

Syracuse University

SURFACE at Syracuse University

Dissertations - ALL

SURFACE at Syracuse University

5-14-2023

Causal effects of green infrastructure on stormwater hydrology and water quality

Caitlin G. Eger
Syracuse University

Follow this and additional works at: <https://surface.syr.edu/etd>

Recommended Citation

Eger, Caitlin G., "Causal effects of green infrastructure on stormwater hydrology and water quality" (2023).
Dissertations - ALL. 1668.
<https://surface.syr.edu/etd/1668>

This Dissertation is brought to you for free and open access by the SURFACE at Syracuse University at SURFACE at Syracuse University. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

Abstract

Applications of green infrastructure to stormwater management continue to increase in urban landscapes. There are numerous studies of individual stormwater management sites, but few meta-analyses that synthesize and explore design variables for stormwater control structures within a robust statistical framework. The lack of a standardized framework is due to the complexity of stormwater infrastructure designs. Locally customized designs fit to meet diverse site conditions create datasets that become messy, non-uniform, and difficult to analyze across multiple sites. In this dissertation, I first examine how hydrologic processes govern the function of various stormwater infrastructure technologies using water budget data from published literature. The hydrologic observations are displayed on a Water Budget Triangle—a ternary plot tool developed to visualize simplified water budgets—to enable direct functional comparisons of green and grey approaches to stormwater management. The findings are used to generate a suite of observable site characteristics, which are then mapped to a set of stormwater control and treatment sites reported in the International Stormwater Best Management Practice (BMP) database. These mapped site characteristics provide site context for the runoff and water quality observations present in the database. Drawing from these contextual observations of design variables, I next examine the functional design of different stormwater management technologies by quantifying the differences among varied structural features, and comparing their causal effects on hydrologic and water quality performance. This stormwater toolbox provides a framework for comparison of the overall performance of different system types to understand causal implications of stormwater design.

CAUSAL EFFECTS OF GREEN INFRASTRUCTURE ON STORMWATER HYDROLOGY AND WATER QUALITY

by

Caitlin G. Eger

B.S., Juniata College, 2008
M.S., The Ohio State University, 2012

Dissertation

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Civil Engineering.

Syracuse University

May 2023

Copyright © Caitlin G. Eger, 2023.

All rights reserved. This material or any portion may not be reproduced or used in any manner without express written permission of the author except for the use of brief quotations with proper citation and attribution.

Printed by Syracuse University, Syracuse, NY.

Syracuse University Press

621 Skytop Road, Syracuse, NY 13210

First printing, 2023.

For my children, who welcomed me on the doorstep of motherhood

Acknowledgements

This research was funded by the National Science Foundation under Grants DGE-1449617 and SES-1444755, the Surdna Foundation under Grant 20140225, a Syracuse University Fellowship and the EMPOWER program at Syracuse University. Gratitude to CUAHSI and Arizona State University for supporting me while learning Python.

This dissertation relies heavily on the hard work of many other scholars, professionals, and students from the past 50 years. It would not exist without their field observations, water budget data and the contextual elements reported in the literature. The International Stormwater BMP Database is a long-standing effort by a consortium of engineers, experts, database administrators, and developers, including the US Environmental Protection Agency, the US Department of Transportation, the Water Research Foundation, the American Society of Civil Engineers Environmental & Water Resources Institute, Geosyntec Consultants, and Wright Water Engineers. The code repository for the causal analysis work is named in honor of Jane Clary, one of the principal investigators who has helped manage the International BMP Database for two decades.

Blood, sweat and tears were shed over the difficulty of learning causal inference methods alone, a feat made possible only through many online tutorials, helpful posts and questions raised on StackOverflow and CrossValidated. I express appreciation to Scott Cunningham, Ben Bolker, Jakob Runge, Richard McElreath, Gary King, Daniel Lüdecke and Brady Neal for their excellent free content, R packages and persevering online presence. Cheers and thank you to Nicholas Hamilton for the development of the ggtern package, which facilitated the production of the ternary plots for chapter 3, and Johannes Textor, who maintains DAGitty.net and the dagitty package used to make the causal diagrams in chapters 4 and 5. God bless my dear friend Linden McBride for her statistical cheerleading.

My deepest gratitude to all the supportive people I've worked with: my advisors and advisory committee; Geoff Millard, and Angelica Huerta for checking in on me regularly; to

Cliff Davidson, Babak Kasaee Roodsari, Kathy Fallon Lambert, Peter Groffman, Neil Bettez, and for their feedback on the initial white paper that became chapters 1, 2 and 3; as well as Thomas Evans, Aditi Padhye and Megan Daley for editing and literature review. A hearty thanks to Telsha Curry, Chris Feikes, and Erica Frederick for copy editing, and to Nick Clarke for keeping my paperwork in order semester after semester. I appreciate the support and many kind words from Heather Flaherty, Deanna McCay and Annie Pennella, who provided steady encouragement and snacks. Thank you to Catherine Grube and all the baristas at St. Inie's Coffee who always let me write until closing time. My heartfelt thanks to Michael Beanland and RTR for encouraging me to finish this, and for always reminding me: "You can do hard things."

Lastly, thank you to my family for your patience and encouragement. Thank you to my husband, who remained calm despite many storms of anxiety. My parents, sisters, in-laws, niece and son believed that I could finish this tome, long before I felt it was possible. Thank you to my little daughter, whose presence gave me so much emotional support during the final chapter. To my dear family I offer the blessing from Lutkin's Benediction anthem singing Numbers 6:24-26: *'May the Lord bless you and keep you, make his face shine upon you and give you peace.'* I love you all very much.

Contents

1	Introduction	1
1.1	Motivation and background	1
1.1.1	Stormwater as an environmental hazard	1
1.1.2	Modern engineering approaches to stormwater management	2
1.2	Research objectives	3
1.3	Document roadmap	4
1.4	Research approach	5
1.4.1	Research complexity in environmental experiments	5
1.4.2	The role of causality in engineering	8
2	Review of stormwater infrastructure technologies	10
2.1	Background information	10
2.2	Defining green infrastructure	10
2.3	Overview of stormwater infrastructure technologies	11
2.3.1	Detention ponds	12
2.3.2	Retention ponds	12
2.3.3	Constructed wetlands	13
2.3.4	Bioretention	14
2.3.5	Grass swales and filter strips	15
2.3.6	Media filters	16
2.3.7	Porous pavement	17
2.3.8	Infiltration basins	17
2.3.9	Manufactured devices	18
2.3.10	Green roofs	18
2.4	Green infrastructure engineering objectives	19
3	Hydrologic processes that govern stormwater infrastructure behaviour	20
3.1	Introduction	20
3.2	Data	23
3.3	Methods	24
3.4	Results	28
3.4.1	Natural and constructed wetlands	28
3.4.2	Retention basins and ‘wet’ ponds	32
3.4.3	Detention basins and ‘dry’ ponds	34
3.4.4	Bioretention cells, infiltration trenches, swales, and rain gardens	35
3.4.5	Pervious pavement	38
3.4.6	Green roofs	41
3.5	Discussion	43
3.5.1	Modelling and diagnostic estimates	43
3.5.2	Data limitations	44
3.5.3	Factors affecting hydrologic performance	46
3.5.4	Factors primarily affecting I–Q trade-off	47
3.5.5	Factors primarily affecting Q–ET trade-off	50

3.5.6	Factors primarily affecting ET–I trade-off	51
3.6	Conclusions	53
4	Urban green infrastructure ecohydrology	55
4.1	Introduction	55
4.1.1	Mechanisms to relate stormwater control design to ecohydrology . . .	55
4.1.2	Bridging model assumptions with observable field conditions	57
4.1.3	Counterfactual questions for design iteration	58
4.1.4	Using directed acyclic graphs to clarify assumptions	60
4.1.5	Use of observational data in causal inference studies	61
4.2	Data and methods	67
4.2.1	Dataset and imputation methods	67
4.2.2	Analytical methods	70
4.3	Results	78
4.3.1	Model Set 1: BMP type effects	79
4.3.2	Model Set 2: Structural feature effects	82
4.3.3	Model Set 3: Media effects	88
4.3.4	Model Set 4: Ponding effects	91
4.3.5	Model Set 5: Vegetation effects	93
4.4	Discussion	95
4.4.1	Features with the greatest magnitude of influence on runoff reduction	95
4.4.2	Practical application of the findings	97
4.4.3	Addressing estimate variability	99
4.5	Conclusions and recommendations	104
4.5.1	Conclusions from linear mixed model estimation of structural feature effects	104
4.5.2	Future recommendations	104
5	Estimation of biogeochemical flux in green stormwater infrastructure	106
5.1	Introduction	106
5.1.1	Understanding controls on solubility and transport in the urban environment	106
5.1.2	Contaminants present in urban waters	107
5.1.3	Chronic exposure and toxic effects	110
5.2	Data and methods	114
5.2.1	Dataset preparation	114
5.2.2	Method 1: Estimates of concentration and mass flux by BMP Type .	115
5.2.3	Method 2: Dissolved fraction estimates	118
5.2.4	Method 3: Matched difference-in-difference estimates for individual structural features.	119
5.3	Results and Discussion	122
5.3.1	Total suspended and total dissolved solids	122
5.3.2	Nutrients	127
5.3.3	Trace metals	130
5.4	Conclusions and recommendations	134

5.4.1	Conclusions from estimation of structural feature effects	134
6	Synthesis and future recommendations	136
6.1	Interpretation and synthesis of model results	136
6.2	Future Recommendations	140
	Appendices	142
A	Data sources used for analysis	143
A.1	Water budget sources from Chapter 3	144
A.2	Description of International BMP Database dataset cleaning	146
A.3	Structural features sources for causal effects	152
A.4	Exploratory data analysis of the BMP database	153
B	Supplementary water budget datasets	157
C	Causal theory for engineered treatment	164
D	Median contaminant concentration and mass flux ranges	173
E	Contaminant concentrations and mass flux reduction estimates	190
F	Shifts in dissolved fractional ratio	204
G	Causal effects in order of magnitude change	212
G.1	BMP type effects on hydrologic performance	213
G.2	Individual effects by structural feature	215
G.3	Covariate balance after matching	216
G.4	Structural treatment effects on contaminant concentrations	219
	References	220
	Vita	244

List of Figures

2.1	Common site sizing across various BMP types	11
3.1	Demonstration of the Water Budget Triangle tool	24
3.2	Water budgets of natural wetlands, lakes, and constructed wetlands	32
3.3	Water budgets of retention (wet) ponds	33
3.4	Water budgets of detention ponds and constructed wetlands	35
3.5	Water budgets of bioretentions, swales and lysimeters	39
3.6	Water budgets of permeable pavements	41
3.7	Water budgets of green roofs	43
3.8	Water balance summary demonstrating I-Q axis tradeoffs	49
3.9	Water balance summary demonstrating primary and secondary axis tradeoffs	50
3.10	Water balance summary demonstrating Q-ET axis tradeoffs	52
4.1	Demonstration of basic DAG structures: chains, forks and colliders	61
4.2	Demonstration of a back-door path and an instrumental variable	66
4.3	Model identification and adjustment strategies for Model Sets 1 and 2	74
4.4	Precipitation distribution	78
4.5	Model 1: Site level predictions by BMP type	84
4.6	Model 2 results: Structural effects	86
4.7	Model 2 predictions: Structural features	87
4.8	Model 3 predictions: Media amendment	89
4.9	Model 3 predictions: Media amendment by depth	90
4.10	Model 4 predictions: Ponding	92
4.11	Model 5 predictions: Vegetation	94
5.1	Controls on contaminant solubility and engineering intervention options	108
5.2	Dissolved fractional ratios of contaminants in untreated stormwater.	124
5.3	Trace metal concentration ranges in untreated stormwater.	132
A.1	Nutrients concentrations in stormwater and rainwater	154
A.2	Metals concentrations in stormwater and rainwater	154
A.3	Raw distribution of inflow and outflow volumes by BMP type	155
B.1	Maturation of vegetation changes water budgets over time	159
B.2	Cistern water budgets	160
B.3	Seasonal retention pond water budgets	161
B.4	Sewer exfiltration water budgets	163
B.5	Continental water budgets	163
C.1	Stormwater ecohydrology DAG	167
C.2	DAG representing the direct and indirect effects of media amendment.	168
C.3	DAG representing the direct and indirect effects of onsite ponding.	169
C.4	DAG representing the direct and indirect effects of vegetation.	169
C.5	General controls on solubility and transport of stormwater contaminants.	170
C.6	Equations describing physical and chemical processes in Figure C.5	171
D.1	Median concentrations and mass flux of TSS and TDS.	175
D.2	Median concentrations and mass flux of total N and NO _x	177
D.3	[Median concentrations and mass flux of dissolved, total and reactive P.	179
D.4	Median concentrations and mass flux of dissolved and total Cd.	181

D.5	Median concentrations and mass flux of dissolved and total Cr.	183
D.6	Median concentrations and mass flux of dissolved and total Cu.	185
D.7	Median concentrations and mass flux of dissolved and total Pb.	187
D.8	Median concentrations and mass flux of dissolved and total Zn.	189
F.1	Shifts in dissolved fractional ratio of solids.	205
F.2	Treatment effects on dissolved fractional ratio of P and N.	206
F.3	Treatment effects on dissolved fractional ratio of Cd.	207
F.4	Treatment effects on dissolved fractional ratio of Cr.	208
F.5	Treatment effects on dissolved fractional ratio of Cu.	209
F.6	Treatment effects on dissolved fractional ratio of Pb.	210
F.7	Treatment effects on dissolved fractional ratio of Zn.	211
G.1	Model 1 results: BMP type effects	213
G.2	Model 1 results: Area-normalized BMP type effects	214
G.3	Model 3 results: media amendment covariate balance	216
G.4	Model 4 results: ponding covariate balance	217
G.5	Model 5 results: vegetation covariate balance	218
G.6	Comparison of individual structural treatment effects on solids, nutrients and trace metals.	219

List of Tables

3.1	Hydrologic function of stormwater technologies vs. natural systems	30
3.2	Diagnostic operational ranges for engineered stormwater systems	31
3.3	Design factors affecting hydrologic performance	48
4.1	Design factors that influence ecohydrology.	56
4.2	Station and record counts across monitoring site type	70
4.3	Model 1: Site level effects by BMP type	82
4.4	Model 1: Area-normalized effects by BMP type	83
4.5	Best estimates of marginal effects	96
4.6	Best estimates of marginal effects compared with water budget data	100
5.1	Dissolved solutes in stormwater and precipitation	107
5.2	Analyte detection limits.	115
5.3	Counts of observations by BMP type.	116
5.4	Summary of treatment effects on dissolved fractional ratios.	125
5.5	Summary of individual structural treatment effects on solids and nutrients. .	126
5.6	Structural treatment effects on Cd, Cr, Cu, Pb and Zn	133
A.1	Counts of observations by BMP type	153
A.2	Imputed distribution of inflow and outflow volumes by BMP type	156
C.1	Summary of contaminant removal mechanisms	172
D.1	Median concentrations and mass flux of TSS and TDS.	174
D.2	Median concentrations and mass flux of total N and NO _x	176
D.3	Median concentrations and mass flux of dissolved, total and reactive P. . . .	178
D.4	Median concentrations and mass flux of dissolved and total Cd.	180
D.5	Median concentrations and mass flux of dissolved and total Cr.	182
D.6	Median concentrations and mass flux of dissolved and total Cu.	184
D.7	Median concentrations and mass flux of dissolved and total Pb.	186
D.8	Median concentrations and mass flux of dissolved and total Zn.	188
E.1	Treatment effects on total suspended solids.	191
E.2	Treatment effects on total dissolved solids.	192
E.3	Treatment effects on TN.	193
E.4	Treatment effects on NO _x	194
E.5	Treatment effects on total phosphorous.	195
E.6	Treatment effects on dissolved and reactive phosphorus.	196
E.7	Treatment effects on dissolved and total cadmium (Cd).	197
E.8	Treatment effects on dissolved and total chromium (Cr).	198
E.9	Treatment effects on dissolved copper (Cu).	199
E.10	Treatment effects on total copper (Cu).	200
E.11	Treatment effects on dissolved and total lead (Pb).	201
E.12	Treatment effects on dissolved zinc (Zn).	202
E.13	Treatment effects on total zinc (Zn).	203
G.1	Model 2: Area-normalized effects by structural feature	215

1 Introduction

1.1 Motivation and background

1.1.1 Stormwater as an environmental hazard

Stormwater runoff currently impairs thousands of waterway reaches within the United States, and millions of reaches globally. Impairment conditions associated with stormwater runoff include exposure to chemical pollutants, salinization, sedimentation, modification of flow regime, extreme temperature variation, erosion, and channel incision. Combined, these changes to habitat degrade hydrologic function and create environmental conditions that are unsuitable for aquatic life. Impaired waterways also threaten the quality of drinking-water supplies and create conditions that put individual property and public infrastructure at risk of flooding, sewage exposure, and foundation damage—all conditions that threaten human life, health, and livelihood.

Under future climate scenarios, stormwater poses a risk to cities, where there is greater runoff and less land available to implement mitigation solutions. Runoff from the urban landscape already causes flooding, combined sewer overflows, and pollution in local tributaries, which negatively impact surface water quality. Contamination from urban runoff chronically degrades surface water quality during nearly every storm event, regardless of event magnitude (Balderas Guzman et al. [2022]). Damage from urban development and point-source pollutant emissions puts surface waters at risk of impairment for recreational and drinking water uses. Degraded conditions are expected to increase in all watersheds with current or planned human development and in many watersheds where sea level rise or changes to rainfall patterns will disrupt current flow regimes. Warmer, wetter conditions carry a greater risk of flooding events in low-lying urban areas. Roadway development introduces hydrocarbons and trace metals. Wide scale vegetation removal and deforestation causes increased erosion, channel incision, and limits oxygenation.

Widespread fertilization of grass increases chemical nutrients in runoff. Warming summer water temperatures sustain lower oxygen levels, making lakes and rivers more sensitive to stormwater nutrient inputs and susceptible to dead zones (Rice and Jastram [2015]).

Increased probability of extreme events, such as hurricanes, wildfires, and drought can have long-lasting pollution effects when they impact urban areas. Pollution caused by extreme events can create chronically degraded aquatic ecosystems, due to trash, contamination, or derelict property (Reible et al. [2006], Burton et al. [2016], Shevah [2019]).

1.1.2 Modern engineering approaches to stormwater management

The design of next-generation stormwater infrastructure is informed by both modern hydraulic engineering models and ‘green’ ecological management tactics. Green methods for managing urban stormwater runoff have grown popular because they can be more effective than traditional, concrete-based drainage infrastructure due to: 1) equal or better pollutant capture, 2) lower maintenance costs, 3) distributed treatment networks, 4) enhanced neighborhood aesthetics, and 5) greater resiliency in the face of a changing climate. Despite the growing popularity of ‘low-impact’ green infrastructure (GI) designs over the past two decades, the hydrologic benefits associated with specific structural features remain largely unquantified. While green infrastructure has proven itself as a useful strategy for frequent, low-to-medium-intensity storm events that affect regional stream ecosystems, it remains largely untested against high-intensity or high-magnitude events that inherently carry more risk to life and property. Green stormwater infrastructure deployed in tandem with traditional ‘gray’ infrastructure helps reduce the hydraulic load on sewershed and wastewater treatment facilities, increasing infrastructure longevity and treatment capacity through ecological services (Shakya and Ahiablame [2021]). Thus, there remains room in the engineering toolbox for a hybrid approach that incorporates both green and gray systems to manage stormwater within the same watershed.

Green infrastructure for stormwater management is generally designed for capture and

treatment of runoff associated with single, short-duration storm events (i.e., maximum design storms between 2.5 and 3.8 cm, as in Reese and Parker [2014a,b]). Generally, this condition is achieved by one or more of three mechanisms: 1) increasing on-site storage to retain stormwater runoff until it can be discharged to sewer infrastructure (e.g., through ponding or tanks); 2) facilitating water loss by evapotranspiration to the atmosphere; or, 3) via infiltration into the near-surface water table or deeper groundwater (Reese and Parker [2014a,b]). The latter two eliminate the need to treat additional runoff volume through wastewater facilities or by surface water discharge. Green infrastructure technologies vary in their ability to process water by each of these mechanisms. In addition to the primary objective of reducing overall runoff volume, a secondary objective of GI is water quality improvement, decreasing the demand for and costs of in-line wastewater treatment and increasing overall surface water health (Denchak [2022]). There are also auxiliary social, environmental, and public health co-benefits to urban GI adoption, but these are considered ancillary to the primary objective of effective stormwater management.

1.2 Research objectives

Many communities have implemented their own unique green infrastructure solutions to address stormwater runoff; but, without rigorous systematic comparison, it is difficult to improve the quality of the engineering design recommendations. In this dissertation, I aim to make such a comparison by teasing out the causal implications of choosing distinct engineering design features. Causal design features are those whose presence within the built structure have a consequential and observable impact on the performance of the system’s hydrology and water quality outcomes. To clarify, they are not simply correlated with better performance, but create conditions that have a measurable effect on

stormwater treatment. Three questions are examined in this body of work:

1. Can existing observations about site ecohydrology and water budgets identify design features that transcend stormwater infrastructure types?
2. What is the estimated causal effect of these design features on site hydrologic performance?
3. Which design features have specific causal effects on water quality (aside from their hydrologic effects)?

1.3 Document roadmap

An overview of the dissertation work is as follows:

- Chapter 1, (this chapter): An introduction to the urban eco-hydrological problem space, and an overview of the concept of causality;
- Chapter 2: A definition of green infrastructure management, and a descriptive review of 10 stormwater control structure technologies, including their most common features;
- Chapter 3: An analysis of site water budgets at many stormwater control structures, and identification of eco-hydrological similarities in the performance of the different technology types;
- Chapter 4: A hypothesis about which structural features affect water budget loss pathways, and several modeled estimates of the effect sizes of these structural features on hydrologic discharge;
- Chapter 5: An extension of the models in chapter 4 to estimate the effect sizes of the most common structural features on a limited set of water quality parameters; and
- Chapter 6: A synthesis integrating recommendations from Chapters 3, 4 and 5.

1.4 Research approach

Throughout this dissertation, I integrate and synthesize diverse stormwater management strategies into a unified set of green and gray stormwater tools, based on publicly available and observational real-world data (see Chapter 3). In this research I am concerned with understanding how small engineering design changes affect hydrologic and biogeochemical performance (e.g., designing for the presence or absence of standing water, or effects due to choice of soil media). I searched for physical properties associated with various hydrologic characteristics from GI technologies to improve future design iterations, retrofits, and diagnostics (summarized in Chapters 3 and 4). Then, I used the structural characteristics to provide context for field observations of flows and water quality samples (Chapters 4 and 5). Instead of conducting a single design case study, I hope to garner as much statistical power and common support as is available by using information present in large public datasets and aggregating data points from existing case studies and sites.

As discussed earlier, there is significant strategic overlap between the goals and outcomes of green and gray infrastructure approaches to stormwater management in an urban landscape. Through this dissertation I aim to identify how the two approaches may be used to complement one another, by integrating civil engineering and ecological engineering techniques. Merging these two disciplines is largely possible by recognizing that the hydrological controls and ecological mechanisms at play in engineered stormwater structures are nearly identical to those phenomena present in natural wetland systems. Doing so takes advantage of data from both non-engineering and non-stormwater research fields, and allows compilation of new, GI-relevant datasets to form a design meta-analysis.

1.4.1 Research complexity in environmental experiments

Most publicly available stormwater field data have been collected with a focus on a single infrastructure site, resulting in an entirely observational monitoring approach, mostly

without controlled study designs or contextual measurements from other water systems with similar characteristics. Three major hurdles to studying this problem in a systematic manner are: 1) complex experimental design obstacles, 2) real-world logistical barriers, and 3) budgetary limitations. It is challenging to design a controlled environmental experiment, particularly one that involves an intervention such as an engineered system within an urban landscape. Finding appropriately matched site locations within the urban environment and funding to purchase land, engineer designs, and install materials are all considerations that limit the practical reproducibility of this type of experiment. Additionally, variability in environmental factors, such as climate, precipitation, and landscape conditions, largely eliminate the possibility of exact experimental replication. Many, (if not most), of the green infrastructure sites that have been built to-date were not installed within the context of an experimental research design, and very few are monitored for relevant hydrological and biogeochemical characteristics needed to evaluate the functionality of the design choices. Many published studies omit or overlook inclusion of relevant contextual site and environmental variables. Real-world constraints alone prevent the implementation of a randomized controlled trial or fully blocked experiment, and the range of ambient environmental conditions where these systems exist adds another layer of complexity that is measurable, but not manipulable (e.g., rainfall depth, temperature).

Since randomization of some variables is largely impossible, data availability is limited to observational study. There is a significant body of observational data available for study on this topic, but the obstacles discussed above suggest that these datasets (including those used here) are highly susceptible to statistical selection bias, stemming from sampling bias (which geographic/climatic site locations are monitored), time interval bias (which storm events are monitored), exposure bias (which site structural treatment and control features are observed and reported), and participation bias (which categories of stormwater control technologies are equipped for monitoring). Bias susceptibility does not mean the data cannot be used to answer questions of causal interest, but rather intensifies

the level of care needed to present a descriptive study of the dataset. The high number of potentially confounding factors present in ecological systems also means that untempered selection bias can mask the magnitude of causal effects. A confounder is a variable that affects both the study treatment (usually called the independent variable) and the outcome of interest (dependent variable). The presence of a confounder magnifies or diminishes the measured effect of the independent variable on the dependent one. There are also spuriously correlated variables that can induce misinterpretation of the relationship between variables, including the direction of the causal relationship (that is, which phenomenon causes another one to occur). Improper use or careless interpretation of a regression model with confounded or spurious model specification can lead to misinterpretation about the sign, magnitude, and importance of causal effects (see further discussion about causal structure in Chapter 4).

However, the effort is worth the struggle. The resulting information has utility for planners, engineers, and decision-makers who must consider small, catchment-scale ecohydrology to effect change at the watershed scale. The overarching goal of this work is to provide a generalized set of engineering tools for designers and inform technical and non-technical users alike about hydrologic and water quality processes occurring in GI systems. Thus far, the breadth of names for various stormwater management structures, practices, and design modifications have not led to a clear understanding of functional differences in performance among different technologies (an overview of terms is presented in Chapter 2). These diverse terms have muddled communication among users, practitioners, and stakeholders. In this research, I strive to provide meaningful ways to communicate about various design modifications, and to bridge the gap between stormwater engineering and ecological engineering, such that design structure can better support intended onsite function. Performance characteristics and design features are used to classify structures based not on regional vocabulary, but on reproducible and quantifiable characteristics, for the benefit of engineers, designers, and urban hydrologists.

1.4.2 The role of causality in engineering

Engineering is inherently causal; it relies on gaining enough understanding of causal factors within a system of variables to explicitly allow manipulation and modification of the features of a system in pursuit of an anthropocentric outcome or design. Without understanding the causal foundations creating observable physical phenomena, engineers cannot apply causal theory to design a technological solution. Discovery of causal connections can be tricky; and many people, both inside and outside of the academy, have been taught that only correlation can be experimentally discerned, never causation. However, as a discipline, engineering has proven that 1) causality can be determined through observation and development of scientific theory, 2) well-developed theory informs the causal direction of effect between a designed intervention and an outcome, and 3) causal understanding may be used to manipulate a desired outcome. Correlation may not indicate causation, but certainly correlation does not disprove causation. Understanding the explicit causal assumptions present in the engineered or experimental system is the key to avoiding confounding and spurious correlations that lead to misinterpretation. A more detailed discussion of this topic is presented in Chapter 4. Through this body of research, I seek to join new understanding in ecological engineering with existing civil engineering knowledge by examining causal questions (taken from Angrist and Pischke [2009]):

- What is the causal relationship of interest, its units and real-world interpretation?
- What is the ideal experiment that could capture the causal effect of interest?
- What questions are fundamentally unidentifiable/unanswerable given the data or system of interest?
- What is the identification strategy? Is it possible to use non-randomized observational datasets to approximate an experiment?
- What is the statistical mode of inference (data and methods, population, sample,

assumptions) for construction of estimators and errors?

Civil engineering is a distinctly interventionist field. History records six millennia of building drainage structures, underground storage and sewers, dikes and dams for water supply, stormwater, and flood control (Hilprecht [1904]). The longstanding history of traditional drainage solutions for stormwater management draws a biased perception that it is less risky and more reliable than green infrastructure strategies. Quite the reverse, the evidence of human reliance on longstanding ecological services throughout the Holocene and Anthropocene is quite robust. Ecological engineering is equally interventionist, although many ecosystem features are challenging to modify (Mitsch [2012]). Using causal connections gleaned by physical and biological scientific theory, ecological engineering seeks to predict, design, construct, restore and manage ecosystems for the benefit of human society within its natural environment (paraphrasing Jørgensen and Mitsch [1989]).

2 Review of stormwater infrastructure technologies

2.1 Background information

International attention to the problem of urban stormwater has resulted in the innovation of many new civil infrastructure solutions over the past forty years. These solutions have developed through iterative engineering design, in which slight modifications to previous designs occur many times at each new treatment site. Through this iterative development pattern, some structures have proven more popular over time, due to regional success or cost-effectiveness. Here, I provide a short review of the language and structures described in the stormwater management literature.

2.2 Defining green infrastructure

Diverse names (and acronyms) describe a similar underlying strategic approach and principles for stormwater management (Fletcher et al. [2015]): ‘green infrastructure’ (GI; US EPA [2023]), ‘low impact development’ (LID; Low Impact Development Center [2023]), ‘best management practices’ (BMPs; WRF [2019]), ‘stormwater control measures’ (SCMs; Fassman-Beck et al. [2016]), ‘sustainable urban drainage systems’, ‘water-sensitive urban design’ (WSUD; Melbourne Water [2023]), ‘blue-green infrastructure’ (BGI; Thorne et al. [2015]), ‘soft-path’ water infrastructure (Brown [2014]), and the latest one, ‘sponge city’ strategy (Chan et al. [2018]). The variety of names for the strategy of using distributed management structures for stormwater control and treatment within an urban landscape does not support clear communication about functional differences or successful outcomes to practitioners and stakeholders. There is also variability in the nomenclature used to refer to different types of individual stormwater treatment structures and their functional design components. The structures are variously described as ponds, cells, basins, swales, wetlands, vaults, filters, or tanks, and are equipped for detention, retention, biofiltration, and infiltration. They are ultimately intended to capture and delay stormwater runoff,

reduce runoff volume and/or improve pollutant concentration levels. Quite often the names describe an intended use, rather than the operational outcome that exists at the field site. Many functionally similar structures have unrelated descriptive nomenclature, which can create confusion about which features are essential for meeting specific performance goals. Therefore, the designations that follow represent specific sets of structural features that have common, but not universally accepted names.

2.3 Overview of stormwater infrastructure technologies

The following descriptions are based on field observations of stormwater control structures and contextual measurements (Figure 2.1) recorded in the International BMP Database (WRF [2019]).

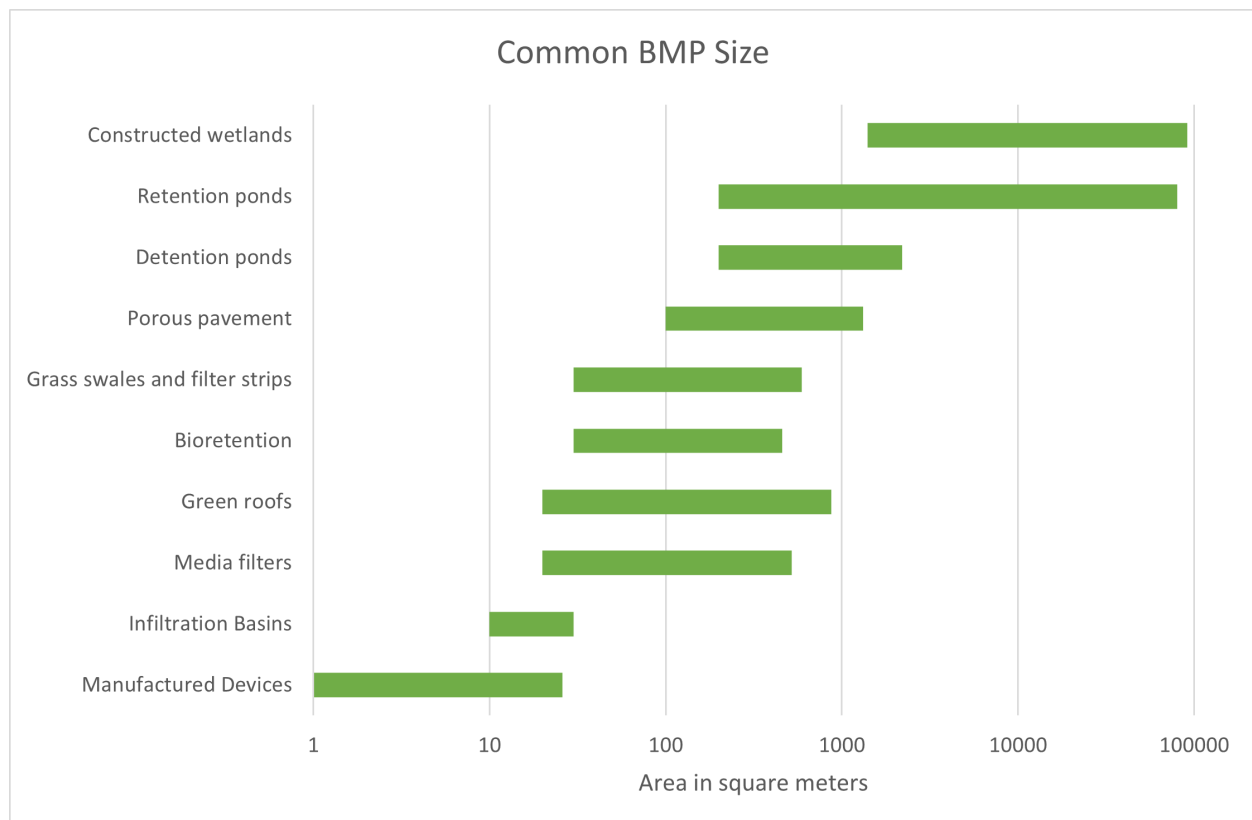


Figure 2.1: Site sizing across various stormwater control structure types found in the International BMP Database (WRF [2019]).

2.3.1 Detention ponds

Detention ponds (or basins) are well-known stormwater management structures designed to attenuate peak flow during a runoff event. Ideally a detention pond will fill during a wet-weather event and then slowly release the collected water during the 12 to 48 hours that follow, providing short-term storage but little permanent capture. Detention basins nearly always have a piped inlet to receive point flow runoff from the surrounding watershed. The inlet is often protected with rip-rap or concrete spillways to prevent erosion. Less frequently, this structure receives water from a pumped source. There are usually one or two surface drains at different stages (heights above the lowest point in the basin). The lower one is usually a smaller pipe or culvert and serves to draw down captured runoff in between events. The upper one serves as an overflow culvert or spillway. Detention ponds are generally grassed or sparsely vegetated, open to air, and are usually unlined and have no amendments to the native soil (occasionally they may be planted with shrubs or lined with concrete, clay, or geotextile fabric). When full of water, detention ponds appear like a pond, and a grassed field with surrounding berms when dry. Because their footprint is moderately large (~ 200 to 2000 sq. meters), they are common in residential and suburban neighborhoods where there is more land available amongst private parcels and structures. Maintenance usually involves periodic mowing and infrequent adjustments to reinforce existing inlet and outlet erosion protection.

2.3.2 Retention ponds

Retention ponds are like detention ponds, and are also very popular for residential areas, but unlike detention ponds they hold permanent standing water. As with detention basins, the maximum capture depth is limited to the difference in height between the overflow/spillway and the draw down outlet (fluxing depth). The capture volume depends on the flux depth, basin area and slope of the basin, and to a lesser extent, the infiltration rate. While many retention ponds are unlined, (thus, allowing for groundwater exchange),

these sites generally do not have soils that support active groundwater exchange. There are a few possible reasons for low groundwater exchange: 1) retention ponds may be purposely sited in places with low drainage rates, 2) they are often located in urban and suburban sites that have been modified, compacted, excavated or scraped using heavy machinery, which disrupts the natural soil structure, and 3) they have standing water, which creates hydric soils and supports a different biotic community than dry basins, likely including fewer rooting structures. Hydric soils also create water quality characterized by reducing conditions. The slopes of the berms surrounding the retention basin are relatively steep, usually around 3:1. Retention ponds are usually quite large (~ 200 sq. meters to 8 ha), which may resemble ponds or small reservoirs. Occasionally, compacted clay liners are installed to limit groundwater exchange.

2.3.3 Constructed wetlands

Constructed wetlands are frequently installed in areas adjacent to known tributaries or seasonal rivulets, or in pockets of low-lying, poorly draining soils. In the field, it can be difficult to differentiate between a constructed wetland and a retention basin, because both may have excavated basin space and/or minimally managed vegetation. Both structures may have pumped inflow and/or outflow volumes. Constructed wetlands are larger or equivalent in scale to many retention ponds (~ 1400 sq. meters to > 9 ha), but often occur alongside existing drainage or waterways, and may have more naturalized (planted but unmanaged), self-colonized, or taller vegetation. When constructed wetlands occupy riparian space, they often have a more distinctly channelized shape, rather than a basin-like shape. Natural wetlands are known to provide flood and inclement weather protection as well as nutrient removal, depending on groundwater exchange and oxidation-reduction conditions. Constructed wetlands also provide additional surface storage during wet-weather events and may facilitate stormwater infiltration, or conversely, act as a conduit for discharge of groundwater, depending on the groundwater table. They

may be built to replace disturbed natural wetlands after the completion of a construction project. The label ‘constructed wetland’ is a good example of nomenclature that poorly describes a set of very diverse structural features. For example, many constructed wetlands are designed to include sets of pools and baffles that encourage recirculating, vertical, or horizontal flow, depending on the purpose of the structure. To understand how each feature affects hydrology and water quality outcomes, rather than assigning a single name to a complex set of sites, it is better to describe the structural features individually for each site and compare them to sites that have documented similarities and differences.

2.3.4 Bioretention

A bioretention cell, also known as a bioretention basin, biofilter, bioswale, or bio-infiltration system, is an infiltrative system that captures water from building and roadway runoff. Stormwater is diverted from the existing urban stormwater pathway and collected at a location upstream from the storm sewer inlet. The names are quite specialized to region, climate, and the individuals responsible for installation. The cells are installed by first excavating a pit or trench that is relatively small (~ 30 to 430 sq. meters). The excavated space is filled with porous, low-nutrient media, e.g., a layer of mixed sand-loam over coarse gravel or stone underdrain and topped with a layer of mulch. Many bioretention sites have a geotextile (semi-permeable) or HDPE (impervious high-density polyethelene) liner; some manufacturers insert a pre-cast concrete box below the surface level of the ground. Nearly all sites are equipped with a subsurface drain, such as drainage tile, perforated HDPE, or an internal water-storage zone with an upturned drainage elbow. Most sites have a shallow ponding basin above the surface of the amended soil. These sites may collect runoff from overland sheet flow, or they may collect water to an inlet point with a flat concrete splashpad or a rock-lined entryway. This type of site is usually planted with a variety of native or horticultural species, flowers for pollinators, native grasses, and/or street trees, which are usually fertilized or given compost only

during an initial establishment period. In contrast, bioswales are generally smaller and receive sheet flow, with one or no outlet points, whereas bioretention basins generally have an inflow and outflow point. Bioswales without surface outlets are usually ‘pocket’ shaped; these structures collect runoff until the ponded head inside the structure forces water to bypass the inlet. Bioswales are occasionally planted with grass mixes, but most bioretention cells are planted with more diverse species, and maintenance usually does not include mowing. The volume of stormwater captured in bioretention cells and bioswales is related to the contributing area, the size and depth of the soil media, the porosity of the underlying native soil, whether the site is equipped with an underdrain and/or liner, and to a lesser extent, the type of plantings.

2.3.5 Grass swales and filter strips

Swales, also called grassed waterways or vegetated filter strips, may be confused with bioretention cells, which are also called bioswales. In general, grassed swales are planted with grass seed or sod and they may resemble a wide, rectangular, gently sloping field with a shallow V-shaped trench (bioretention cells tend to be more bathtub shaped). Although they tend to be longer and narrower when adjacent to roadways or structures, grassed swales are roughly the same size as bioretention cells (~ 30 to 560 sq. meters). Swales with sod have limited connectivity with underlying groundwater, although the amount of infiltration is highly dependent upon the slope, vegetation cover, underlying soil type, and wetting/drying cycle periodicity (Duley and Kelly [1939], Duley and Domingo [1949]). Vegetated swales are often maintained by periodic mowing or bush-hogging (on monthly or seasonal intervals), whereas bioretention cells are usually not mowed and have much more porous soil. The soil in a grassed swale may be excavated and replaced with a better-draining media replacement or, more commonly, native soil re-graded to create a trench that facilitates stormwater collection. There are also subtle differences between detention ponds and swales that are related to the purpose of the structure. Detention

ponds almost always have distinct culverts or pipes serving as point inlets and outlets, while swales typically receive water as overland flow, along the length of the swale. The sheet flow may enter the swale directly or through a level spreader to slow the velocity of flow and allow for the removal of suspended solids from runoff via overland sheet flow. These ‘level-spreader’ type swales are common in agricultural and transportation applications as well as residential developments. The term ‘vegetated filter strip’ is more common when adjacent to a roadway.

2.3.6 Media filters

A media filter uses a substrate to remove suspended solids and clarify water as it passes through the filter. Media filters rely on a wide variety of substrates, such as sand, peat, geotextile fabric, crushed rock or glass, carbon, shredded paper, rubber pellets, and foam. They may be designed to remove dissolved pollutants, especially ammonia and nitrate. Media filters are sometimes called ‘biofilters’ because they provide microbiological habitat on the media surface. Ordinarily, media filters have an inlet that allows water to enter from a single point inlet, although many are adjacent to a tank or pond that collects water before overflowing into the filter media. The media filter area is usually small (~ 20 to 500 sq. meters), but the tank or pond can be much larger, generally, on the scale of a small retention basin (as large as 1 ha or more). There also may be a shallow basin or ponding volume above the surface of the filter. Most media filters have a small storage volume with limited connectivity to natural groundwater tables, either because they have a small footprint or because they are contained inside a precast concrete structure. Therefore, they may impact water quality without changing overall stormwater volume. Some filters are designed for high-rate flow or re-circulation. Nearly all media filters are open to air, but some may be enclosed, and nearly all have a subsurface drain or a drain with an up-turned elbow to create an internal water storage zone. When they are contained inside an enclosed, air-tight tank, media filters are usually marketed as branded, manufactured

devices and called by their brand name.

2.3.7 Porous pavement

There are many different types of porous pavements, including porous asphalt, porous concrete, permeable/grassed interlocking pavers, cobblestone, permeable friction coarse surfacing and other materials used for vehicle and pedestrian traffic. Porous pavement is usually installed over layered courses of gravel or crushed stone and nearly always has a subsurface underdrain. Porous pavement installations usually cover between ~ 100 and 1225 sq. meters. The volume of capture depends upon the contributing area, permeable surface area, and infiltration rate. The spacing gap between paver blocks controls the permeable surface area and infiltration rate, having a strong effect on evaporative losses in between storm events. After a storm event, a significant fraction of the stormwater captured onsite is diverted to infiltrate to groundwater or evaporate up through the pavement surface. Evaporation from permeable asphalt surfaces is influenced by solar radiation, permeable surface color and the planting scheme of adjacent land (Starke et al. [2010, 2011]).

2.3.8 Infiltration basins

Infiltration basins, or dry wells, are more common in very dry environments. They are small areas (~ 10 to 20 sq. meters) intended to collect runoff from paved urban spaces and to re-infiltrate it to replenish groundwater. Their most defining characteristic is a very deep well (3 meters or more) that has been back-filled with highly porous material, such as gravel or crushed stone. The sites do not have surface vegetation, and there is usually a surface ponding basin fed by curb cuts. Occasionally there is a liner on the sides of the well, but not on the boundary between the base of the amended media and the native soil. The small infiltrating surface area, large pore sizes and deep well create conditions with lower evaporation rates, evaporating a smaller proportion of the captured water than a site

with ordinary bioretention cell or porous pavement site dimensions.

2.3.9 Manufactured devices

Manufactured stormwater devices come in many designs and purposes. Thus, their performance should not be analyzed as a group, but with careful description of each structural feature to accurately assess how changing one design feature affects the overall performance of the device. Several examples of these devices include:

- **catch basin inserts**, which are geotextile filters that are placed into a storm drain to catch trash, leaves and debris from entering the storm sewer
- **hydrodynamic separators**, which force the runoff into a vortex flow pattern to remove suspended particulate matter. Particulates settle into a central tank and the tank is pumped on a periodic maintenance schedule.
- **filtration cartridges**, which are barrel-shaped tanks that force stormwater through a media filtration unit using a hydraulic pressure gradient. The filter media can be adjusted to target removal of specific particulate and dissolved contaminants, depending on the type.

2.3.10 Green roofs

A green roof collects water from the top of a built structure, often an institutional building or parking garage. The design consists of several layers of liners to protect the building structure from water damage. The liners are overlaid with a lightweight soil matrix, and planted with a variety of drought-tolerant plants, such as sedums, native or pioneer species, and occasionally, edible, or horticultural plants. Most green roofs are between ~ 20 and 850 sq. meters; some on commercial buildings are quite large (Syracuse OnCenter is > 5500 sq. meters, Save the Rain [2020]). The volume of stormwater captured by a green roof is directly related to the roof area and depth of the soil media. If the structure remains

un-planted, containing only a layer of gravel or crushed stone, it may be referred to, colloquially, as a ‘blue roof.’ Blue and green roofs are effective in dense, urban landscapes; their primary mechanism to reduce runoff is evapotranspiration. LEED-certified green roofs are popular for their architectural novelty, aesthetics and visibility.

