCD Unpublished Data. (1997). *Shop Creek Wetland-Pond (1995-97)*. Denver, CO: International Stormwater BMP Database.
- UNHSC. (2005). *University of New Hampshire Stormwater Center 2005 Data Report*. Durham, NH: University of New Hampshire. Retrieved from <http://www.unh.edu/unhsc/publications>.
- UNHSC Unpublished Data. (2006a). *Bioretention DI*. International Stormwater BMP Database.
- UNHSC Unpublished Data. (2006b). *Tree Filter*. International Stormwater BMP Database.
- UNHSC Unpublished Data. (2006c). *Sand Filter*. Durham, NH: International Stormwater BMP Database.
- UNHSC Unpublished Data. (2006d). *Vegetated Swale*. Durham, NH: International Stormwater BMP Database.
- UNHSC Unpublished Data. (2006e). *Porous Pavement (J-PAS)*. Durham, NH: International Stormwater BMP Database.
- UNHSC Unpublished Data. (2006f). *University of New Hampshire Retention Pond*. International Stormwater BMP Database.
- Urbonas, B., Carlson, R., and Vang, B. (1994). *Joint Pond-Wetland System Performance in*

- Colorado, U.S.A. Denver, CO: Urban Drainage and Flood Control District.
- USEPA. (1983). *Results of the Nationwide Urban Runoff Program. Volume 1 - Final Report*. Washington, D.C. Retrieved from https://www3.epa.gov/npdes/pubs/sw_nurp_vol_1_finalreport.pdf.
- Van Seters, T., Smith, D., and Macmillan, G. (2006). Performance Evaluation of Permeable Pavement and a Bioretention Swale. In *8th International Conference on Concrete Block Paving* (pp. 161–170).
- VanWoert, N. D., Rowe, D. B., Andresen, J. A., Rugh, C. L., Fernandez, R. T., and Xiao, L. (2005). Green Roof Stormwater Retention. *Journal of Environment Quality*, 34(3), 1036. doi:10.2134/jeq2004.0364.
- Walsh, P. M., Barrett, M. E., Malina, J. F., and Charbeneau, R. J. (1997). *Use of Vegetative Controls for Treatment of Highway Runoff. CRWR Online Report 97-5*. Austin, TX: Center for Research in Water Resources. Retrieved from <http://www.crwr.utexas.edu/reports/pdf/1997/rpt97-5.pdf>.
- Washington DOE Unpublished Data. (2005). *Embankment at SR 167 MP 16.4*. International Stormwater BMP Database.
- Welborn, C. T., and Veenhuis, J. E. (1987). *Effects of Runoff Controls on the Quantity and Quality of Urban Runoff At Two Locations in Austin, Texas. Water-Resources Investigations Report 87-4004*. Austin, TX: U.S. Geological Survey. Retrieved from <http://pubs.usgs.gov/wri/1987/4004/report.pdf>.
- Wenck Associates Inc. (2009). *Villa Park Wetland Management Plan*. St. Paul, MN: Capitol Region Watershed District. Retrieved from <http://www.capitolregionwd.org/press/crwd-reports/>.
- WERF, EWRI, USEPA, U.S. DOT, APWA, Wright Water Engineers, and Geosyntec Consultants. (2014). International Stormwater BMP Database. Retrieved January 1, 2016, from <http://www.bmpdatabase.org/index.htm>.
- Willenbring, P. R. (1985). Wetland treatment systems-why do some work better than others? In *Lake and Reservoir Management Proceedings of the Fourth Annual Conference and International Symposium* (pp. 234–237). McAfee, NJ: North American Lake Management Society.
- Wilson, C. E., Hunt, W. F., Winston, R. J., and Smith, P. (2014). Assessment of a rainwater

- harvesting system for pollutant mitigation at a commercial location in Raleigh, NC, USA. *Water Science and Technology: Water Supply*, 14(2), 283–290. doi:10.2166/ws.2013.200.
- Winston, R. J., Davidson-Bennett, K. M., Buccier, K. M., and Hunt, W. F. (2016). Seasonal Variability in Stormwater Quality Treatment of Permeable Pavements Situated Over Heavy Clay and in a Cold Climate. *Water, Air, & Soil Pollution*, 227(5), 140. doi:10.1007/s11270-016-2839-6.
- Winston, R. J., Dorsey, J. D., and Hunt, W. F. (2016). Quantifying volume reduction and peak flow mitigation for three bioretention cells in clay soils in northeast Ohio. *Science of the Total Environment*, 553, 83–95. doi:10.1016/j.scitotenv.2016.02.081.
- 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*, 138(1), 101–111. doi:10.1061/(ASCE)EE.1943-7870.0000454.
- Winston, R. J., Hunt, W. F., Osmond, D. L., Lord, W. G., and Woodward, M. D. (2011). Field Evaluation of Four Level Spreader – Vegetative Filter Strips to Improve Urban Storm-Water Quality. *Journal of Irrigation and Drainage Engineering*, 137(March), 170–182. doi:10.1061/(ASCE)IR.1943-4774.0000173.
- Wu, J. S., Holman, R. E., and Dorney, J. R. (1996). Systematic Evaluation of Pollutant Removal by Urban Wet Detention Ponds. *Journal of Environmental Engineering*, 122(11), 983–988.
- Yousef, Y. A., Hvitved-Jacobsen, T., Wanielista, M. P., and Harper, H. H. (1987). Removal of contaminants in highway runoff flowing through swales. *Science of the Total Environment*, 59(C), 391–399. doi:10.1016/0048-9697(87)90462-1.
- Yu, S. L., Fitch, G. M., and Earles, T. A. (1998). *Constructed Wetlands for Stormwater Management*. VTRC Report No. 98-R35. Charlottesville, Virginia: Virginia Transportation Research Council. Retrieved from <https://trid.trb.org/view.aspx?id=536718>.
- Yu, S. L., and Stopinski, M. D. (2001). *Testing of Ultra-Urban Stormwater Best Management Practices*. Charlottesville, Virginia: Virginia Transportation Research Council.