2.4 Green infrastructure engineering objectives

Overall, the primary objective of engineered stormwater management is to improve downstream watershed conditions; it is considered a cost-effective strategy to mitigate runoff associated with short to medium sized storm events (i.e., 1 to 10 year event size or less) (VA Stormwater Management Program [2016]). These events occur frequently, at least once every two years, but generally on a monthly or weekly basis (depending on location and seasonal climate). Small storm events have a chronic impact on local watersheds; treating them can improve water quality in receiving surface waters.

Treatment may involve reducing the total volume of runoff from a landscape, or removing contaminants from the runoff, or both. Some structure types may also have flood control as a primary design objective. Many site designs aim to accomplish these stormwater management goals, each with unique advantages and limitations. The effectiveness of their application depends on localized stormwater management needs, operating and design conditions, as well as the surrounding built landscape and surface configuration. In general, better downstream water quality can be accommodated by reducing overall volume of runoff. Runoff reduction can be achieved through two mechanisms: 1) increasing landscape storage to retain stormwater until it can be processed by the existing sewer and drainage network, or 2) diverting water from overland runoff by facilitating water loss through evapotranspiration to the atmosphere or infiltration to the groundwater system. Both strategies reduce the need to process stormwater runoff through wastewater treatment; the latter limits surface water discharge. Green infrastructure technologies vary in their ability to process water by these two mechanisms.

3 Hydrologic processes that govern stormwater infrastructure behaviour

3.1 Introduction

Stormwater treatment structures are used to capture urban runoff, reduce wastewater inputs, eliminate combined sewer overflows, and meet total maximum daily pollutant load goals; however, there are many co-benefits of green infrastructure (GI) (CNT [2010], Zahmatkesh et al. [2015]). U.S. regulators have attempted to credit best management practices (BMPs) in state stormwater design standards by incorporating a ‘runoff reduction’ (RR) method into design manuals and spreadsheet calculators (CNT [2010], Hartigan et al. [2009], Hirschman et al. [2008], NERR [2016], NYS DEC [2023], VA Stormwater Management Program [2016]). This method helps practitioners select appropriate BMP options from a suite of choices based on projected hydrologic function. Application of this method stems from recognition that water quality benefits from stormwater control structures are largely volume-driven (Ballesterio et al. [2012], Eger [2012], Hirschman et al. [2008]). Although the runoff reduction method credits ‘green’ BMPs, it neither credits nor discredits the selection of conventional stormwater (grey) structures over GI (e.g., runoff reduction worksheets in NYS DEC [2023]). This approach presumes conventional ‘grey’ technologies are either environmentally benign or hydrologically superior to green engineering strategies. Often, neither is the case. Therefore, these two broad stormwater management approaches are implemented on unequal terms, rather than as complementary technologies that should be evaluated on the same assessment scale. This perspective unfortunately limits the breadth of information available to urban hydrologists, engineers, planners, and civic decision makers when choosing among GI and conventional structures for stormwater management.

The advantages of GI over conventional grey infrastructure for stormwater abatement

are widely reported (De Sousa et al. [2012], Lucas and Sample [2015]). However, runoff reduction values for GI often have wider operating ranges than conventional grey systems (Driscoll et al. [2015]). As a result, implementation of GI has met resistance from regulatory barriers, which mandate inflexible standards or prescribe specific performance metrics, and from communities with fractured or complex stormwater regulation (UNH Stormwater Center [2014], US EPA [2013], Worstell [2013]). In some cases, civil infrastructure professionals and permitting organizations have expressed concern over the uncertainty of adopting green infrastructure BMPs for runoff mitigation (Matthews et al. [2015], Thorne et al. [2015]). Community leadership is right to seek evidence of GI benefits prior to investing public dollars in major projects. However, technical concerns from the engineering community about performance uncertainty and undefined operational ranges can be interpreted by non-technical decision makers as increased risk for implementation of BMPs relative to conventional stormwater infrastructure (Hu and Shealy [2020]). In contrast, risk of implementing grey infrastructure is less commonly addressed, despite established social, economic, and ecological impacts (Vineyard et al. [2015], Walsh et al. [2005]). Concerns about inconsistent performance stem from an absence of clear metrics to compare and contrast green and grey systems in straightforward, meaningful ways. The proliferation of field studies on GI systems has been accompanied by greater availability of performance data and range of metrics in the literature, including volumetric reduction, peak flow reduction, and delayed time-to-peak, among others (Stovin et al. [2017]). However, some of these metrics are not well suited for comparison of GI with grey systems. For example, recent reports indicate that GI often outperforms grey infrastructure on a per cent volumetric or mass reduction basis; however, this metric is not commonly used for conventional stormwater infrastructure monitoring (Bhaskar and Welty [2012], Driscoll et al. [2015]). Peak flow reduction metrics are routinely used in stormwater reporting, but this measure is generally inappropriate for monitoring subterranean sewer systems, green roofs, and porous pavement installations. The range of function among BMP technologies

and designs has impeded efforts to gather consensus on the benefits associated with these practices. Moreover, the lack of traditional descriptive metrics acts as a barrier to decision-makers whose options may be restricted by regulatory code. For example, retention and detention ponds generally exhibit favorable time-to-peak delay but poor overall volumetric reduction, which caps the benefits realized for downstream water quality (Driscoll et al. [2015]). Further, comparison among similar designs is complicated by different climatic conditions and scales (Driscoll et al. [2015]). Without standardized metrics for comparison, it is difficult to choose which type of stormwater control structure is appropriate to meet the stormwater goals of individual sites and catchments. Common reporting methods are also necessary to synthesize datasets and identify which physical factors have the greatest influence over hydrologic and water quality variables. In this assessment, I conduct a quantitative comparison among BMP types and evaluate the hydrologic processes occurring within green systems (primarily infiltrative and evaporative) alongside grey systems (primarily conveyance) using the common water budget metric and a ternary plot tool, the Water Budget Triangle.

3.2 Data

Water budget data from stormwater management sites were discovered and retrieved through systematic internet searches on each type of structure by examining Google results that contained the BMP structure name and the words ‘stormwater’ and ‘water balance’ or ‘water budget.’ After discovering sources that described water balance measurements, reported values of discharge, evaporation and infiltration were recorded. It was not uncommon that infiltration or evaporation values were not reported. For sites that did not report all three values, estimates of the missing value were made by subtracting the measured quantities from the total inflow volume. If a water balance source did not include at least one measurement of inflow and two outflow measurements, it was not included in the dataset. Additional sources were discovered by reviewing the references of sources with water budget data. Data were aggregated across the longest time period that was practical based on the source report, usually at least one month up to one water year, and measurements made over shorter time periods or single events were reported separately. The dataset includes water balance measurement estimates for natural wetlands ($n = 21$), natural lakes (13), constructed wetlands (8), models of natural wetlands (5), models of constructed wetlands (2), retention basins (9), detention ponds (7), lysimeter cells (3), lined and unlined bioretention cells with underdrains (1 each), and unlined cells without underdrains (2), a bioretention model (1), porous pavement (15), models of porous pavements (12), grassed pavement (2), impervious surface (1), lined porous pavement (3), green roofs (59), lab-scale green roofs (7), control roofs (6), blue roofs (4), and green roof models (10). References for the water balances are listed by type in Appendix A.1.

3.3 Methods

The Water Budget Triangle (Figure 3.1) was developed to address the fundamental question: “How does stormwater leave the unit volume of a control structure?” The tool facilitates comparative assessment of dissimilar systems by providing graphical depiction of simplified hydrologic budgets exiting the control volume of an engineered or natural stormwater structure. It is intended to visually represent the fractional distribution of volumetric (or mass) outflow among discharge (Q), percolation (I), and evaporation (ET) on a ternary diagram (Eger et al. [2014]). The tool assumes that after influent stormwater enters a system, there are three potential pathways for water loss:

- 1. Discharge to a pipe or surface water (Q, right axis);
- 2. Evaporation or transpiration into the atmosphere (ET, top axis); or
- 3. Drainage into soil pores/groundwater (I, left axis).

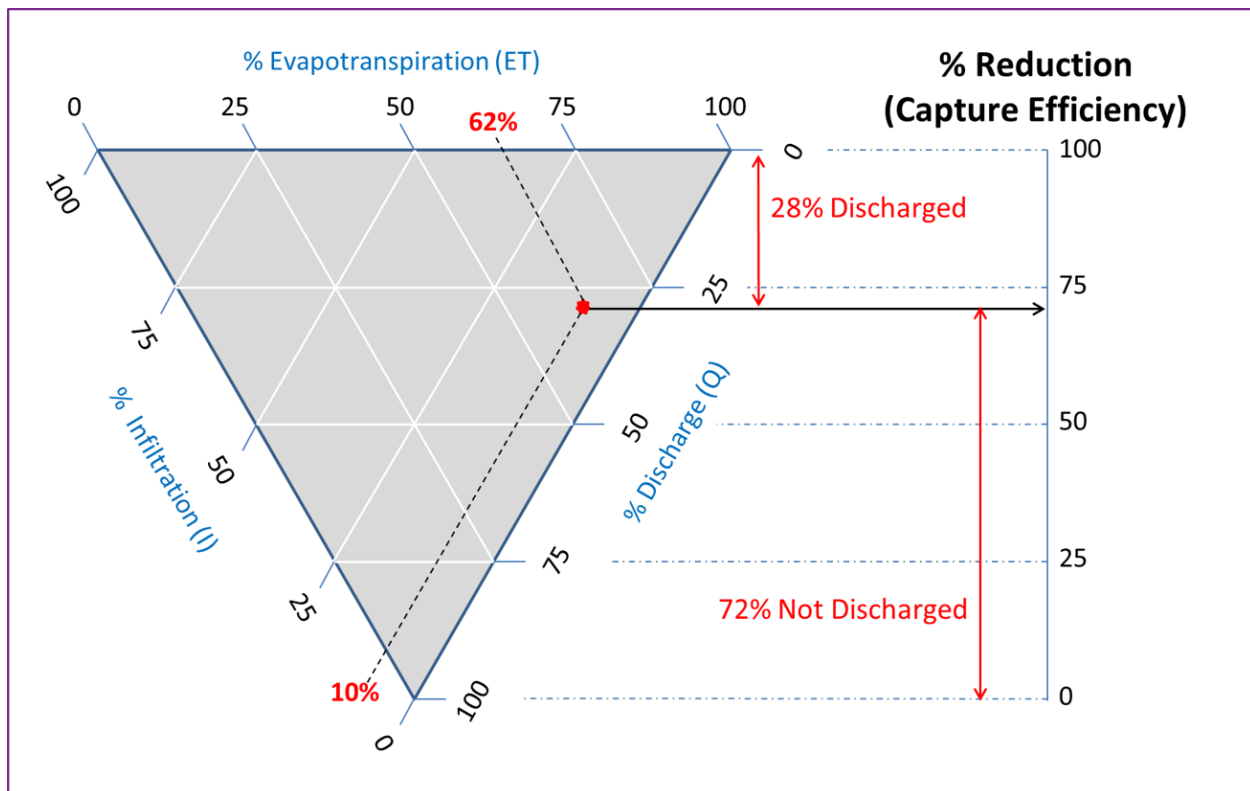


Figure 3.1: Example usage of the Water Budget Triangle.

This analysis and visualization approach is similar to the Piper plot diagrams developed by Lent et al. [1997] to characterize and visualize hydrologic indices for wetlands. However, the Water Budget Triangle is inverted from the Piper plot (to emphasize the importance of prioritizing ET and percolation in GI design) and does not account for influent sources of water. Other simplified water budget visualization tools also depict water budgets for both individual structures and whole watersheds (see Askarizadeh et al. [2015]). The Water Budget Triangle method was developed to lower communication barriers that limit GI implementation, including (a) conveying technical information about various stormwater devices to technical and lay stakeholders; (b) providing a systematic visualization tool to compare performance of dissimilar systems; and (c) eliminating ambiguity in the description of BMPs for non-technical stakeholders. The methodology uses a water balance approach to account for fractional fluxes of water leaving the boundaries of the stormwater control structure along each pathway (Q, ET, and I). The mass water flux of any given system may be represented by the water balance equation:

$$R + P = Q + ET + I + \Delta S$$

R represents influent water or run-on, P represents direct precipitation input and Q, ET, and I (defined above) are calculated for the time step of interest. This approach may include stored water released after the event hydrograph. Note that water stored (ΔS) in the system control volume can be depleted after a storm event; however, the time step represented on the triangle must be uniform for all axes. The change in water storage (ΔS) within the system is not explicitly represented in the ternary diagram, because it is not a flux. The tool makes a distinction between changes in stores and fluxes because it is used to quantify the relative importance of different loss processes occurring over the time step of interest. Following increases in inputs from a storm event, there is an increase in water storage for a period,

$$P + R > 0 \text{ (a storm event occurs)}$$

and

$$\Delta S \geq 0 \text{ (the structure fills).}$$

After the event, losses will eventually exceed inputs, resulting in decreasing storage,

$$P + R = 0 \text{ (the event ends and antecedent dry period begins)}$$

and

$$\Delta S < 0 \text{ (the structure empties);}$$

therefore,

$$-\Delta S = ET + I + Q \text{ (change in storage must equate to total losses).}$$

It is convenient to choose the time scale for analysis as the period over which $\Delta S = 0$, when storage returns to the initial condition prior to a runoff event and the mass balance is fully described by the loss terms in the diagram (i.e., steady-state). However, steady-state condition is not a requirement for application of the tool, as long as the time step remains constant across all loss pathways. Over increasingly long timescales, the magnitude of ΔS approaches zero in comparison to the other terms in the water balance equation, leaving:

$$R + P = Q + ET + I.$$

Creating complete water budget triangle datasets. Water budget data for each loss pathway (Q, ET, I) was plotted in R using the `ggtern` package (Hamilton and Ferry

[2018]). Full datasets and R scripts for plots and statistical analysis are available online at: <https://github.com/cgeger/WaterBudgetTriangle>. It is relatively uncommon for complete water budgets to be reported in the peer reviewed literature for individual LID systems. Most researchers studying GI systems solely measure runoff (Q), fewer measure percolation (I) and almost none report evaporation or transpiration (ET). Few studies report a closed water budget (all three loss pathways), with the most common measurements being Q, P, R, and either ET or I. Missing loss pathway variables were calculated by assuming a closed water balance for each system, and solving for the estimate using the equation above. The ecohydrology of natural wetlands has been an ongoing field of study for more than four decades, thus, water budget data are most readily available for natural and constructed wetlands (Lent et al. [1997], Mitsch et al. [2014]). Complete water budget data for green roofs are numerous, owing to the fact that percolation is zero and ΔS approaches zero over long time scales. The most comprehensive reports of complete water budgets over months or years of monitoring are generally available through engineering reports and student theses or dissertations (Appendix A.1). Water budget data are scant in the peer-reviewed literature for more conventional systems, including sewer conveyance networks, detention (dry) ponds and retention (wet) ponds. In this analysis, I show data visualizations for both individual studies and summaries from several review papers. Many prior analyses model peak discharge from retention (wet) and detention (dry) ponds, but few comprehensive water budgets partition evapotranspiration or permanent groundwater recharge in these systems. Water budget data available for conventional stormwater control structures come primarily from post-project monitoring reports compiled by engineering firms and government planning agencies.

3.4 Results

It is essential that stormwater infrastructure designers understand how physical design, drainage media preparation, long-term maintenance, and plant species affect the water budget of a built system. To explore this, I present a series of case studies from stormwater management technologies found in the literature, including runoff reduction calculations; hypothetical assessment of dynamic behavior; comparison of modelled and measured behavior for both constructed and natural BMPs; and comparison of multiple GI and conventional technologies. This synthesis supports reasonable expectations that modifying contributing catchment area, basin area, hydraulic retention time, media depth or soil particle characteristics, rooting depth, and other ecohydrologic characteristics will change the water budgets of engineered stormwater systems. Equipped with a quantitative understanding of the hydrologic function of stormwater technologies, the application of watershed models should allow for projections of the stormwater management actions needed to achieve water resource objectives. Using the datasets collected from the literature, I calculate acceptable operational ranges for these structures and understand design factors that influence hydrologic processes (Tables 3.1 and 3.2). This data-driven approach supports the development of a ‘sliding scale’ performance credit system (Brown et al. [2011]).

3.4.1 Natural and constructed wetlands

The hydrologic function of natural wetlands has been explored for more than 50 years (Crisp [1966]) and offers a good reference for comparison with constructed systems. Wetlands exhibit a wide range of hydrologic behavior, due to (a) varying hydrogeomorphic controls for natural wetlands and (b) diverse design objectives for constructed wetlands (Lent et al. [1997], Nungesser and Chimney [2006]). Dominant hydrologic fluxes in wetlands may change over short time scales throughout the period of surface runoff and may reverse during baseflow, flood levels, or tidal extremes (Choi and Harvey [2000],

Hughes et al. [1998]). For example, the major hydrologic fluxes from kettle wetlands are ET and infiltration, which may exhibit diel or seasonal fluctuation, depending on temperature and precipitation (Hollands [1989]). Natural wetlands typically have more complex water budgets than small constructed wetlands. It may not be possible to depict water budgets for estuaries and other coastal wetland systems with tidal forcing using the Water Budget Triangle (see Hughes et al. [1998]). Lent et al. [1997] presented a Piper plot water budget method to classify natural wetlands and lakes, summarized in Figure 3.2 alongside data from constructed wetlands and wetland models. Natural wetlands show greater ET fractions than constructed wetlands (46% vs. 8% in Table 3.1), but similar proportions of flux to groundwater (5–10%). Models for natural and constructed wetlands also reflect this difference. The compiled data (Figure 3.2, Table 3.1) also show constructed wetlands produce 22–26% more runoff on average than natural wetlands. This observation makes a good case for preservation of naturally occurring wetlands during landscape development rather than building ‘replacement’ constructed wetlands, an intervention pertinent to developers and land managers in watersheds struggling to control downstream flooding. The difference may arise from (a) reduced ET in constructed wetlands associated with lower vegetation density; (b) lower ET related to soil carbon content (or different humic material structure), which affect relative infiltration and evaporative fluxes from the system; or (c) seasonally high groundwater surfaces that are shorter in duration for constructed wetlands. In comparison, lakes (Figure 3.2) exhibit a slightly greater fraction of infiltration than natural wetlands.

Table 3.1: Comparative summary of hydrologic function of stormwater technologies and natural systems. Mean and median values from the compiled datasets for each system type, calculated for each fractional variable of the water budget (discharged runoff [Q], percolated or infiltrated drainage [I], and evapotranspiration [ET]). Note that the values represent the mean or median of each fractional loss pathway calculated independently from the other fractions; totals may not sum to exactly 100, and may not match plotted versions.

System	N	Mean			Median		
		Q	I	ET	Q	I	ET
Porous pavement	23	34	54	11	20	64	10
Continent ^a	6	37	7	55	38	7	54
Natural lake	13	38	15	47	45	12	43
Green roof	59	39	0.30	61	36	0	64
Bioretention cell ^a	10	46	16	37	52	17	28
Cistern ^a	36	43	55	3	41	58	1.5
Detention pond ^a	7	43	37	21	48	28	24
Natural wetland	19	46	11	42	45	5	46
Constructed wetland	8	68	19	13	71	10	8
Retention pond	9	73	5.3	21	85	0.20	8.4
Sewer pipe section ^a	13	76	24	0	83	17	0
Sewershed network ^a	12	88	12	0	91	9	0

^acalculation includes at least one model estimate

Table 3.2: Diagnostic operational ranges for engineered stormwater systems. Estimates of short-term operational range for each engineered system, calculated as the 5th and 95th quantile of the dataset, rounded to the nearest 5th percentile. Estimates of long-term performance, which is the 5th and 95th percentile confidence interval range about the calculated mean, estimated by resampling the distribution of means 1,000 times and rounding to the nearest percentile. Confidence intervals and quantiles were calculated for each loss variable separately and may not sum to 100. Bold values represent worst-case performance scenarios, underlined values represent good performance design benchmarks.

5th-95th%ile Short-term operating range					5th-95th CI about the mean Long-term performance														
System	N	Q	I	ET	Q	I	ET												
Sewershed network ^a	12	65	—	100	0	—	35	0	81	—	93	7	—	19	0				
Retention pond ^a	9	30	—	95	0	—	25	0	—	70	58	—	87	0	—	13	8	—	37
Sewer pipe section ^a	13	50	—	90	10	—	50	0			68	—	83	17	—	32			0
Constructed wetland	8	25	—	95	0	—	60	0	—	40	51	—	81	6	—	35	6	—	22
Bioretention cell	10	10	—	75	0	—	30	20	—	65	33	—	59	11	—	22	29	—	47
Detention pond ^a	7	10	—	70	15	—	70	10	—	35	29	—	57	24	—	50	15	—	27
Cistern ^a	36	20	—	70	30	—	80	0	—	10	39	—	48	50	—	59	2	—	4
Green roof	59	20	—	70	0			30	—	80	36	—	43		0		57	—	64
Porous pavement	23	0	—	80	5	—	90	5	—	20	26	—	44	45	—	63	10	—	13

^acalculation includes at least one model estimate

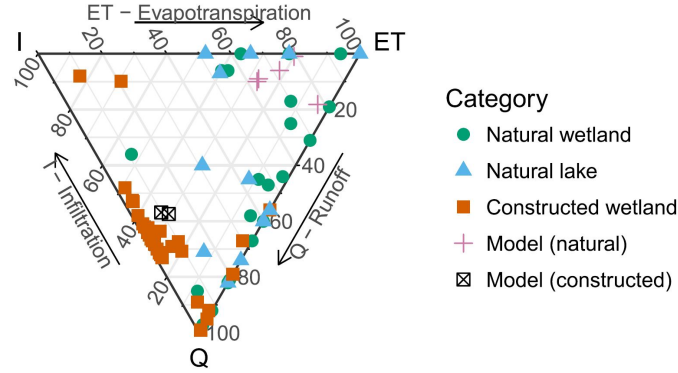


Figure 3.2: Variability of hydrologic performance of natural wetlands and lakes ($n = 21$ wetlands and 13 lakes) with data reported for eight constructed wetlands ($n = 34$ annual or greater measurements). Results from two modelling efforts were used to estimate modelled water budgets in natural wetlands ($n = 5$ model estimates) and constructed wetlands ($n = 2$ model estimates). Note that the eight constructed wetlands are broken out by water year for the purposes of visualization, there are several overlapping points on the Q vertex. Data compiled from Ayub et al. [2010], Caldwell et al. [2007], Choi and Harvey [2000], Crisp [1966], Daniels et al. [2000], Hemond [1980], Hey et al. [1994], Lent et al. [1997], Mitsch et al. [2014], Nungesser and Chimney [2006], Strosnider et al. [2007].

3.4.2 Retention basins and ‘wet’ ponds

Like many constructed wetlands, retention basins are designed to maintain permanent standing water. As with wetlands, low seepage and limited groundwater exchange in retention ponds is controlled by subsurface hydrology: natural groundwater table, impermeable soils, compaction, and/or presence of a liner (Hartigan et al. [2009], PA DEP [2006]). Water budget data were obtained for measured values from seven retention ponds in Florida and 2 years of modelled wet pond water budgets from the City of Austin Stormwater Treatment Section (Harper [2010a,b,c, 2011], Harper et al. [2003], Hartigan et al. [2009], Teague et al. [2005]). Cumulative monthly water budgets sized by the monthly precipitation (Figure B.3), show that retention ponds may behave as zero-discharge systems during seasonally dry conditions, but typically discharge 85–95% of influent. The remaining water is lost to ET (8–13%) or groundwater ($< 5\%$). Hydraulic retention time (average length of time a unit of water spends in the basin storage volume)

is a significant predictor of the fraction of inflow occurring as runoff ($R^2 = 0.81$, $p = .0006$) and ET ($R^2 = 0.98$, $p = 1.32\text{e-}7$), but not infiltration, in retention basins (color gradient in Figure 3.3 shows hydraulic retention time varying across the ET axis, but not I axis). Increased hydraulic retention time increases evaporation but not infiltration. Infiltration losses are explained more by site location than other variables, indicating this pathway is controlled by site groundwater hydrology, not surface water inputs (Figure B.3). These observations are in line with the design assumption that percolation is not an important sink for retention ponds, although the Poppleton, Palm Bay, and Tampa sites show seasonal groundwater connectivity (Figure B.3).

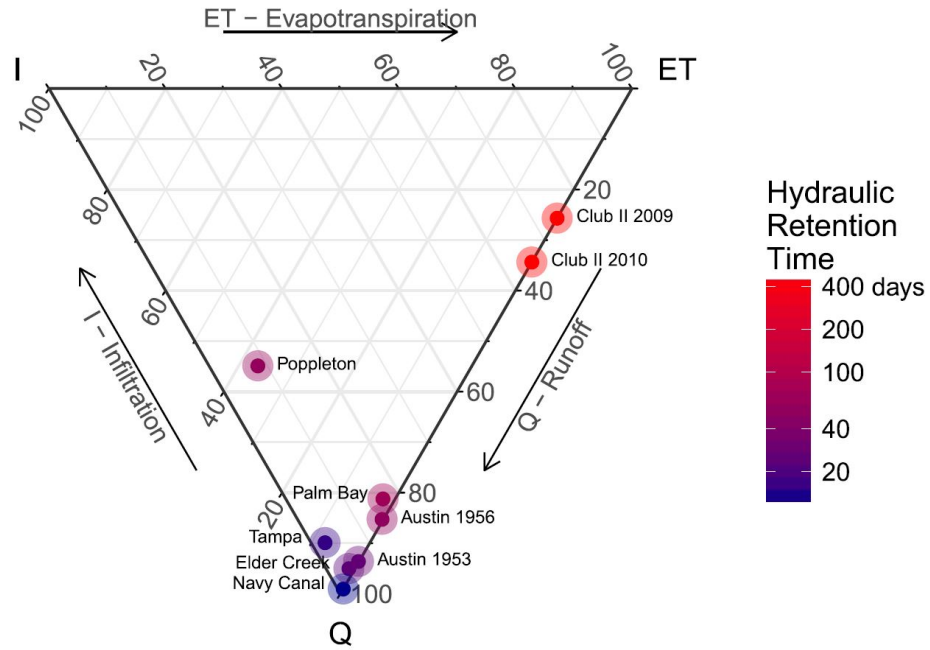


Figure 3.3: Cumulative water budgets for seven retention ponds and modelled wet pond performance in Austin, Texas for 1953 and 1956 ($n = 9$). Symbol color represents measured or estimated hydraulic retention times calculated from basin volume and average pond influx per day. Data compiled from Harper [2010a,b,c, 2011], Harper et al. [2003], Hartigan et al. [2009], Teague et al. [2005].

3.4.3 Detention basins and ‘dry’ ponds

Detention ponds are commonly engineered for 6–72 hr of transient storage to attenuate peak flows. Unlike retention ponds, they are not designed to maintain permanent standing water, so are typically dry except for periods of wet weather. Despite their widespread use, detention structures are rarely studied from an ecohydrological perspective; consequently, retention and detention ponds are rarely considered to be types of constructed wetlands. However, this is an artificial classification, because hydrologic behavior of retention and detention systems places them alongside constructed wetlands on the same continuum (compare Figures 3.2 and 3.4). Much of the literature reporting hydrologic performance of detention and retention basins has focused on the event-scale (Geosyntec Consultants and Wright Water Engineers [2011]), which overlooks the longer-term roles of ET or groundwater recharge (I) from detention ponds (WEF and ASCE [1998]). Use of the triangle tool to study detention ponds requires defining an appropriate time scale to partition infiltration (I) and ‘event runoff’ (Q). Event water detained during the period of surface runoff is considered beneficial to watershed function if it is released gradually during baseflow and is a comparable quantity with percolated drainage. The water balance of detention basins is more variable than retention ponds (compare Figures 3.3 and 3.4). Analysis of five detention basins from California and Nevada and 11 grass lined detention basins from the International Stormwater BMP Database suggests volumetric reduction to be between 8% and 33% (i.e., $67 < Q < 92$; 2nd Nature, LLC [2006], Geosyntec Consultants and Wright Water Engineers [2011]). This pattern indicates these systems may behave similarly to retention basins, or may produce substantially less runoff. Unlike wet ponds, dry detention basins are thought to have good hydrologic connectivity with groundwater, depending upon the infiltration area, soil hydraulic conductivity, and unsaturated depth (2nd Nature, LLC [2006]). The assumption of good groundwater connectivity overlooks ET as a loss pathway for detention basins, as indicated in Figure 3.4. The model represented (black dot) seems to underestimate both hydrologic connectivity

and ET when compared with actual field observations (green and blue dots). The presence of mowed vegetation in detention basins also contributes to greater long-term ET losses. Further study of detention basin water budgets may indicate specific design criteria (area to depth ratio), or management techniques (mowing, aeration, and planting strategies) that may improve long-term stormwater retention by increasing I or ET losses. Analysis of detention and retention pond design characteristics from a water budget standpoint may lead to improved hydrologic or water quality performance of constructed wetlands.

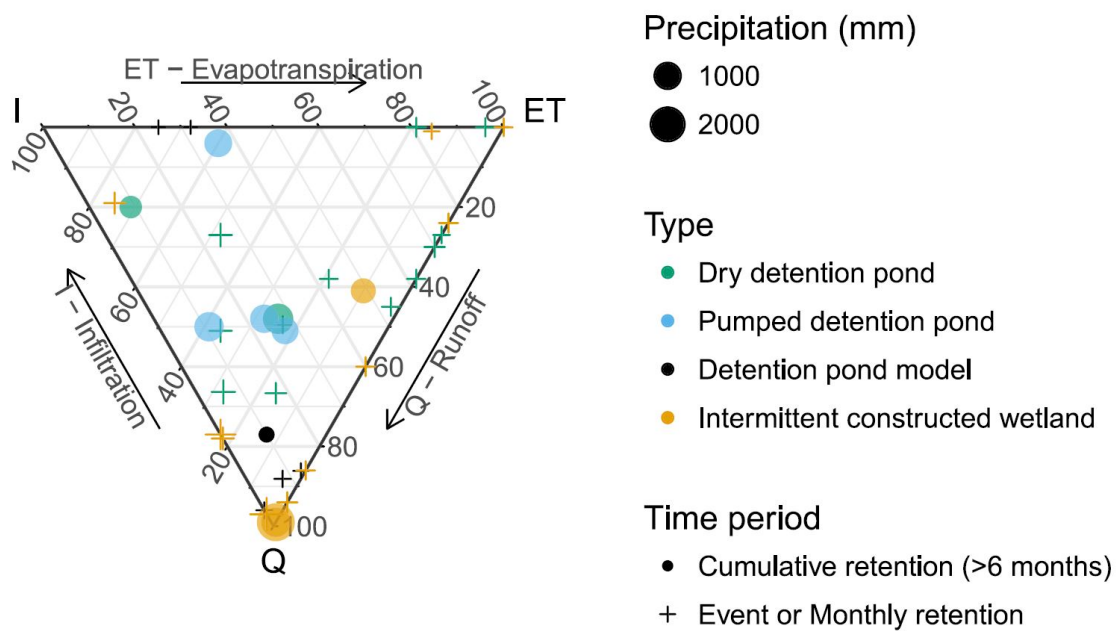


Figure 3.4: Circles display cumulative water budgets from gravity-fed detention ponds ($n = 2$) and pump-fed detention ponds ($n = 4$, two annual measurements for two ponds), a Green-Ampt detention infiltration model ($n = 1$) and constructed wetland systems that report having detention facilities ($n = 3$). Crosses show the single-event or monthly retention variation for the same detention systems ($n = 48$). Data compiled from Ayub et al. [2010], Daniels et al. [2000], Emerson [2003], Harper et al. [1999b, 2002], Shukla et al. [2015].

3.4.4 Bioretention cells, infiltration trenches, swales, and rain gardens

Analysis of precipitation and runoff by Traver and DeBarry [2003] in southern Pennsylvania indicated that 80% of total annual precipitation volume can be captured by retaining the first 25 mm of each rainfall event. Previous studies have found that rain

gardens and bioretention cells reduce runoff volume by 30–99% (Schlea [2011], Strauch et al. [2016]). Newcomer et al. [2014] modelled volumetric reduction of 58–79% for infiltration trenches, compared with 8–33% reduction on irrigated grass lawn. The wide range in performance arises from several design factors that significantly affect the long-term water budgets of swale systems. These design factors include: presence of an underdrain, liner, or internal water storage zone; contributing catchment area ratio; direct connection of impervious surfaces; ponding depth; media depth, composition, and particle size distribution; and plant density and species composition (Bratieres et al. [2008], Li et al. [2009], Roy-Poirier et al. [2010]). Environmental factors that affect water budget include: native sub-base drainage and water table height; event depth and intensity; season and temperature; age and maturity of the planted system; and particle clogging. To my knowledge, there are no known studies or reviews that examine all of these factors in controlled experiments and prioritize their relative importance to hydrologic performance. However, there are published case studies, multiple models, and a general intuitive understanding about how these factors affect performance at individual sites (Wardynski and Hunt [2012]). Winston et al. [2016] observed that the underlying sub-base conductivity is a key factor for volume reduction in bioretention cells, reporting that even poorly drained soils can be effective for events smaller than ~ 6 mm if design allows for an internal water storage zone. Thus far, discussion of internal water storage zones has mostly assumed that better hydrologic performance arises from increases in exfiltration (I increases). However, lysimetry studies indicate that bioretention ET can become water-limited under dry conditions (Wadzuk et al. [2015]). ET from bioretention cells decreases as a function of decreasing soil moisture below field capacity: drier soils evaporate less water than wet soils (Buckingham [1907]). ET is greatest following a rain event and decreases over subsequent days, resulting in water limitation of bioretention cells. Therefore, internal water storage zones may also maintain system capacity for ET by limiting plant water stress and maintaining sufficient capillary conductance and

connectivity to the soil surface. Wadzuk et al. [2015] demonstrate ET limitation by water availability using weighing lysimeters with and without an internal water storage zone.

Lysimetry data from Hess [2014] clearly implicate ET as an important loss pathway for bioretention (Figure 3.5, in green). Several models of bioretention estimate ET at or below 5% of the water budget, which is much less than the estimates from lined bioretention cells (19%) and weighing lysimeters (40–78%), and less than the estimates for unplanted porous pavement (~ 10 –20%). The discrepancy could be due to the time step used in model calculations, which is narrowed to the event scale plus 24 hr, over which little ET occurs. However, 50-year climate simulations using DRAINMOD also largely underestimate the evaporative fraction of long-term water budgets (Wardynski et al. [2011]; Figure 3.5, orange box, lower left). This pattern is corroborated by Hess et al. [2017] and Wadzuk et al. [2015], who reported that using the Penman–Monteith equation tends to underestimate ET while using the Hargreaves equation tends to overestimate ET. The yellow circles in Figure 3.5 represent two cells where ET was estimated using Penman–Monteith. Additional work is needed to constrain annual ET estimates more closely for swale systems before making long-term performance projections under varying climate conditions. The most comprehensive work on ET in bioretention thus far has come from three weighing lysimeters with differing soil types (Hess [2014], Hess et al. [2015, 2017], Hickman [2011], Hickman et al. [2011], Wadzuk et al. [2015]. Trials from Hess et al. [2015, 2017] indicate that soil composition controls whether percolation or ET is a more important loss pathway over the long-term (i.e., right to left I–ET axis on the triangle). Using lysimetry, Denich and Bradford [2010] reported summer ET rates of 4.2–7.7 mm/d in Ontario; Hess et al. [2017] reported average ET rates of 2.9–4.3 mm/d during the growing season and 1.5 mm/d in winter, with an annual ET of approximately 600 mm in Pennsylvania. For context, the annual rainfall depth in Guelph, Ontario is approximately 958 mm, and 1230 mm in Villanova, Pennsylvania. These numbers indicate that ET losses may accommodate half (or possibly more) of the annual precipitation budget, even in

northern continental climates with cold winters.

Peak flow reduction is a major goal in mitigating downstream flooding and is used as a common assessment metric for bioretention cells, retention, and detention ponds. However, for infiltrative systems like bioretention cells and porous pavement, volumetric reductions (ET and I) drive peak flow reduction, whereas temporary storage (ΔS) accounts for peak flow mitigation in retention/detention systems. This distinction is significant for understanding both site-level and watershed-scale impacts of engineered stormwater systems. Also, unlike peak flow reduction, volumetric reduction is not related to event intensity. Modest increases in volumetric reduction seem to drive large peak flow attenuation in bioretention and porous pavement systems but less so for grassed swales, detention, and retention ponds. For instance, researchers at NC State and Ohio Department of Natural Resources reported runoff reduction of 36–60% but median peak flow reduction of 97–100%, with maximum flow rates decreasing by at least 29% (NERR [2016]). Strauch et al. [2016] reported that only 39 out of 255 events produced measurable runoff at a bioretention facility in Nebraska; volumetric reduction was 33–100% on an event scale and mean peak flow reduction was 63%. Additional research is needed at the event scale to determine if there is a predictable relationship between peak flow reduction and volumetric reduction for different stormwater technologies (under uniform climate conditions).

3.4.5 Pervious pavement

Porous or pervious pavement includes permeable asphalt or concrete, interlocking pavers, grassed paver surfaces, and many proprietary mixes for walking, driving, and parking surfaces. Infiltration rates of engineered porous surfaces can vary widely, ranging from 2.4 to 4.0 mm/min, greater than double the infiltration rates of natural soils or grassed surfaces (Valinski and Chandler [2015]). Most designs and models attribute runoff reduction volumes to infiltration ($I \approx 90\%$, $Q = 10$), but ignore evaporation (Drake et al.

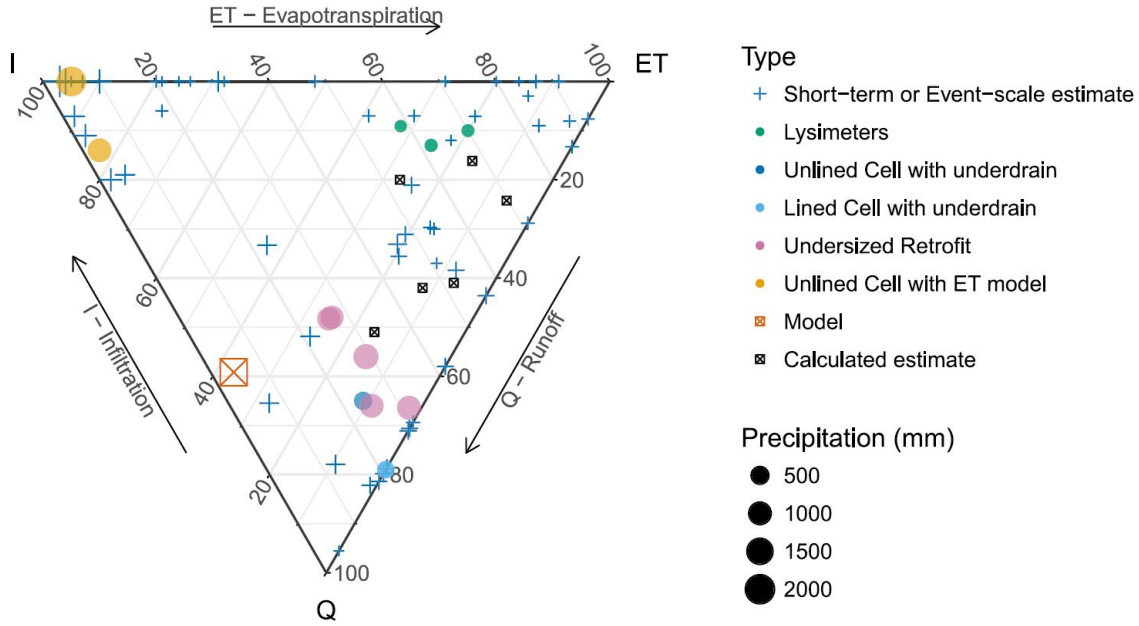


Figure 3.5: Water budgets for three lysimeters in Pennsylvania (green), a pair of lined and unlined cells with underdrains from North Carolina (light and dark blue), measurements from two sets of undersized, unlined retrofits with underdrains in Ohio (pink), and two unlined cells from Nebraska with ET fraction estimated using the Penman–Monteith method (yellow). Several shorter-term estimates from the same locations are also presented (blue crosses, $n = 59$), along with calculated estimates using volumetric moisture content constraints (black boxes, $n = 6$) and a DRAINMOD estimate with very low ET (orange box, $n = 1$). Data from Hess [2014], Kosmerl [2012], Li et al. [2009], Pitt et al. [2007], Strauch et al. [2016], Wardynski et al. [2011].

[2014]). However, Pratt et al. [1995] reported lined porous pavement systems equipped with underdrains reduced runoff by 20–50% due to increased evaporation ($I = 0$, $Q = 50$ –80%) at rates between 0.2 and 5.5 mm/day. Evaporation loss estimates for unplanted porous pavement range from 3% to 44% and are heavily influenced by the time step of the monitoring period. In general, surface runoff from porous pavements is more sensitive to rainfall intensity than rainfall depth, so results from event-scale studies are more common than long-term cumulative water balances. Event-scale studies frequently assume evaporative losses are negligible because values less than 0.5 mm/day are common between March and November (Göbel et al. [2013]). This assumption is likely reasonable at the event scale, because porous pavement can have a runoff threshold of up to 7 mm and

because low-intensity, small precipitation events (less than 2 mm) are sometimes excluded from observations. However, evaporative losses over longer timescales are substantial, with annual values reaching 150 mm, easily 10–20% or more of an annual water budget in North America or Europe (Figure 3.6; Göbel et al. [2013], Hein et al. [2013]). Martin and Kaye [2014] indicate that approximately 1 mm/day is a conservative ET estimate for porous pavements without underdrains. Göbel et al. [2013] estimate cold-weather ET rates from porous pavement during December through February around 0.24 mm/day. Similarly, winter evaporation can be substantial during cold, dry weather: Drake et al. [2014] reported > 20 mm cumulative ET over a winter, accounting for $\sim 9 - 13\%$ of a winter water budget in Ontario, Canada. Ignoring winter measurements further contributes to the underestimation of ET on an annual basis, especially at sites with intermittent snow cover. A 2-year water balance study on three types of lined ($I = 0$) porous pavement measured 95 mm of evaporation and estimated ET losses to be between 2.4% and 7.6% of annual precipitation (Brown and Borst [2015]); however, the authors conclude that the design could be modified to enhance evaporation to between 7% and 12%. This is a conservative range for long-term model estimates. Surface color is a key factor affecting ET losses for porous pavements, because the energy for evaporation is conducted to pore water through the thermal conductivity of the paver; dark-colored pavements may increase ET by up to $\sim 20\%$ (Göbel et al. [2013], Starke et al. [2010]). Seam area is also an important factor influencing both infiltration and evaporation rates (Starke et al. [2010]). Using grassed pavers or pairing unplanted porous pavements with street trees increases transpiration and rainfall interception (Vico et al. [2014]). Vegetated grass pavers may evaporate ~ 1.5 mm/day, accounting for more than 50% of annual precipitation (Göbel et al. [2013]). Like green roofs, evaporative losses from both grassed pavers and unvegetated pavement may become water-limited during dry periods (ET may decrease in periods of low pore moisture content; Brown and Borst [2015], Pratt et al. [1995]). Thus, eliminating underdrains or including an upturned elbow for internal water storage can increase exfiltration time (I

increases) and prevent water-limiting conditions from occurring between storms (ET increases).

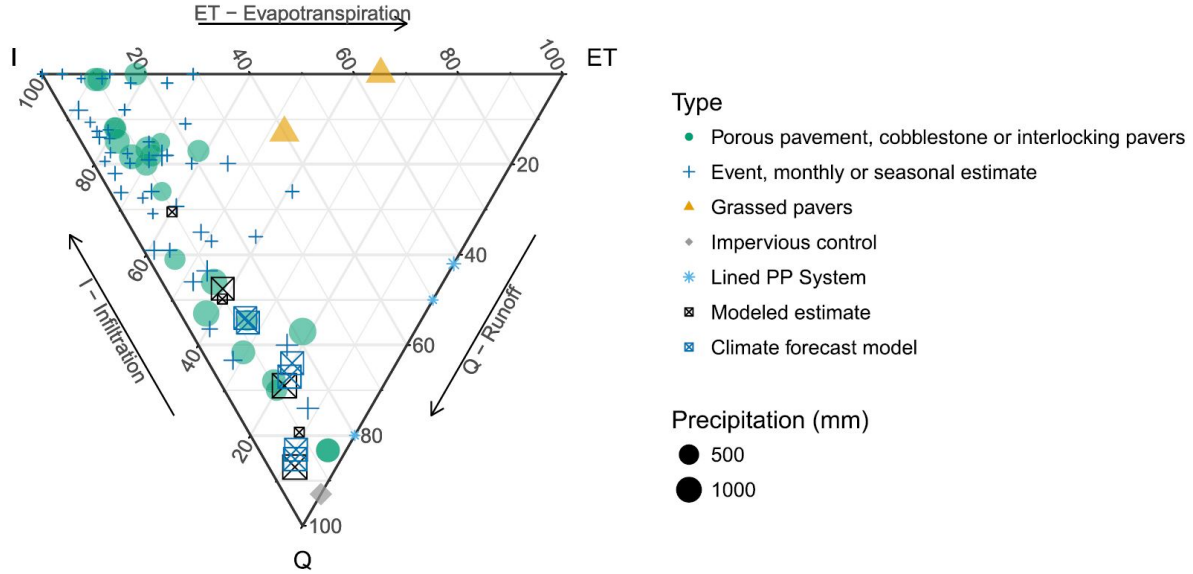


Figure 3.6: Reported for water budgets from unplanted porous asphalt, permeable concrete, cobblestone and interlocking or tongue-and-groove pavers ($n = 15$, some estimated more than once) alongside models ($n = 12$) and short-term measurements ($n = 43$) for the same locations. For reference, estimates for grassed pavers ($n = 2$), an impervious surface ($n = 1$) and lined porous pavement systems ($n = 3$) are also presented. Data compiled from Brown and Borst [2015], Drake et al. [2014], Göbel et al. [2013], NERR [2016], Pratt et al. [1995], Rim [2011].

3.4.6 Green roofs

The primary water sinks for green roof systems are ET and discharge, with minimal permanent storage and no infiltration occurring across the impermeable membrane below the growth media (Wadzuk et al. [2013]). Because green roofs are disconnected from ground infiltration, below a minimum event threshold they can operate as zero-discharge systems (Figure 3.7; top-right corner of triangle), with larger events plotting progressively closer to the lower vertex. Every green roof has a maximum water retention limit; progressively larger and more intense events retain and evaporate/transpire proportionally less water. The hydrologic function of a green roof is greatly affected by media depth.

Deeper media cells capture incrementally more water; Fioretti et al. [2010] show study roofs with greater than 15 cm of substrate retain more water than shallower systems (2–15 cm). However, they also show that a green roof with a modest 2-cm media depth retained more than 400% more precipitation than a conventional roof. Soil media characteristics also play an important role, because particle size distribution determines water holding capacity and retention (Graceson et al. [2013]). The chemical properties of green roof soil media are less well studied, but agronomic and soil science studies have demonstrated that some soil media characteristics enhance water retention capacity (Bleam [2016]).

Vegetation increases retention by enhancing transpiration losses; the blue (unplanted) roofs presented in Figure 3.7 show lower ET than planted roofs. Morgan et al. [2013] reported that a minimum of 20–25% vegetated roof coverage is needed to increase stormwater retention beyond the capacity of the growth media alone. Lab-scale roofs do not seem to capture the full range of performance shown in full-scale green roofs (Figure 3.7). This discrepancy is possibly due to shorter monitoring periods, lower vegetation density, or greater soil moisture range, and edge effects in smaller lab-scale roofs. The roof media storage volume is also dependent upon ambient temperature and to a lesser extent on the carbon content and root biomass. Green roofs presented in Figure 3.7 represent performance across a range of seasons and climates; green roofs in warmer seasons and climates generally capture and evaporate more water than cold-climate green roofs. However, cold-climate green roofs still perform well in comparison with conventional roofs. The mean and median values of the 59 green roof water budgets from the literature are very similar (Figure 3.7), Q is approximately 36–39% and ET around 61–64%.

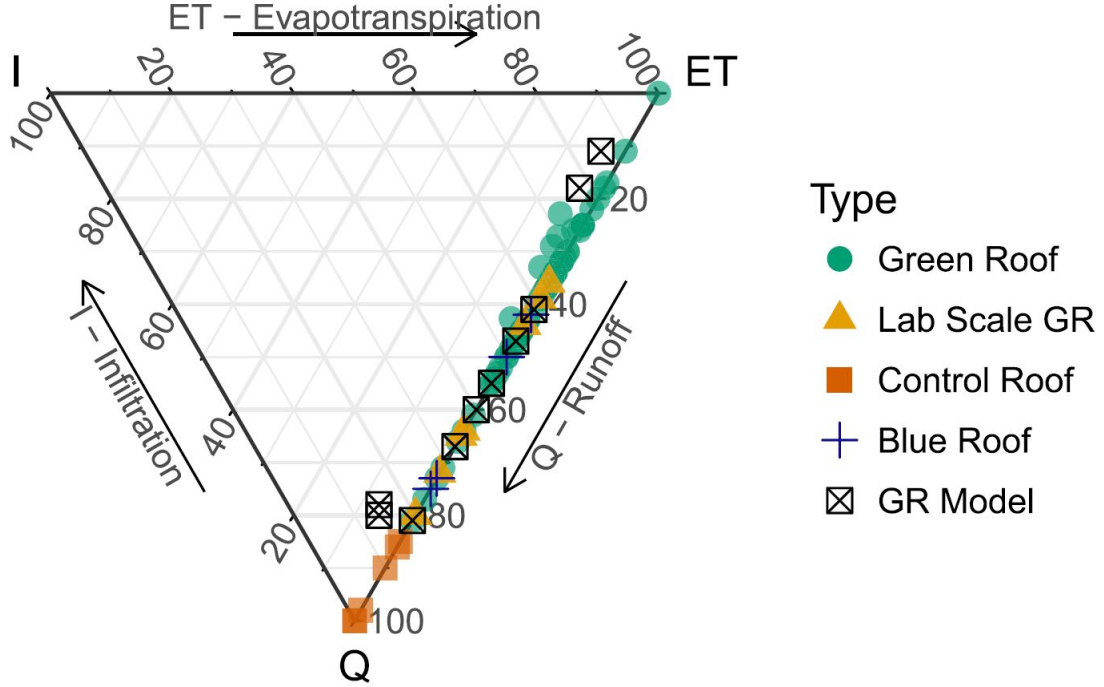


Figure 3.7: Eighty-seven water budgets representing installed green roofs ($n = 59$, in green), lab-scale study roofs ($n = 7$, in yellow), measured control roofs ($n = 6$, in orange), blue roofs (consisting of gravel or other unplanted substrate, $n = 4$, in blue), and green roof models ($n = 10$, in black). In some cases, the initial abstraction value was reported; this value was used to estimate I for green roofs and ET for control roofs, otherwise the I is assumed to be 0 because it is minimal over long timescales. Data compiled from Ahiablame et al. [2012], Berghage et al. [2007, 2009], Buccola and Spolek [2011], Czemieli Berndtsson [2010], Carpenter and Kaluvakolanu [2011], Carson et al. [2013], Carter and Rasmussen [2005, 2007], Fassman-Beck et al. [2013], Fioretti et al. [2010], Getter et al. [2007], Gregoire and Clausen [2011], Hathaway et al. [2008], Hoffman et al. [2010], Hutchinson et al. [2003], Liu and Minor [2005], Mentens et al. [2006], Moran et al. [2005], Nawaz et al. [2015], Palla et al. [2011], Stovin [2010], Stovin et al. [2012, 2013], Toronto Regional Conservation Authority [2006], Teemusk and Mander [2007], Van Seters et al. [2009], Vanuytrecht et al. [2014], VanWoert et al. [2005], Villarreal and Bengtsson [2005], Wadzuk et al. [2013].

3.5 Discussion

3.5.1 Modelling and diagnostic estimates

Basic water budget statistics from results presented in the previous sections are summarized in Table 3.1. This data summary can be used for modelling efforts and as a diagnostic benchmark to identify under-performing sites. Mean and median estimates of

runoff are within 5% for many systems, apart from bioretention, porous pavement, natural lakes, retention ponds, and sewer pipe sections. Given the variability of event-scale measurements and site-to-site comparisons, operational ranges are likely more useful for both modeling and diagnostic applications than individual summary statistics. Estimates of short-term operational range and long-term performance range for each engineered system type are displayed in Table 3.2. These water budget ranges are intended as a benchmark for comparing hydrologic function among unlike technologies. The discharge numbers in bold on Table 3.2 represent a suggested worst-case performance value for each technology. Correspondingly, it would be prudent for designers to aim for designs that meet or exceed the underlined values. For example, bioretention cells that discharge more than 59% on an annual basis should be examined for retrofit or design changes that can improve performance (Table 3.2). Similarly, sewershed networks that discharge less than 81% of the water conveyed should be examined for leakage (Table 3.2). These diagnostic ranges provide a baseline that should help to optimize designs in the future. However, a larger and more representative water budget dataset would provide more robust certainty about reasonable expectations for retention ponds ($n = 7$ ponds and 2 models), constructed wetlands ($n = 8$), detention ponds ($n = 6$ ponds and 1 model), and bioretention cells ($n = 3$ lysimeters, 5 undersized retrofits, 1 lined cell with underdrain, and 1 unlined cell with underdrain). (Appendix B has additional discussion about systems not shown here.) Additionally, there may be future innovations that improve volumetric capture; performance ranges may change over time to reflect this improvement.

3.5.2 Data limitations

The water balances reported here are not uniform among sites—specifically, the various time steps stem from differences in study questions and approaches. Hydrologic monitoring at some sites was conducted over multiple water years, whereas most studies of GI collected 6 to 9 months of data. These differences are notated to the extent possible to

provide the greatest context for comparison. The data reported also represent diverse field techniques due to a wide range of approaches used in measurement and reporting among institutions and study cases. Although these differences are not unimportant, the datasets represented in the literature show clear patterns for summary and interpretation, despite varied methods. Q is typically measured directly using a tipping bucket gage or structure calibrated with a stage-discharge relationship. Infiltration measurements are estimated using a linear scale or pressure transducers to quantify the height of the water table and/or soil moisture sensors to estimate pore water saturation. Evaporation estimates are generally modelled by energy balance techniques typical of micrometeorology studies (Wadzuk et al. [2015]). A few studies use pan evaporation or other direct measurement techniques, but this approach is less common. As a result, the estimated proportional importance of ET within the water balance is likely incorrect and probably underestimated over long time and large spatial scales. Underestimation of ET is likely because many models project by upscaling short-term/small-scale measurements during periods when ET is known or assumed to be minimal. The relative simplicity of a green roof water budget (as compared with systems with complex groundwater hydrology) highlights the effect that divergent measurement and analysis techniques have on conclusive outcomes. Several of the green roof case studies presented demonstrate both time-and scale-dependent results (Czemiel Berndtsson [2010], Fioretti et al. [2010], Stovin et al. [2012], Voyde et al. [2010]) that contribute to the broad array of performance (Figure 3.7). For example, lab-scale studies generally result in more runoff than full-scale green roofs, and event-scale monitoring results in more variable ranges than cumulative studies.

Event-scale analysis tends to focus on storms above a specific threshold, creating a cumulative water budget that is less representative of overall annual performance, usually reflecting less ET and more runoff. However, using event-scale analysis water budgets can help elucidate the role of antecedent soil moisture and event precipitation depth and may be appropriate for investigating other design variables, such as site sizing, particle size

distribution or roof slope. Studies collected over a single season are less valuable than those collected for longer durations, and water budgets for a single season are rarely indicative of annual performance. Monthly observations vary widely due to seasonal changes in hydrologic and plant function, temperature, and variability in precipitation and runoff events. Therefore, I recommend collecting at least a full water year of daily data to monitor hydrologic function. Green roofs, bioretention, and constructed wetlands are dynamic living systems; there is evidence that water retention and evaporation increase along with vegetation extent and density following installation, which may take more than one or two growing seasons to develop (Figure B.1). There are few studies of GI function in the literature with more than 5 years of data. It is likely inappropriate to use values collected in the first year to represent long-term performance of dynamic systems, as they require time for plant maturation, soil settling, and particle loss/accumulation to equilibrate (Figure B.1). For long-term (decadal) modelling applications, I recommend collecting continuous data for at least 5 years for water budget analyses, including the winter season—shorter duration datasets are recommended to be analyzed at the event scale, where median performance is likely more representative than mean. Fortunately, the ternary diagram is not particularly sensitive to small values changes in the dataset; measurements with the correct order of magnitude provide an adequate level of precision. As a result, users can obtain an accurate ‘sense’ of hydrologic performance by collecting ‘ballpark’ measurements, despite the field challenges of collecting complete water budgets for a site.

3.5.3 Factors affecting hydrologic performance

Using the triangle to compare dissimilar systems allows study of how individual design factors affect water budget partitioning. Although some design changes may affect all three water budget variables simultaneously, there are several design factors that primarily influence the trade-off between two water budget variables while remaining isometric in proportion to the third variable. This feature is represented visually on the Water Budget

Triangle by observations tracking parallel to one side of the triangle; compare Figure 3.6 with Figures B.2 and B.4, which show observations tracking parallel to the left side of the triangle, indicating a trade-off between infiltration and runoff (I–Q axis). For example, the most simplified systems showing infiltration–runoff (I–Q) axis tradeoffs are sewers and cisterns (Figures B.2 and B.4). Understandably, increasing the tank volume of a cistern will increase the fraction of water harvested (I for cisterns), resulting in less runoff. In the same way, increasing pipe volume in a sewershed network will increase the total capture volume, generally leading to more infiltration and decreased runoff, whereas ET remains negligible. By synthesizing the data presented above and examining engineered systems as a suite of green and grey tools, it is possible to parse which design factors affect these trade-offs between site water budget fractions when other design variables are held constant (Table 3.3).

3.5.4 Factors primarily affecting I–Q trade-off

The structural design factors that show I–Q trade-off behaviour are volume, contributing area, and the presence of an underdrain or liner. Larger tank or basin volumes increase the initial amount of water captured by a system, usually leading to higher infiltration rates and lower runoff; this is especially apparent at the event scale. Structural analogues that control capture volume include above ground empty basin volume (constructed wetlands, retention, and detention ponds); ‘ponding’ basin volume or depth (bioretention and swales); tank volume (cisterns), subsurface collection boxes, pipes or tanks (porous pavement, bioretention, infiltration trenches, tree boxes, sewers, etc.). Increasing the volume of subsurface storage increases infiltration losses but also may affect ET if the system is water-limited by inducing either standing water or an internal water storage zone. Nearly all engineered systems have design variations that include options for a subsurface drainage outlet (‘underdrain’) or an impermeable liner; systems may have one of these features, both of them, or neither. Designs that may include underdrains or subsurface

Table 3.3: Design factors affecting hydrologic performance. Design factors that primarily drive a trade-off between two water budget variables while remaining isometric in proportion to the third variable (holding all other design variables constant). Arrows represent visual direction of influence when data is plotted on a water budget triangle.

Factors affecting tradeoff between two water budget variables	I - Q axis	Q - ET axis	ET - I axis
Physical, Structural & Internal Design Factors			
	<ul style="list-style-type: none"> • System capture volume or ponding depth • Contributing catchment area • Direct connection of impervious surfaces • Presence of drain • Presence of liner 	<ul style="list-style-type: none"> • Presence of internal water storage zone or standing water • Particle size distribution • Particle surface chemistry • Media depth 	<ul style="list-style-type: none"> • Planting density & species composition • Site management practices
External, Site & Environmental Design Factors	<ul style="list-style-type: none"> • Hydraulic conductivity of sub-base • Plant Establishment • Particle clogging • Event depth & intensity 	<ul style="list-style-type: none"> • Season & temperature • Groundwater table height 	<ul style="list-style-type: none"> • Surface roughness or Initial abstraction

outlets exist for constructed wetlands, bioretention, porous pavement, green roofs, and retention and detention ponds. Liners may be present in wetlands, retention and detention ponds, porous pavement, and bioretention. Sealed pipe joints between sewer sections act as a liner analogue and affect the proportion of I-Q fractions. The presence of a drain increases runoff (through the drain), decreases stored capture, and reduces infiltration; ET from drained bioretention, constructed wetlands, and porous pavement systems remains low ($< 30\%$). The presence of an impermeable liner, geotextile, or compaction layer impedes infiltration and constrains hydrologic performance to the right-hand axis ($I = 0$; see Figure 3.5, lined bioretention cell; lined porous pavement in Figure 3.6 and green roofs in Figure 3.7). Holding other design variables constant and changing the contributing area to a system will tend to affect the balance between runoff and infiltration. Auxiliary (but

related) factors such as the land use of the contributing area or direct connection of impervious surfaces will also affect system performance along the I–Q axis. Wetlands, retention, detention, bioretention, porous pavement, and cisterns show performance changes along the I–Q axis when contributing area changes. External environmental factors affecting infiltration fraction include the hydraulic conductivity of sub-base, plant root establishment, particle clogging, compaction or cementation, event depth, and intensity.

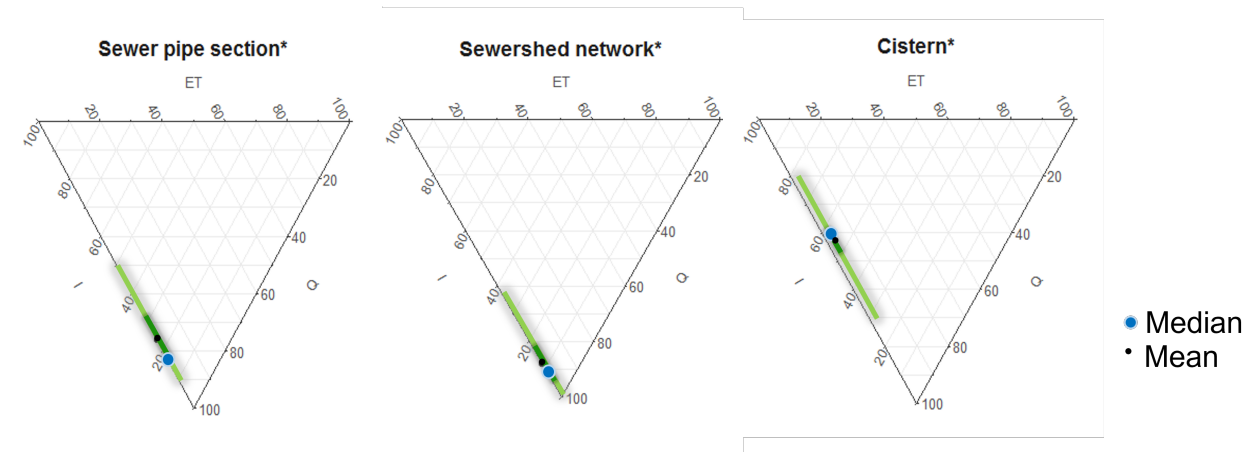


Figure 3.8: Water balance summary demonstrating I–Q axis tradeoffs. Pipe sections and cisterns are enclosed, which prevents evaporative losses. This limits the water balance shifts to the I–Q axis. Values are from Tables 3.1 and 3.2. Asterisk (*) indicates inclusion of modeled estimates.

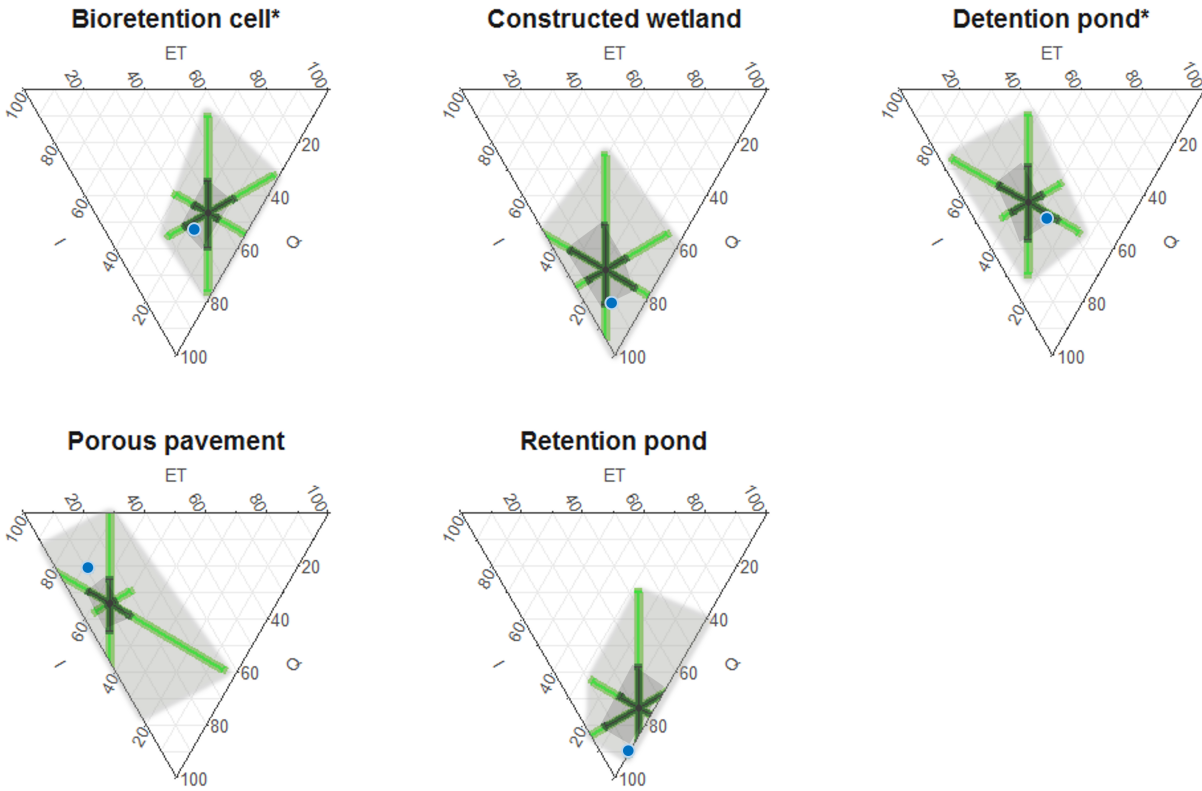


Figure 3.9: Water balance summary demonstrating primary and secondary axis tradeoffs. Water balances for several BMP structures may display both primary and secondary trade-off axes, resulting in an expected operational zone, shaded in green. The mean value is represented by the black dot at the central intersection of the star in the dark green zone, median represented in blue (see Table 3.1). The arms of the star extending out from the mean show the range of estimates from Table 3.2. The lines should be read from the axis with the vertex at 100 (e.g., vertical lines are estimates of Q). Asterisk (*) indicates inclusion of modeled estimates. The primary axis is the one with the longest arm in the star, indicating the loss pathway with the greatest variability. For example, detention ponds and porous pavement are I- Q dominant because the longest dimension of the shaded region is on the I axis, and the next longest dimension is in the Q direction. Bioretention and retention ponds are Q -ET dominant because the longest dimension is in the Q direction and the next longest is in the ET direction.

3.5.5 Factors primarily affecting Q -ET trade-off

The most simplified systems showing Q -ET axis trade-offs are green roofs and other lined systems. The structural design factors that affect Q -ET trade-off behaviour are presence of an internal water storage zone or standing water, amended media depth, and particle size

distribution and surface chemistry. Much of the soil media research that has previously been published for green roofs is relevant for ground-based infiltration systems, although weight load limitations are not. The amended media depth and particle characteristics strongly affect the volume of moisture retained at field capacity, once all ponded water has drained through soil media (~ 24 hr after the end of the rain event). Soil moisture retained within the media matrix after the system has drained to field capacity likely leaves the system through evaporative losses, not through infiltration into the sub-base. Shallower amended media depths retain less soil moisture, allow systems to become water limited and decrease the importance of ET as a loss pathway. Likewise, designs that use an upturned elbow drain or raised outlet elevation to promote internal water storage prevent systems from becoming water limited and increase ET. Preventing internal water limitation is a design consideration that is relevant for nearly all types of engineered stormwater systems to promote ET. Maintaining a small amount of residual soil moisture between events also allows for higher hydraulic conductivity at the soil surface, reduces hydrophobicity, and promotes infiltration rates at the onset of subsequent rain events. However, sites that do not effectively evaporate soil moisture between storm events remain saturated and result in increased runoff. External environmental factors affecting soil moisture and evaporative losses include season, temperature, and natural depth to groundwater onsite.

3.5.6 Factors primarily affecting ET–I trade-off

The addition of plants to an unplanted system increases loss by ET; there is a shift from left to right visible in Figure 3.6 for grassed pavers, as compared with unplanted permeable pavement. Site planting density, species, and root density also likely affect systems along the ET–I axis; this is a factor for constructed wetlands, bioretention cells, green roof, and vegetated pavers. Site management of emergent vegetation also plays a role in limiting or encouraging ET from retention ponds and constructed wetlands. Mowing frequency, surface aeration, and site management likely influence relative losses of ET versus I in

detention ponds and grassed swales. Environmental factors affecting the ET-I trade-off include the surface roughness of contributing areas or initial abstraction of precipitation.

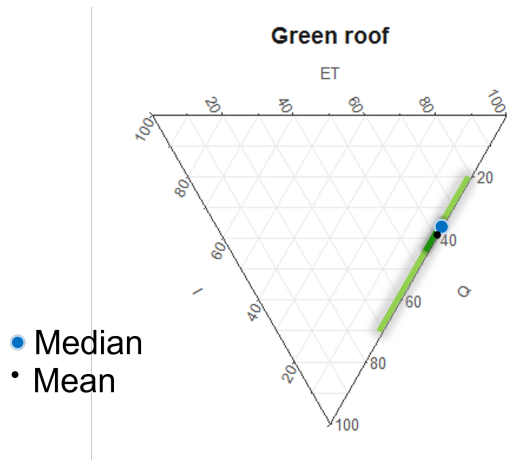


Figure 3.10: Water balance summary demonstrating Q-ET axis tradeoffs. Because infiltration is not possible, water balance ranges for green roofs shift along the Q-ET axis. See Tables 3.1 and 3.2 for values.

3.6 Conclusions

The benefits of using the Water Budget Triangle are fourfold:

1. Provides a visual aid to compare green and grey stormwater tools as an integrated suite of management options;
2. Eliminates non-technical uncertainty in language in favor of comparisons based on observable hydrologic behavior;
3. Facilitates communication of detailed technical information to both scientific and lay audiences; and
4. Illuminates how environmental factors and site design affect hydrological performance and allows simplified (two-pathway) systems to act as proxies for analysis of more complex systems.

The results of this study indicate:

1. Event-scale understanding of stormwater systems is not representative of long-term performance for GI or conventional systems; short-term monitoring underestimates ET, especially during dry and cold periods. Experimental studies for green infrastructure should collect measurements appropriate for the spatial and temporal scales of interest, and long-term modelling should not simply upscale event-scale measurements.
2. cursory estimation of water budget variables may be adequate to provide an understanding of constructed system water budgets. However, more accurate estimation of ET is necessary for both living and nonliving technologies to account for discrepancies between current model estimates and weighing lysimetry studies.
3. Constructed green infrastructure ecosystems may not adequately replace natural ones, as indicated by differences in natural and constructed wetlands;

4. Modest increases in volumetric reduction (30%) can achieve large peak flow attenuation (90%). Employing permanent reduction pathways (ET, I) instead of temporary storage (ΔS) is a more effective management strategy for mitigation of stormwater.

Future study recommendations:

1. Additional measurements of water budgets are necessary to better predict hydrologic performance of green infrastructure, especially retention and detention ponds.
2. A water year of daily data is a good starting dataset for this method. Shorter duration or intermittently collected datasets are recommended to be analyzed at the event scale.

4 Urban green infrastructure ecohydrology

4.1 Introduction




4.1.1 Mechanisms to relate stormwater control design to ecohydrology

Engineered control strategies drive efforts to improve urban stormwater management. Stormwater control measures (SCMs) are commonly included as best management practices (BMPs) for low impact development (LID). They consist of both green infrastructure (GI) and other conventional concrete structures. Sound SCM design decreases runoff inputs to sewers, wastewater facilities, and surface waters by increasing available storage and distributing flows across the landscape. There are three main pathways to distributed storage and runoff attenuation: 1) temporarily retaining water in a pond or structure (increasing storage, ΔS), and allowing it to release slowly, thereby reducing runoff (Q); 2) removing excess water from the landscape through evapotranspiration (ET) into the atmosphere, or 3) infiltration (I) through the surface into the vadose zone and/or recharge to the groundwater table (see Section 3.3 or Eger et al. [2017]). All three processes co-occur in most SCM designs. However, when considering among SCM designs, differences in the way design components affect local hydrology can result in substantially different catchment-scale performance estimates.

A set of hypotheses about analogous performance effects of structural design features is presented in Table 4.1. These hypotheses are based on work from Chapter 3, which explicated how differing designs affect site water budgets. Each row of the table aligns a design feature with a logically equivalent mechanism that affects the onsite water budget of different green infrastructure technologies. The primary axis tradeoff column denotes the two dominant water budget components most affected by each structural feature. For example, the water budget data from Chapter 3 supports the hypothesis that increasing the total capture volume of a system affects the tradeoff between runoff and infiltration, or

the I-Q axis on the water budget triangle (holding all other structural features constant). The modification of capture volume occurs slightly differently in distinct BMP types, but most systems have a type of basin or tank for which size can be adjusted. While only structural features are represented in the table, non-structural features (both known and intangible) also affect site water budget, including: the hydraulic conductivity of the native soil or underlying sub-base material (I-Q axis on the water budget triangle), the land use and curve number of the contributing watershed (I-Q), the directly connected impervious area (I-Q), the event depth (I-Q), the seasonal groundwater elevation (Q-ET), seasonal temperature and climate (Q-ET), particle clogging (I-Q) and the maturity of the system (ET-I).

Table 4.1: Design factors that influence ecohydrology.

Design factor affecting water budget		PRIMARY AXIS TRADEOFF	Constructed Wetland	Retention Pond	Detention Pond	Bioretention	Porous Pavement	Green Roof	Sewer	Cistern
(suggested analogous design factor specific to engineered stormwater system)										
Physical, Structural & Internal Design Factors	Volume		empty basin volume	empty basin volume	basin volume	ponding volume & sub-surface storage volume	sub-surface storage volume	media void volume	pipe volume	tank volume
	Contributing area	I - Q axis 	contributing area, land use	contributing area, land use	contributing area, land use	contributing area, directly connected impervious area	contributing area, directly connected impervious area	roof area	sewershed	contributing area
	Media depth		depth of media/soil to sub-base (vertical flow wetlands)	-	-	depth of media/soil to sub-base	depth of aggregate courses to sub-base	intensive vs extensive media depth	-	-
	Presence of drain		underdrain	outlet size	outlet size	underdrain	underdrain	underdrain	outfall	spigot
	Presence of liner		liner or geotextile	compaction liner or geotextile	compaction liner or geotextile	liner or geotextile	liner or geotextile	roof liner	pipe joint seal	tank
	Presence of internal water storage zone	Q - ET axis 	permanent pool depth	permanent pool depth	outlet elevation	uptumed elbow	uptumed elbow	roof drain elevation	standing water in joints	drain config.
	Particle size distribution		media size distribution	degree of compaction	degree of compaction	media size distribution	aggregate size distribution	media size distribution	-	-
	Surface chemistry		media chemistry	soil chemistry	soil chemistry	media chemistry	aggregate chemistry	media chemistry	-	-
	Planting density & species	ET - I axis 	presence & density of emergent vegetation	grass coverage, presence of emergent vegetation	grass coverage	plant species palette & coverage	presence of grassed pavers or paver spacing	plant species palette & coverage	-	-

4.1.2 Bridging model assumptions with observable field conditions

There are different approaches used to describe how to choose and design catchment appropriate green and gray stormwater infrastructure. Most publicly accessible information is based on hydrologic models (e.g., SWMM, EPA Green Infrastructure Modeling Toolkit), design handbooks (state and city stormwater management design manuals), and spreadsheet calculators, which provide rule-of-thumb guidelines to inform design (Hartigan et al. [2009], NYS DEC [2023], PA DEP [2006], Rosa et al. [2015], Rossman and US EPA [2010], US EPA [2014]). These engineering models are based on assumptions on the physics and hydraulics that occur in field settings, but they can overlook the roles of ecology and ecohydrology within a complex urban setting. For example, stormwater design models often use scaled terms to represent environmental conditions like land use, climate, or soil parameters, and they may make model adjustments by re-scaling the relevant term within the empirical or deterministic equations (Hofmann and Hofmann [1992]). While this approach is reasonable and can be effective in terms of model fit, it can obscure the underlying catchment-scale processes and make the relevant variables dimensionless (Jajarmizadeh et al. [2012]). This approach makes it difficult to assess which primary assumptions prompt a model to perform accurately or inaccurately when compared to field observations (Pons et al. [2023]).

Predicting how a particular SCM design affects local site and catchment hydrology is difficult; the SCM is a localized intervention that both affects and is affected by surrounding ecohydrology. Equally challenging is confirming that an installed SCM has met both design criteria and catchment-scale hydrology goals. The non-normal and skewed distributions of water flows that occur in observational monitoring data are not always accurately fitted by models, leading to questions about whether green technologies are performing adequately, particularly if the design was intended to meet a narrow performance benchmark. Deterministic models based on conceptual understanding of

physical phenomena may not produce performance benchmark ranges that realistically allow for the stochastic processes occurring in the field. Empirical models may allow for stochastic behavior, but obscure or misinterpret actionable engineering insights. In most engineering fields, there is better clarity about underlying causal mechanisms of actionable interventions. Additionally, the installation guidelines available for stormwater control and treatment are based on the anticipated performance of each SCM design model, rather than on field measurements that indicate how designs actually perform within the urban landscape mosaic. This subtle but observable difference may create performance expectations that do not align with the limitations posed by the capabilities of the type of green infrastructure. Understanding the functional assets and limitations of SCM features in the context of field ecohydrology supports creation of achievable performance benchmarks, more nuanced models, and appropriate remedies for designs that have failed to meet catchment goals.

Thus far, it is evident that the causal effect of a single BMP site remains unclear in the field. Therefore, the causal effect of a modeled BMP design is also unclear; structural features have yet to be decomposed into their individual causal effects. Likewise, using deterministic models (or other models that are neither stochastic nor causal in nature) to forecast site performance creates heavy reliance on black-box assumptions regarding the individual engineering and design choices of a given facility. Black-box assumptions can obscure which model components a) accurately reflect field conditions; b) are accurate but irrelevant in the magnitude of their effect; and c) are erroneous. *Answers to questions about causal effects of design choices are limited by the constraints of performing fully replicated and controlled catchment-scale experiments in urban systems.*

4.1.3 Counterfactual questions for design iteration

Empirical models can be tuned to make accurate predictions without correctly representing the causal mechanisms that underpin the system they represent. McElreath [2020]

emphasizes: ‘*Models that are causally incorrect can make better predictions than those that are causally correct.*’ This tendency toward empirical models that do not fully explain causal mechanisms occurs in hydrological engineering because it is largely impossible to fully represent the multidimensional problem space and stochastic processes in real-world ecohydrological systems. However, current models can provide good predictions of existing system performance without the ability to make counterfactual estimates (counterfactual estimates are predictions of how a BMP site would perform with a given set of potential modifications or environmental changes). Employing models that lack causal explanation limits the ability to assess how a site structural change or retrofit of a stormwater facility would affect future SCM performance, and whether the gains in onsite performance are worth the cost and effort of changing the facility. The iterative design process that is essential to good engineering practice relies on the ability to make counterfactual estimates. Therefore, development of a causally accurate model that can answer counterfactual questions falls squarely within the responsibility of the engineering community.

Each structural feature in Table 4.1 has an individual effect on the overall hydrologic performance of the stormwater facility. For example, many professionals describe a retention pond with the following features: a basin-shaped structure with permanent standing water and a single control structure at the outlet, low groundwater connectivity, grassed sides, limited or absent aquatic vegetation, and relatively steep basin slopes. These structural features are common among many retention pond sites. Each feature has individual and differing effects on the magnitude of the water budget components of the retention pond. Other SCM types, however, are more loosely defined. Some professionals use ‘bioretention’ and ‘swale’ interchangeably, whereas others discern between the two terms based on differences in structural features (planted vegetation, soil media, basin/channel shape, etc.). For detailed descriptions of each SCM type, see Chapter 2.

Engineering choices drive which combination of features and local characteristics of

these features are used in the construction of a stormwater facility during the design process. SCM terminology generally refers to a stormwater system as it was designed, as opposed to its function in the field. Questions arise regarding the combination of onsite features: if a retention pond was designed to retain permanent standing water, but rarely does, does it need intervention? What intervention is appropriate? Would it need intervention if it still met the catchment-scale hydrology goals? Should it still be called a retention pond if it no longer holds standing water? What should it be called after the intervention is retrofitted? The key to answering these questions lies in decomposing the overlapping structural features of various SCM sites to estimate the causal effect of each feature on the overall hydrologic performance of a stormwater facility.

4.1.4 Using directed acyclic graphs to clarify assumptions

The ability to estimate a potential or counterfactual outcome requires a causal framework. Systems diagrams are common in hydrology, ecology, engineering, and sustainability research –all fields trying to answer complex questions involving dynamic systems. The classic water cycle diagram and other mapped biogeochemical cycles, energy diagrams, and systems engineering models all represent relevant variables within a nodal structure with directed arrows. However, it is rare to see such diagrams in these fields analyzed using causal methods by treating the systems diagram as a directed acyclic graph (DAG, Figure 4.1) or structural causal model.¹

DAGs carry the conceptual value of a systems diagram in addition to standard mathematical notation to describe the presence of causes, effects, confounding factors, backdoor paths, instrumental variables, and other causal inference tools. The underlying nodes and edges within the diagram provide essential information about the researcher’s

¹Note, throughout the remainder of this text I try to avoid using the term ‘structural causal model’, and its common abbreviation, SCM. This is an attempt to avoid confusion, and a conscious choice to reserve the term ‘structure’ to indicate built engineering features, and the acronym ‘SCM’ to abbreviate ‘stormwater control measures.’

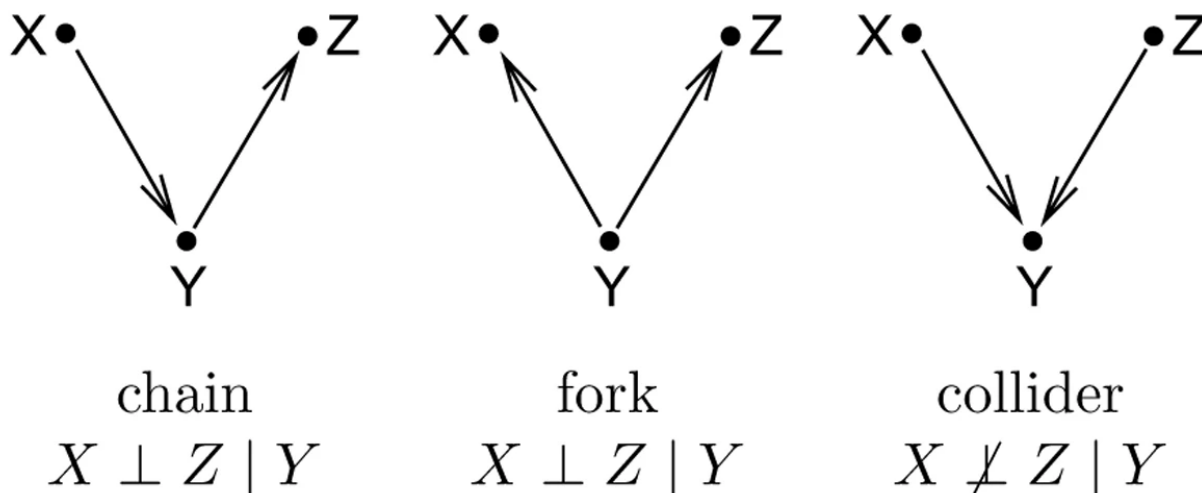


Figure 4.1: Demonstration of the three most basic DAG structures and the standard logic underlying the conditional dependency relationships between parent and child variables in each configuration. Z is independent of X conditional on Y in both chain and fork configuration of variables; Z is not independent of X conditional on Y in the collider variable configuration. Figure originally published in Markowitz and Spang [2007].

logical assumptions, including which variables must precede each other in time, as well as which variables are certain to have no causal relationship. The DAG helps identify which causal estimates are measurable (using regression or other inferential methods) and the adjustments that should be made to appropriately specify the statistical model. The hydrologic performance of any given SCM is driven by both natural and engineered factors, creating the need to differentiate between structural effects on local ecohydrology and environmental effects from the surrounding landscape or site climate. The roles of structural features were identified and isolated from site and environmental factors using a modified systems diagram (Figure C.1). Environmental variables are much more likely to affect inflows than outflows at the scale of individual storm events; these two types of effects are differentiated on the DAG in Figure C.1.

4.1.5 Use of observational data in causal inference studies

Causal inference tools are particularly appropriate for the natural observational experiments that are common in environmental fieldwork because they help identify and

remedy sources of statistical bias that may be present in a dataset and provide researchers with tools that help to discern correlation from causation. Nearly all environmental science questions rely on observational data; only a few types of studies allow true controls or formal randomization. It is impractical, costly, or impossible to design a fully blocked stormwater management experiment, largely because site locations cannot easily be randomized, and storm events cannot be replicated. Observational studies do not have the benefit of randomization or pre-selection into a study for balance, which introduces statistical bias into the underlying dataset. For example, observations of storm events at BMP sites in arid locations are ‘underrepresented’ in observational hydrologic datasets because they occur in locations where precipitation is less frequent. It seems obvious to state that it rains less in the desert, but the implication is that using a purely observational dataset for studying structural BMP features statistically biases performance results toward locations that receive more frequent precipitation events. This limitation can be corrected by choosing some level of group clustering to prevent sites with more observations (or climates with more sites) from biasing the results. This example is one of many untreated biases and confounded variables in the observational environmental data literature. Some causal relationships are much more logical or obvious than others. To give an example of the standard causal logic underlying the relationships between parent and child variables, three hydrologic examples of standard DAG formations are discussed here:

Example 1, FORKS: In a hypothetical stream gauge dataset monitoring water velocity and water temperature, it is reasonable that the two variables would display a correlation, not because they are strongly causally linked (although there could be a minor causal effect of temperature on velocity because of changes in fluid density), but because the observed measurements of these variables have a common unmeasured parent cause: surface water input. Large pulses of surface water runoff would increase the velocity at the gauge, and simultaneously shift the temperature toward the temperature of the runoff, creating, for instance, a summertime correlation between higher velocity and warmer water. This

correlation effect occurs for any two variables with a common cause (this DAG formation is called a fork, see Figure 4.1, center), and the correlational relationship can be tested for a direct, causal link between the two child variables (X , Z) by controlling for the parent variable (Y). In the example above, the parent variable (Y) is surface water input, and child variables (X , Z are temperature and velocity). Temperature is independent of velocity conditional on surface water volume.

Example 2, CHAINS: A mediating variable represents the middle of a chain of effects from one cause, through a second (mediating) cause, and onward to an outcome of interest; similar to a chain of three events or three (or more) dominoes in a chain reaction (Figure 4.1, left). The correlation between the outcome variable (Z) and the first cause in the chain of effects (X) can seem surprising or unpredictable because they are not directly causal, but only connected through the mediating variable (Y). For example, if we observe that stormwater control structures on the north side of a building have strong shading, we may also observe that the soil temperature is cooler in shaded north-side sites than in sites with full sun on the south side of the same building. Temperature is a well-known and direct cause of several relevant variables, including evaporation rate, chemical solubility, and microbial activity. We could hypothesize that despite receiving nearly identical precipitation, the sites on the south side export more nitrate in effluent than the sites on the north side. Without a causal understanding of the system, the correlation between shade and nitrate concentration might seem baffling because shade does not influence the influx of nitrogen into the system, and is not a direct cause of nitrate efflux. However, the mediating variable, temperature, directly affects the nitrogen cycle because it has an effect on the populations of nitrifying and denitrifying microbial flora, the soil moisture available in their habitat, and the rates of their nitrogen transformation activities. If, for some reason, there is uncertainty about the direct effect of shade on nitrogen efflux, the direct causal relationship can be tested by controlling for the mediating variable, temperature. When controlling for the mediator, it would quickly become clear that shade has no direct

causal influence on nitrogen concentration, and the effect is entirely mediated by temperature.

Example 3, COLLIDERS: Lastly, it is possible to mistakenly induce a spurious correlation between two variables that have no causal relationship. Take, for example, the simplified conceptual model that watershed runoff volumes are caused by both precipitation depth and area of impervious surface in each watershed. The conceptual diagram would show runoff with two parent variables (this DAG formation is called a collider, see Figure 4.1, right). If, for any reason, the runoff variable becomes stratified, the two parent variables (X , Z) will show an induced correlation. That is, for values of runoff within a narrowed window of magnitude, precipitation and area of imperviousness display a discernible pattern of relationship, and the stronger the stratification, the more discernible the correlation. The stratification of the collider variable could occur through data analysis methods, by dropping zero/low flow values and days with extreme conditions, or stratifying runoff by rough orders of magnitude. It could also occur through field monitoring limitations, wherein the monitoring station can only record within a known band of the stage-discharge relationship, ignoring other conditions. In this example, it is obvious that there is no evidence that precipitation depth in a watershed has a causal effect on the area of imperviousness, or vice versa. However, there are many other environmental variables that have less obviously defined causal relationships, where spurious relationships may appear to be legitimate results.

The DAG presented in Appendix C, Figure C.1 was developed through initial examination of the stormwater problem space. It clearly displays a variety of parent and child node relationships, and includes several nodes arranged in forks, chains and collider formations. Causal logic is subject-agnostic and based solely on the nodal relationships present in the DAG. Therefore, the formations in Figure C.1 have the same underlying inferential relationships as those described in Figure 4.1 and the examples above. There are

also additional DAG formations present in Figure C.1, including several backdoor paths and a few instrumental variables, which are discussed in brief below.

One interesting clarification that can be made using DAGs is the deconvolution of variables of interest and their related interventions (Hernán and Taubman [2008]). For example, in studying the effect of groundwater influx on green infrastructure performance, a method chosen to intervene on the groundwater variable could be the introduction of an impervious liner (Figure 4.2). The observed effect would be attributable to the intervention itself (presence of a liner), and not directly to the effect of eliminating groundwater influx. In this case the liner does prevent groundwater influx, but it also prevents infiltration of runoff into the native soil profile. Introduction of the liner variable in the study opens a backdoor path between groundwater influx and infiltration. Since these are two opposing effects in the water balance of the structure, and since they are generally unobservable variables without specialized equipment, it is better to change the frame of the question to focus on the engineering problem, which is the effect of the liner itself. Another option is to collapse the question to use the liner as an intervention on groundwater connection as a whole effect (both influx and infiltration), in which case, the liner variable becomes an instrumental variable for the effect of this connection on runoff.

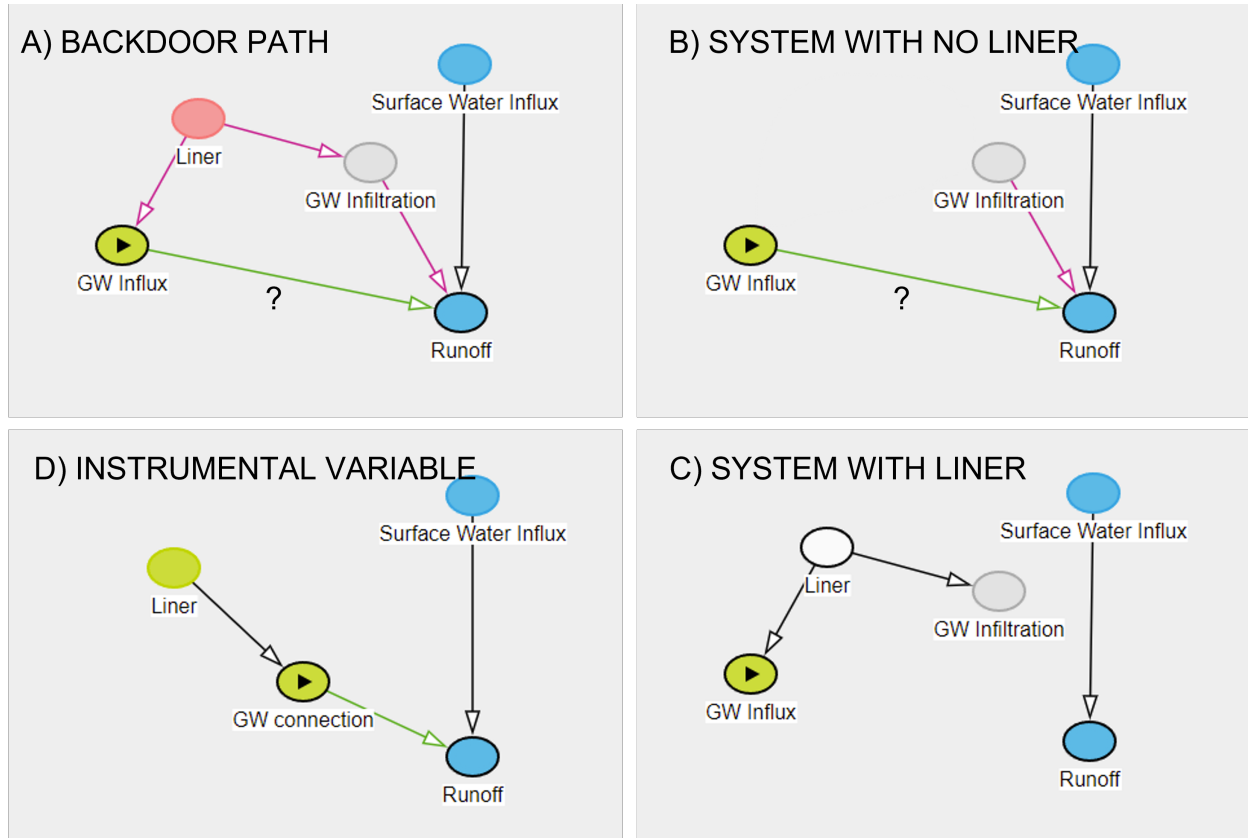


Figure 4.2: The inclusion of the liner intervention variable opens a backdoor path between groundwater influx and runoff (Box A, in pink). Additionally, because of the nature of the binary variable ‘Liner’, the flow of causal information in a system with no liner (Box B) is not equivalent to the causal information present in the model of a system with a liner (Box C). Therefore, the effect measured is attributable to the presence of a liner, but not to the effect of interest (green line). As a result, the liner is not a suitable intervention for testing the effect of groundwater influx on runoff. However, the liner works as an instrumental variable for estimating the causal effect of a two-directional groundwater connection (Box D).

4.2 Data and methods

4.2.1 Dataset and imputation methods

The International BMP Database (BMPDB) contains a large compilation of SCM case studies with information about structural design, hydrologic activity, and water quality performance for more than three dozen BMP types (WRF [2019]). The cataloged information includes observations of precipitation, surface inflow volume, surface runoff volume, categorical BMP Type, and onsite design information at more than 700 sites in North America, Asia and Oceania (the majority are in the US). Measurements from the database were matched to observe the precipitation depth (cm) and cumulative inflow and outflow volumes (L) for 10,011 distinct rainfall events at 370 sites. Using documentation in the database, site technical reports and virtual site visits, observations for 18 structural feature variables were collected for each of the sites, including: BMP Type, watershed area, site surface area, presence of liner, vegetation, amended media (and depth), exposure to air, internal water zone, permanent ponding, flow path, overflow path, irrigation, basin shape, underdrain, and groundwater connection, slope, and hydrologic surface fluxing depth. These structural features were chosen to standardize and facilitate evaluation of the hydrologic function of the SCM facilities in the database; the choice of variables was informed by the observations present in Table 4.1, Figure C.1 and contextual field data that existed in the BMPDB tables for each BMP type.

The raw BMPDB dataset is very broad, but it is missing many unobserved values. Code for readying the 2019 version of the BMP database for statistical inference is available at www.github.com/cgeger/clary. This code modifies a flat file version of the original 2019 MS Access database; out of respect for the publishers of the original dataset, the extracted version of the data itself is not redistributed on GitHub. Records from the 2019 version of the BMP database were cleaned, converted to common units, and matched across site for statistical analysis. Appendix A.2 contains a textual description of the data

cleaning process for each of the tables used. After all the cleaning steps, the BMPDB dataset contained 625 SCM sites in 13 BMP categories and 32 control structures. From these sites, 11,260 precipitation events were identified at 374 stations, which covered 375 individual SCMs (not including control and reference sites). There were 144 SCMs in the initial dataset without observed precipitation data. From among the sites with hydrologic data, 543 SCMs had observed contributing watershed areas, and 60 SCMs did not have observations of watershed area. Clearly, even after data cleaning, a considerable number of missing observations remained in this dataset (missingness). The counts of gaging stations at sites and the number of events with raw and cleaned data (all BMP site types) are summarized in Table 4.2. Although some sites were eliminated because they had no observed flow events, special care was taken to try to maintain as many of the partially observed sites as possible, in order to maintain common support and statistical power across the multidimensional problem space (Honaker and King [2010], Jakobsen et al. [2017]). Where appropriate, reference values from paired reference site stations within the BMPDB were used as a proxy for influent or effluent flow volumes. After cleaning and matching, there was approximately 12% missingness of individual observations of event hydrological variables (Inflow, Outflow, Precipitation depth), with 7113 complete cases out of 10375 remaining observations. A complete case has all experimental measurements observed, an incomplete case has one or more missing values.

After identifying sites with hydrologic measurements, each of the remaining 387 BMP sites was classified by its structural features. The structural features for each site were encoded from narrative and descriptive information contained in the design table of the BMPDB and from mapped observations on the BMP database mapping tool, Google Maps and Google Earth (WRF, EWRI, US EPA, US DOT, Geosyntec Consultants, Wright Water Engineers [2023], Google [2023a,b]), as well as white paper references available through internet search (see Appendix A.3 for reference list of technical resources used for these virtual site visits). The BMP database mapping tool and dataset acted as the

starting point for locating and observing most sites, but some sites had little or incomplete structural information recorded in the BMPDB. At sites with scanty structural information, categorical variables were assigned based on best-known information according to BMP Type or left blank for imputation. A few sites were eliminated entirely due to missing contextual information. After data collection, there was approximately 1% missingness of site feature variables, with 324 complete cases out of 381 sites. After merging the site information with the event observations, the dataset had just 4% missingness, but only 6,090 complete cases out of 10,375. There is evidence that dropping results that are missing at random by selecting only complete cases biases results and reduces statistical power (Honaker and King [2010], Jamshidian and Mata [2007], Madley-Dowd et al. [2019]). To avoid this problem, the missing data were imputed 5 times using the Amelia package for multiple imputation (Honaker et al. [2010, 2011]). Multiple imputation estimates the missing cells in an incomplete dataset and creates m complete versions of the original dataset for analysis. After multiple imputation, the m imputed datasets are analyzed in parallel, producing m estimates. The m estimated values from analyzing each dataset are pooled into a single final estimate using Rubin’s rules (Heymans and Eekhout [2019]). The Amelia package is R software that uses the three-step iterative EMB algorithm for multiple imputation; each iteration computes these steps in order: **Expectation - Maximization - Bootstrapping**. The first two steps (EM) represent the classical iterative computational approach to finding the local maximum likelihood of a distribution (Dempster et al. [1977]). The EMB algorithm combines the classic EM algorithm with a bootstrap approach to take draws from the distribution. For each of n draws, the data is bootstrapped to “simulate estimation uncertainty and then run the EM algorithm to find the mode of the posterior for the bootstrapped data” (quote from Honaker et al. [2022], see Honaker et al. [2010] for details of the EMB algorithm). Imputation for each variable was bounded by the minimum and maximum observed values, and the hydrologic and area variables were log-transformed to normal distributions. After

imputation, the final model-ready dataset contained 10,011 observations at 370 sites.

Table 4.2: Station and record counts used for this study from the BMPDB.

Station Type	# stations	# raw records	# stns w/ clean events	# clean events
Rain Gage	412	11543*	374	11260
Inflow	404	10247	324	8101
Inflow Estimate		32		
Outflow	443	11766	345	8825
Subsurface	6	513	3	111
Bypass	7	45	26	459
Reference Outflow	21	2687**	18	644

4.2.2 Analytical methods

To mathematically treat causal inference for observational datasets, Runge [2022] suggests the following steps:

1. Formalize the underlying structural causal model;
2. Define the causal effect(s) of interest and hypothetical interventions; and,
3. Establish criteria to decide if and how causal effects can be estimated *post-hoc* from data alone, without experimental intervention onsite in the field, randomization or other study design changes (US Clivar [2021]).

Following these guidelines, the first step in this analysis was a detailed examination of the stormwater ecohydrology problem space, represented in a generalized DAG presented in Figure C.1 and discussed in Appendix C. The outflow runoff volume (effluent) is the effect of interest (dependent variable), and the interventions (independent variables) are individual structural treatment features, such as planted vegetation, soil media amendments, and other design features. The site location is an important confounding variable that defines which features are present at each observed location, but it opens a back-door path that biases estimation of the unique effect of each treatment feature on the

hydrologic performance of the site (see Figure 4.3, Box A). As an example, consider a single retention pond: it is impossible to attribute the runoff attenuation of the pond to the presence of vegetation or standing open water surface because the observed features overlap at that location, producing a combined effect signal in the dataset. Stratification (on inflow volume) and matching (on site features) were both identified as potential options for blocking the confounding backdoor path to measure the marginal causal effects of the treatment features on runoff volume (see Figure 4.3, Boxes B and C).

Five sets of generalized linear mixed models were identified to try to estimate the causal effects of different types of structural features on effluent stormwater volume; two models relied on only stratification, and three models also relied on matching. All five sets employed generalized linear mixed models run using the R package `glmmTMB`, a generalized linear mixed modeling package (Brooks et al. [2017]). Two different multi-level model sets were specified by stratifying on site (see Figure 4.3, box B). The backdoor path is blocked by controlling for site and the resulting estimates represent the effect of the combination of features at each site. The linear mixed (or multi-level) model approach utilizes all the observations available in the dataset and does not require observations from the same site to be combined into a summary statistic. However, to reduce oversampling bias, the linear mixed model sets the group level variable (in this case, individual BMP site) as a random effect and groups measurements from the same location. The effect of inflow was also allowed to vary on its own random slope at each site, with sites having correlated intercepts (Bolker et al. [2009], see sections below for more explanation for individual model specifications). This model specification allows for the estimation of structural effects while controlling for the bias associated with individual effects associated with sites and variable event inflow volumes.

Model Set 1 (BMP Type): The first set of models focuses on identifying effects of categorical BMP site types, not individual structural features. This stratified model

approach assumes similar structural features are present at each BMP Type. Therefore, the causal effects are associated with the labeled BMP Types from the BMP Database, which represent a combined observation of standard structural features across all sites of the same BMP type (see Chapter 2 for standard structural features). The DAG for this model does not have an open backdoor path because each site falls into only one BMP Type category, and the categories are assumed to have identical feature sets (this is a simplifying assumption for Model Set 1). However, it is still appropriate to control for the individual site by adding it as a random effect because there are multiple observations nested at each site—the sites do not experience identical storm events and the sampling is unbalanced (different numbers of events were observed at each site). Marginal effect estimates of the order of magnitude of each BMP Type were made using both raw, log-transformed outflow and area-normalized log-transformed outflow as the outcome variables. The marginal effect estimates were estimated at the mean. The area-normalized model was also used to generate runoff predictions for a notional 1-cm storm event at each site type, which is near, but not at the mean. Both methods were run on each of the 5 imputed datasets, and the model effect results were pooled using Rubin’s rules for mixed models in the mice and broom.mixed packages (Bolker 2021, Brooks et al. [2017], Bolker et al. [2009], Heiss [2018], van Buuren and Groothuis-Oudshoorn [2011], van Buuren [2018]). The predictions for each imputation are reported unpooled to show the variation within and between the prediction ranges. Model Set 1 was represented by the following glmmTMB specification in R:

```
glmmTMB(Outflow ~ 0 + BMPType + (1 + Inflow|Site) + offset(Inflow) )
```

Model Set 2 (Structural Features): The second model set is identical to Model Set 1, except that in Model Set 2 the marginal effects of structural features are examined individually. For this model, it is assumed that there is no causal interaction process between unique structural features (this is the simplifying assumption for Model Set 2). For this assumption to hold, variables that were regarded as similar or causally related

were eliminated so that the model would avoid causally-obvious cases of multicollinearity and indirect effects. Therefore, the `Liner` variable was included in the regression model in lieu of the `Groundwater Connection` variable, `Irrigation` was excluded because it correlated strongly with `Vegetation`, and `Internal Water Storage Zone` was included as a category within the `Underdrain` type variable. The DAG for this model has an open backdoor path, which is blocked by controlling for the site as a random factor. Once the regressors had been selected, all were encoded as categorical factors, with the reference structure representing a pipe (lined, unplanted, channel-shaped, enclosed structure that receives stormwater from a pointflow source with no bypass structure, no media amendments present, no internal sieves, screens or baffles, no permanent or intermittent ponding, and drainage at a surface outfall). These reference features were chosen to characterize the kinds of pipes and ditches that are most closely associated with urban drainage systems as described by Kaushal and Belt [2012]. As above, the model was used to generate marginal effect estimates of the order of magnitude of the effect of each structural feature, and runoff predictions for a notional 1-cm storm event based on area-normalized effect estimates. The model was run on each of the 5 imputed datasets and the results were pooled using the `mice` package to obtain a pooled effect estimate. Predictions for a notional 1-cm storm event were also estimated. The second model identifies all individual structural features in the `glmmTMB` specification:

```
glmmTMB(Outflow ~ 0 + Liner + VegetationType + AmendedMediaType +
AirExposure + UnderdrainType + PondingType + InflowSourcePath +
BypassRouting + BasinShape + (1 + Inflow|Site) + offset(Inflow) )
```

The last three sets of models examine more carefully specified DAGs to measure the causal effects of three individual structural features, the effect of amended media, vegetation, and surface ponding in accordance with their underlying causal processes. These models acknowledge causal interactions between structural features of interest. The

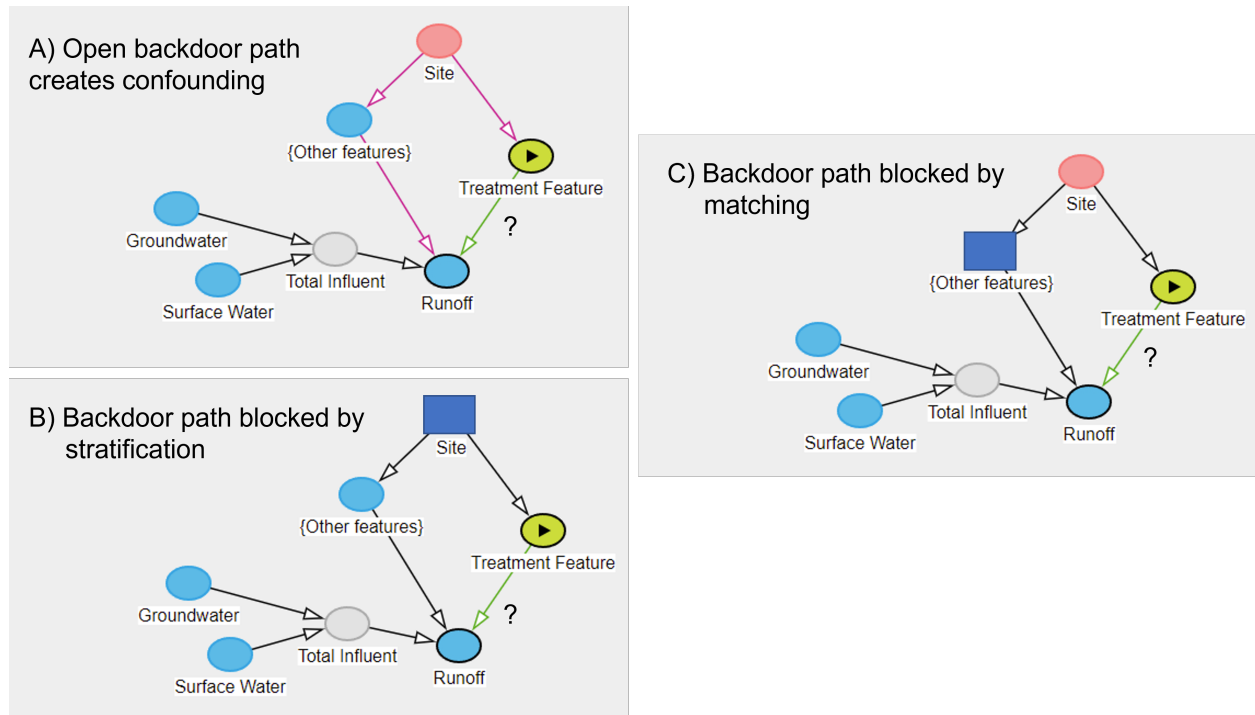


Figure 4.3: Box A displays the open backdoor path (pink lines) that confound the estimation of the causal effect of the treatment feature on runoff (green line). Box B displays the backdoor path blocked by controlling for individual site through stratification. Box C displays the same backdoor path blocked by controlling for the other relevant features through matching.

models use matching to balance the data and counteract a confounding backdoor path between the matched variable and the treatment variable. From Mansournia et al. [2013]: “Balanced matching forces the distribution of the matching factors to be identical across groups of individuals.” Under this approach the data from the International BMP database are treated as repeated measures in a quasi-experimental design to estimate effects of known structural features. Müller and Levy [2019] describe this type of natural experiment: ‘specific attributes of observation[al] units allows their *ex post* assignment to appropriately randomized control and treatment groups.’ In this case, the *ex post* assignment to treatment or control group is based on structural features present at the site, such as presence of standing water, planted vegetation and other features regarding the design of the structure. Matching involves de-biasing the dataset by selecting one control observation for each observation in the treatment subset. Ideally, the control observation has all the same observed characteristics (other relevant structural features) as those recorded for the treatment observation, except on the treatment variable. To reduce dimensionality, only the features that interact as confounders of the treatment estimate are used for matching; trying to match on every structural feature increases the dimensional requirements of the problem beyond the capacity of the observed dataset, and the treatment and control groups lose common support. In practice, an exact match is sometimes not possible, but the dataset can still be trimmed for balance in the non-treatment variables by using nearest neighbor matching, propensity score weighting, or ‘coarsened exact’ matching (Ho et al. [2011]). The information recorded for each SCM site allows coding the dataset by the presence or absence of many types of structural features that affect the individual performance of that site, as well as the range of environmental conditions that influence the site.

Model Set 3 (Amended Media): The DAG for this model is specified in Figure C.2, which indicates that the media amendment (**M**) is not directly affected by many other site structural features, but that it is still affected by site location (**Site**). One way that it

is specifically affected is that sites with large areas are less likely to have a media amendment applied. The liner (L) affects the depth of the soil media profile that is available for infiltration. Therefore, the sites were matched using coarsened exact matching on (**Area**) and on presence/absence of a liner. Coarsened exact matching cuts continuous variables into strata (e.g. quintiles) and matches treatment and control observations within the same stratum. The matched set from each imputation contained approximately 4500 observations of control-treatment pairs. The coarsened exact match estimate was also compared to a rough match between bioretention cells and detention basins. This comparison was done using the same Model Set 3 specification and by subsetting the dataset to include only bioretention cells and detention basin BMP Types. Bioretention and detention ponds have generally similar shapes and sizes but differ as bioretention cells have media applied to the surface (a good rough match for the estimate of media). A third variation of this model examined the effect of amended media depth, using the same matched dataset and model specification, except that the amended media variable was continuous instead of binary. Predictions and pooled effect sizes were estimated as described in Model Sets 1 and 2. After coarsened exact matching on area and presence/absence of a liner, the specified multilevel regression model in glmmTMB was:

```
glmmTMB( Outflow ~ 0 + Liner + AmendedMedia + (1 + Inflow|Site) +
offset(Inflow) )
```

Model Set 4 (Surface Ponding): The DAG for this model is specified in Figure C.3. The degree of ponding is directly affected by the presence of an underdrain (or other drain type) and a liner, as well as other site effects and total influent. The MatchIt package uses a logit regression to match treatment and control observations, which means that the treatment variable must be binary for it to match data properly (Ho et al. [2011]). The ponding variable (P) was originally encoded as a multi-level variable (no ponding, intermittent surface ponding, permanent ponding), so the variable was re-coded as binary

(0: no ponding, 1: any ponding) for the purposes of matching only. The dataset was then matched using coarsened exact matching on Liner (L), Drain type (D), and area-normalized log-transformed surface water inflow (SW). The regression model was run on the matched dataset using the original coding, so that the marginal effects of surface ponding could be estimated. The matched dataset was used to run the specified model:

```
glmmTMB( Outflow ~ 0 + Liner + Ponding + DrainType + (1 + Inflow | Site)
+ offset(Inflow) )
```

Model Set 5 (Vegetation): The DAG for this model is specified in Figure C.4, which indicates there are several auxiliary features that affect both vegetation (V, in green) and runoff (Q, in blue). These features include: media amendment (M), surface ponding (P), irrigation (I) and additional (Site) effects. There were only 9 sites with irrigation and 7 sites with surface water aeration (e.g., pond fountains). These sites did not provide enough common support to match and analyze alongside the other two confounding factors, so these sites and their events were eliminated from the analysis, leaving 354 sites. As with the Ponding model (Model Set 4), the multi-level vegetation variable (unplanted, grass and sedum, trees and shrubs) was recoded as binary (0: unplanted, 1: vegetated) for the purposes of matching. Based on the results of Model Sets 3 and 4, the ponding and media variables were also recoded as binary to simplify the matching and model adjustments. Ponding was binarized to 1: permanent standing water and 0: intermittent or absent surface ponding. Media amendment was binarized to 1: media amended, 0: baffles/screens or no media at all. Therefore, the dataset was matched based on binarized ponding, media and area-normalized log inflow using coarsened exact matching. The matched dataset was used to run the specified regression model:

```
glmmTMB( Outflow ~ 0 + Vegetation + Liner + Ponding + Media + (1 +
Inflow | Site) + offset(Inflow) )
```

4.3 Results

The distribution of 10,375 precipitation events in the clean BMPDB dataset is presented in Figure 4.4, and the clean counts and pre-imputation distributions of inflow and outflow volumes for each of these events are displayed by BMP type in Table A.1 and Figure A.3. A post-imputation distribution summary is shown in Table A.2. Note that some BMP types show a clear pattern of generating zero-discharge events, particularly bioretention (BR) and green roofs (GR). This is a desirable outcome, so these observations were not dropped from the dataset, but they make model fitting complicated. Also, note that on an event basis, some BMP types are generally designed to receive more inflow than others. The sites with the highest mean inflow volumes per event are underground concrete detention structures (DU: 5,000,000 L), layered filtration systems (FL: 1,700,000 L), retention ponds (RP: 1,200,000 L) and wetland channels (WC: 2,800,000 L). The BMPs that receive the smallest mean volumes of inflow per event are bioretention (BR: 4,200 L), other swale-type structures (BW: 2,200 L), and green roofs (GR: 3,300 L).

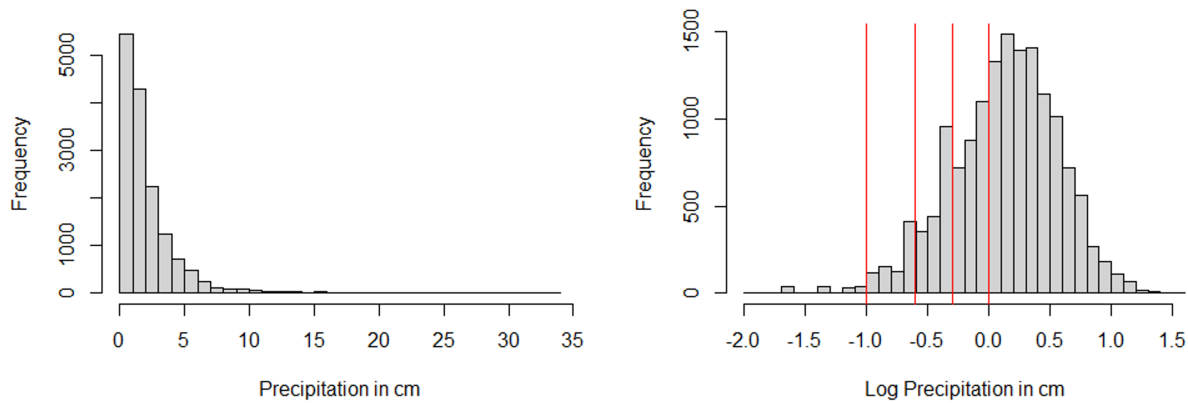


Figure 4.4: Distribution of precipitation depth for individual events in the BMPDB. Red lines indicate possible breakpoints in the dataset due to measurement rounding and detection limits.

4.3.1 Model Set 1: BMP type effects

Model Set 1 represents the most general model, which assumes a linear fit of log-transformed flow data with correlated group intercepts and random slopes associated with inflow across individual BMP sites. This specification was chosen according to suggestions in Barr et al. [2013] and Bolker et al. [2009], and also coincided as the model specification that presented the lowest AIC and BIC values compared with several other specifications. AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are heuristics for model selection, that is, choosing which parameters to use as variables in the regression model.

In the model specified for Model Set 1, the zero term ‘0’ sets the global y-intercept to zero, and the `offset(Inflow)` term allows measurement of the regression slope against the 1:1 (Inflow:Outflow) line (as opposed to the horizontal axis). This model is the most general because it assumes that all BMP sites with the same BMP type represent a common set of structural features, including groundwater connection. This representation is an oversimplification of the actual dataset, but remains useful because it provides an estimate of the causal effects of groups of features that commonly co-occur at the range of BMP sizes that are observed in the field (see Table A.1 and Chapter 2 for descriptions). The `(1 + Inflow|BMPID)` term groups multiple observations at the same site and controls for the effect of each individual BMPID by letting its slope vary randomly with log-transformed Inflow. The units of marginal effect reported in Table 4.3 for each structural type are expressed as orders of magnitude (OM) change, meaning that for a BMP type with an effect size of -0.55, a BMP site of that type would be expected to reduce inflow by 0.55 orders of magnitude. For an event with 10,000 L (10^4 L) of inflow, the site would be expected to discharge $\sim 2,818$ L ($10^{3.45}$ L) of outflow, which is equivalent to 72% runoff reduction (RR). These estimates represent the order of magnitude change effect of the average site with that BMP type classification, which accounts for random effect of

inflow volume at the site but does not explicitly account for site area. However, it is more practical to have the area-normalized effects, so that predictions can be made to estimate the runoff per square meter. Therefore, the same model specification was also run using area-normalized data, which resulted in similar marginal effect outcomes. The area-normalized marginal effects and equivalent runoff reduction values are represented in Figure G.2 and reported in Table 4.4; predictions are reported in Figure 4.5.

The Model Set 1 results, from the area-normalized model, indicate the BMP types with the strongest runoff reduction (RR) effects are permeable pavements (-1.18 to -1.7 orders of magnitude; 93-98% RR), bioretention (-1.44 orders of magnitude, 96% RR), green roofs (-0.68 orders of magnitude, 79% RR), grass strips and swales (-0.36 to -0.39 orders of magnitude, 56-60% RR), and detention basins (-0.28 orders of magnitude, 47% RR). The model estimates that these structure types generate outflow volumes that are statistically less than the inflow volume (represented by the dashed line at 0 orders of magnitude, Figure G.2). The other effect types have discharge volumes that are not statistically different than the inflow volume. Several types of BMPs have pooled standard errors that range above the 0 line, indicating that there is an expectation for these types to discharge more outflow than inflow during some events, even when accounting for habitually poor performance at some individual sites (this is accounted for in the random effect of site within the mixed model). A few site types had mixed estimates, notably, retention ponds show a statistically slight reduction in outflow on an area-normalized basis, but when analyzed at the site level, the average retention pond site does not have a runoff reduction effect that is statistically different than inflow. This pattern is also the case for high-rate media filters, underground detention tanks and pipes, and media filters with unspecified media types (not sand, peat, layers of media or geotextiles).

There are a few types of BMPs that were somewhat poorly represented in the dataset, which also have wide ranges due to large amounts of uncertainty both among and within

sites. Four are structures that are expected to have minimal runoff reduction effects due to their design characteristics: concrete-lined detention basins at the surface, volume control/attenuators, hydrodynamic swirl concentrators, and catch basin inserts. These BMP types were included as important control structures to ensure that the model was performing as expected, and indeed they all have estimates very close to 0 orders of magnitude (0% RR). The diversity of structures represented is important in the subsequent models because they provide important edge cases for treatment and control feature matching (this is called common support). The remaining structures that have relatively poor estimates are, wetland vegetation biofilter/swales, wetland channels, peat and sand media filters, and vertical filters. The difficulty associated with obtaining stable estimates for these BMP types is: 1) that they are characterized by heterogeneous designs and 2) there are a limited number of sites represented in the dataset; meaning 3) none of the individual sites have the same sets of identical structural features.

The model fits for Model Set 1 show that this linear mixed model approach does not fit particularly well for this dataset; it displays non-linearity, over-dispersion, and outliers (analyzed using Lüdecke et al. [2021]). One of the fundamental assumptions of a linear mixed model is that the means of the random groups must be normally distributed—this assumption does not hold at the distribution tails. There are a couple of reasons for this characteristic: the glmmTMB diagnostic indicates that the model has unusually large z-statistics for some BMP types. Paraphrasing the glmmTMB error description—this is primarily due to a large number of zeroes at the low end of the predictor scale, which occurs when a site has a zero-discharge event (`Inflow` > 0 but `Outflow` = 0).

Table 4.3: Pooled Model Set 1 mean site effect estimates and std errors for each BMP type, units are in orders of magnitude. Pooled T-statistic and degrees of freedom for pooled calculation are also reported. Estimates that are statistically significant are marked with ‘***’ ($p < 0.01$) or ‘*’ ($p < 0.05$) depending on their level of significance, poor estimates are marked with ‘—’ in red based on unusually large Z-statistics ($|x| > 5$), which may indicate that the parameters are near the edge of the range and the (site-specific) random effects are near zero for the sites represented. Equivalent runoff reduction estimates for average installations of each BMP type are also reported in the RR and RR range columns; values with better statistical confidence are in bold, those with less statistical confidence are labeled in gray. Positive RR values indicate the site discharges less runoff than inflow, negative values indicate discharge may exceed inflow volume.

BMP Type	n sites	n obs	estimate	std error	t statistic	df	p value	RR	RR range	Description of structure
BI	43	750	-0.42	0.14	-2.88	1226	0.004 ***	62%	46% - 73%	Biofilter - grass strip
BR	47	3253	-1.49	0.13	-11.30	3471	0.000 ***	97%	96% - 98%	Bioretention
BS	25	313	-0.41	0.19	-2.18	1087	0.029 *	61%	40% - 75%	Biofilter - grass swale
BW	1	21	-0.15	0.92	-0.17	1347	0.867 --	30%	-488% - 92%	Biofilter - wetland vegetation swale
CBI	11	114	-0.03	0.26	-0.11	9619	0.916 ----	6%	-72% - 49%	Catch basin insert
DB	29	391	-0.20	0.17	-1.19	7675	0.235	37%	7% - 57%	Detention basin - open surface grass-lined basin (dry)
DC	4	74	-0.04	0.45	-0.09	6161	0.925 ----	9%	-156% - 68%	Detention basin - open surface concrete or lined tank/basin (dry)
DU	2	23	-0.53	0.71	-0.76	9976	0.449	71%	-48% - 94%	Detention - underground vault, tank or pipe(s) (dry)
FL	3	44	0.07	0.56	0.12	206	0.905 --	-17%	-328% - 68%	Media filter - combination of media or layered media
FO	4	81	-1.53	0.48	-3.18	7229	0.001 ***	97%	91% - 99%	Media filter - other media types
FP	2	18	0.00	0.62	0.00	9976	1.000 ----	0%	-315% - 76%	Media filter - peat mixed with sand
FS	23	392	-0.06	0.19	-0.31	7934	0.758	13%	-35% - 43%	Media filter - sand
FV	1	24	0.00	0.87	0.00	9976	1.000 ----	0%	-640% - 86%	Vertical filter - geotextile fabric membrane
GR	15	613	-0.44	0.27	-1.64	5522	0.101	63%	32% - 80%	Green roof
HDS	19	350	0.02	0.22	0.11	1049	0.911 --	-6%	-74% - 36%	Hydrodynamic swirl concentrator or separation system
HRBF	3	74	-0.20	0.51	-0.40	9844	0.687	38%	-100% - 81%	High rate biofiltration
HRMF	15	363	-0.30	0.23	-1.27	2197	0.204	49%	13% - 70%	High rate media filtration
IB	3	125	-0.83	0.71	-1.18	9971	0.239	85%	25% - 97%	Infiltration basin
MCTT	2	21	-0.38	0.71	-0.54	467	0.591	59%	-113% - 92%	Multi-chambered treatment train
OGS	10	167	-0.23	0.30	-0.75	560	0.451	41%	-18% - 70%	Oil/grit separators and baffle boxes
PC	5	258	-1.09	0.39	-2.81	8796	0.005 ***	92%	80% - 97%	Porous pavement - pervious concrete
PF	2	38	-1.70	0.61	-2.80	9873	0.005 ***	98%	92% - 100%	Permeable friction course pavement
PM	11	480	-1.12	0.27	-4.16	5135	0.000 ***	92%	86% - 96%	Porous pavement - modular blocks
RP	58	1053	0.01	0.14	0.04	392	0.969 --	-1%	-40% - 27%	Retention pond - Open surface pond with a permanent pool (wet)
RV	3	108	-0.34	0.49	-0.69	9912	0.490	54%	-42% - 85%	Retention underground vault or pipes (wet)
VC	3	31	0.00	0.51	0.00	9976	0.997 ----	0%	-221% - 69%	Volume control/attenuation structures
WB	20	713	-0.10	0.21	-0.46	1521	0.648	20%	-30% - 51%	Wetland - basin with open water surfaces
WC	6	119	0.09	0.40	0.22	4903	0.825 --	-23%	-209% - 51%	Wetland - channel with wetland bottom

4.3.2 Model Set 2: Structural feature effects

Model Set 2 uses the same specification as Model Set 1, but doesn’t consider the BMPDB BMP type classification, and instead uses the individual structural features that were mapped for each of the 370 sites as covariates. The structural features were mapped as factors (mostly binary, but a few categorical), and the reference conditions are those consistent with an urban storm drain. The effects of the reference and the treatment conditions are indicated in Table G.1. The features with the greatest runoff reduction effects are shrub/tree vegetation (-0.84 orders of magnitude, 98% RR), the presence of an up-turned elbow or internal water storage

Table 4.4: Pooled Model Set 1 area-normalized effect estimates and std errors for each BMP type, units are in orders of magnitude. T-statistic and degrees of freedom for pooled calculation are also reported. Estimates that are statistically significant are marked with ‘***’ ($p < 0.01$) or ‘*’ ($p < 0.05$) depending on their level of significance, poor estimates are marked with ‘-’ in red based on unusually large Z-statistics ($|x| > 5$), which may indicate that the parameters are near the edge of the range and the (site-specific) random effects are near zero for the sites represented. Equivalent runoff reduction estimates for each BMP type on an area normalized basis are also reported in the RR and RR range columns; values with better statistical confidence are in bold, those with less statistical confidence are labeled in gray. Positive RR values indicate the site discharges less runoff than inflow, negative values indicate discharge may exceed inflow volume.

BMP Type	n sites	n obs	estimate	std error	t statistic	df	p value	RR	RR range	Description of structure
BI	43	750	-0.36	0.11	-3.13	3651	0.002 ***	56%	43% - 66%	Biofilter - grass strip
BR	47	3253	-1.44	0.11	-13.12	3700	0.000 ***	96%	95% - 97%	Bioretention
BS	25	313	-0.39	0.17	-2.28	492	0.023 *	60%	40% - 73%	Biofilter - grass swale
BW	1	21	-0.17	0.80	-0.21	659	0.835 ---	32%	-332% - 89%	Biofilter - wetland vegetation swale
CBI	11	114	-0.01	0.23	-0.06	9201	0.948 ----	3%	-62% - 42%	Catch basin insert
DB	29	391	-0.28	0.14	-2.01	9709	0.044 *	47%	28% - 62%	Detention basin - open surface grass-lined basin (dry)
DC	4	74	-0.25	0.37	-0.70	3494	0.487	44%	-29% - 76%	Detention basin - open surface concrete or lined tank/basin (dry)
DU	2	23	-0.72	0.52	-1.38	9976	0.167	81%	37% - 94%	Detention - underground vault, tank or pipe(s) (dry)
FL	3	44	-0.34	0.46	-0.73	118	0.464	54%	-33% - 84%	Media filter - combination of media or layered media
FO	4	81	-0.64	0.38	-1.70	9742	0.089	77%	46% - 90%	Media filter - other media types
FP	2	18	0.00	0.52	0.00	9976	1.000 ----	0%	-229% - 70%	Media filter - peat mixed with sand
FS	23	392	-0.09	0.15	-0.59	6954	0.558	19%	-16% - 43%	Media filter - sand
FV	1	24	0.00	0.74	0.00	9976	1.000 ----	0%	-446% - 82%	Vertical filter - geotextile fabric membrane
GR	15	613	-0.68	0.20	-3.40	8916	0.001 ***	79%	67% - 87%	Green roof
HDS	19	350	0.00	0.18	0.01	3581	0.994 ----	0%	-51% - 33%	Hydrodynamic swirl concentrator or separation system
HRBF	3	74	-0.19	0.41	-0.47	9967	0.637	36%	-65% - 75%	High rate biofiltration
HRMF	15	363	-0.28	0.19	-1.46	7430	0.146	48%	18% - 67%	High rate media filtration
IB	3	125	-0.64	0.65	-0.99	9976	0.320	77%	-1% - 95%	Infiltration basin
MCTT	2	21	-0.53	0.64	-0.83	84	0.409	70%	-28% - 93%	Multi-chambered treatment train
OGS	10	167	-0.18	0.25	-0.73	509	0.464	35%	-17% - 63%	Oil/grit separators and baffle boxes
PC	5	258	-1.18	0.32	-3.71	9801	0.000 ***	93%	86% - 97%	Porous pavement - pervious concrete
PF	2	38	-1.70	0.51	-3.36	9947	0.001 ***	98%	94% - 99%	Permeable friction course pavement
PM	11	480	-1.51	0.22	-6.93	9927	0.000 ***	97%	95% - 98%	Porous pavement - modular blocks
RP	58	1053	-0.14	0.10	-1.36	7143	0.173	27%	8% - 42%	Retention pond - Open surface pond with a permanent pool (wet)
RV	3	108	-0.36	0.41	-0.88	9968	0.380	56%	-12% - 83%	Retention underground vault or pipes (wet)
VC	3	31	0.00	0.42	0.00	9976	0.996 ----	0%	-162% - 62%	Volume control/attenuation structures
WB	20	713	-0.11	0.17	-0.66	3351	0.508	23%	-14% - 48%	Wetland - basin with open water surfaces
WC	6	119	-0.03	0.31	-0.11	9762	0.913 ----	7%	-87% - 54%	Wetland - channel with wetland bottom

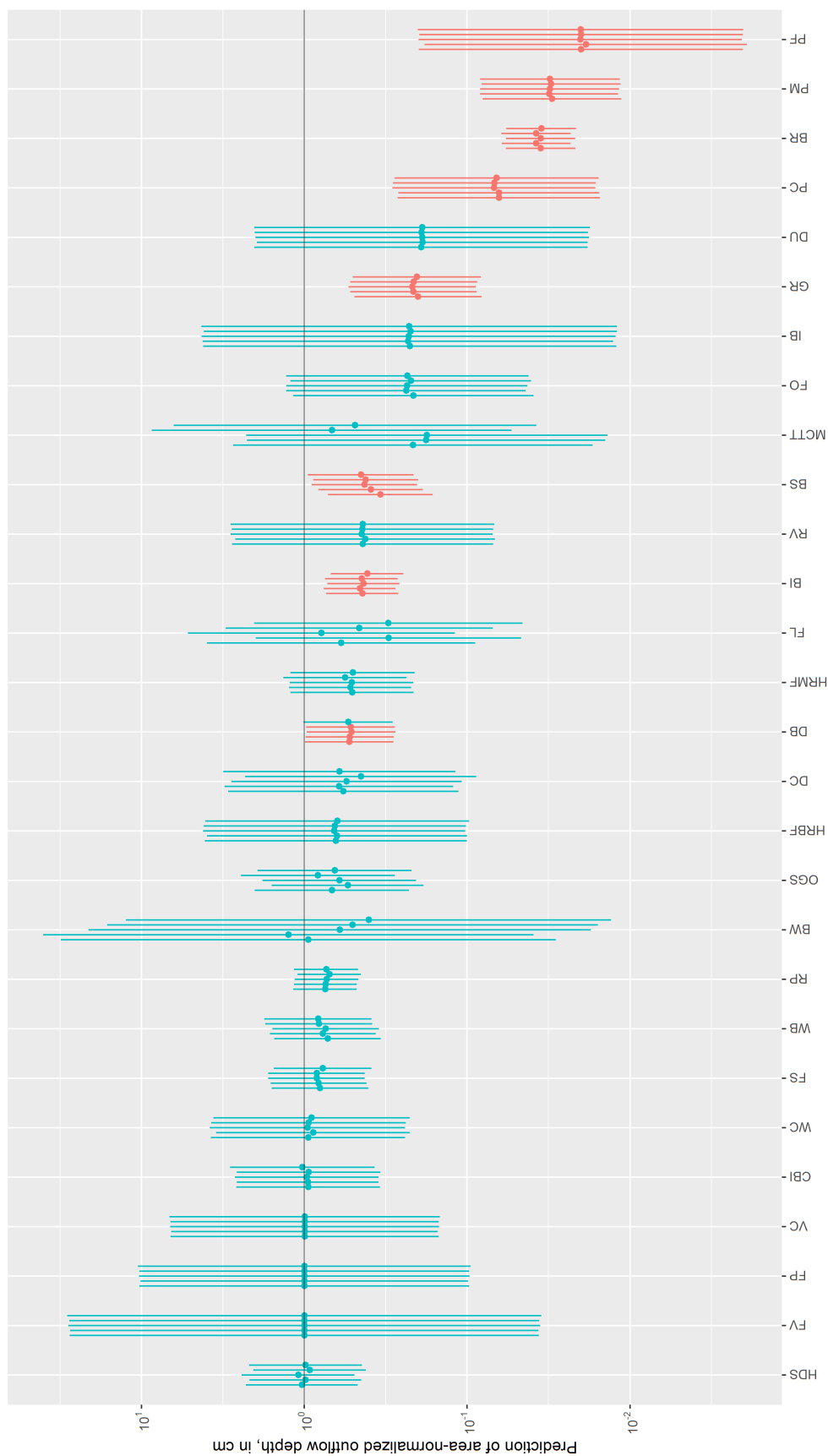


Figure 4.5: Unpooled predictions for each BMP type, based on a 1-cm notional storm event (reference line at 10^0 cm).

zone (-0.44 orders of magnitude, 64% RR), absence of a liner (-0.39 orders of magnitude, 59% RR), and soil media amendment (-0.32 orders of magnitude, 52% RR). The effect of permanent standing water on site had a poor effect on site performance, indicating that sites with permanent pools may discharge +0.6 orders of magnitude more water than the reference structure, likely due to groundwater seepage, this is an equivalent increase of 300% runoff.

Model Set 2 produced fewer poor-quality estimates than Model Set 1. This improvement is largely because the treatment and reference conditions have more observational support across the dataset than in the previous models. Only two features had potentially large Z statistics in two out of five imputations: the presence of a sieve or screen (which only had two site observations) and designed intermittent ponding (a broad category of sites that had neither permanent ponding, nor complete lack of surface ponding).

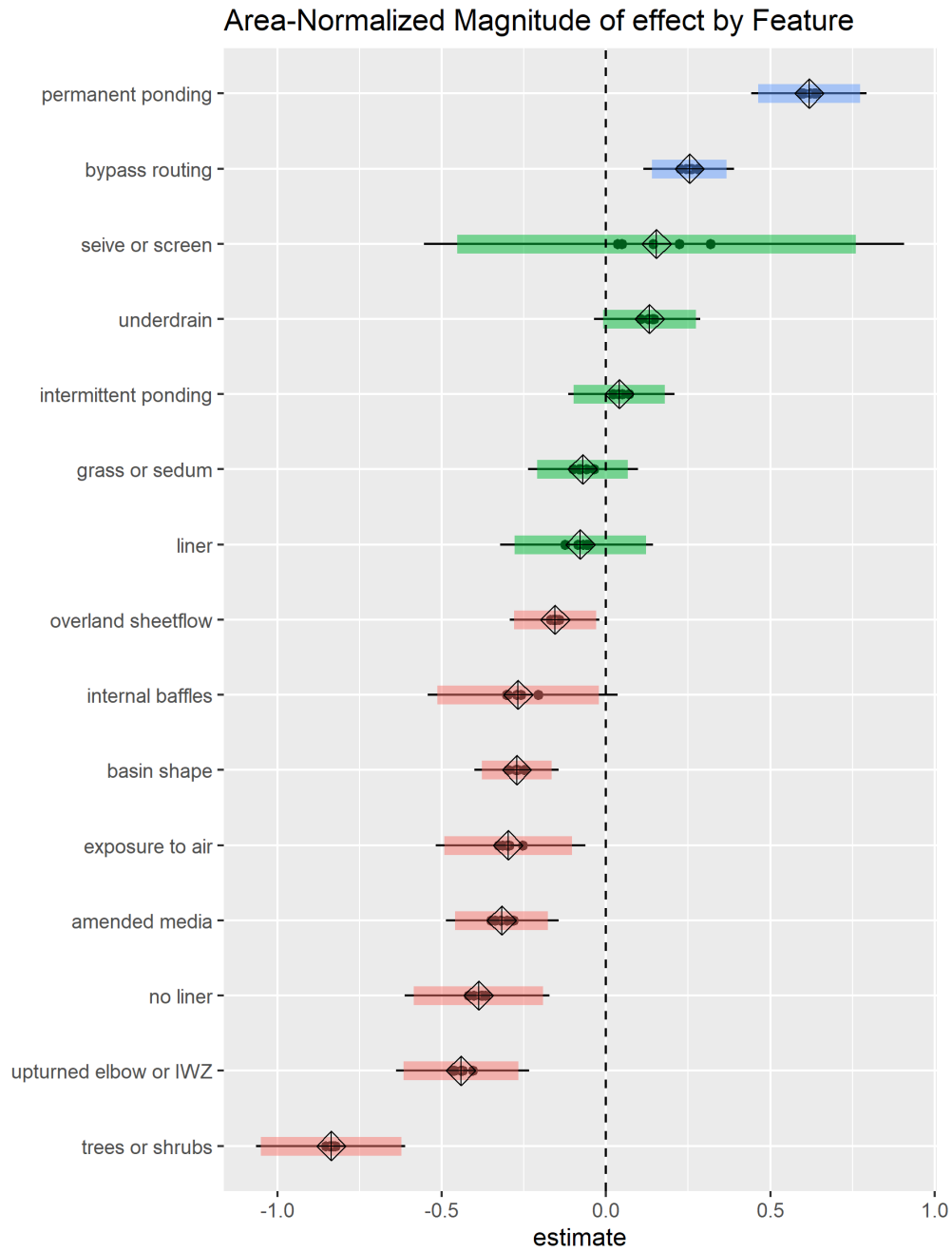


Figure 4.6: Effect size estimates for each structural feature (units of OM change). Imputed estimates (small black points) are pooled (black diamond) with color-coded pooled standard error range. Colors indicate structural effects that reduce effluent (red, to the left of the dotted line), net export (blue, to the right of the dotted line), or have no effect (green, crossing dotted line).

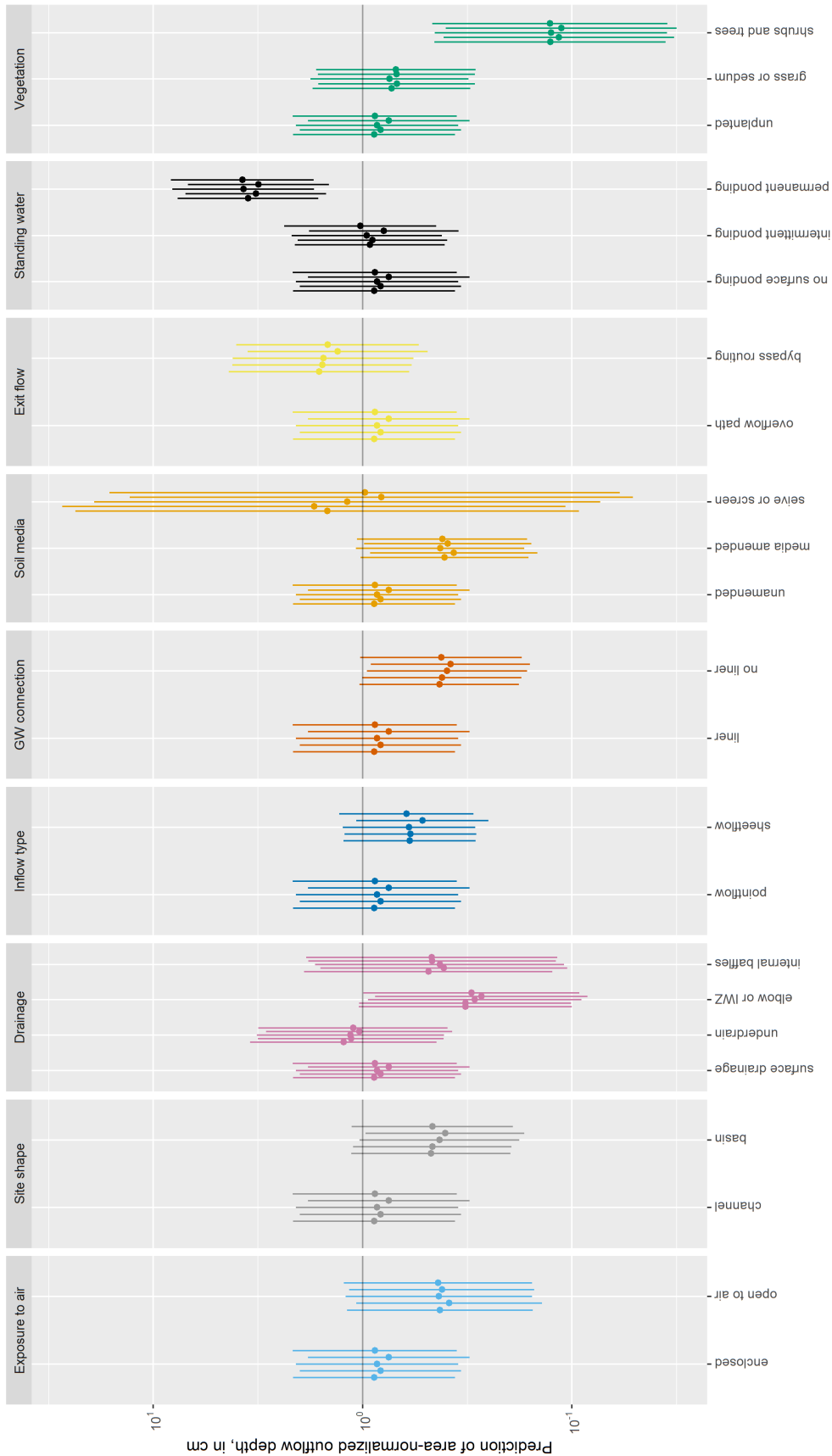


Figure 4.7: Area-normalized outflow predictions for each structural feature based on the 1 cm inflow depth (represented by the horizontal line).

4.3.3 Model Set 3: Media effects

In Model Set 3, the structural effect of soil media amendments are not isolated from other site structural features, nor can the qualities associated with media be flattened into a binary presence/absence variable. Therefore, the pooled marginal effect of soil media amendment is estimated to be -0.61 orders of magnitude (75% RR), which is nearly double the estimate made in Model Set 2. The matched model estimates are considered more accurate because they are unbiased and the model specification has been adjusted for backdoor paths. An example Love plot showing the effect of matching on standardized mean difference for the media amendment model is presented in Figure G.3. However, because the matching step specifically matches the dataset for a treatment of interest, the covariate effects are not valid estimates for the non-treatment covariate features. For example, the covariate estimate for removing a liner in Model Set 3 was -0.42 orders of magnitude, which is nearly identical to the Model 2 estimate. However, a matched model balancing liner presence and absence as the treatment variable gives an unbiased effect estimate of the marginal effect of liner absence at just -0.13 orders of magnitude (26% RR), which is likely more accurate. Model Set 3 indicates there is a strong combined effect of media amendment and absence of a liner, which could be explained by infiltration into native soil, or invertebrate or root intrusion into the amended soil media. Predictions for site performance with and without media amendment and liners are presented in Figure 4.8. The absence of a liner and approximately 1 meter of soil media amendment is sufficient to generate good runoff reduction at a BMP site ($\sim 60 - 70\%$ RR, Figure 4.9). This model infers that to engineer an equivalent system *with* a liner, about 3 meters of soil media would be necessary to establish similar hydrologic performance.

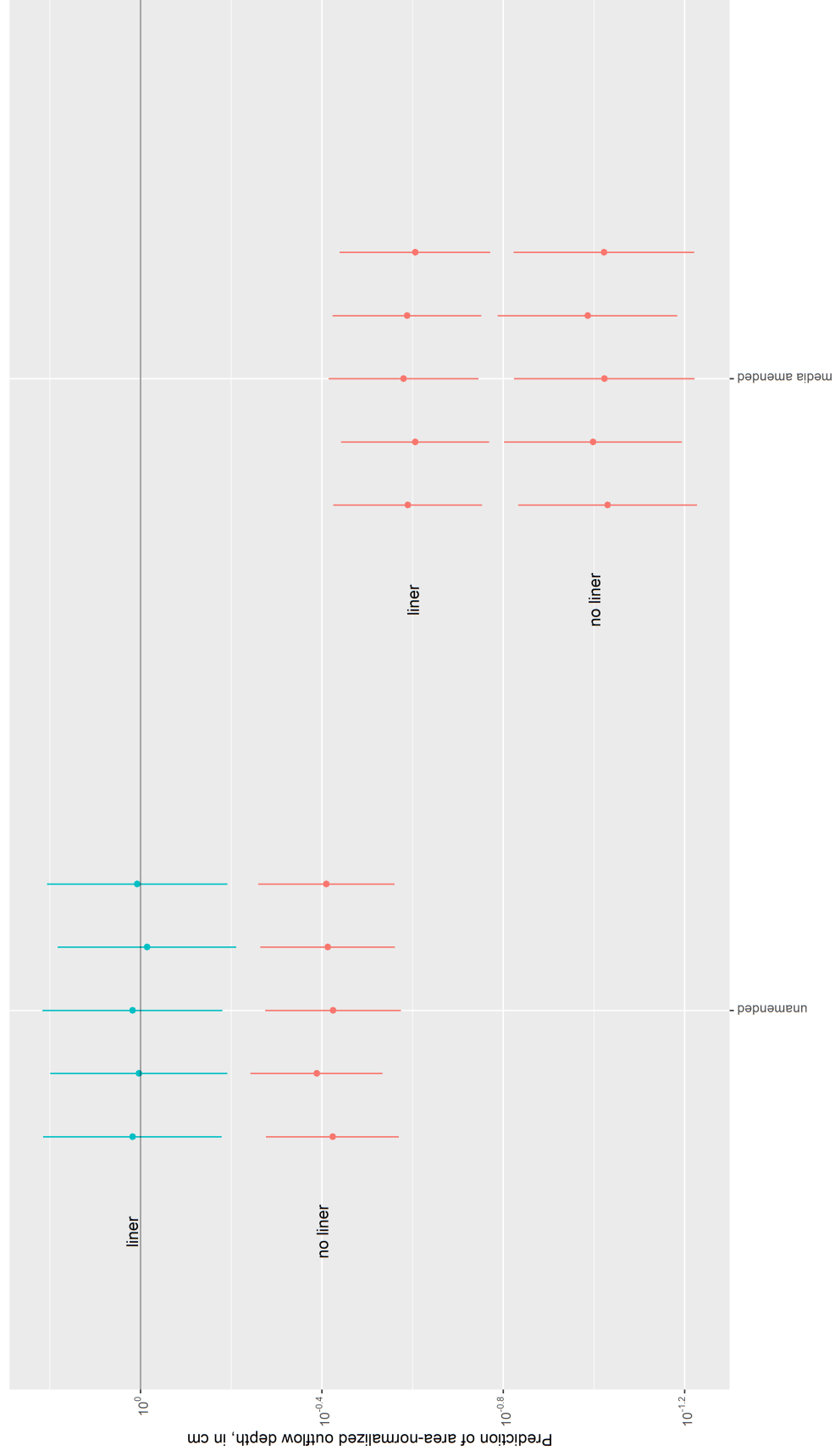


Figure 4.8: Area-normalized outflow predictions for media amendment and liners based on the 1 cm inflow depth

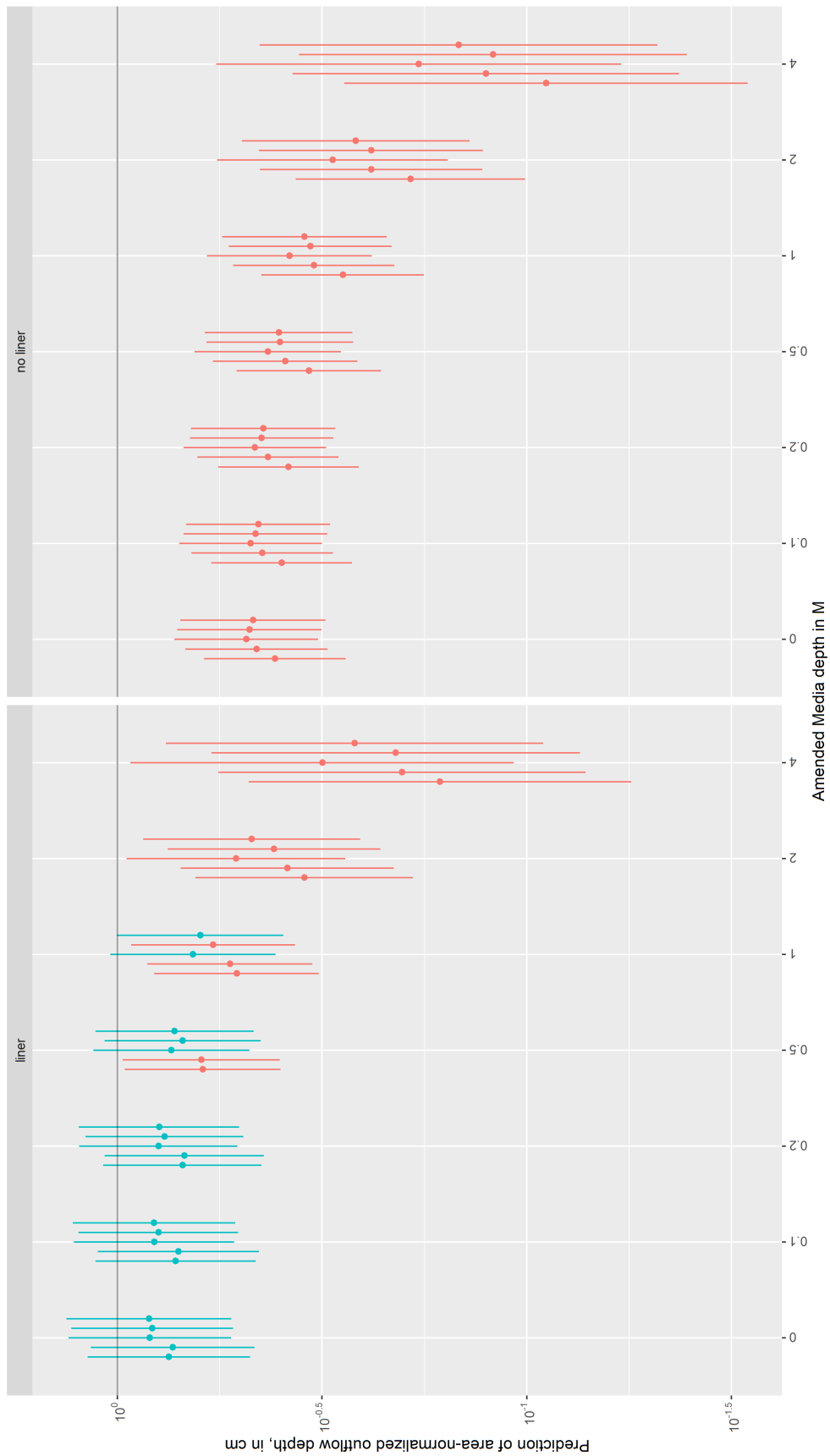


Figure 4.9: Area-normalized outflow predictions for media amendment and liners across the media depth gradient, based on the 1 cm inflow depth

4.3.4 Model Set 4: Ponding effects

In Model Set 4, the structural effect of permanent standing water is not isolated from other site structural features. It is particularly affected by the presence of a liner and the placement of drainage structures. The pooled marginal effect of permanent ponding is +0.56 orders of magnitude (in comparison to a reference of no standing water, which is close to the estimate in Model 2 (+0.62, Figure 4.6)). This value represents the marginal difference between the Model Set 4 effect estimates for permanent ponding and no surface ponding. Sites with no surface ponding discharge roughly half an order of magnitude less water than equivalent sites with permanent ponding (eliminating surface ponding is associated with 72% RR). The marginal effect of intermittent ponding is approximately +0.08 orders of magnitude as compared with no surface water ponding and a marginal benefit of -0.47 orders of magnitude (66% RR) associated with intermittent ponding as compared to permanent standing water. The rough match estimate comparing retention basins and detention basins estimates a marginal difference between permanent and intermittent ponding to be closer to -0.24 orders of magnitude (45% RR), which is similar to the estimates produced in Model Set 1. As in Model Set 3, the matched model estimates for the effect of interest are considered more accurate because they are unbiased and the model specification has been adjusted for backdoor paths. However, the accuracy of the marginal effect estimates for the other covariates are not to be used out of context because they were not the target variable during matching. An example Love plot showing the effect of matching on standardized mean difference is presented in Figure G.3. Predictions for site performance with different underdrain configurations with and without liners are presented in Figure 4.10. Adjusting a site design to accommodate intermittent surface ponding instead of permanent ponding likely has a stronger effect on runoff reduction than changing underdrain configuration or removing a liner alone.

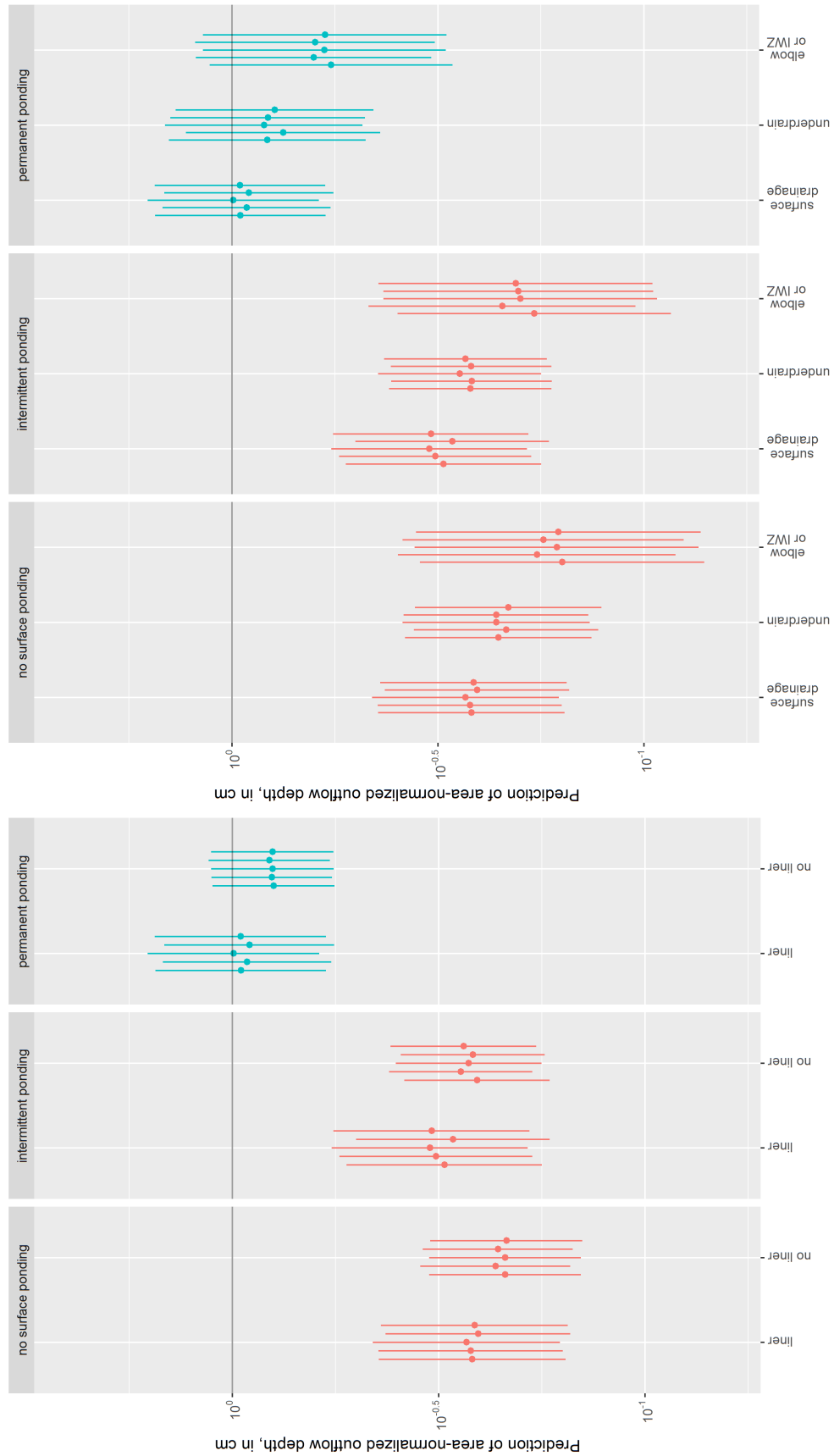


Figure 4.10: Area-normalized outflow predictions for ponding, underdrains and liners based on the 1 cm inflow depth.

4.3.5 Model Set 5: Vegetation effects

Model Set 5 acknowledges that the structural effect of vegetation is affected by the presence of media amendments and standing water, as well as the liner and other site effects. The estimated the pooled marginal effect of trees and shrubs is -0.72 orders of magnitude at the event scale (81% RR), which is slightly more conservative than the estimate from Model Set 2 (-0.84 orders of magnitude, 86% RR). This value represents the marginal difference between the Model Set 5 effect estimates for shrubs and trees versus an equivalent unplanted site. The marginal estimate for grass and sedum is slightly net positive compared with an unplanted site (0.12), indicating sites with short, mowed vegetation or shallow rooting discharge slightly more water than bare ground ($\sim 30\%$ increase). This is not a large effect, but it may be due to rapid saturation of the grassed surface, by ponding and discharge, rather than infiltration through the rooting zone. As in Model Sets 3 and 4, the matched model estimates for the vegetation effects are considered more accurate than the model 2 estimate, because they are unbiased and the model specification has been adjusted for backdoor paths. However, the accuracy of the marginal effect estimates for the other covariates are not to be used out of context because they were not the target variable during matching. An example Love plot showing the effect of matching on standardized mean difference is presented in Figure G.3. Predictions for site performance with different planting schemas are presented in Figure 4.11 (note that there is still net runoff reduction, even for the ‘grass and sedum’ category).

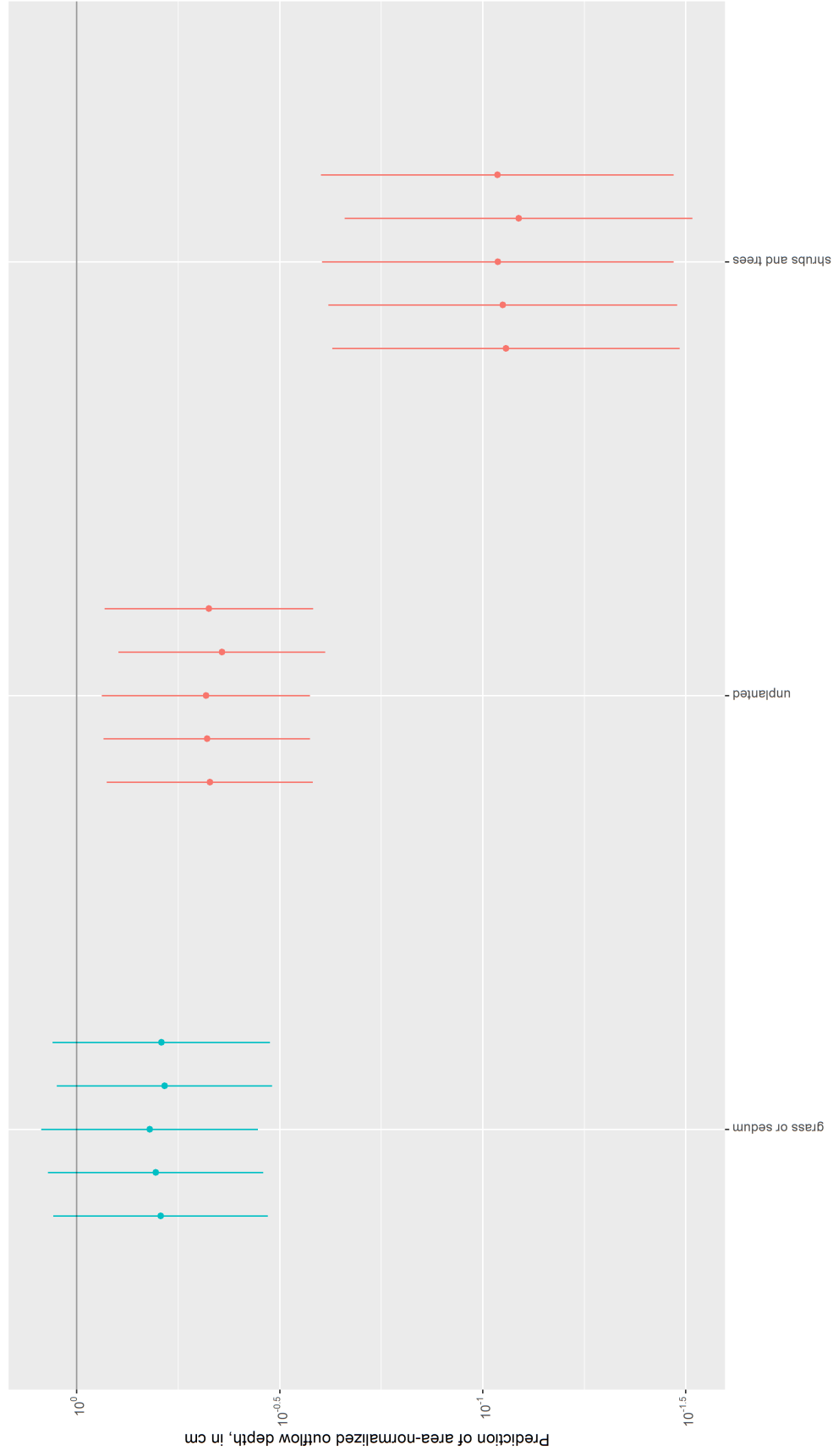


Figure 4.11: Area-normalized outflow predictions for vegetation based on the 1 cm inflow depth.

4.4 Discussion

4.4.1 Features with the greatest magnitude of influence on runoff reduction

A summary of the ‘best guess’ estimate for the marginal effect of each structural feature and its equivalent runoff reduction (RR) value is given in Table 4.5. Features with the greatest influence on runoff reduction are vegetation (-0.74 to -1.1 orders of magnitude or 86 to 92% RR), media amendment (-0.3 to -0.6 orders of magnitude, or 52 to 75% RR), upturned elbow drains or internal water storage zones (approximately -0.50 orders of magnitude, 68% RR), and unrestricted groundwater connection (at least -0.13 orders of magnitude, but up to -0.40 orders of magnitude, depending on the underlying native soil type, 28 to 60% RR). Features that show a smaller, but observable difference include: basin-shaped site design (-0.20 orders of magnitude, or 37% RR), as well as exposure to air, sometimes known as daylighting (up to -0.3 orders of magnitude, but likely closer to -0.1 or -0.15, 20-50% RR), and overland sheetflow (0.15 orders of magnitude, or 29% reduction). It seems that planting with mowed grass or shallow rooted vegetation can be helpful or harmful (in some cases up to -0.07 or as much as +0.08 orders of magnitude change, 15% to -20% RR). The variable effect is likely due to differences in landscape management, vegetation height, mowing, foot traffic, overall soil compaction and other unobserved factors. Permanent standing water and bypass routing structures show marginal discharge gains (+0.54 to +0.62 and +0.25 orders of magnitude respectively). Removal or redesign to change these features would have an effective RR of 71 to 76% for standing water and 43% for bypass routing.

Table 4.5: Synthesized results showing the best estimate of the marginal effect of each type of system and its equivalent runoff reduction (RR), with reference to a storm drain.

Best estimate of marginal effect						
Feature	estimate	std. error	RR	RR range		
liner	0					
no liner	-0.14	0.10	28%	9%	-	43%
no media amendment	0.00					
amended media	-0.60	0.11	75%	68%	-	80%
sieve or screen	0.00					
no surface water	0.00					
intermittent ponding	0.08	0.11	-20%	-56%	-	7%
permanent ponding	0.56	0.11	-261%	-367%	-	-179%
surface drain	0.00					
underdrain	-0.06	0.11	12%	-13%	-	32%
elbow or IWZ	-0.50	0.15	68%	55%	-	78%
internal baffles	-0.02	0.24	5%	-64%	-	46%
no vegetation	0.00					
grass or sedum	0.13	0.13	-35%	-80%	-	-1%
shrubs and trees	-0.72	0.21	81%	70%	-	88%
enclosed	0.00					
open to air	-0.30	0.19	50%	21%	-	68%
pointflow	0.00					
sheetflow	-0.25	0.11	44%	28%	-	56%
overflow	0.00					
bypass routing	0.23	0.11	-70%	-119%	-	-32%
channel shape	0.00					
basin shape	-0.20	0.10	37%	21%	-	50%

4.4.2 Practical application of the findings

The marginal estimates represented in Tables 4.3, 4.4, and 4.5 allow practitioners to make estimates of hydrologic performance for various BMPs prior to construction. The following are three worked examples using the estimates for detention basins (DB).

Generalized site estimate:

- The mean event inflow into a detention basin is $3 * 10^5$ L (Table A.2, DB).
- The marginal estimate for detention basins in Table 4.3 is -0.20 orders of magnitude (OM) on a site-averaged basis.

$$Inflow = 3 * 10^5 = 10^{5.49} L$$

$$Outflow = 10^{(5.49-0.20)} = 10^{5.29} = 1.95 * 10^5 L$$

The expected outflow is estimated at $1.95 * 10^5$ L (37% RR) compared with $1.8 * 10^5$ L (42% RR) in Table A.2.

Area-normalized site estimate:

- The area-normalized marginal estimate for detention basins in Table 4.4 is -0.28 orders of magnitude (OM) on an area-normalized basis.
- Assume a watershed area of 3 ha
- Assume a rainfall-runoff event depth of 1 cm from the watershed
- Assume a BMP site area of 1000 sq. meters

$$Inflow = 3ha * 10000m^2/ha * 0.01m/cm = 300m^3$$

$$Depth_{Inflow} = 300m^3/1000m^2 = 0.3m = 30cm$$

$$\log_{10}(30) = 1.48$$

$$Depth_{Runoff} = 10^{(1.48-0.28)} = 10^{1.2} = 15.8cm$$

The expected area-normalized outflow depth is 15.8 cm per square meter of BMP site, based on the physical assumptions above and the estimate for detention basins, 0.28 OM (47% RR, compare with 42% RR from the first worked example).

Structure-specific estimate: The marginal estimates from Table 4.5 are additive and can be calculated given a known set of structural features at a site. The structural estimates allow estimation of individual site performance, rather than the average performance associated with a given BMP type. To create an estimate, first one value is selected for each of the nine feature categories. The detention pond in Chapter 2 is described as having the following features:

- no liner (-0.14 OM)
- no media amendment (0 OM)
- intermittent ponding (+0.08 OM)
- surface drainage (0 OM)
- grassed surface (+0.13 OM)
- open to air (-0.15, assuming less than maximum)
- pointflow inlet (0 OM)
- overflow (0 OM)
- basin shape (-0.2 OM)

$$DB_{estimate} = -0.14 + 0 + 0.08 + 0 + 0.13 - 0.15 + 0 + 0 - 0.2 = -0.28$$

$$Inflow = 10^{5.49} = 3 \times 10^5 L$$

$$Outflow = 10^{(5.49-0.28)} = 10^{5.21} = 1.62 \times 10^5 L$$

The runoff reduction estimate by indicating features for a detention basin with the

structures described is approximately -0.28 orders of magnitude (47% RR), compared with $1.8 * 10^5$ L in Table A.2 and previous worked examples. This type of calculation is applicable to any known set of structural features, either at an installed site, or prior to construction. The information in Table 4.5 also allow estimation of the performance improvement value of isolated structural retrofits. A good rule of thumb for calculating an error range for an effect estimate based on a set of structural features is to use the feature with the largest error value, or 0.1, whichever is higher. In this case, the feature with the highest standard error is exposure to air, 0.19. Therefore, the estimate for a detention pond with the features described is -0.28 +/- 0.19 (19 to 66 % RR).

Using the workflows presented above, estimates for several widely used BMP types were compared with the Water Budget Triangle findings from Chapter 3 in Table 4.6. The site types with the best agreement between the two sets of models were detention basins (DB) and green roofs (GR). These site types are also likely more narrowly defined than bioretention (BR) or porous pavement (PP). Retention ponds have an obvious difference in effect estimate because the water budget triangle results account for runoff reduction calculation differently than the linear mixed models that generated the marginal estimates. The water budget data assumes that Q represents total discharge from a site over time, and $RR = 100 - Q$, where $0 \leq Q \leq 100$. The earlier model does not compare inflow to outflow, and the estimates of runoff reduction cannot be negative. The water budget triangle numbers from Chapter 3 tend to represent longer-term performance estimates, whereas the linear mixed models in this chapter are event-based. This means the events in the imputed dataset end once discharge is complete, which specifically ignores ET and soil moisture drawdown that are captured in the water budget dataset.

4.4.3 Addressing estimate variability

There are several components that represent variability in the dataset, which overlap in the results seen in Tables 4.5, 4.6 and G.2. First, the mapped structural features observed for

Table 4.6: Effect estimates for five BMP types (gray rows), and water budget triangle (WBT) runoff reduction values from Tables 3.1 and 3.2 (in bold). Equivalent effect sizes and error estimates for WBT values were back-calculated based on mean inflow for each system type reported in Table A.2.

BMP Type	Source Estimate	OM Est	Error	RR	RR range
BR	Structural Estimate	-1.00	0.21	90	84 - 94
	Area-normalized Estimate	-1.44	0.11	96	95 - 97
	Site Estimate	-1.49	0.13	97	96 - 98
	WBT mean	-0.34		54	
	WBT median	-0.28		48	
	WBT event scale		0.44		25 - 90
	WBT long term		0.13		41 - 67
DB	Structural Estimate	-0.28	0.19	48	19 - 66
	Area-normalized Estimate	-0.28	0.14	48	28 - 62
	Site Estimate	-0.20	0.17	37	7 - 57
	WBT mean	-0.37		57	
	WBT median	-0.32		52	
	WBT event scale		0.42		30 - 90
	WBT long term		0.15		43 - 71
GR	Structural Estimate	-0.64	0.19	77	65 - 85
	Area-normalized Estimate	-0.68	0.20	79	67 - 87
	Site Estimate	-0.44	0.27	64	32 - 81
	WBT mean	-0.41		61	
	WBT median	-0.44		64	
	WBT event scale		0.27		30 - 80
	WBT long term		0.04		57 - 64
PP	Structural Estimate	-1.26	0.19	95	91 - 96
	Area-normalized Estimate	-1.46	0.20	97	95 - 98
	Site Estimate	-1.30	0.27	95	91 - 97
	WBT mean	-0.47		66	
	WBT median	-0.70		80	
	WBT event scale		2.33		20 - 100
	WBT long term		0.11		56 - 74
RP	Structural Estimate	0.34	0.19	-119	-239 - -41
	Area-normalized Estimate	-0.14	0.10	28	9 - 42
	Site Estimate	0.01	0.14	-2	-41 - 26
	WBT mean	-0.14		27	
	WBT median	-0.08		17	
	WBT event scale		0.25		5 - 70
	WBT long term		0.09		13 - 42

each site show that there is considerable variability within the same BMP type categories. The minimum standard error observed in any category is approximately ~ 0.1 orders of magnitude (RP); this minimum value seems to hold throughout the causally inferred estimate findings. The limitation on precision is likely due to poor model fit on zero-discharge events. However, it may be possible to overcome the limitation and reduce the error associated with estimates by using a two-step hurdle model in the linear mixed model imputation workflow or using a difference-in-differences approach (see Chapter 5). The maximum standard error can be quite high (0.80-0.92 OM), which occurs for sites with few observations or diverse structural design implementations. None of the BMP type categories has homogeneous characteristics, and categories with more heterogeneity in structural features also have less well-defined effect estimates. For example, constructed wetlands are particularly difficult to describe because they have more complex combinations of structures, flow paths, and vegetation than many other types of BMPs. This condition translates to poor estimation for sites that have wetland-like characteristics (BW, WB, WC), with little agreement between the estimates for each wetland category.

In addition to differences in structural feature sets, there is also variability in the implementation of the structures that are mapped to each structural feature category. For example, there is considerable variation within the types of media amendments that are represented in the binary **Media** covariate, as well as different plant communities represented within each categorical level of the **Vegetation** covariate. These details were intentionally generalized to allow for representation in a model that depicts a large, ecosystem-scale problem space with the level of detail that stormwater designers are likely to be able to reproduce in their designs. The bulk of the literature that exists on engineering better stormwater BMPs has focused on narrowly specific engineering changes within a single design factor, such as choosing the optimum depth, particle size distribution, or chemical composition for green roof media. These findings have important implications for later phases of engineering design, but do not help address whether a BMP

site design will meet future catchment hydrology goals, or if a current site needs to be retrofitted. Therefore, similar features were grouped intentionally as an attempt to compare the functional similarities of BMP systems with different colloquial names and structures with similar hydrologic activity. The results from Model Sets 1 through 5 show the general ranges that designers can expect from their systems, including the potential variability in performance. With additional research it may be possible to estimate the structural feature from a sliding scale for a more nuanced predictive model. The comparison of variability between the linear mixed model estimates and the water budget triangle datasets from Chapter 3 show relatively good agreement, with a clear decrease in variability for longer term water budget observations when compared with ranges representing measurements at the event scale. This level of agreement is expected; the effect estimates from the linear mixed model are event-based observational results that liken most closely to the event-based WBT results.

As described in the results, the model fits are poor in general, but particularly for Model Sets 1 and 2. This condition is an artifact of the data at some sites having many real zeroes in the response variable (**Outflow**) below a threshold value in the predictor variable (**Inflow**). In practical terms this condition occurs when a site captures all the inflow associated with an event and generate no runoff. This situation is visible in Figure A.3 BMP Type ‘BR’ (Bioretention) and is a common characteristic at specific types of sites. Several attempts were made to fit a model that improves on this problem, but the challenge was too great; the appropriate solution must be able to handle semi-continuous multilevel data log-transformed to base 10, not base e (natural log). The approach that will give the best fit is to treat the zeroes as part of a two-step hurdle model (first estimate the likelihood of a discharge occurring, then estimate the magnitude of discharge). However, I was not able to find a package that would accommodate this two-step generalization while grouping by site (maintaining a mixed model format). Apparently, the `glmmTMB` package cannot easily treat semi-continuous data for a two-step hurdle model.

The GLMMadaptive package likely can address this condition, but the model syntax is not the same as the lme4/glmmTMB standard. I was not able to successfully build my own two-part model template format from scratch. Additionally, the computational load for GLMMadaptive is also much greater than the glmmTMB approach, which makes multiple imputation and pooling steps slower and harder to troubleshoot. Therefore, I determined the best course of action was to use the glmmTMB model and make predictions between the middle quartiles of the precipitation and inflow distributions, where the model fits are the most acceptable. The distribution of precipitation events is shown in Figure 4.4 at the measurement scale and log-transformed scale. All predictions were made for a 1-cm notional event, which is represented at 0 on the base 10 log-transformed scale; this is quite close to the distribution mean of the log-transformed distribution, (0.13) and is a practical point of interest. The matched model fits (Model Sets 3-5) are also relatively poor in the tails, but they have fits that are considerably better than Model Sets 1 and 2.

4.5 Conclusions and recommendations

4.5.1 Conclusions from linear mixed model estimation of structural feature effects

Features with the greatest influence on runoff reduction at the event scale are:

- vegetation: -0.74 to -1.1 orders of magnitude (OM), equivalent to 86 to 92% runoff reduction (RR)
- media amendment: -0.3 to -0.6 OM; 52 to 75% RR
- upturned elbow drains or internal water storage zones: ~ -0.50 OM; 68% RR, and
- unrestricted groundwater connection: -0.13 to -0.40 OM; 28 to 60% RR

Adjustments to design and retrofits should prioritize these factors, as they are the most cost-effective ways to improve event-scale hydrologic performance. Additionally, features that have a negative impact on hydrologic performance are:

- permanent standing water: +0.54 to +0.62 OM; more than 250-320% increase in discharge), and
- bypass routing structures: +0.25 OM; 78% increase in discharge)

Including these characteristics in the design of a BMP site should be closely examined to understand why their use is necessary and if there are other ways to mitigate the marginal gains associated with including them in the design.

4.5.2 Future recommendations

Several recommendations regarding the use of these model findings:

- Where possible, practitioners should monitor the actual performance of their existing sites for no less than 10 events at each site, including: inflow volume; outflow volume; and, precipitation depth.

- Field observations can be used to calculate actual performance and generate an area-normalized order of magnitude reduction distribution.
- These values can be compared to the expected performance ranges estimated by BMP type and by structural feature set to evaluate how the site performs against the modeled benchmarks.
- Adjustments to design should be iterative, and focus on small, incremental changes to improve each of the structural features that exist onsite; and confirmed by ongoing monitoring.

Future refinement of this modeling technique should employ a two-step hurdle model to account for the discontinuity in the dataset associated with zero-discharge events. This approach should improve overall model fit and reduce the error associated with individual marginal estimates.

5 Estimation of biogeochemical flux in green stormwater infrastructure

5.1 Introduction

5.1.1 Understanding controls on solubility and transport in the urban environment

Precipitation in urban locations transforms into stormwater through a complex and diverse pattern of water interacting with human-made materials and artificially constructed surfaces under hyper-localized physical and chemical conditions. The chemical transformation from precipitation to stormwater is observable; urban precipitation and surface water have distinct profiles of suspended particles and dissolved solutes (Table 5.1, Figures A.1 and A.2), even within the same catchments (apparent from the summaries from the BMP database subset used in this study). The chemical distinction is not simply due to physical separation of contaminant generation processes between the atmosphere and landscape; localized physico-chemical conditions differentiate contaminant transport in these two ecospheres. Localized conditions occur because urban drainage systems characteristically display a complex interconnection between natural and built drainage structures. Urban drainage paths interweave rill and gully formations with roofs, downspouts, gutters, culverts, and subsurface drainage pipes before emptying into urban streams. The localized set of physico-chemical factors that act as controls on the solubility and transport of contaminants in the urban environment are the same as those that have been observed in the laboratory: pH, temperature, dissolved oxygen, exposure to ultraviolet radiation, as well as stormwater kinetic energy, and introduction of biological material and organisms. These physico-chemical factors may seem simple, but they act as fundamental controls on the mobility and transport of contaminants and are highly locally diverse along the drainage pathways that run across the urban mosaic. These control

factors also interact to govern solubility and transport; careful observation of the precipitation-to-stormwater transformation can facilitate construction of a causal understanding of how the physical and chemical conditions interact to control contaminant mobility and transport in the urban ecosystem (Figure 5.1, in blue). The resulting causal information allows engineered intervention on these factors to manipulate the physico-chemical system and treat stormwater contaminants more effectively (Figure 5.1, in green). Ecological engineering practices seek treatment methods that intervene on specific variables to modify surface water runoff toward an outcome that supports good water quality, human health, and aquatic habitats.

Table 5.1: Comparison of dissolved solute concentrations in stormwater and precipitation in $\mu\text{g/L}$.

Contaminant ($\mu\text{g/L}$)	Stormwater (as reported in WERF 2019)				Stormwater (this study subset)			Rainwater (this study subset)		
	low	Bca median range		high	10th %ile	50th %ile	90th %ile	10th %ile	50th %ile	90th %ile
TDS	24000.	46000.	-	390000.	1600000.	20000.	70000.	310000.	29000.	270000.
TSS	10000.	26000.	-	77000.	160000.	8000.	46000.	250000.	3000.	420000.
TP	64.	99.	-	320.	690.	40.	200.	860.	10.	660.
Dissolved P	20.	48.	-	130.	230.	10.	70.	320.	40.	380.
PO4 as P	8.6	17.	-	260.	510.	3.2	20.	250.	2.	110.
TN	450.	710.	-	2800.	4300.	610.	1500.	3600.	300.	3500.
TKN	370.	760.	-	2000.	3000.	-	-	-	-	-
NH3 as N	33.	66.	-	1200.	2200.	-	-	-	-	-
NOx as N	72.	270.	-	530.	1300.	40.	380.	1400.	60.	1100.
Total Cd	.061	.13	-	.5	1.	.1	.48	2.	.1	.25
Dissolved Cd	.012	.031	-	.3	.5	.04	.17	.97	.04	.05
Total Cr	1.2	1.7	-	7.	12.	1.	4.	14.	.73	2.4
Dissolved Cr	.29	.5	-	2.7	1.4	.5	1.6	7.9	.25	.5
Total Cu	3.8	7.4	-	25.	52.	3.1	13.	54.	1.	8.
Dissolved Cu	1.8	3.9	-	12.	24.	1.6	6.2	27.	3.2	6.
Total Pb	1.3	3.8	-	17.	43.	1.	8.	54.	.75	4.
Dissolved Pb	.044	.094	-	1.3	7.8	.09	1.	6.9	.13	.5
Total Zn	14.	27.	-	110.	390.	15.	65.	280.	10.	52.
Dissolved Zn	4.6	10.	-	190.	380.	6.	27.	130.	2.5	17.

5.1.2 Contaminants present in urban waters

Urban stormwater chemistry is dependent on localized precipitation patterns and land use. Some solutes and particulates are observed to associate more closely with urbanization (Cl, N, P, K, synthetic plastics and oils), anthropogenic activities (Zn, Cd, Cr, pesticides, dissolved residual medication and pharmaceutical breakdown products in wastewater), and built structures (Cu, Ni, Zn, resins and asphalt breakdown products), as opposed to

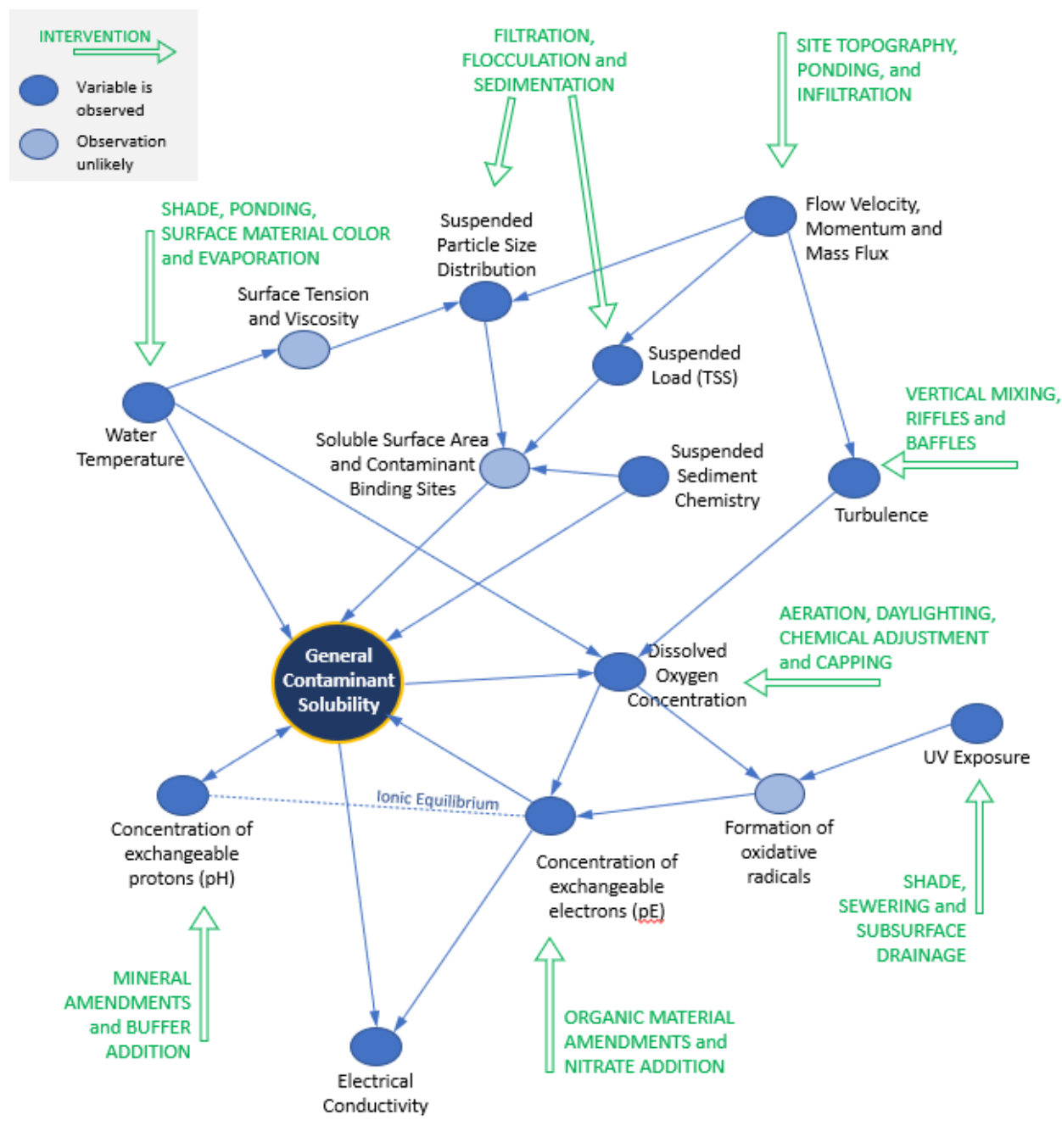


Figure 5.1: A directed acyclic graph (DAG) conceptual model showing causal hypothesis on the interaction of physico-chemical processes as controls on the general solubility and transport of most contaminants in the urban environment, represented in blue. General solubility rules affect the solubility of salts and other ions, dissolved metals, nutrients, and organic molecules. Solubility of gases is excluded due to a different set of constraints on gaseous solubility. A suggested set of engineered treatment interventions for distinct variables are labeled in green. Biological processes have been excluded for the sake of clarity. A more detailed version of this figure with empirical and observational equations is included in Appendix C.

natural landscape biogeochemistry (HCO_3 , Fe, As, DOC, F) (Drever [1997]). While many solutes are considered benign or beneficial to humans (Fe, Na, K, Ca, Mg), others are known to cause negative health effects at nearly any concentration (Pb, methylmercury, waterborne pesticide residues, PFAS and PCB family chemicals). Yet more may present little threat to human health but cause environmental degradation and aquatic habitat loss (Al, Cu, N, P, Na) (Drever [1997]). Freshwater ecosystems are particularly sensitive to soluble nutrients (N, P) in runoff; for example, the ecological limits for soluble reactive forms of phosphorus range widely depending on the receiving water body type, salinity and locale. Concentrations considered low in one water body may cause eutrophication in another. For example, New York State has ambient water quality guidance values of 20 $\mu\text{g/L}$ total phosphorus (TP) in ponds, lakes, and reservoirs, but half that for Lake Ontario. These concentrations contrast with urban stormwater runoff TP concentrations that frequently occur between 60 and 700 $\mu\text{g/L}$. Likewise, the EPA's ecoregional nutrient criteria for total nitrogen (TN) in water bodies of the Western forested mountain regions is more than an order of magnitude lower than the limit for those in the Southern Florida coastal plain (2.18 mg/L) (US EPA Office of Water [2013]). Nutrients are well known as the most widespread stressor impacting rivers and streams and the residents who rely on them for drinking water, recreation, and economic activity (Fox [2022]). There are more than 1.1 million kilometers of impaired stream reaches in the US (Kelderman et al. [2022]). Due to chemical forms or biophysical partitioning, specific contaminants vary in mobility and and/or bioavailability. Dissolved solutes are more readily transported than particulate-bound ones, and contaminants that are bound with organic or particulate ligands interact differently with biota than inorganic forms (Drever [1997]). Interactions between ions and organic molecules cause changes in molecule structure and shape, due to effects of altered polarity, bond angles, and ionic radii. For example, organo-metallic complexation can have a protective effect on metal toxicity, while methylation or phenylation of metals can increase toxicity by bioconcentration and biomagnification,

interrupting normal enzyme, protein or cell signaling functions. As a general rule, mobile and labile forms pose a greater threat to human and environmental health than those forms that are non-reactive or isolated.

5.1.3 Chronic exposure and toxic effects

Exposure to toxic substances in stormwater is poorly understood as a potential threat to humans and aquatic ecosystems alike; most of the concern in US cities has been focused on reducing and treating bacterial loads from combined sewer overflow events (Clary et al. [2014, 2022], Ergas Sarina J. and Fassman-Beck Elizabeth [2023]). However, the ongoing presence of complex chemical mixtures from urban stormwater has chronic degrading effects on aquatic habitats and may pose a risk to humans even at low concentrations (Bradley et al. [2023]). My previous work on this topic with Balderas Guzman et al. [2022] shows that concentrations of nutrients, trace metals and other contaminants are above ecological effects limits following nearly all urban storm events, and more than a quarter of events have concentrations of single contaminants within the mixture that approach or exceed levels of concern for human health. Cousins et al. [2022] recently found that exposure to the PFAS family of chemicals occurs globally during nearly all precipitation events and exceed both US and European health advisory limits. Previous analyses and regulations have focused mainly on the status of a single analyte or contaminant family. In any given stormwater runoff event, some individual contaminants from within the mixture will exceed limits of concern and others will not. Even if no single contaminant regularly exceeds an acute ecological exposure limit, there is often at least one contaminant from within the stormwater mixture that does, and many different contaminants commonly exceed a chronic exposure limit during pulsed hydrologic events. The concern over toxic exposure to urban stormwater is twofold: 1) chronic occurrence becomes insidious and unavoidable for human and non-human inhabitants; and 2) the likelihood of any one analyte exceeding a chronic exposure limit increases along with the diversity of

contaminants within the chemical mixture. Contaminant mixtures have become so complex that they are difficult to adequately monitor, regulate and treat (Bradley et al. [2023]). Toxic solute mixtures are the norm rather than the exception for stormwater; the fifty largest American urban metropolises typically experience between 59 (San Jose, CA; relatively arid) and 154 (Pittsburgh, PA; abundant precipitation) precipitation days per year (Arguez et al. [2010]). This frequency of precipitation and potential runoff events suggests that ecologically toxic levels of any one contaminant within the chemical mixture occurs in urban watersheds approximately 1 to 3 times per week, and human health limits are likely exceeded between 1 to 3 times per month. These calculations are simple averages based on the annual frequency of rainfall; toxic conditions likely occur more frequently in regions where annual precipitation is compressed into a rainy season. Stormwater toxicity is higher when the weather switches from a long dry period to intense rainfall, and in locations with frequent seasonal flooding or degraded surface contamination. While actual human exposure to contaminants from stormwater effluent is difficult to gauge, more than 182 million people live in the fifty largest US metropolises (this accounts for 55% of the US population) (US Census Bureau [2022b]). These metropolitan areas account for approximately 8% of the US land area, and all include reaches with chronically degraded watershed conditions (US Census Bureau [2022a]). Human health is affected by chronic activities that occur with regularity (e.g., diet, drug or alcohol consumption). Weekly or biweekly stormwater exposure that has high levels of lead or other trace metals would negatively impact both human and animal inhabitants of urban ecosystems.

As mentioned previously, much of the stormwater treatment literature tends to focus on treatment to reduce concentrations of individual contaminants. For example, the US EPA and state environmental agencies use the total maximum daily limit (TMDL) approach to regulate contaminants on an analyte-by-analyte and waterbody-by-waterbody basis. This approach is probably the appropriate method for municipal drinking water regulation, but it is not suitable for trying to manage many different contaminants in

urban runoff for a wide range of waterways. The approach stems from the limitation of detecting single analytes or analyte families for analysis, and it neglects the fact that stormwater contains a diverse array of chemical constituents in a complex mixture. Successful water quality improvement relies on overall mass reduction of contaminant loads, which is heavily dependent on volumetric reduction of effluent. Focusing design of stormwater control measures (SCMs) on volumetric reduction of runoff also reduces the mass of all contaminants in the complex mixture of effluent, not just a single analyte targeted for treatment. Even modest volumetric reduction can often result in very good contaminant mass reduction (80% or more) (Driscoll et al 2017). General reduction of all contaminants in urban receiving waters leads to less chronically toxic aquatic habitats and better overall downstream water quality. Sometimes SCM design will focus on single analyte percent-based concentration reductions (e.g., targeting 90% reduction of TN concentrations from 2 ppm to 0.2 ppm). This approach places the emphasis of the reference benchmark on the initial concentration value, which may be elevated (compare with 90% reduction of TN from 20 ppm to 2 ppm). The target reduction is the same, but the benchmark is being measured against the initial concentration instead of the ecoregional nutrient criteria. In both cases, the effluent would be above the ecoregional nutrient criteria for the western forested mountains (0.12 mg/L), including urban streams surrounding Seattle, WA, the foothills of Denver, CO, Santa Barbara, CA, and many other smaller municipalities in the western US (US EPA Office of Water [2013]). Using the ecoregional nutrient criteria as the relevant benchmark for BMP design will help municipal managers align local treatment efforts with state and federal guidelines.

Chronic exposure to toxic substances can be decreased in two ways: 1) reduction of the frequency of exposure, and 2) reduction in the dose of exposure. Humans are largely unable to control the frequency of chronic stormwater exposure, since this is based on local precipitation and runoff patterns. As a result, management of chronic surface water toxicity is limited to mitigating the severity of the chronic conditions. In addition to volumetric

reduction as a driver of mass removal of contaminants, engineers can reduce exposure of chemicals to humans and the environment by manipulating key physico-chemical features that govern contaminant transport (Table C.1). Nutrient mobility and particulate retention is controlled by intervening on the causes of contaminant solubility and transport (Figure 5.1): filtration of suspended particles, pH and pE adjustment through soil media amendments, changing sunlight exposure patterns to adjust temperature and ultraviolet exposure, changing airflow patterns to improve or reduce dissolved oxygen, designing site topography to modulate kinetic energy and appropriate ponding characteristics. In addition, biochemical and biological mechanisms are also key removal strategies, including biologically-mediated complexation, flocculation or transformations of dissolved metals, particulate and dissolved organic matter. Understanding how these factors can be engineered and implemented appropriately will help improve the contaminant treatment capacity of SCMs in general and help tailor the function of individual facilities to meet water quality goals in the urban environment. Urban stormwater managers have little control over contaminant generation processes; yet, by careful engineering design they may succeed in controlling solubility and transport conditions to minimize contaminant flux in stormwater as it acts as a solvent and transport mechanism.

5.2 Data and methods

5.2.1 Dataset preparation

As described in Chapter 4, the International BMP Database (BMPDB) contains a large compilation of stormwater control structures and best management practices (BMPs). These are compiled observational data from case studies with information about structural design, hydrologic activity, and water quality performance for more than three dozen BMP types (BMP Database, 2019). The 2019 version of the database contains more than 374,000 observations of approximately 600 unique analytes, among these are details about whether the samples represent dissolved, suspended (particle-bound) fractions, or total concentrations. The raw water quality dataset covers one or more observations at over 500 unique stormwater control structures and includes information about the location of the monitoring station for the sample (inflow, outflow, rain gauge, etc.). There are approximately 24 different BMP types with water quality observations represented in the raw database. The sampling plans are unique to each site; overlap among common analytes, detection limits and methods is irregular (Table 5.2). Despite these inconsistencies, there are enough observations in the dataset to gain a general understanding of how BMP structural choices affect the retention of some analytes on an event timescale. Data missingness was high, but complete observations were not required for inclusion in the study dataset to maintain as much common support and statistical power as possible. The concentrations of eleven of the most commonly monitored analytes were selected for analysis: N (dissolved and total), NO_x (dissolved), P (dissolved and total), total soluble reactive P (SRP, sometimes labeled PO₄ in the figures; the records in the BMPDB use these terms exchangeably), Cd (dissolved and total), Cr (dissolved and total), Cu (dissolved and total), Pb (dissolved and total), Zn (dissolved and total), total suspended solids (TSS) and total dissolved solids (TDS). The initial sub-selection of water quality data comprises 122,010 unpaired observations of inflow and outflow concentrations

for 9802 unique storm events at 536 BMP sites. In a minority of cases, a reference sample stands in for the inflow value, depending on the BMP type. The number of observations with reported laboratory detection limits, and the percentage of analytical measurements greater than the detection limit are presented in Table 5.2. Observations that were known to be below the detection limit were set to half the reported detection limit. Where possible, the volume of water associated with the storm event was also paired with the observed water quality value (about 55% of observations) to allow calculated estimates of analyte masses. Each of the following methods described employs and limits this dataset in different ways.

Table 5.2: Crosstable of observations, detection limit ranges and estimated number of samples above a known detection limit. Green values indicate analytes with the greatest fraction of observations above a reported detection limit, red values indicate lower proportions of observations above the recorded detection limit.

Analyte (ug/L)	Reported detection limit ranges			Count of Observations			% Detects (estimated)		
	Inflow	Outflow	Rainwater	Inflow	Outflow	Rainwater	Inflow	Outflow	Rainwater
TSS	0.5 - 45000	0.5 - 40000	500 - 10000	7257	6957	852	99	92	98.9
TDS	100 - 50000	100 - 50000	unreported	2049	1773	122	94	95.3	
TP	0.01 - 1500	0.01 - 2500	5 - 500	7207	6758	1002	96.4	95.2	99.6
PO4 as P	0.003 - 2500	0.003 - 2500	1.6 - 225	2798	2526	358	82.1	86.4	86.3
Dissolved P	2 - 500	2 - 500	10 - 500	2342	2237	399	91.9	93.1	91.6
TN	2 - 500	2 - 500	100 - 100	3441	3204	653	97.6	99.8	100
NOx as N	0.05 - 25000	1.3 - 10000	4 - 10000	6186	5749	930	94.8	92	95.8
Total Cd	0.004 - 500	0.0006 - 10	0.02 - 2	2827	2619	636	55.1	39.3	24.4
Dissolved Cd	0.01 - 10	0.01 - 10	0.05 - 0.5	1817	1604	160	33.3	31.2	18.1
Total Cr	0.1 - 50	0.1 - 250	0.5 - 5	1804	1573	194	74	63.6	62.1
Dissolved Cr	0.1 - 1000	0.1 - 1000	0.1 - 2	1449	1184	146	65.6	65.9	50.7
Total Cu	0.01 - 100	0.002 - 100	0.1 - 10	4492	4348	863	93.4	87.7	82.6
Dissolved Cu	0.01 - 10	0.01 - 10	0.1 - 10	2921	2755	235	92.5	92.6	100
Total Pb	0.003 - 73	0.002 - 73	0.08 - 10	3838	3649	801	78.1	58.9	53.7
Dissolved Pb	0.003 - 73	0.003 - 73	0.05 - 2	2208	2040	153	33.7	33.5	8.2
Total Zn	0.005 - 500	0.003 - 500	0.5 - 50	5114	4987	902	95.3	88.7	75.8
Dissolved Zn	0.005 - 30	0.005 - 30	0.5 - 10	2916	2767	222	92.6	86.2	88.5

5.2.2 Method 1: Estimates of concentration and mass flux by BMP Type

The dataset for Method 1 was limited to BMPs with at least 3 unique observations, and chemical analytes represented by at least 3 BMP sites and a pool of 100 total observations

or more. The trimmed Method 1 dataset contained 94,909 observations of 14 different BMP types reporting results for between 2 and 11 analytes (Table 5.3). This subset is referred to as the ‘clean’ starting water quality dataset for Methods 1a and 1b, summarized in Table 5.3.

Table 5.3: Summary of data coverage for BMP types with at least 100 pooled observations at 3 or more unique BMP sites. Each site had at least 3 local observations.

BMP Type	n analytes	n sites
Grass strip biofilter (BI)	11	44
Bioretention (BR)	10	53
Grass swale biofilter (BS)	9	39
Detention basin (DB)	10	39
Sand media filter (FS)	10	26
Green roof (GR)	2	11
Hydrodynamic separator (HDS)	9	27
High rate biofilter (HRBF)	2	6
High rate media filter (HRMF)	8	18
Oil and grit separator (OGS)	4	15
Modular block porous pavement (PM)	5	15
Retention pond (RP)	11	67
Wetland basin (WB)	7	32
Wetland channel (WC)	8	14

Method 1a: Concentration estimates. Cleaned concentration measurements of each analyte and sample fraction were grouped by BMP type and sampling location (inflow, outflow), but remained unpaired at the event and site levels. The distributions of concentrations were right-skewed and bounded at zero; these values were normalized using a log base 10 transformation and a small amount of jitter introduced as a tiebreaker for identical values. The log concentration distributions of analytes monitored at inflow and outflow stations were tested for mean equivalence using an unpaired non-parametric two-sample Wilcoxon test. The estimate from the non-parametric unpaired two-sample Wilcoxon test for log transformed concentrations represents the order of magnitude (OM) change associated with treatment by each BMP type. A 95th percentile confidence interval range for the mean OM change estimate was calculated for each analyte sample fraction

concentration, along with its corresponding p-value. The OM confidence interval bounds were used to calculate a percent reduction value range as described in Chapter 4.

Next, the distribution of (untransformed, jittered) concentration observations was resampled 10,000 times using a non-parametric bias-corrected and accelerated (bca) bootstrap (npctest package, method ‘bca’) to find the median inflow and outflow values for each BMP type and a credible interval representing 90th percentile of the median bootstraps. This method is similar to the 2020 BMP Database summary methods, which also used a bca bootstrap. The credible intervals were compared for median inflow and outflow concentrations to assess whether the median ranges overlapped.

Method 1b: Mass flux estimates. The cleaned concentration dataset from Method 1a was matched with hydrologic flux (stormwater flow volume) measurements at the same station during the same event using cleaned hydrologic data from Chapter 4 (about 38% of the water quality data were missing event-matched volumes). The missing hydrologic values were each imputed 5 times (using the Amelia package), and the five imputed datasets were used to create five estimates of analyte mass flux in grams for each of the rows in the clean Method 1 concentration dataset (see Chapter 4 methods for more discussion about multiple imputation). An estimate of the mean order of magnitude (OM) change mass flux for each analyte sample fraction was calculated using the same process as Method 1a. Rubin’s rules for pooling confidence intervals do not apply for nonparametric tests, so the confidence interval range estimates were pooled crudely by calculating the mean of the bounds. The median of the five p-values was chosen as a representative estimate of significance.

Next, each of the five imputed mass flux datasets was resampled using the same bca bootstrapping procedure described in Method 1a to calculate the credible interval of the median mass flux of each analyte sample fraction at each BMP type. The median mass flux estimates were pooled by averaging them as described for imputed concentration

datasets above. A simple difference-in-differences estimate was made between the concentration estimates (M1a) and the mass flux estimates (M1b) and used to compare and group which BMP technologies show a boosted mass removal (or export) effect in comparison to their estimate of concentration effects for each analyte. Difference in differences is a simple difference between mass flux and concentration OM change. The DID estimates were used to qualitatively group technologies into 4 groups that have 1) net mass capture associated with a hydrologic treatment effect on water quality through volumetric runoff reduction, 2) an effect on concentration, with equal or lesser effect on mass flux 3) no effect on mass flux or concentration, and 4) mass export.

5.2.3 Method 2: Dissolved fraction estimates

The cleaned concentration measurements for Method 2 were limited to only those measurements with a total concentration value above the reported detection limit, then total and dissolved fractions were paired for each event. This method has a more limited subset of data, because only observations with both paired measurements were included in the analysis ($n = 27,453$ paired observations). Next, a simple ratio was calculated to estimate the dissolved fractional ratio (dissolved concentration \div total concentration) during each event. The precision of the observations varied by laboratory and study, pairs of values that were imprecise enough to estimate the dissolved concentration above the total concentration value were rounded to 1, which means they were interpreted to be samples where the dissolved fraction represented the entire total concentration. The median dissolved fractional inflow and outflow ratios were bootstrapped using the same process as Method 1a. Then the highest probability density representing 51% of the data (shortest half) and median ranges were compared to assess whether the median dissolved fractional ratios overlapped.

Next, the single point and range summary statistics were plotted on the inflow and outflow empirical cumulative distribution functions (for each BMP Type and each chemical

sample fraction with more than 20 observations. The empirical cumulative distribution functions (ECDFs) were tested for equivalence using a nonparametric Kolmogorov-Smirnov test, a measure of the maximum vertical distance between two CDFs. Mean differences in inflow and outflow dissolved fractional ratio were calculated to estimate the shift between particulate and dissolved forms of each chemical analyte during stormwater treatment by each BMP Type.

Lastly, a simple difference-in-differences estimate was made between the effect of each technology on treating solids versus all other contaminants. As in Method 1, this value was used to qualitatively assess whether the treatment of each trace metal or nutrient contaminant was greater or less than the treatment of solids for a given BMP type. This qualitative grouping represents technologies that have a chemical, biochemical or biological treatment effect on water quality through one or more of the solubility controls from Figure 5.1.

5.2.4 Method 3: Matched difference-in-difference estimates for individual structural features.

The inflow and outflow observations from the cleaned concentrations dataset were paired by event and then joined with the same structural features dataset from Chapter 4. Next, the paired inflow and outflow concentration observations were used to calculate the OM change in concentration for all storm events with complete data. The DAGs and model specifications from Chapter 4 were used to duplicate the matching process used in Model Sets 3, 4 and 5 and control for the same sets of confounders in order to infer the causal effects of soil media amendments, liners, standing water and vegetation on individual water quality parameters. Next, the data were matched using the Matchit package in R using the coarsened exact matching method and a logit link. The matched datasets were used within a difference-in-differences regression model framework to determine the causal effect estimates for each of the four structural features. A separate model was used for each

contaminant, but only one model specification was used to estimate the causal effects of each feature. Using regression for difference-in-difference estimation results in the same effect estimates as simple subtraction:

$$DD_{treatment} = (T_{outflow} - T_{inflow}) - (U_{outflow} - U_{inflow}) \quad (1)$$

Where T represents concentrations of contaminants observed at sites in the treatment group, and U represents concentrations of contaminants observed at untreated sites. In this context, treated sites include the structural feature of interest: subsurface liner, standing water, soil media amendment or vegetation. Using regression to make the estimation supports automated calculation of associated standard errors and p values. Details for modeling each feature follow below.

Soil media. To block confounders on the effect of soil media amendments, the media treatment features (amended, unamended) were matched on the presence of a liner, the log area of the site, and the log influent concentrations (see Appendix C.2 for details). After matching, difference-in-differences regression was used to estimate order of magnitude change as predicted by presence of soil media amendment.

Liners. To block confounders on the effect of liners, the liner treatment features (lined, unlined) were matched on the log area of the site, and the log influent concentrations (see Appendix C.2 for details). After matching, difference-in-differences regression was used to estimate order of magnitude changes as predicted by presence of a liner for each contaminant.

Ponding. To block confounders on the effect of standing water, the ponding treatment feature (no standing water, intermittent ponding, permanent ponding) was binarized into two groups (no standing water, any ponding) then matched on the presence of a liner, underdrain type, and the log influent concentrations (see Appendix C.3 for details). After matching, difference-in-differences regression was used to estimate order of magnitude

changes as predicted by presence of standing water for each contaminant (using the original three factor levels in the ponding feature).

Vegetation. To block confounders on the effect of planted vegetation, the vegetation treatment feature (unplanted, grass and sedum, shrubs and trees) was binarized into two groups (unplanted, any planting) then matched on the binarized ponding feature (from method above), presence of amended media, and the log influent concentrations (see Appendix C.4 for details). After matching, difference-in-differences regression was used to estimate order of magnitude changes as predicted by presence of planted vegetation (using the original three factor levels from the vegetation feature).

5.3 Results and Discussion

The following discussion relies results presented in Appendices C, D, E, F and G.

5.3.1 Total suspended and total dissolved solids

Results of the paired median bootstrap ranges for TDS and TSS concentration and mass estimates are presented numerically in Table D.1 and visually for each BMP type in the pair plots in Figure D.1. Non-parametric two-sample test estimates of the order of magnitude (OM) change in concentration and mass flux, and the corresponding percent reduction ranges derived from the confidence intervals around the OM change for each BMP type are presented in Tables E.1 and E.2. These tables also contain text describing the interpreted difference-in-difference results, qualitatively grouped into four categories: 1) net removal due to reduction in concentration (green), 2) a boost in net removal likely caused by volumetric reduction (blue), 3) net export due to hydrologic or other processes (yellow), or 4) no net effect (gray). The distributions of dissolved fractional ratios for raw stormwater and for each BMP type are reported in Figures 5.2 and F.1 and the shifts between inflow and outflow for each technology is summarized in Table 5.4. Effects of individual structural features are represented in Table 5.5 and visualized in Figure G.6.

All BMP types display reduced concentrations of TSS in effluent compared with untreated stormwater. Most technologies also effectively reduce some portion of influent TSS mass, notably constructed wetlands are the only type that may or may not be effective sinks. This pattern for constructed wetlands is likely site dependent and may involve complex surface and groundwater inflows or interaction with urban streams. Effective removal mechanisms that were observed for suspended solids include physical filtration or straining (sand media filters), settling (retention ponds), sedimentation and saltation across grassy or mulched basin surfaces (detention basins, bioretention, grass biofilter strips). Catch basin inserts and hydrodynamic separators show reduction of TSS concentrations,

but surprisingly do not appear to shift the ratio of particulates and dissolved solids. Technologies that appear to display volumetric reduction effects that boost net capture are: grassed bioswales and biofilter strips, porous pavement, bioretention and high-rate biofiltration. The matched difference in difference estimates of the effects of individual structural features on TSS were mixed. Soil media amendments (physical filtration and straining) decrease TSS concentrations by 30-40% ($p \ll 0.001$). Subsoil geotextile and clay liners cause no effect or marginally slight increase in TSS concentration (5-10%, $p = 0.0503$). Intermittent ponding causes a decrease in TSS concentration by 35-45% ($p \ll 0.001$) and permanent ponding a slightly smaller reduction of 25-35% ($p \ll 0.001$). Sedum and low grassy vegetation has little or no effect on TSS concentration, but taller plantings of shrubs and trees do reduce TSS concentrations by 25-40% ($p \ll 0.001$).

In contrast, most stormwater technologies show little effect on TDS concentrations or mass flux. Three features showed a causal effect that reduced TDS concentrations: shrubs and trees (30-50%, $p = 0.001$) and both permanent and intermittent ponding (25-40% and 25-35% respectively, both $p \ll 0.001$). Organic leaf litter could affect TDS concentrations directly through adsorption/protein binding or indirectly affect TDS concentrations through redox or pH pathways. The unknown mechanism behind the ponding effects could be related to redox, biological or chemical factors. It seems counter-intuitive that ponding shows a causal reduction of TDS while retention ponds show increased TDS concentrations in effluent, as do sand media filters. The net export of dissolved solutes on a mass basis is likely due to evaporative concentration in between precipitation events or groundwater influx during the observed storm event period.

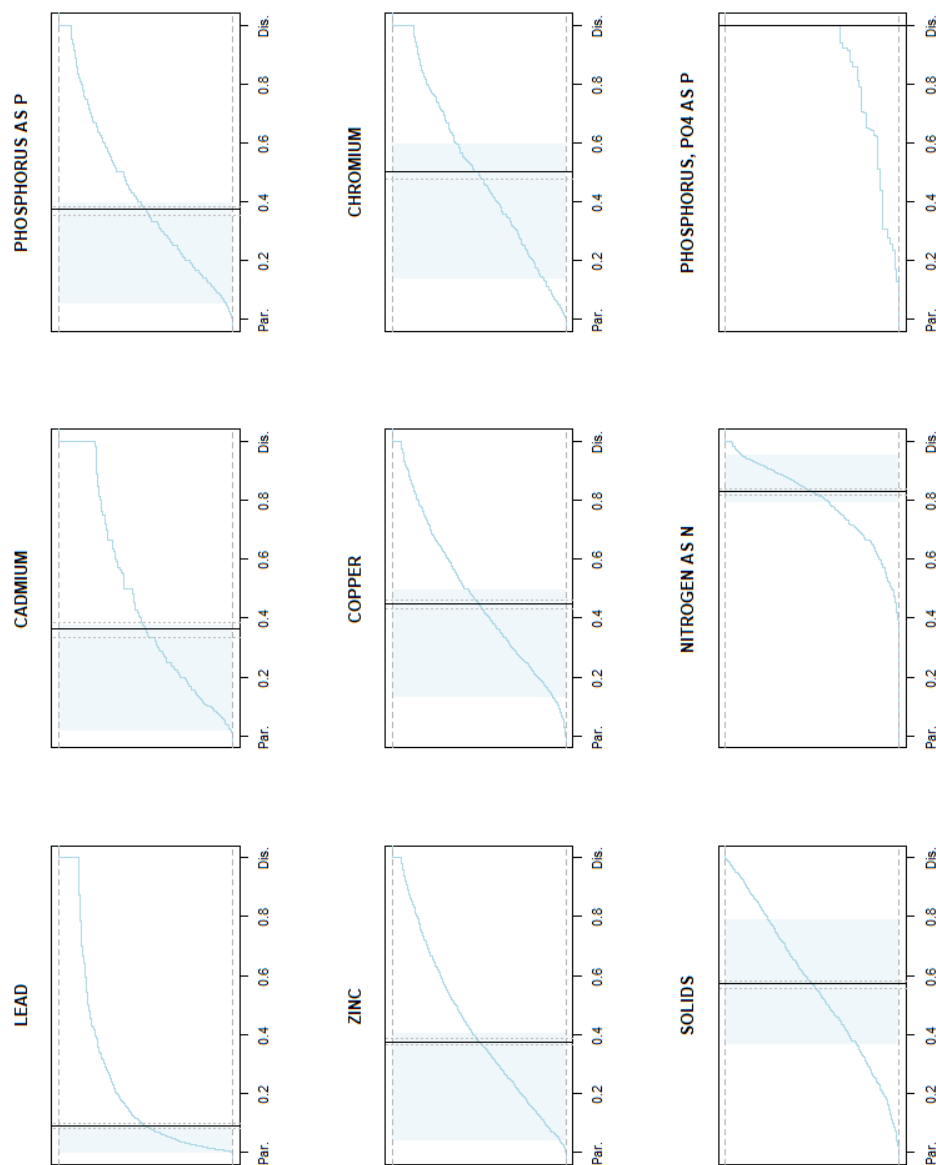


Figure 5.2: Empirical cumulative distribution functions for the ratio of dissolved fractions of contaminants to total measurements in untreated stormwater. Values on the left represent 100% particulate fraction, right indicates 100% dissolved. Bootstrapped median ranges are marked with vertical lines; shaded ranges represent the shortest half of the data distribution.

Table 5.4: Treatment effects on dissolved fractional ratios of solids, trace metals and total phosphorus. Number indicates the mean shift between the inflow ecdf and the outflow ecdf for each contaminant and treatment technology. Dissolved fractional ratio color gradient uses green to indicate reduction of particulate matter (increase in DFR), and red to indicate higher proportion of particulates (decreased DFR), non significant results are labeled "ns". Difference in differences (DID OM estimate) color gradient uses red to indicate worse removal performance compared to solids estimate, blue to indicate better.

BMP Type	Dissolved Fractional Ratio Shift (Inflow to Outflow)							Difference in Differences (vs Solids)						
	Solids	Cd	Cr	Cu	Pb	Zn	TP	Cd	Cr	Cu	Pb	Zn	TP	
media filter (sand) FS retention pond RP grass filter strip BI	-0.40	-0.33	-0.23	-0.27	-0.37	-0.13	-0.20	6.9%	17%	13%	3.0%	27%	21%	
	-0.26	-0.28	-0.24	-0.19	-0.22	-0.15	-0.11	-2.5%	2.1%	7.2%	3.7%	11%	15%	
	-0.19	-0.10	-0.12	-0.17	-0.054	-0.085	-0.25	9.5%	7.9%	2.9%	14%	11%	-5.8%	
high rate media filter HRMF bioretention BR constructed wetland basin WB	-0.18	ns -0.064	ns -0.049	-0.11	ns -0.056	-0.14	ns -0.040	12%	13%	6.8%	12%	3.6%	14%	
	-0.17	-0.26	-0.29	-0.33	-0.23	-0.34	-0.32	-8.6%	-12%	-16%	-6.0%	-17%	-15%	
	-0.15	ns -0.26		-0.20	ns -0.14	ns -0.043	0.034	-10%		-4.5%	1.4%	11%	19%	
detention basin (grass) DB grass swale BS catch basin insert CBI	-0.14	ns -0.14	-0.20	-0.25	-0.12	-0.23	-0.13	0.10%	-6.4%	-12%	2.1%	-8.9%	1.1%	
	-0.08	ns -0.10	ns -0.065	-0.12	ns -0.080	-0.080	-0.13	-2.0%	1.8%	-3.6%	0.3%	0.3%	-4.9%	
	ns -0.066	0.080	0.053	-0.21	0.33	-0.14	ns +0.080	15%	12%	-14%	40%	-7.4%	15%	
hydrodynamic separator HDS pervious concrete PC constructed wetland channel WC	ns -0.036	ns -0.025	ns -0.011	ns -0.014	ns -0.022	ns -0.056	ns +0.029	1.0%	2.5%	2.2%	1.4%	-2.0%	6.5%	
	ns -0.028			ns -0.0068	ns 0.092	ns +0.011				2.1%	12%	3.9%		
	ns -0.016		ns 0.066	ns +0.022	ns -0.026	ns -0.088	ns -0.074		8.2%	3.9%	-1.0%	-7.2%	-5.7%	
concrete detention basin DC media filter (peat and sand) FP green roof GR				ns -0.056	ns -0.0013	ns -0.00076	ns -0.14							
				ns -0.17	-0.45	ns -0.065								
			ns 0	ns -0.016		ns -0.065	ns 0							
high rate biofilter HRBF oil and grease separators/baffles OGS modular block porous pavement PM		ns -0.100		-0.23		ns -0.10	ns -0.11							
				ns -0.078	ns -0.14	ns -0.12	ns -0.12							
			-0.24	ns -0.049	-0.16	ns -0.065	-0.26							
underground retention vault RV				ns -0.020		ns -0.073								

Table 5.5: Structural treatment effects on TSS, TDS, TN, NO_x, TP, DP, and reactive P.

Analyte	Sample Fraction	Causal OM			% Reduction	Feature
		Estimate	Std error	p value		
TDS	DISSOLVED	-0.219	0.066	0.001	30 - 48	* shrubs and trees
		-0.177	0.033	0.000	28 - 38	* permanent ponding
		-0.159	0.029	0.000	26 - 35	* intermittent ponding
		-0.053	0.036	0.138	4 - 19	media
		0.001	0.027	0.970	no effect	liner
TSS	SUSPENDED	0.018	0.028	0.512	no effect	grass and seedum
		-0.238	0.028	0.000	38 - 46	* intermittent ponding
		-0.194	0.022	0.000	33 - 39	* media
		-0.164	0.044	0.000	24 - 38	* shrubs and trees
		-0.156	0.027	0.000	26 - 34	* permanent ponding
NITROGEN AS N	TOTAL	0.030	0.020	0.144	ns export	grass and seedum
		0.044	0.022	0.050	ns export	liner
NITROGEN, NOX AS N	TOTAL	-0.258	0.033	0.000	40 - 49	* permanent ponding
		-0.218	0.034	0.000	34 - 44	* intermittent ponding
		-0.053	0.020	0.009	7 - 16	* grass and seedum
		-0.037	0.019	0.051	4 - 12	media
		-0.010	0.020	0.611	no effect	liner
PHOSPHORUS AS P	TOTAL	0.046	0.030	0.122	ns export	shrubs and trees
		-0.360	0.025	0.000	54 - 59	* permanent ponding
		-0.243	0.024	0.000	40 - 46	* liner
		-0.213	0.023	0.000	36 - 42	* grass and seedum
		-0.131	0.043	0.002	18 - 33	* shrubs and trees
PHOSPHORUS AS P	TOTAL	0.089	0.027	0.001	export	* intermittent ponding
		0.214	0.025	0.000	export	* media
DISSOLVED	TOTAL	-0.417	0.044	0.000	58 - 65	* permanent ponding
		-0.171	0.050	0.001	24 - 40	* intermittent ponding
		-0.098	0.085	0.250	3 - 34	shrubs and trees
		0.013	0.029	0.655	no effect	grass and seedum
		0.013	0.034	0.700	no effect	media
PHOSPHORUS AS P	TOTAL	0.072	0.037	0.052	ns export	liner
		-0.472	0.023	0.000	64 - 68	* permanent ponding
		-0.290	0.023	0.000	46 - 51	* intermittent ponding
		0.017	0.019	0.365	no effect	media
		0.037	0.020	0.055	ns export	liner
PHOSPHORUS AS P	TOTAL	0.056	0.035	0.116	ns export	shrubs and trees
		0.131	0.018	0.000	export	* grass and seedum
		-0.445	0.062	0.000	59 - 69	* permanent ponding
		-0.074	0.058	0.202	4 - 26	intermittent ponding
		-0.029	0.054	0.585	no effect	liner
PHOSPHORUS AS P	TOTAL	0.280	0.060	0.000	export	* media
		0.380	0.051	0.000	export	* grass and seedum
		0.400	0.073	0.000	export	* shrubs and trees

5.3.2 Nutrients

Results of the paired median bootstrap ranges for nitrogen (TN, NO_x) concentration and mass estimates are presented numerically in Tables D.2 and visually for each BMP type in the pair plots in Figure D.2. Non-parametric two-sample test estimates of the order of magnitude (OM) change in concentration and mass flux, and the corresponding percent reduction ranges derived from the confidence intervals around the OM change for each BMP type are presented in Tables E.3 and E.4. These tables also contain the interpreted difference-in-difference results described earlier. Phosphorus results (TP, DP and reactive P) are presented numerically in Appendix D.3, and visually in Appendix D.3, with OM change and percent reduction ranges presented in Appendix E.5 and E.6. Dissolved and total fractions for TN and TP were available for a limited set of BMPs, the ECDFs showing the shift in distributions of dissolved fractional ratios is presented in Figure F.2 and summarized in Table 5.4. Individual structural feature effects are represented in Table 5.5 and visualized in Figure G.6.

Nitrogen in stormwater is mostly dissolved (dissolved fractional ratios of raw stormwater are at least 80% of total N load, Figure 5.2). Most treatment technologies either decrease total nitrogen concentrations in effluent or have no effect. For those that decrease TN concentrations in effluent (sand media filters, constructed wetland basins, bioretention, bioswales and retention ponds), the reduction is moderate, about 0.10 OM or 20% reduction, not a very large chemical/biochemical effect. Bioretention and constructed wetland basins reduce TN mass flux, but based on their nitrate mass flux, they likely nitrify TN into nitrate, as do sand media filters and grass biofilter strips. The total mass flux of nitrate in all of these systems is net neutral (no effect), but this is most likely due to offset by volumetric reduction. The systems that export the highest concentrations of nitrate are porous pavement and sand media filters, but the highest net export by mass is from constructed wetlands due to very high volumes of influent water, which may include

groundwater influx. Both retention and detention ponds are moderately effective in decreasing nitrate concentrations, likely through denitrification. Green roofs and grass biofilter strips also decrease nitrate, but do not receive large runoff volumes or high N concentrations, so their application is somewhat limited for stormwater nutrient abatement.

Permanent and intermittent ponding both show a strong causal effect on decreasing concentrations of TN in stormwater at the event scale. Permanent ponding reduces TN by 40-50% ($p \ll 0.001$); intermittent ponding by 35-45% ($p \ll 0.001$); and grass and sedum by a small amount ($p = 0.009$). As discussed in the previous paragraph, the nitrogen cycle is complex, and there are multiple mechanisms simultaneously occurring in these stormwater treatment systems. Where permanent ponding causes 55-60% nitrate reduction, intermittent ponding and media amendments cause increases in NO_x concentrations (both $p \leq 0.01$). It is apparent that technologies with rapid draw-down and very dry conditions in the vadose zone between storm events affect nitrification and net export of NO_x , whereas systems that drain intermittently and remain moist or saturated facilitate N removal by denitrification. This pattern is clear because, in addition to permanent ponding, liners and vegetation also both decrease nitrate concentrations in effluent. These features point to moisture as a key facilitator of the appropriate redox conditions and biological activity driving denitrification. Nutrient-poor mulch and vegetative shading can be used to even out evapotranspiration over the inter-event dry period, which can increase denitrification and decrease nitrate export.

Phosphorus in stormwater is primarily particulate (dissolved fractional ratios are about 35% of total P load). Although not statistically significant, there is a consistent decrease in mass of dissolved and total phosphorus across all technologies, about 0.1 OM, or 20% reduction. Many technologies achieve net TP capture, and the mechanisms may be diverse: particle settling or sedimentation (porous pavement, constructed wetland basins, retention ponds), physical filtration or straining (sand media filters), as well as soil adsorption and

biological uptake. Bioretention, grass biofilters and constructed wetland channels show evidence of mobilizing total reactive P, possibly due to fertilizer application or mobilization of iron due to reducing conditions. Notably, although bioretention, grass biofilter strips and detention basins boosted capture of TP and reactive P via volumetric reduction, and shifted TP from particulate to dissolved fractions, neither show evidence of net mass capture. This is consistent with the causal estimates: permanent and intermittent ponding cause very good decreases of TP and DP, but only permanent ponding reduces PO_4 concentrations during the event scale timeframe. This is likely because intermittent ponding activates the redox cycle of alternating reducing and oxidizing conditions, causing mineralization and demineralization of phosphate with iron. Permanent ponding may support biological intervention or sedimentation after complexing with colloids in the water column. Technologies with vegetation tended to perform better at shifting the dissolved fractional ratio as compared with removal of solids. There is a causal export associated with plantings, but it is unknown whether that is due to fertilizer application or the vegetation itself. Future studies of phosphorus dynamics should collect data from the total and dissolved fractions, and soluble reactive phosphorus should be measured with a detection limit no higher than $3 \mu\text{g}/\text{L}$.

Vegetation and microbiota are well-known to immobilize various forms of N and P, therefore it is somewhat surprising that there is not more biotic uptake observed for these essential nutrients. However, it would be erroneous to conclude that bioretention and other vegetated stormwater control structures do not have significant effects on nutrient capture. The event-based sampling timescale is likely too short to observe the biotic effects associated with plants and soil microbiome. There is relevant selection interval bias in the choice of precipitation events—most plants utilize nutrients for growth during sunny weather, not during the rain events used in this dataset. It is likely that the effects reported here underestimate the nutrient capture value of vegetation, and highlight the importance of increasing both hydrologic capture volume and hydraulic retention time of stormwater within GI systems.

5.3.3 Trace metals

Results of the paired median bootstrap ranges for trace metal (Cd, Cr, Cu, Pb, Zn) concentrations and mass estimates are presented numerically in Appendices D.4 through D.8, and visually for each BMP type in the pair plots in Appendices D.4 through D.8. Non-parametric two-sample test estimates of the order of magnitude (OM) change in concentration and mass flux, and the corresponding percent reduction ranges derived from the confidence intervals around the OM change for each BMP type are presented in Tables E.7 through E.13. These tables also contain the interpreted difference-in-difference results described earlier. Dissolved and total fractions for trace metals were available for limited sets of BMPs, the ECDFs showing the shift in distributions of dissolved fractional ratios are presented in Figures F.3 through F.7 and summarized in Table 5.4. Individual structural feature effects are represented in Table 5.6 and visualized in Figure G.6.

The distribution of dissolved and total trace metals observed in this stormwater dataset spans 3 to 3.5 orders of magnitude (Cd: 10^{-2} to 10^1 $\mu\text{g/L}$; Cr: 10^{-2} to 10^1 $\mu\text{g/L}$; Cu: $10^{-0.5}$ to $10^{2.5}$ $\mu\text{g/L}$; Pb: 10^{-1} to $10^{2.5}$ $\mu\text{g/L}$; Zn: 10^0 to 10^3 $\mu\text{g/L}$). The concentrations of observed trace metals were often below the EPA's drinking water standards, with the notable exception of lead, which has a drinking water standard set at 0 $\mu\text{g/L}$ (no known safe level). The proportion of storm events with concentrations above toxicity limits that are known to be ecologically damaging is between 25% and 85% for any of these 5 trace metals individually. On average, the cumulative probability of observing any one trace metal at concentrations exceeding EPA recommended drinking water limits is about 45%; the probability of exceeding ecologically harmful levels is approximately 99% (see Figure 5.3). Trace metals results are often below the method detection limit, especially for dissolved fractions.

Unsurprisingly, trace metals are very sensitive to reduction-oxidation cycling caused by soil moistening and drying cycles. Physical filtration (sand media filters, grass biofilter

strips and bioretention) and settling (retention ponds) are effective methods for particulate trace metals removal, and volumetric reduction boosts net capture of many types.

Volumetric reduction is also effective for net capture even when influent concentrations are low, and is the only mechanism for capturing dissolved lead. When compared to total solids capture rates, vegetation shows an additional biological or biochemical removal effect for all five trace metals types. Roadway-adjacent grass strip biofilters receive some of the highest concentrations of Cu, Zn, and Cr, and also show good net capture. Creative vegetation maintenance of highway systems may be able to increase their efficacy or retrofit new stretches of high speed, high traffic roads.

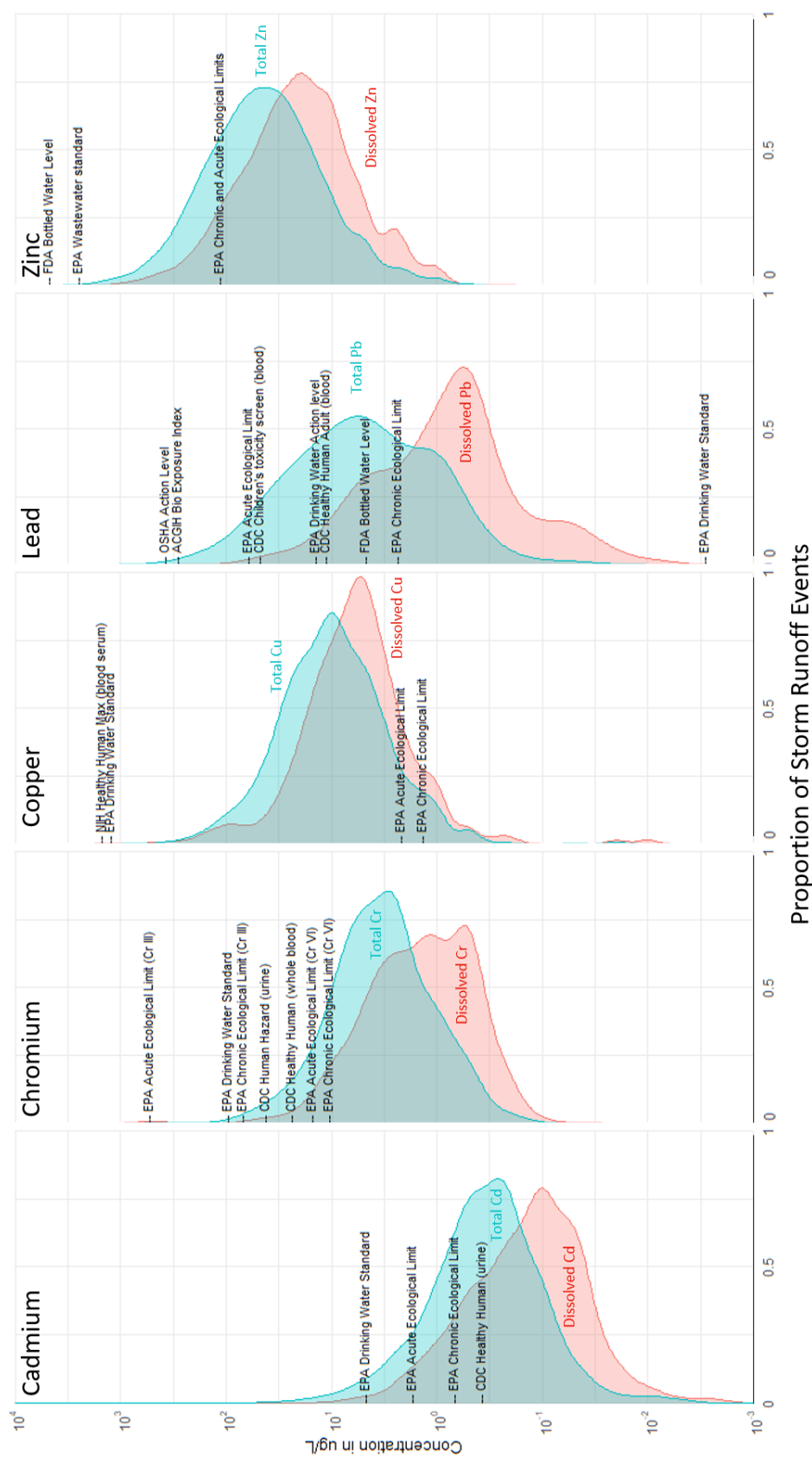


Figure 5.3: Trace metal concentration ranges in untreated stormwater and relevant toxicity labels for context. Toxicity levels are from EPA National Primary Drinking Water Regulations, EPA National Recommended Water Quality Criteria Freshwater Aquatic Life Criteria, CDC Agency for Toxic Substances and Disease Registry Toxic Substances Portal and Case Studies.

Table 5.6: Summary of structural treatment effects on total and dissolved Cd, Cr, Cu, Pb and Zn.

Analyte	Sample Fraction	Causal OM				Causal OM						
		Estimate	Std error	p value	% Reduction	Feature	Estimate	Std error	p value	% Reduction	Feature	
CADMIUM	DISSOLVED	-0.067	0.024	0.006	9 - 19	* grass and seedum	LEAD	-0.027	0.027	0.313	no effect	intermittent ponding
		-0.022	0.024	0.359	no effect	intermittent ponding		-0.005	0.020	0.817	no effect	grass and seedum
		0.015	0.023	0.525	no effect	liner		0.007	0.035	0.848	no effect	permanent ponding
		0.022	0.061	0.725	no effect	shrubs and trees		0.009	0.032	0.782	no effect	media
		0.028	0.042	0.503	no effect	permanent ponding		0.010	0.021	0.652	no effect	liner
	TOTAL	0.046	0.035	0.181	ns export	media		0.031	0.057	0.579	no effect	shrubs and trees
		-0.023	0.025	0.375	no effect	liner		-0.181	0.073	0.014	22 - 44	* shrubs and trees
		0.009	0.061	0.883	no effect	shrubs and trees		-0.136	0.029	0.000	22 - 32	* media
		0.026	0.027	0.332	no effect	permanent ponding		-0.093	0.027	0.001	14 - 24	* permanent ponding
		0.045	0.022	0.039	export	* grass and seedum		-0.088	0.029	0.003	13 - 24	* intermittent ponding
CHROMIUM	DISSOLVED	0.053	0.028	0.054	ns export	media	ZINC	0.020	0.023	0.384	no effect	grass and seedum
		0.060	0.029	0.037	export	* intermittent ponding		0.052	0.025	0.034	export	* liner
		-0.267	0.065	0.000	37 - 53	* shrubs and trees		-0.122	0.025	0.000	20 - 29	* grass and seedum
		-0.150	0.033	0.000	23 - 34	* intermittent ponding		-0.095	0.063	0.129	7 - 30	shrubs and trees
		-0.128	0.023	0.000	22 - 29	* grass and seedum		-0.086	0.026	0.001	13 - 23	* media
	TOTAL	-0.092	0.025	0.000	14 - 23	* liner		-0.080	0.029	0.006	11 - 22	* intermittent ponding
		-0.059	0.232	0.799	no effect	permanent ponding		-0.025	0.025	0.317	no effect	liner
		0.001	0.032	0.987	no effect	media		0.065	0.032	0.041	export	* permanent ponding
		-0.133	0.031	0.000	21 - 31	* permanent ponding		-0.173	0.047	0.000	25 - 40	* shrubs and trees
		-0.130	0.034	0.000	20 - 32	* intermittent ponding		-0.094	0.023	0.000	15 - 24	* intermittent ponding
COPPER	DISSOLVED	-0.043	0.026	0.096	4 - 15	liner	TOTAL	-0.079	0.021	0.000	13 - 20	* media
		0.037	0.026	0.161	ns export	grass and seedum		-0.075	0.018	0.000	12 - 19	* grass and seedum
		0.065	0.032	0.042	export	* media		-0.039	0.021	0.054	4 - 13	liner
		0.125	0.367	0.734	ns export	shrubs and trees		0.012	0.023	0.580	no effect	permanent ponding
		TOTAL	-0.109	0.038	0.004	15 - 29	* shrubs and trees	TOTAL	-0.075	0.018	0.000	12 - 19
	-0.101		0.023	0.000	16 - 25	* permanent ponding	-0.039		0.021	0.054	4 - 13	liner
	-0.052		0.021	0.015	7 - 15	* intermittent ponding	0.012		0.023	0.580	no effect	permanent ponding
	-0.039		0.016	0.018	5 - 12	* grass and seedum	-0.075		0.018	0.000	12 - 19	* grass and seedum
	-0.010		0.017	0.571	no effect	liner	-0.039		0.021	0.054	4 - 13	liner
	0.017	0.019	0.366	no effect	media	0.012	0.023		0.580	no effect	permanent ponding	
TOTAL	-0.177	0.039	0.000	27 - 39	* shrubs and trees	TOTAL	-0.075	0.018	0.000	12 - 19	* grass and seedum	
	-0.062	0.018	0.001	10 - 17	* permanent ponding		-0.039	0.021	0.054	4 - 13	liner	
	-0.058	0.016	0.000	9 - 16	* liner		0.012	0.023	0.580	no effect	permanent ponding	
	-0.056	0.019	0.003	8 - 16	* intermittent ponding		-0.075	0.018	0.000	12 - 19	* grass and seedum	
	-0.001	0.015	0.937	no effect	grass and seedum		-0.039	0.021	0.054	4 - 13	liner	
0.027	0.019	0.151	ns export	media	0.012		0.023	0.580	no effect	permanent ponding		

5.4 Conclusions and recommendations

Nearly all storm events create runoff with contaminant concentrations above limits that are known to cause ecological degradation and/or harm to human and aquatic life. Special attention should be paid to treating runoff in locations where roadways cross streamlines. BMP designers should be cognizant of the effects of wetting and drying cycles on the redox sequence, which drives effects on many contaminants simultaneously. Some contaminants are largely unaffected by green or gray stormwater infrastructure systems, notably total dissolved solids and dissolved lead concentrations. These two types of contamination may benefit from watershed-level policy changes, rather than catchment-scale treatment.

5.4.1 Conclusions from estimation of structural feature effects

Features with the greatest influence on suspended solids reduction at the event scale are:

- intermittent ponding: -0.2 orders of magnitude (OM), equivalent to 38 to 46% TSS concentration reduction
- media amendment: -0.2 OM; 33 to 39 % TSS concentration reduction
- shrubs and trees: -0.16 OM; 24 - 38% TSS concentration reduction

Redox cycling occurs during wetting and drying cycles and has a strong influence on solubility and mobility of nutrients and trace metals. Permanent and intermittent ponding have the greatest influence on nutrient reduction at the event scale:

- permanent ponding: -0.26 OM, equivalent to 40 to 49% TN concentration reduction
- permanent ponding: -0.47 OM, equivalent to 64 to 68% TP concentration reduction

Vertical flow through soil media profiles that alternately wet and dry has the greatest influence on nutrient export at the event scale:

- soil media: +0.21 OM increase in NO_x concentration

- soil media: +0.28 OM increase in reactive P concentration

Systems with multiple treatment zones that have different hydrologic action in sequence are likely the most effective (these are sometimes called multi-chamber treatment trains). The cause of this is likely decoupled redox cycling patterns in each treatment cell. Vegetation management likely plays a role in whether systems export nutrients, and management choices that affect soil moisture in the near-surface vadose zone will affect trace metals.

Adjustments to design and retrofits should prioritize hydrologic reduction, as it is a driving factor in total contaminant removal for many contaminant types. Roadside biofilter strips, grass swales, and bioretention are very cost-effective ways to capture trace metals and improve event-scale contamination. Additional landscape management experimentation may improve contaminant removal in existing retention, detention and constructed wetland systems.

6 Synthesis and future recommendations

6.1 Interpretation and synthesis of model results

The overarching lesson that emerges from the models in this body of research is that improving urban stormwater hydrology and water quality means adopting structural features that mimic those present in natural ecosystems. Focusing interventions on decreasing the total volume of effluent from stormwater control structures will decrease downstream flooding, channel incision in the urban environment, and improve overall surface water quality. Allowing vegetation and soil composition to develop to maturity will improve both volumetric reduction and biochemical intervention, and mimic natural evaporation and transpiration patterns. Specifically, four natural conditions to emulate in the urban environment are:

1. better surface-to-groundwater exchange,
2. enhanced vadose zone storage,
3. flatter event discharge peaks, and
4. less reactive wetting-drying cycles.

Precipitation in the urban environment quickly outpaces infiltration rates, due to compaction and impermeable paved surfaces, creating a saturated and poorly hydraulically connected surface layer above the vadose zone. This results in 'flashy' runoff patterns, often with little more than 5 mm of rain. Improving the infiltration of surface water through the unsaturated zone and into the water table requires less compact soil with good porosity and continuity between the surface and the groundwater table. Vertical connection can be achieved by using at least 1 meter of amended media that is more porous and less compacted than native soil. Deep-rooted vegetation also provides vertical connectivity from the surface along the rooting tunnels. Roots help prevent compaction by constantly growing and dying, leaving new infiltration pathways through the amended

zone, and into the surrounding native soil. Liners isolate exchange from an amended zone into the natural groundwater table, they should be used specifically for sites that have concerns about contamination in the native soil profile, but otherwise should be used sparingly. Liners should not be used to isolate stormwater from infiltration in locations with a high water table. Liners rarely remain completely watertight, and application as a barrier prevents evaporation from wet sites in between events. Removing the liner allows groundwater to wick up through the amended soil profile and evaporate or transpire. Evapotranspiration between events reduces soil moisture and increases the subsurface storage volume available for stormwater capture during the next event. Media amendments, vegetation and surface mulching help create and maintain less compact soil surface profiles with better continuity with the native soil profile.

In addition to the measures described for improving vertical connectivity, vadose zone storage can also be improved by slightly increasing the latent soil moisture between events, and widening the soil media pore distribution to make it larger and less uniform. Surface layers become baked dry in urban heat islands, creating hydrophobic surface conditions and preventing the wetting front from forming across the soil profile. This results in an infiltration delay, and contributes to flashy discharge peaks. A site that has rapidly infiltrated water during a storm event will continue to seep or wick water into the surrounding soil profile for several days as it equilibrates with the local groundwater table. A site that has surface compaction or hydrophobic crust will remain relatively dry in the vadose zone, generate runoff quickly and have limited subsurface storage. Better shading conditions at the surface, and sorptive leafy carbon matter in the soil profile can help maintain soil moisture slightly above the minimum value necessary for pore moisture to maintain readily wettable conditions and reduce air entrainment. Natural wetlands achieve this condition through repeated seasons of growth and deposition of leafy organic material, which increases soil carbon content. The organic content helps the soil maintain latent moisture between storm events, which promotes prolonged periods of evaporation, and

quick re-wetting. Shading of the infiltrating surface can be achieved with direct planting, stones, mulch, or adjacent overhanging structures/vegetation.

Dense, short-mowed grass can also create a muddy, saturated mat at the surface, preventing infiltration through the unsaturated zone to the water table. Turf aeration, taller/deeper vegetation, reduction of mowing frequency, and mixed species grasses are good ways to create a vadose zone that is quick to infiltrate, increases surface roughness coefficients and initial abstraction of precipitation. It is key to continue maintenance and mowing around the inlet and outlet structures to make sure vegetation does not prevent water from entering from sheetflow or pointflow inlets and prevent outlet clogging. However, the common practice of mowing entire detention basins can be replaced with simply clearing the inlet and outlet structures. Decreased mowing improves soil compaction, vegetation generates more diverse infiltration pathways, and both help to maintain latent moisture during antecedent dry periods and prepare soil pores to fill quickly during a storm event.

Thus far, this discussion has focused primarily on improving subsurface storage and vertical connectivity. However, the discharge peak can also be flattened by increasing surface roughness and vertical 'complexity' above the vadose zone. Trees and shrubs provide excellent surface complexity – they enhance runoff reduction through initial abstraction and delay inflow through throughfall and stemflow. Woody vegetation also increases evaporative losses, which creates additional subsurface pore storage in root zones where shallower vegetation does not grow roots. The role of vegetation cannot be overlooked for most stormwater management sites, even for BMP types that are traditionally planted with short-mowed grass. Landscape management experimentation is key to understanding how best to enhance the hydrologic performance of the many detention basins, retention ponds and vegetated strips that already exist across the United States. Trees may not be appropriate for every location, but it does not mean that

hydrologic performance cannot be enhanced with vegetation. For example, some transportation conduits require mowing of adjacent vegetation to maintain adequate driver visibility. The results of Model Set 5 in Chapter 4 indicate there is likely a good improvement in runoff reduction at a site that is mowed quarterly or annually in comparison with a location that is mowed weekly and will maintain sufficient visibility requirements. The results of Chapter 5 indicate that grass biofilter strips are very effective for water quality improvement. If short grass is necessary to protect adjacent structures or roadways, it is possible the mowing strategy can be adjusted to allow for a vegetation gradient, with the tallest grasses growing 2-5 meters away from structures, particularly on slopes and low-lying areas. At sites with standing water, mowing should never extend to the waterline, but leave as much 'riparian' vegetation as possible (no less than ~ 3 meters). Vegetation gradients and no-mow zones should be demarcated clearly using flags, stones, or other visual aids to prevent mower creep, which occurs when a mowed space grows successively larger with each mowing. Roadside vegetation and biofilter strips are some of the cheapest and most effective ways to reduce trace metals from highway road runoff. Reduced mowing and taller, more complex vegetation may be a simple and cost effective method for low density urban and suburban areas to improve water quality.

In dry climates where evaporative losses are less desirable, vertical connectivity from the surface to the water table and shading of the site surface is equally important. Vertical connections can be artificially induced by using large stones at the surface, and throughout the amended media profile, allowing fast percolation to a depth below the local root zone (usually more than 2-3 meters). These types of infiltration basins should be unlined to promote drainage into the water table. Another alternative approach is to use a closely-spaced field of wooden or bamboo stakes, roughly 10-20 cm apart, and driven 20-50 cm into the vadose zone. The above-ground portion of the field will provide some surface shading, and the lower portion provides some limited artificial stemflow, allowing water to seep into the ground along the interface between the stake and the soil.

6.2 Future Recommendations

- Stormwater control design should prioritize volumetric reduction through: 1) limiting the use of subsurface liners and geotextiles and adding upturned elbow drains; 2) increasing the height, diversity, and density of vegetation; and 3) periodically adding 0.5-1 m of nutrient-poor organic matter or mixed-density soil media.
- Urban precipitation events produce runoff that nearly always contain levels of contaminants that are harmful to aquatic ecosystems and/or human health. To minimize the negative impact of these contaminants, it is crucial to 1) treat the contaminated stormwater as close to the source as possible, and 2) recognize that the most effective treatment will remove all contaminants within the complex and dynamic stormwater mixture. Use of spatial data to identify locally-specific point sources of contamination in areas that are usually managed for non-point source pollution will open new opportunities for treatment. Urban culverts, bridges, and overpasses, for instance, are identifiable intersections between transportation and stream networks that could benefit from biofiltration.
- Water quality is closely tied to local land use, including unhoused and vulnerable populations who may lack access to municipal waste services and regularly experience exposure to urban stormwater. Municipalities may achieve water quality benefits by implementing creative incentives that link housing access with watershed cleanup efforts (e.g., Butler [2023]). Effective stormwater control can be achieved through creative landscape management experiments and retrofits, particularly for traditional retention and detention basins. Incentivizing lower mowing frequency and/or increased vegetation density can be more effective when integrated with local stormwater credit and fee systems across various land use types.

- Constructed wetlands are not sufficient replacements for natural wetlands that have been disturbed. The conservation of existing natural wetlands is more effective in providing stormwater control through ecological services and should be incentivized over constructed wetlands or stream restoration.
- To achieve more accurate estimates of the causal effects of individual features, data collection efforts for stormwater monitoring should take two actions: 1) introduce randomization to reduce observational selection bias in the dataset, and 2) deliberately include reference sites for all relevant design features to increase common support. Additionally, data collection efforts should report a standard set of contextual metadata for all BMP system types. For better design, metadata should help clarify the local interaction between the constructed site and its ecohydrological loss pathways (groundwater hydrology, atmospheric interaction, surface water characteristics). Useful metadata includes: site and watershed area, site flowpath length and site perimeter length, liner material, density and height of vegetation, depth of soil media amendment, permanent ponding depth, freeboard fluxing depth, local hydraulic head elevation, surface aspect and shading.

Appendices

A Data sources used for analysis

A.1 Water budget sources from Chapter 3

Water budget data sources by type; data files are available in the github repository :

<https://github.com/cgeger/WaterBudgetTriangle/tree/master/data>.

Bioretention and Lysimeter estimates and observations. Hess et al. [2015], Kosmerl [2012], Li et al. [2009], Pitt et al. [2007], Strauch et al. [2016]

Bioretention model. Wardynski et al. [2011]

Cistern models and observations. Guizani [2016], Millar et al. [2003], Steffen et al. [2013], Zhang and Hu [2014]

Continental scale estimates. Jones et al. [2012]

Detention basins. Harper et al. [1999b, 2002], Shukla et al. [2015]

Detention basin models. Emerson [2003]

Green roofs. Ahiablame et al. [2012], Berghage et al. [2009], Carpenter and Kaluvakolanu [2011], Carson et al. [2013], Carter and Rasmussen [2005, 2007], Czemieli Berndtsson [2010], Fassman-Beck et al. [2013], Fioretti et al. [2010], Getter et al. [2007], Gregoire and Clausen [2011], Hathaway et al. [2008], Hutchinson et al. [2003], Liu and Minor [2005], Mentens et al. [2006], Moran et al. [2005], Nawaz et al. [2015], Palla et al. [2011], Riley et al. [2009], Stovin [2010], Stovin et al. [2012], Teemusk and Mander [2007], Toronto Regional Conservation Authority [2006], Van Seters et al. [2009], Villarreal and Bengtsson [2005], Voyde et al. [2010], Wadzuk et al. [2013]

Green roof models. Berghage et al. [2007], Stovin et al. [2013], Vanuytrecht et al. [2014]

Blue roofs. Carpenter and Kaluvakolanu [2011], Mentens et al. [2006], VanWoert et al. [2005]

Control roofs. Berghage et al. [2009], Carpenter and Kaluvakolanu [2011], Mentens et al. [2006], Van Seters et al. [2009], Vanuytrecht et al. [2014]

Lab-scale roofs. Buccola and Spolek [2011], Carson et al. [2013], Fassman-Beck et al. [2013], VanWoert et al. [2005], Villarreal and Bengtsson [2005]

Porous pavement and controls. Brown and Borst [2015], Drake et al. [2014], Göbel et al. [2013], Pratt et al. [1995], Rim [2011], Winston et al. [2016]

Porous pavement model. NERR [2016]

Retention ponds. Harper et al. [2003], Harper [2010a,b,c, 2011], Teague et al. [2005]

Retention ponds models. Hartigan et al. [2009]

Natural lakes. Lent et al. [1997]

Natural wetlands. Crisp [1966], Daniels et al. [2000], Hemond [1980], Hey et al. [1994], Lent et al. [1997], Mitsch et al. [2014]

Natural wetlands model. Caldwell et al. [2007]

Constructed wetlands. Ayub et al. [2010], Choi and Harvey [2000], Daniels et al. [2000], Mitsch et al. [2014], Hey et al. [1994], Nungesser and Chimney [2006]

Constructed wetlands model. Strosnider et al. [2007]

Sewers and sewer sections. Amick and Burgess [2000], Ellis et al. [2003], Guizani [2016], Rieckermann et al. [2007], Rutsch et al. [2006, 2008], Selvakumar et al. [2004]

A.2 Description of International BMP Database dataset cleaning

Preprocessing the International BMP Database. This section outlines the data loading, cleaning and joining process for data tables extracted as flat files from the International BMP database. The description matches the scripts called in the DRIVER.R file in the github repository associated with this project (<https://github.com/cgeger/clary>).

bmp table was created from **BMPInfo** and **tblBMPTypes** tables: These tables were joined on the **BMPTType** and **BMPTTypeCode** fields to identify which structural BMP sites were suitable for this analysis. Nonstructural BMPs were dropped from the analysis. ID Fields were converted to character strings so they would work more reliably for joins. Dates were converted to date format. Several fields were refactored to improve data integrity and uniformity: **BypassorOverflow** was collapsed into three categories: 'Bypass', 'Overflow' and 'NULL'. **Installation_Descr** was collapsed into three categories: 'No oversight identified', 'Installed per engineering design', 'Installed as designed, no eng oversight', 'Not installed as designed, no eng oversight'. **BMPTType**, **BMPGroup**, **BMPCategory_Code** and **X_BMPAnalysisGroup** were all factored from strings. Fields with yes/no factors were collapsed into 'N' and 'Y' and 'NULL'. The **SiteID** field was converted to uppercase to match cross-table data format. **X_HasMonitoringData** and **LastRehabDate** were dropped from the dataset due to mostly missing values. Five records with incorrect **WSIDs** and **SITEIDs** were corrected to match the tables containing watershed and site information. Empty cells, cells with errant whitespace, and blank strings were replaced with missing values. Missingness in the **bmp** table fields is approximately 41%, but no missing values are present in the fields required to match records {**BMPID**, **SITEID**, **WSID**} or in fields that indicate details about the structural features present at bmp type categories {**BMPTType**, **BMPGroup**, **BMPCategory_Code**, **BMPTType_Desc**, **BMPCategory_Desc**, **X_BMPAnalysisGroup**},

which were used as the reference basis for site-by-site dummy variable encoding.

climat table was created from **tblClimateStation** table: The raw data was trimmed of whitespace and blank cells, and converted to metric units. The latitude and longitude at one site in Alaska was corrected to represent the same format as the other locations. Missingness in the **climat** table is 1%, three records are missing climate summary information.

site table was created from **TestSite** table: The **SiteID**, **ClimateID** and **USGSHUC8** fields were converted to character strings to act reliably as join fields. All character fields were trimmed of whitespace and empty strings were replaced with missing values. Missingness values "ZZ" and "" in **State** and **ZipCode** columns were refactored as NA or NULL values. **Elevation** and **ElevationUnit** fields were converted to meters (m). Missingness in the **site** table was approximately 26% before dropping all columns containing information about documentation. In the remaining fields, {**SITEID**, **SiteName**, **City**, **County**, **State**, **ZipCode**, **Country**, **Latitude**, **Longitude**, **CLIMATID**, **Elevation**, **Elevation_Unit**, **USGSHUC8**, **Comment**} missingness was reduced to 12%, including the **CLIMATID** field, which is missing for 38 sites.

stn table was created from **MonitoringStation** table: The **SiteID**, **WSID**, **BMPID**, **MSID**, and **PDFID** fields were converted to character strings to act reliably as join fields. All character fields were trimmed of whitespace and empty strings were replaced with missing values. Fields with yes/no factors were collapsed into 'N' and 'Y' and 'NULL'. **SiteID** was changed to capital case. Two fields {**HydroFlag** and **WQFlag**} were initialized to keep track of which stations referenced hydrologic and/or water quality information, and which stations should be eliminated because they contained no data, or were identified to record poor quality data. {**nPraw**, **nFraw**, **nWQraw**} were populated with a preliminary count of the number of precipitation, flow and wq records present at each station. Since hydrology and water quality are marked at the site level,

the stations were checked to make sure they were properly assigned the right bmp or eliminated from analysis by marking the HydroFlag field "N". All Sediment/Solids stations were marked to be dropped from analysis. Two Underdrain stations were reclassified as Subsurface. Overflow from two retention sites was reclassified as bypass or eliminated from the analysis, since it was not clear if the water actually entered the BMP. Stations measuring at intermediate points within the BMP were eliminated from analysis. Two porous pavement bypass stations were reclassified as outflow, since bypass from porous pavement is surface water runoff. A few stations were marked with the wrong BMPID, and this was corrected when the correct BMP was apparent. Missingness in **stn** data from the original MonitoringStations table was approximately 14%.

wshed table was created from **Watershed** table: The SiteID, WSID, fields were converted to character strings to act reliably as join fields. All character fields were trimmed of whitespace and empty strings were replaced with missing values. Area and length units were converted to metric units. The original Watershed table in the MS Access database has 75% missingness, therefore, only the {SITEID, WSID, WSName, Type, LandUse_Descr, Area_Descr, Area, Area_unit, AreaImpervious_pct, DOT_ActivityType_flag} fields were included in the analysis set. Two WSIDs that had been swapped were corrected to match their flow and precip data. The areas for two watersheds were added from external sources, and one from the BMP description. The final analysis set has approximately 12% missingness, including 61 watersheds that are missing area estimates.

env table is created to represent a record of all bmps in the database, and act as as comprehensive and essential metadata for each BMPID-SITEID-WSID combination. **env** is compiled from the five tables listed above (bmp, site, wshed, stn, climat). **env** provides a lookup source of information for any {SITEID, CLIMATID, BMPID, WSID,

or MSID} field to enable the analyst to quickly get a snapshot of the site in order to try to answer questions or fix incorrectly joined data. It includes fields with the preliminary counts of the number of precip, flow and wq records (nPraw, nFraw, nWQraw), a list of relevant MSIDs (MSIDls), as well as a count of the number of bmp design details present in the BMPDesign table (ndesnotes). Overall missingness in the **env** metadata for 771 bmps in the env table is approximately 24%.

envkeys table provides a set of keys to match precipitation, flow and water quality information. The script drops sites and bmps from the study dataset via the exclusion vectors in (**exc** if it was not clear how the observed flows should match up for that location, or if something about the site's data seemed otherwise questionable. Sometimes this occurred when multiple bmps were present at one site, and it was not clear which inflow/outflow pairs represented each bmp. There are currently 24 BMPs excluded, some have suggested FIXMEs which could correct the problem. BMPs with the labels "CO", "LD", "OT", "UN", or "DIS" were excluded from the study dataset because they are sites with combined structures, general LID practices or other well-defined structures. The **stnkeys** script identifies a clean list of monitoring stations that belong to the BMPs in the study dataset and matches up each station id with the other location id keys. The **refkeys** script searches plausible reference stations in cleaned stnkeys object and the cleaned stn object to locate station IDs that act as reference flows at control sites.

precip table is based on the 'qry4PrecipFlatFile_Web' query in the MS Access version of the BMP Database, and created from the **Precipitation** table in the BMP database: The precipitation data was read in and column names, dates and factor formattings were corrected to match the other tables. Empty precipitation values marked as -99999 or other negative values were dropped from the dataset, and so were records that had been previously flagged for elimination in the UseForAnalysis_flag. All units were

converted to metric (cm/hr or cm). Some sites with mixed rain gauge station IDs were recoded to simplify the matching process, for example, a station which had recorded on two different gauges at the same location was recoded to have a single MSID. A few records with dates that did not match their flow dates were recoded to correct typos. Sites were flagged as having one MSID or multiple MSIDs at the site. Four sites with single MSIDs were identified with duplicate events, or ones that had more than one event on the same date. About 9600 events were monitored at a single station, and were non-duplicates. The duplicate events were eliminated, and the ones that were split across multiple measurements with the same event id were added together to represent the total precipitation depth for a single storm event. Sites with multiple MSIDs were identified as those with multiple stations at multiple BMPs, or multiple stations at a single BMP. Where multiple rain gauge stations were present at the same site, the average precipitation value was calculated for the event, and a single MSID was assigned for the each bmp. After all the cleaning steps, 11550 precipitation unique SITEID-MSID-EVENTID records in the precip table.

flow table is a modified version of the **Flow** table, and the site-event-flows pairings are based on the 'qry5FlowFlatFile.Crosstab' in the BMP database: the Flow table was corrected for missing values which were coded as -99999 and other negative numbers. Cells that contained empty white space were trimmed and changed to values reflecting missingness. Fields that contained categorical data were re-factored to contain uniform factor levels. Fields were renamed to match the naming conventions set in the previous tables. All measured volumetric fluxes (**Volume_Total**, also coded as **TOTFLOWVOL**) and flow rates were converted to metric values (L or L/sec). The dates for about 20 records were corrected to match other paired observations of flow and/or precipitation. Fluxes from one irrigated site were marked elimination. Fluxes that were measured at a single station and assumed to be equal at both inflow and outflow sites were marked for elimination from the inflow/outflow volume comparison. Initial missingness in the

variable indicating the total volume of flux (inflow, outflow, or other types of flow) was about 9%. Missing values were dropped and the data in the flow dataset was checked for duplicate entries and split records (see above in precip).

A.3 Structural features sources for causal effects

The structural features for each site were encoded from narrative and descriptive information contained in the design table of the BMPDB and from mapped observations on the BMP database mapping tool (WRF [2019], WRF, EWRI, US EPA, US DOT, Geosyntec Consultants, Wright Water Engineers [2023]). The BMP database mapping tool and dataset acted as the starting point for locating and observing most sites, but some sites had little or incomplete structural information recorded in the International Stormwater BMP Database (BMPDB). To reduce missingness in the BMPDB regarding contextual metadata and measurable structural features, virtual site visits were conducted to try to verify ground conditions, view the site location and take measurements remotely using Google [2023a] and Google [2023b]. Technical documents and white paper references referenced in the database and identified through internet search were also used to try to capture as many observable structural characteristics as possible. The complete list of technical references used to identify structural features is presented alongside the data in the github code repository at www.github.com/cgeger/clary. A subset of relevant data sources are listed here: BWE, Inc [2021], Bateman et al. [1999], Bean et al. [2007], CALTRANS Division of Environmental Analysis [2003, 2004], City of San Diego [2023], City of Tacoma, Washington [2012, 2015], Contech ES [2012, 2014, 2023a,b], Corsi et al. [1999], CPWJ [2012], Dally et al. [1983], Davis et al. [2012], ETV, US EPA, NSF [2004, 2008], Field et al., Glass and ETEC [2007], Harper et al. [2004, 1999a], Horwath and Bannerman [2008, 2010], Hussain et al. [2005], HydroLogic Solutions, Jensen Precast [2020], KCI Technologies [2015], Knight et al. [2013], Lenhart and Hunt [2011], Line [2006], Lin et al. [2007], Luell [2011], Messamer [2011], Owen et al. [2015], Orange County Public Works [2014], Prokop [2003], Rushton [2006], Selbig and Balster [2010], Sharkey [2006], Simon [2016], Stanley [1994], Tellessen and Allen, Terre Hill Stormwater Systems [2017], Walch [2008], Wanielista et al. [1986], Welborn and Veenhuis [1987], Winston et al. [2011], Yu et al. [1998]

A.4 Exploratory data analysis of the BMP database

Table A.1: Hydrologic observations in each imputed dataset by BMP Type.

BMP Type	n sites	n obs	Description of structure
BI	43	750	Biofilter - grass strip
BR	47	3253	Bioretention
BS	25	313	Biofilter - grass swale
BW	1	21	Biofilter - wetland vegetation swale
CBI	11	114	Catch basin insert
DB	29	391	Detention basin - open surface grass-lined basin (dry)
DC	4	74	Detention basin - open surface concrete or lined tank/basin (dry)
DU	2	23	Detention - underground vault, tank or pipe(s) (dry)
FL	3	44	Media filter - combination of media or layered media
FO	4	81	Media filter - other media types
FP	2	18	Media filter - peat mixed with sand
FS	23	392	Media filter – sand
FV	1	24	Vertical filter – geotextile fabric membrane
GR	15	613	Green roof
HDS	19	350	Hydrodynamic swirl concentrator or separation system
HRBF	3	74	High rate biofiltration
HRMF	15	363	High rate media filtration
IB	3	125	Infiltration basin
MCTT	2	21	Multi-chambered treatment train
OGS	10	167	Oil/grit separators and baffle boxes
PC	5	258	Porous pavement - pervious concrete
PF	2	38	Permeable friction course pavement
PM	11	480	Porous pavement - modular blocks
RP	58	1053	Retention pond – Open surface pond with a permanent pool (wet)
RV	3	108	Retention underground vault or pipes (wet)
VC	3	31	Volume control/attenuation structures
WB	20	713	Wetland - basin with open water surfaces
WC	6	119	Wetland - channel with wetland bottom

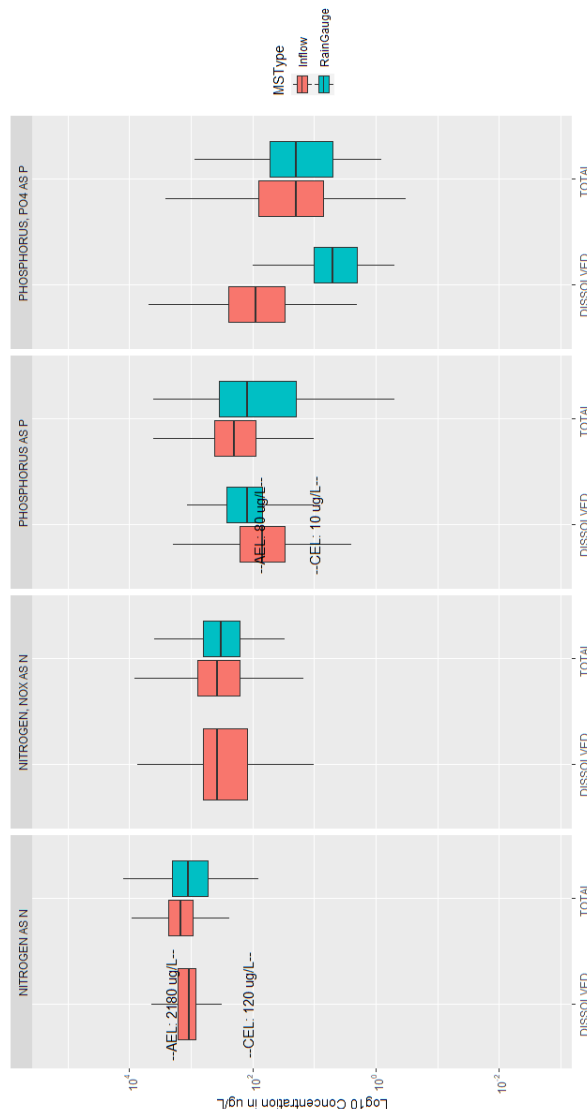


Figure A.1: Boxplots display a comparison of fractionated concentrations of nitrogen and phosphorus in stormwater and rainwater. Annotations 'CEL' and 'AEL' indicate a reference by displaying the acute ecological limit (AEL) and chronic ecological limit (CEL) values when relevant.

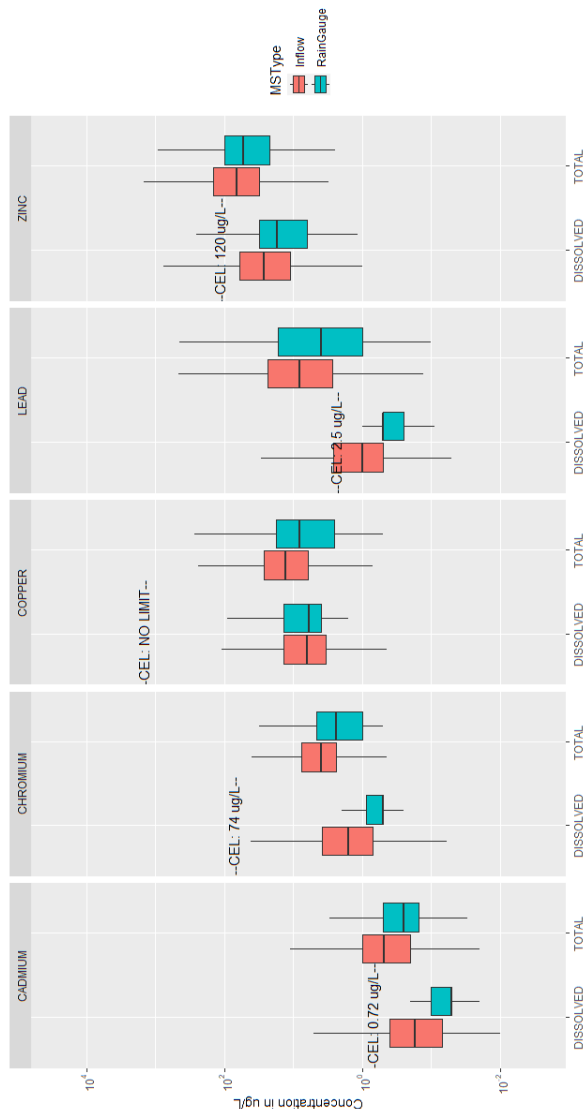


Figure A.2: Boxplots display a comparison of fractionated concentrations of cadmium, chromium, copper, lead and zinc in stormwater and rainwater. Annotations 'CEL' and 'AEL' indicate a reference by displaying the acute ecological limit (AEL) and chronic ecological limit (CEL) values when relevant.

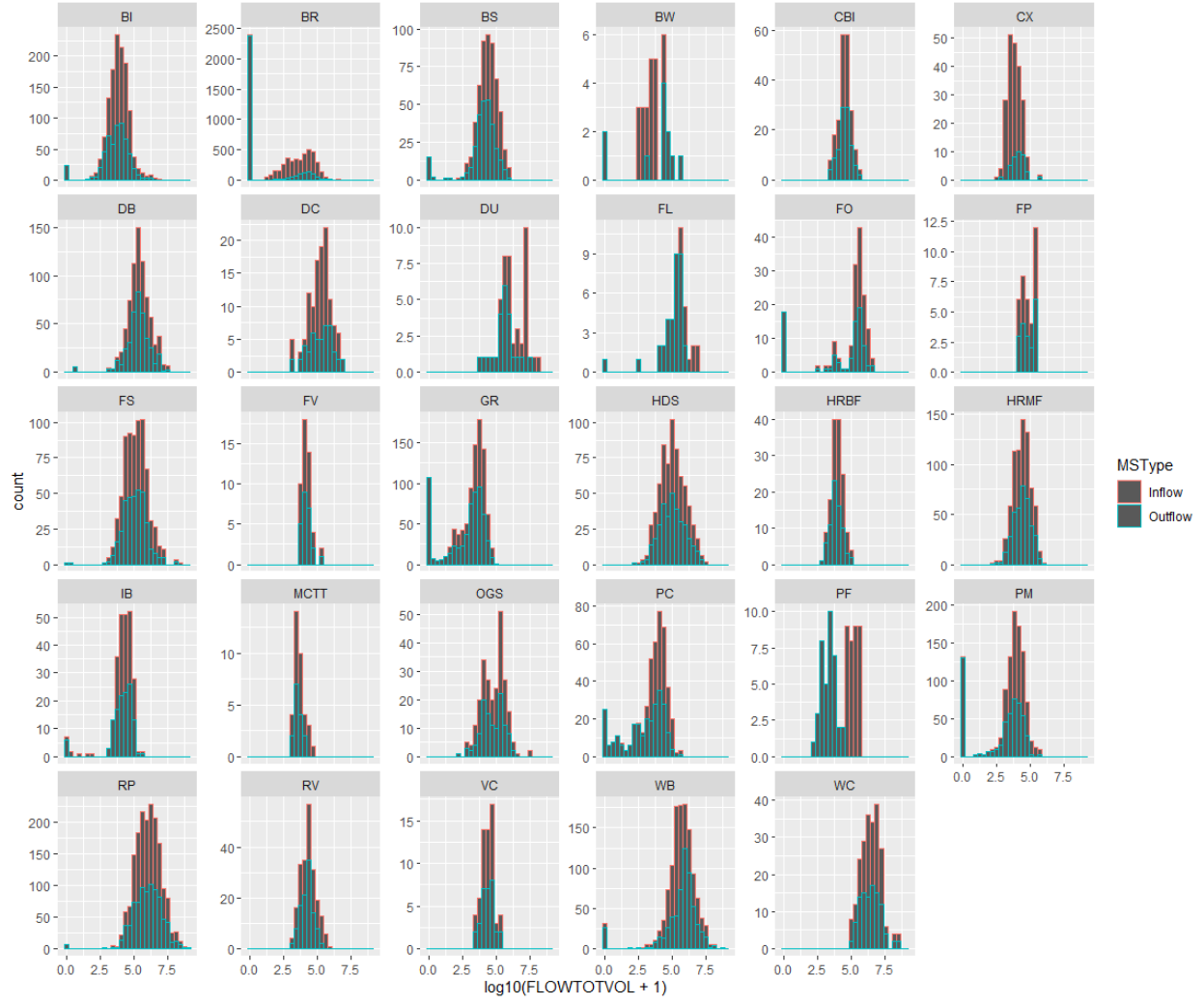


Figure A.3: \log_{10} of distribution of inflow and outflow volumes (all units converted to L) by BMP Type from the TOTFLOWVOL variable in the BMP database. The symbol key for the BMP Type is given in Table A.1.

Table A.2: Number of observations and \log_{10} of distribution of inflow and outflow volumes (in L) by BMP Type. Color gradient shows magnitude of mean stormwater flux (red: 10^6 L; white: 10^4 L; blue: 10^2 L). The symbol key for the BMP Type is given in Table A.1

BMP Type	MSType	n_obs	mean	min	q01	q10	q25	q50	q75	q90	q99	max
BI	Inflow	750	9.1E+03	9.3E+01	2.3E+02	1.2E+03	3.4E+03	8.0E+03	2.4E+04	6.0E+04	1.6E+06	6.1E+06
	Outflow	750	4.2E+03	0.0E+00	0.0E+00	4.6E+02	1.4E+03	6.1E+03	1.8E+04	4.7E+04	5.7E+05	2.3E+06
BR	Inflow	3253	4.2E+03	0.0E+00	2.0E+01	1.4E+02	5.7E+02	4.7E+03	3.3E+04	9.1E+04	4.3E+05	4.3E+06
	Outflow	3253	8.2E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.8E+04	2.4E+05	4.2E+06
BS	Inflow	316	3.4E+04	2.0E+02	7.1E+02	3.6E+03	1.2E+04	4.3E+04	1.1E+05	2.4E+05	8.8E+05	1.0E+06
	Outflow	316	9.4E+03	0.0E+00	0.0E+00	1.2E+03	5.9E+03	1.5E+04	3.9E+04	1.1E+05	4.5E+05	1.3E+06
BW	Inflow	21	2.2E+03	3.8E+02	3.8E+02	3.9E+02	7.6E+02	2.4E+03	5.0E+03	8.1E+03	2.9E+04	3.2E+04
	Outflow	21	4.8E+03	0.0E+00	0.0E+00	0.0E+00	5.1E+03	2.5E+04	5.4E+04	7.5E+04	3.3E+05	3.8E+05
CBI	Inflow	114	3.1E+04	2.2E+03	2.3E+03	7.4E+03	1.8E+04	3.0E+04	6.2E+04	9.6E+04	3.5E+05	3.8E+05
	Outflow	114	3.1E+04	2.2E+03	2.3E+03	7.4E+03	1.8E+04	3.0E+04	6.2E+04	9.6E+04	3.5E+05	3.8E+05
DB	Inflow	421	3.1E+05	1.8E+03	7.2E+03	3.5E+04	8.5E+04	2.6E+05	1.0E+06	4.3E+06	2.5E+07	4.3E+07
	Outflow	421	1.8E+05	2.0E+00	2.6E+00	2.0E+04	7.9E+04	2.1E+05	5.6E+05	2.0E+06	8.8E+06	2.7E+07
DC	Inflow	74	1.7E+05	1.5E+03	1.5E+03	2.3E+04	6.9E+04	2.2E+05	5.0E+05	1.7E+06	4.2E+06	5.0E+06
	Outflow	74	1.6E+05	1.1E+03	1.1E+03	1.0E+04	3.9E+04	1.6E+05	7.5E+05	1.5E+06	6.6E+06	7.6E+06
DU	Inflow	23	5.0E+06	1.9E+05	2.2E+05	5.0E+05	1.1E+06	1.2E+07	1.5E+07	1.9E+07	1.1E+08	1.2E+08
	Outflow	23	4.0E+05	5.6E+03	7.1E+03	3.6E+04	2.3E+05	4.1E+05	8.8E+05	3.3E+06	2.1E+07	2.8E+07
FL	Inflow	44	1.7E+06	3.3E+05	3.5E+05	5.7E+05	8.0E+05	1.3E+06	4.2E+06	5.8E+06	7.4E+06	7.6E+06
	Outflow	44	9.2E+04	0.0E+00	5.6E+00	1.4E+04	5.3E+04	2.1E+05	3.8E+05	4.9E+05	8.6E+05	9.5E+05
FO	Inflow	81	3.0E+05	3.4E+02	8.1E+02	9.4E+04	2.2E+05	3.7E+05	8.0E+05	1.5E+06	3.1E+06	4.5E+06
	Outflow	81	1.1E+04	0.0E+00	0.0E+00	0.0E+00	2.4E+03	1.7E+05	4.8E+05	8.0E+05	3.0E+06	4.1E+06
FP	Inflow	18	6.0E+04	1.3E+04	1.3E+04	1.5E+04	2.4E+04	6.3E+04	1.8E+05	2.0E+05	2.1E+05	2.1E+05
	Outflow	18	6.0E+04	1.3E+04	1.3E+04	1.5E+04	2.4E+04	6.3E+04	1.8E+05	2.0E+05	2.1E+05	2.1E+05
FS	Inflow	392	1.6E+05	1.1E+03	2.6E+03	1.1E+04	3.4E+04	1.5E+05	6.1E+05	2.3E+06	1.5E+07	3.2E+08
	Outflow	392	1.2E+05	0.0E+00	1.7E+03	1.1E+04	3.1E+04	1.2E+05	4.4E+05	1.0E+06	1.6E+07	1.6E+08
FV	Inflow	24	1.6E+04	4.3E+03	4.4E+03	5.5E+03	1.0E+04	1.6E+04	2.3E+04	3.8E+04	1.5E+05	2.1E+05
	Outflow	24	1.6E+04	4.3E+03	4.4E+03	5.5E+03	1.0E+04	1.6E+04	2.3E+04	3.8E+04	1.5E+05	2.1E+05
GR	Inflow	613	3.3E+03	3.2E+01	4.8E+01	1.9E+02	1.5E+03	4.7E+03	1.1E+04	1.9E+04	5.3E+04	6.8E+04
	Outflow	613	4.4E+02	0.0E+00	0.0E+00	0.0E+00	6.2E+01	1.7E+03	5.5E+03	1.2E+04	4.2E+04	8.6E+04
HDS	Inflow	350	1.1E+05	3.0E+02	1.3E+03	7.6E+03	2.7E+04	1.0E+05	5.2E+05	2.3E+06	1.3E+07	3.9E+07
	Outflow	350	1.0E+05	1.7E+02	1.1E+03	7.4E+03	2.4E+04	9.7E+04	4.5E+05	1.8E+06	1.1E+07	1.9E+07
HRBF	Inflow	74	1.1E+04	1.0E+03	1.2E+03	2.7E+03	6.0E+03	1.1E+04	2.2E+04	3.6E+04	1.1E+05	1.1E+05
	Outflow	74	7.4E+03	6.7E+02	1.0E+03	2.0E+03	4.0E+03	7.0E+03	1.4E+04	2.5E+04	7.9E+04	8.0E+04
HRMF	Inflow	428	2.0E+04	2.0E+02	7.3E+02	3.3E+03	7.1E+03	2.0E+04	5.7E+04	1.4E+05	4.6E+05	8.0E+05
	Outflow	428	2.2E+04	2.0E+02	7.2E+02	3.6E+03	7.9E+03	2.3E+04	5.7E+04	1.4E+05	4.4E+05	8.0E+05
IB	Inflow	125	1.7E+04	2.8E+01	6.3E+01	6.6E+03	1.0E+04	2.0E+04	5.2E+04	8.7E+04	1.6E+05	3.6E+05
	Outflow	125	1.2E+04	0.0E+00	0.0E+00	2.5E+03	6.2E+03	1.8E+04	4.9E+04	1.0E+05	1.6E+05	3.6E+05
MCTT	Inflow	21	5.7E+03	1.7E+03	1.7E+03	2.0E+03	2.9E+03	5.1E+03	8.6E+03	2.0E+04	3.8E+04	4.2E+04
	Outflow	21	3.8E+03	1.7E+03	1.7E+03	1.8E+03	2.4E+03	3.4E+03	6.5E+03	8.4E+03	9.0E+03	9.1E+03
OGS	Inflow	167	7.5E+04	9.3E+02	1.1E+03	5.2E+03	1.5E+04	9.7E+04	3.0E+05	6.2E+05	1.7E+07	3.9E+07
	Outflow	167	5.1E+04	1.9E+02	6.0E+02	4.9E+03	1.4E+04	5.6E+04	2.8E+05	5.6E+05	1.8E+06	2.1E+06
PC	Inflow	258	1.3E+04	2.4E+01	6.0E+02	2.7E+03	4.5E+03	1.3E+04	3.1E+04	6.9E+04	2.3E+05	4.3E+05
	Outflow	258	7.2E+02	0.0E+00	0.0E+00	5.5E+01	1.0E+02	1.9E+03	1.1E+04	2.5E+04	8.2E+04	4.2E+05
PF	Inflow	38	1.5E+05	4.3E+04	4.3E+04	4.5E+04	7.2E+04	1.5E+05	2.9E+05	4.9E+05	5.9E+05	6.2E+05
	Outflow	38	2.1E+03	2.0E+02	2.6E+02	4.6E+02	8.6E+02	2.3E+03	4.7E+03	8.3E+03	2.6E+04	3.2E+04
PM	Inflow	517	9.2E+03	0.0E+00	4.0E+02	1.8E+03	3.7E+03	8.6E+03	2.2E+04	5.5E+04	3.1E+05	4.3E+05
	Outflow	517	5.6E+02	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.9E+03	1.2E+04	2.7E+04	7.8E+04	6.0E+05
RP	Inflow	1053	1.2E+06	0.0E+00	9.5E+03	9.2E+04	2.8E+05	1.2E+06	4.5E+06	1.8E+07	1.5E+08	1.2E+09
	Outflow	1053	1.0E+06	0.0E+00	3.1E+03	5.7E+04	2.2E+05	1.2E+06	5.0E+06	1.9E+07	2.0E+08	9.0E+08
RV	Inflow	108	2.9E+04	1.2E+03	1.5E+03	4.1E+03	9.9E+03	2.6E+04	9.6E+04	2.5E+05	5.5E+05	6.7E+05
	Outflow	108	1.8E+04	1.2E+03	1.5E+03	4.1E+03	8.6E+03	1.7E+04	3.2E+04	7.1E+04	2.0E+05	2.9E+05
VC	Inflow	31	2.5E+04	2.6E+03	2.8E+03	5.6E+03	1.1E+04	1.9E+04	6.4E+04	7.1E+04	2.0E+05	2.3E+05
	Outflow	31	2.5E+04	2.6E+03	2.8E+03	5.6E+03	1.1E+04	1.9E+04	6.4E+04	7.7E+04	1.9E+05	2.1E+05
WB	Inflow	717	4.0E+05	0.0E+00	1.2E+03	4.6E+04	1.4E+05	3.5E+05	1.4E+06	6.3E+06	5.3E+07	1.6E+08
	Outflow	717	4.0E+05	0.0E+00	0.0E+00	2.6E+04	1.9E+05	8.3E+05	2.2E+06	6.6E+06	4.6E+07	5.8E+08
WC	Inflow	119	2.8E+06	8.3E+04	1.2E+05	3.6E+05	1.0E+06	3.0E+06	7.8E+06	1.4E+07	2.1E+08	2.3E+08
	Outflow	119	2.7E+06	1.3E+05	1.4E+05	3.4E+05	7.8E+05	2.6E+06	8.8E+06	2.0E+07	2.0E+08	2.2E+08

B Supplementary water budget datasets

Maturation of vegetation changes water budgets over time. The Water Budget Triangle can be used to visualize hysteresis of water budgets across seasonal time scales, look at variable and to clarify ranges of expected summer and winter performance. A time series of water budgets for the first five years following construction of the Everglades Nutrient Removal Project (ENRP) is shown for each wetland treatment cell in a multi-cell series (Figure B.1, data from Nungesser and Chimney [2006]). The water budget for the constructed wetland shows a reduction in surface water runoff through improvements in operation and the maturation of vegetation. Increased infiltrative losses accounted for 80% of the decrease in surface discharge whereas the remaining 20% was associated with increased ET.

Rainwater Harvesting: Rain Barrels and Cisterns. Two important variables affecting the effectiveness of long-term stormwater retention of cisterns and rain barrels are barrel volume and usage pattern. Undersized and under-used cisterns overflow more frequently (Q increases). Cistern sizing is based on regional climate, roof or collection area, expected demand (usage), as well as cost. Real and modeled water budget estimates for cisterns over a range of sizes (190 L to 900 m³) and climatic conditions (36.5 to 1092 mm rainfall) are shown in Figure B.2. Some measurements estimate an initial loss on the roof of the capture structure (used to estimate ET) or include a first flush to eliminate particulates from harvested water (used to estimate Q).

Sewered watersheds versus natural hydrologic function. Urban development and implementation of sewers fundamentally change the hydrologic budget of a watershed. Undeveloped watersheds generally evaporate about half or more of incoming precipitation, even in wet, energy-limited ecosystems such as coastal Maryland, Maine and Ontario (Jones et al. [2012]). Jones et al. [2012] estimate that the most energy-limited systems, such as temperate wet rainforest in Washington State, evaporate about 20% of incoming precipitation. Water-limited systems often evaporate greater than 80% of incoming

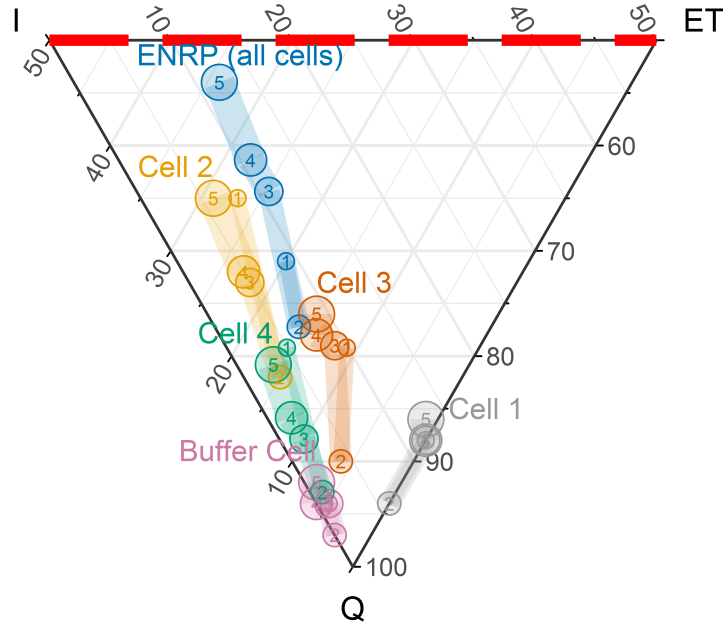


Figure B.1: Five years of water budgets for five treatment cells: Buffer Cell (pink), Cell 1 (gray), Cell 2 (yellow), Cell 3 (orange), Cell 4 (green) and overall wetland water budget performance (ENRP, in blue). Numbers and size of circle indicates progressive series of performance during years 1 through 5. Temporal patterns of water budgets from the first five years of monitoring the Everglades Nutrient Removal Project (ENRP), a constructed wetland site in Florida, USA. Circled numbers represent the year of operation. Each coloured line series represents data from one of the five cells in the constructed wetland. Note that the scale is expanded to show only the bottom 50% of the ternary diagram. The second year shows a decrease for all cells, likely due to a combined ecosystem establishment period and higher total influx in year 2. Overall, the ENRP's volumetric reduction of stormwater improves approximately 25% over time (in blue); 20% is attributable to increased groundwater infiltration and 5% attributable to greater ET. Data from Nungesser and Chimney (2006).

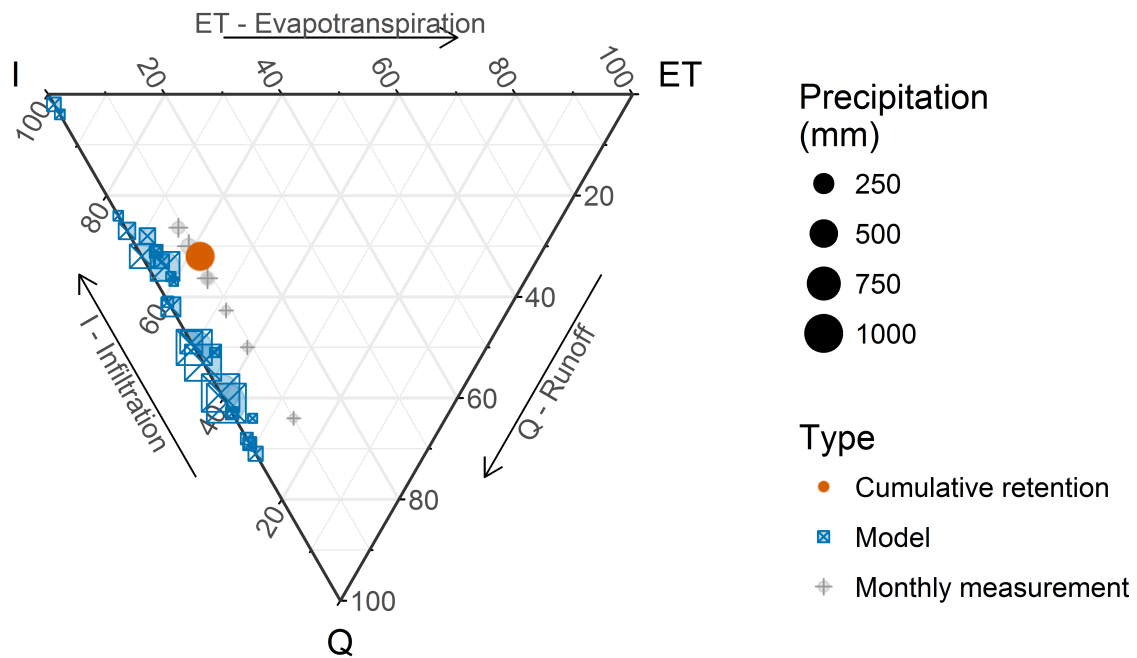


Figure B.2: Reported cumulative and monthly water budgets from a cistern in Queensland, Australia ($n = 1$ long-term measurement and 8 monthly measurements), and model estimates for 17 locations in the US, China, Saudi Arabia and Australia ($n = 27$ model estimates with varied climates). Data from Millar et al. 2003, Steffen et al. 2013, Zhang and Hu 2014, Guizani 2016. Thirty sewer exfiltration water budgets measured in pipe sections ($n = 13$), whole sewersheds ($n = 12$), a long-distance water supply pipeline in Saudi Arabia ($n = 1$), estimates from salt tracer models ($n = 3$) and an experimental laboratory model ($n = 1$). Evaporation from pipe sections and sewershed networks is assumed to be 0. Data compiled from Amick et al. 2000, City of Detroit Water and Sewerage Department and Michigan Department of Environmental Quality 2001, Ellis et al. 2003, Amick and Burgess 2003, Rieckermann et al. 2005, Rutsch et al. 2005, Rutsch 2006, Xu et al. 2014, Guizani 2016.

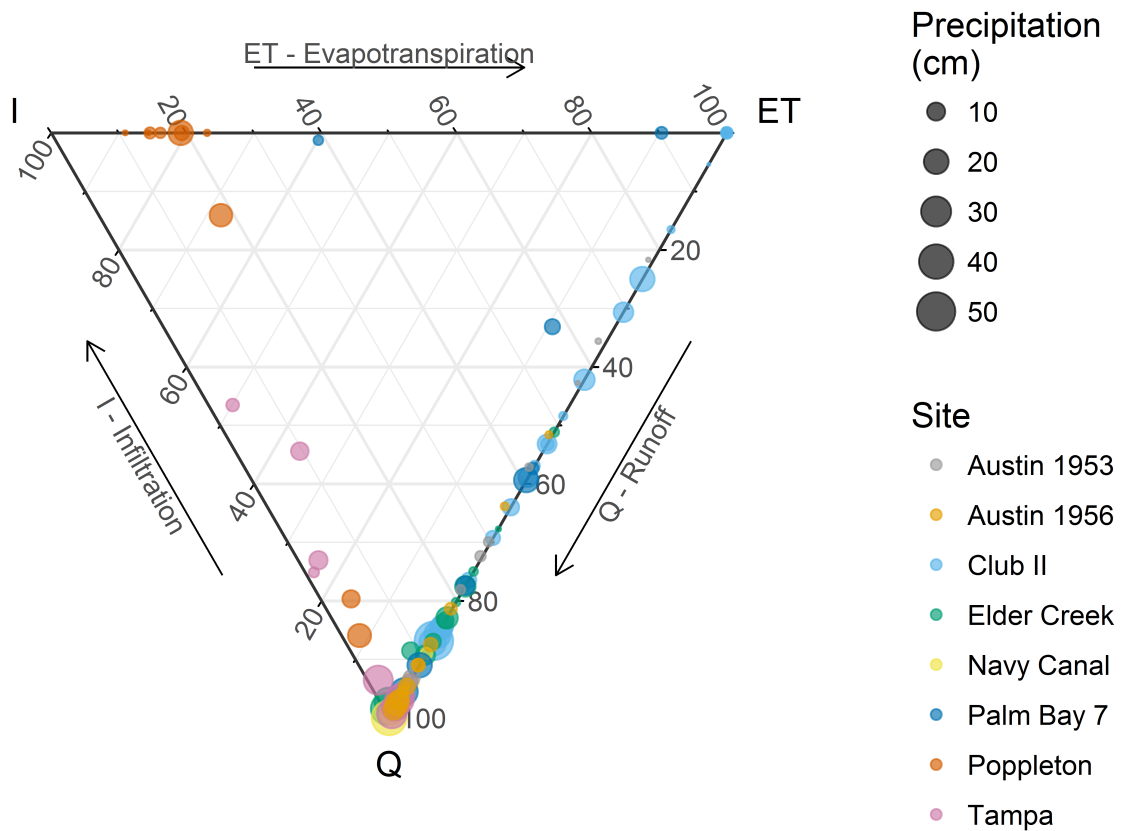


Figure B.3: 106 monthly water budgets from monitoring reports of seven retention ponds in Florida (Club II, Elder Creek, Navy Canal, Palm Bay 7, Poppleton and Tampa) and two years of modelled wet pond performance in Austin, Texas (Austin 1953, Austin 1956). Symbols are sized by monthly precipitation depth. Data compiled from Harper et al. 2003, Teague and Rushton 2005, Hartigan and Kelly 2009, Harper 2010a, 2010b, 2010c, 2011.

precipitation, even in low-temperature alpine or high-desert regions including Alaska, California and New Mexico. Annual continental-scale water balances from Rodell et al. [2015] estimate the average distribution of water budgets around 25 to 55% runoff, 40 to 70% ET and 2 to 10% contribution to groundwater flux (not including Antarctica, $n = 6$; Figure B.5). An estimate for non-landscaped vegetation in semi-humid temperate climates is approximately $Q = 10\%$, $ET = 60\%$, $I = 30\%$ (Starke et al. [2010]), supporting the concept that vegetated surface roughness plays a role in capturing and infiltrating runoff. In contrast, sewer networks in developed watersheds may discharge more than 75% ($Q = 55$ to 98%) of incoming precipitation into receiving waterways (Figure B.4, yellow triangles). The hardscape development in urban sewered catchments reapporitions evaporative losses primarily to runoff (compare Figures B.4 and B.5). Starke et al. [2010] report that increasing impervious surfaces from zero to a range of 10-20% can double the volume of runoff delivered to waterways. This level of hydrologic alteration often results in flow regimes that are outside ecological flow limits, contributing to urban stream syndrome and flooding (Walsh et al. [2005], Poff and Zimmerman [2010]).

Estimated water budgets from exfiltration studies of catchment-scale sewersheds and individual sewer sections show very little loss to ET, but may exfiltrate up to half the conveyed volume of water to groundwater in very dry ecosystems (Figure B.4). On average, water budgets for sewersheds in the literature report 88% runoff, 12% infiltration and 0% ET ($n = 12$). As expected, estimates for individual pipe sections tend to be a bit more leaky (i.e. infiltration is greater) than for entire sewersheds. Due to scant data and the complexity of urban water budgets, we have not presented case studies representing watershed-scale water budgets for urban watersheds; however, they are expected to plot near the lower vertex of the Water Budget Triangle. For reference, an urban watershed budget was developed for Baltimore by Bhaskar and Welty [2012].

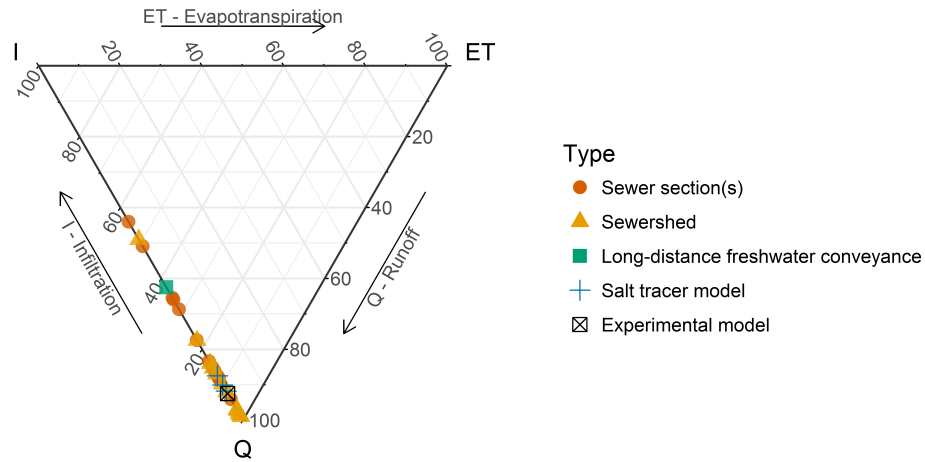


Figure B.4: Thirty sewer exfiltration water budgets measured in pipe sections ($n = 13$), whole sewersheds ($n = 12$), a long-distance water supply pipeline in Saudi Arabia ($n = 1$), estimates from salt tracer models ($n = 3$) and an experimental laboratory model ($n = 1$). Evaporation from pipe sections and sewershed networks is assumed to be 0. Data compiled from Amick et al. 2000, City of Detroit Water and Sewerage Department and Michigan Department of Environmental Quality 2001, Ellis et al. 2003, Amick and Burgess 2003, Rieckermann et al. 2005, Rutsch et al. 2005, Rutsch 2006, Xu et al. 2014, Guizani 2016.

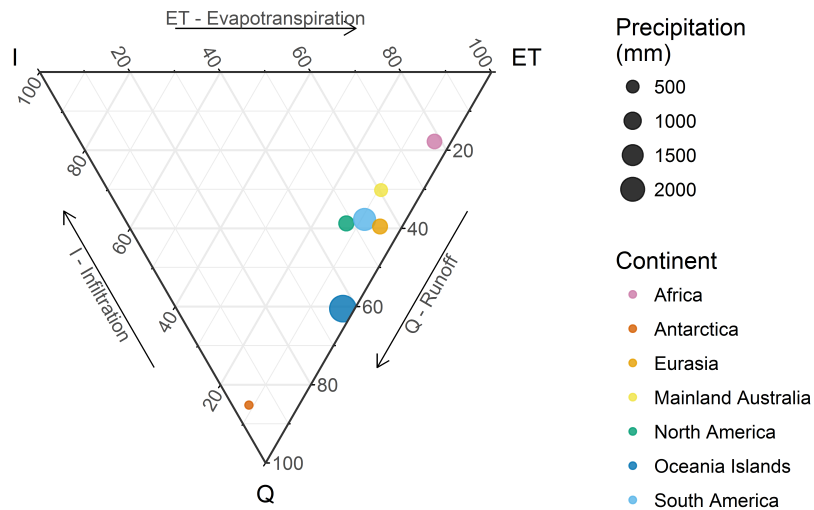


Figure B.5: Continental water budget estimates from Rodell et al. (2015), where infiltration is estimated as the magnitude of vertical groundwater flux ($n = 7$).

C Causal theory for engineered treatment

Figure C.1 displays a high-level systems diagram describing causal flow of information for one storm event at a BMP site. From this diagram, the factors affecting inflow diverge from those affecting outflow; the following textual description denotes individual DAG variables in `codefont`. One way arrows point in the direction from cause to effect. Double-headed arrows indicate a connection for which the causal direction is not obvious. Pink area shows environmental variables, which are mostly constrained to have observational, non-replicable data. (These are largely unavailable for randomization, experimentation, or intervention in the traditional experimental sense.) Blue area shows structural variables, which decision makers may have some ability to modify. Effect of interest is marked in orange, (effect of **Structural Features** on **Site Runoff**), which represent design features that engineers can directly control. Variables underlined in blue have cleaned data from the BMP Database. Direct onsite precipitation is a minimal input compared to influent surface and groundwater; the connection between the variable **Precipitation** and **Total Influent Water** was eliminated in the interest of simplifying the diagram as much as possible. Likewise, the evapotranspiration portion of the water budget was eliminated from the systems diagram because of the short, rainy time window represented. The **Site** variable defines which features are present at each observed location, but it opens a back-door path that confounds estimation of the unique effect of each treatment feature on the hydrologic performance of the site (see Figure 4.3, Box A). Formal DAGs created with dagitty.net (Textor et al. [2016]).

Variables that affect **Total Influent Water** (**Surface Water Runoff** and **Groundwater Influx**) are due to environmental inputs (**Precipitation**), or watershed conditions (**Watershed Characteristics**, **Slope**, **Land Use** and **Area**). In contrast, variables affecting **Site Runoff** (**Structural Features**, **Site Characteristics**, **Total Effluent**) are primarily influenced by structural choices made by engineers, which are indirectly affected by climate and watershed conditions. The complexity of the problem represented by the DAG means that it is easy to assume that variables in the pink zone

(e.g., **Watershed Area**) have a causal effect on site runoff because they are connected by multiple mediating variables and/or backdoor paths. However, if the mediating variable **Inflow** is being measured directly, as it is in the BMP Database, indirectly connected variables can be eliminated as control variables in a regression model for which the predictor is site outflow. By focusing on the variables from the blue region of the systems diagram in Figure C.1, three narrower DAGs were constructed to examine the model identification problem for specific structural features of interest (Figure 4.3).

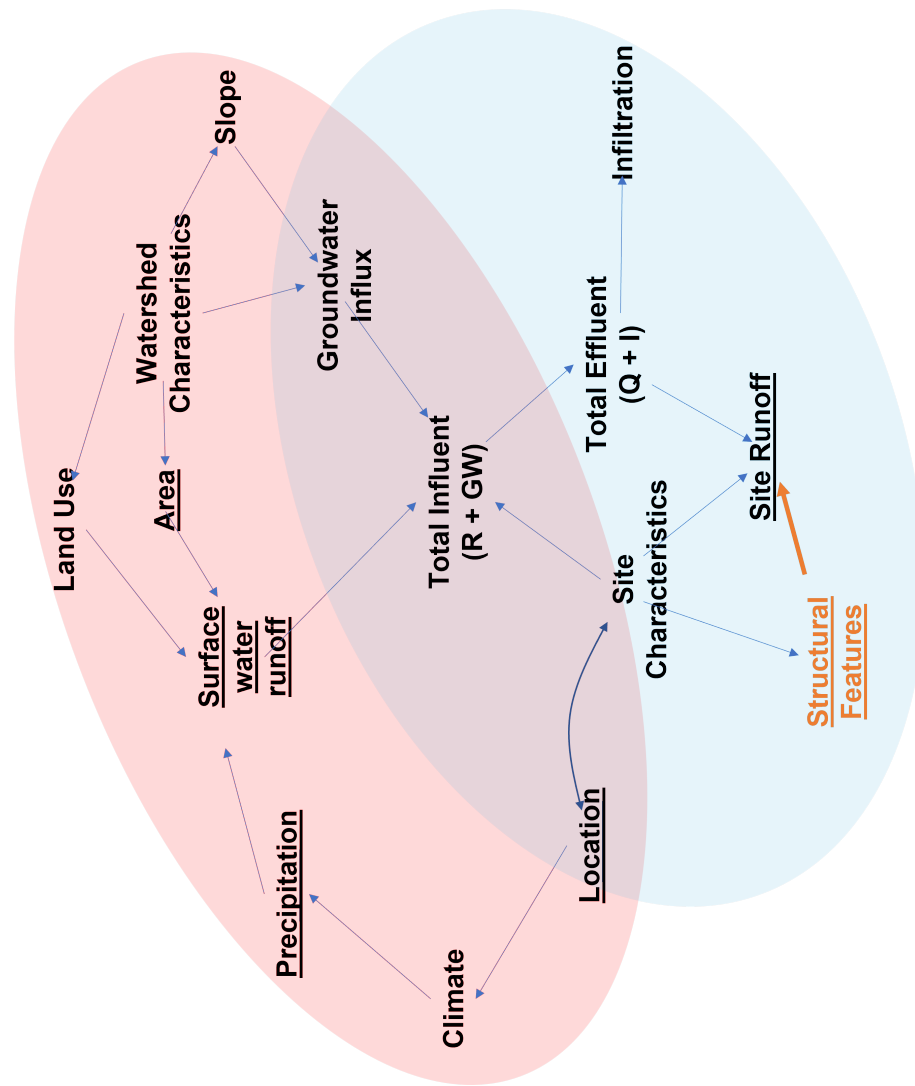


Figure C.1: A generalized systems diagram of the stormwater ecohydrology problemspace, represented as variables in a directed, acyclic graph (DAG). Single-headed arrows are in the causal direction, double headed arrows indicate a connection with an unclear relationship. The effects of structural features on runoff is highlighted in orange. Variables in the pink zone are primarily environmentally driven, variables in the blue zone are ones that engineering design can influence. Underlined variables are commonly observed in BMP field studies.

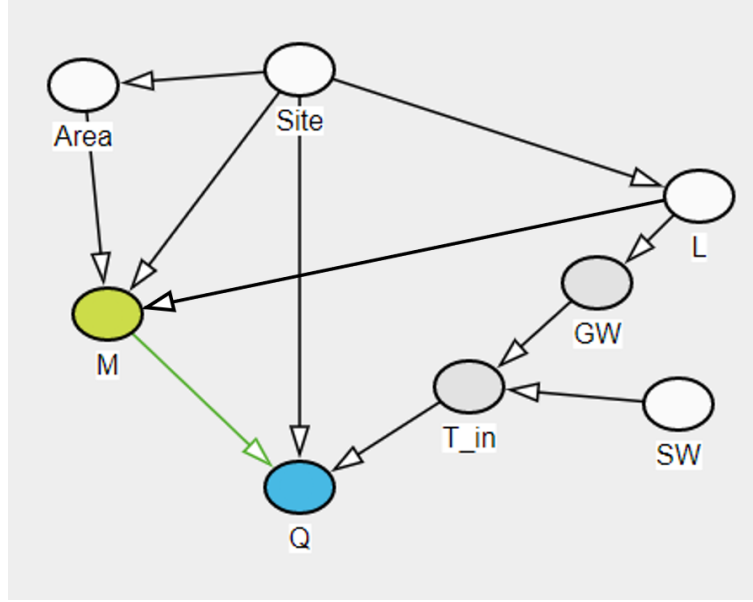


Figure C.2: DAG describing the model specification for identifying the effect of media amendment (M, in green) on Runoff (Q, in blue). First, coarsened exact matching on Area and Liner (L) were used to re-balance the dataset and isolate the treatment effect on the effect of interest (green arrow). The matching step counteracts the $Area \rightarrow M$ and $L \rightarrow M$ paths. Next, the model specification set Site as a random effect to block the remaining backdoor paths between $Q \rightarrow Site \leftarrow M$, and from $Q \leftarrow T_{in} \leftarrow GW \leftarrow L \leftarrow Site \rightarrow M$ with Inflow (SW) as a random covariate and offset, similar to Model Sets 1 and 2.

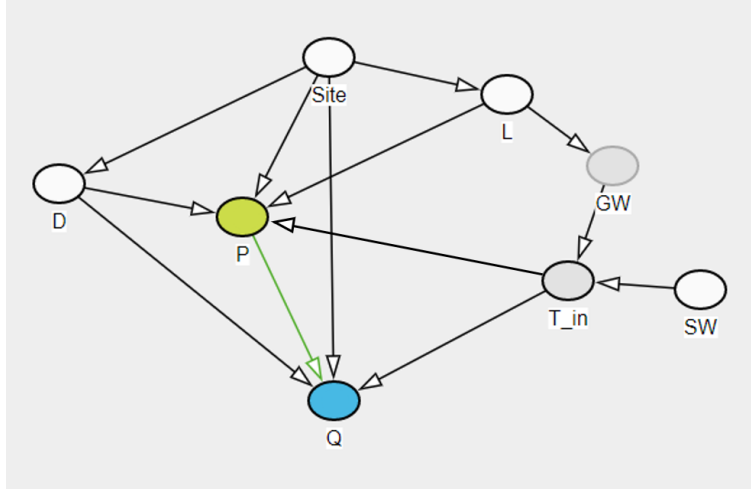


Figure C.3: DAG describing the model specification for identifying the effect of ponding (P , in green) on runoff (Q , in blue). Matching on D , L and SW inflow disrupts the $D \rightarrow P$, $L \rightarrow P$ and $SW \rightarrow T_{in} \rightarrow P$ paths. D and L remain covariates in the regression model, and (similar to earlier models) surface water inflow (SW) and $Site$ are included as random effects.

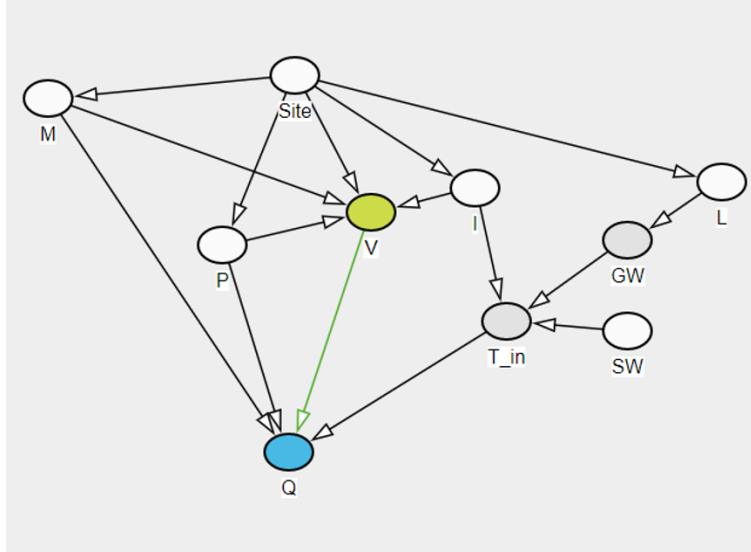


Figure C.4: DAG describing the model specification for identifying the effect of vegetation (V , in green) on runoff (Q , in blue). Matching on M and P disrupts the $M \rightarrow V$, $P \rightarrow V$ paths, and irrigation (I) is subsetting to include uniform treatment (no irrigation). M , P , and L remain covariates in the regression model, because they are ancestors of Q , and (similar to earlier models) surface water inflow (SW) and $Site$ are included as random effects.

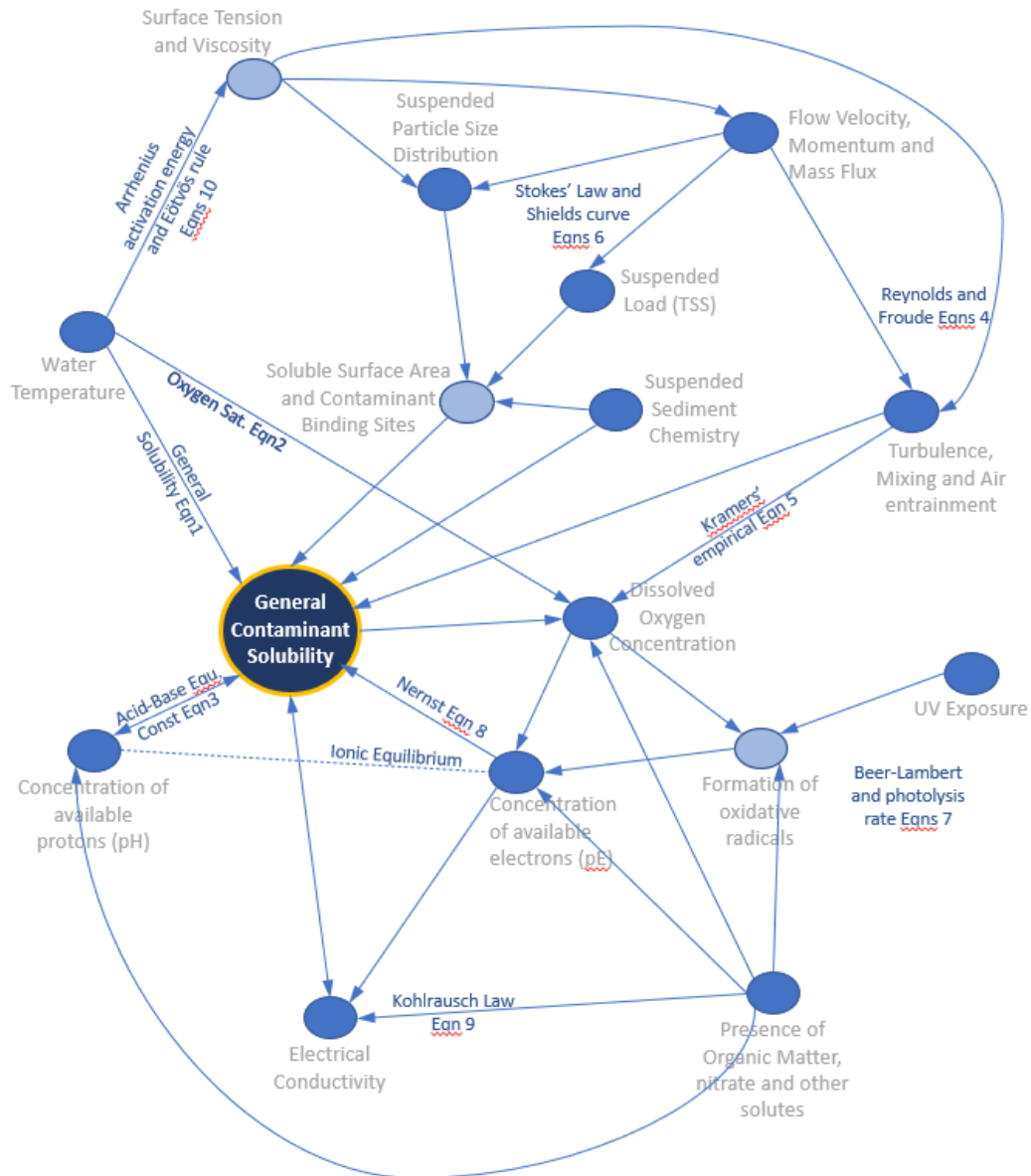


Figure C.5: A DAG displaying a generalized causal theory of solubility. Structural controls on general solubility and transport of contaminants in stormwater. See named equations listed on the following page.

Mathematical models of described physical and chemical processes

- Eqn 1: General solubility equation
 - $\ln(S_2/S_1) = -\Delta H/RT + \Delta S/R * (T_2 - T_1)$
- Eqn 2: Oxygen saturation curve
 - $C = C_{sat} * (1 - (\alpha * (T - T_{ref})))$
- Eqn 3: Acid-Base equilibrium constant
 - $K_a = [H^+][A^-]/[HA]$
- Eqns 4: Reynolds number and Froude number
 - $Re = \rho v L / \mu$
 - $Fr = v / \sqrt{g D}$
- Eqn 5: Kramers empirical mass transfer coefficient
 - $K_{La} = k * u^b$
- Eqns 6: Stokes' Law and Shields empirical curve
 - $v = (2/9) * (\rho_p - \rho_f) * g * d^2 / \mu$
 - $\tau c / \rho g D = 0.3 + 0.056(1 - e^{-20R * \sqrt{\tau c / \rho g D}})$
- Eqns 7: Beer-Lambert and photolysis rate laws
 - $I = I_0 * e^{-\epsilon c L}$
 - $d[C]/dt = -k[C]$
- Eqn 8: Nernst redox potential equation
 - $E_h = E_o + (RT/nF) \ln([oxidized]/[reduced])$
- Eqn 9: Kohlrausch's electrical conductivity law
 - $\kappa = k * \sum[C_i * z^2]$
- Eqn 10: Arrhenius activation energy and Eötvös Rule
 - Arrhenius: $\mu = \mu_0 * \exp(E_a / (RT))$
 - Eötvös: $\gamma_{water} = 0.07275 N/m * (1 - 0.002 * (T - 291 K))$

Figure C.6: Equations describing physical and chemical processes in Figure C.5

Table C.1: Summary of contaminant removal mechanisms

Dissolved Solids		Suspended Solids			
0.45 um		2 um		75 um	4750 um
Salts	Colloids	Fines	Coarse Solids	Gross Solids	
Removal Mechanisms for Dissolved and Particulate Solids					
Mechanisms for Treating Dissolved Solids			Mechanisms for Treating Suspended Solids		
Evaporation			Skimming		
Chemical reaction			Straining/Filtration		
Chemical precipitation			Sedimentation		
Reverse osmosis			Impaction		
Removal of solids to change partial solubility product			Interception		
Phytoremediation			Adhesion		
			Flocculation/coagulation		
			Chemical adsorption		
			Physical adsorption		
			Biological transformation		
			Periodic removal/cutting/harvest/excavation		
Removal Physical and Chemical Controls on Solubility and Particle Transport					
Solubility Controls			Particle Transport Controls		
Temperature			Temperature		
pH			Electrical charge		
Organic Matter or surfactant			Viscosity of the carrier		
Presence of complexing ligands			Particle size distribution		
Alkalinity			Solute and particle densities		
Ionic strength			Kinetic energy of media		
Mixture of solutes present					
Dissolved oxygen					
Turbulence or mixing					
Contaminant Fractionation in Stormwater					
Primarily Dissolved			Primarily Suspended		
Potassium			Chromium		
Sodium			Fats, Oils, Greases		
Calcium			Organic N		
Cadmium (also colloidal)			Orthophosphate		
Carbon dioxide			Total Nitrogen		
Copper			Total Phosphorus		
Magnesium			Lead		
Carbonate and Bicarbonate			Zinc		
Oxygen					
Soluble reactive phosphorus					
Inorganic Nitrogen (Nitrate, Nitrite, and Ammonia)					
Sulfate					

D Median contaminant concentration and mass flux ranges

Table D.1: Observed median concentrations and median mass flux estimates of total suspended solids (TSS) and total dissolved solids (TDS) entering and exiting various stormwater BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow n sites	Inflow n obs	Outflow n sites	Outflow n obs	Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
TDS	RP	14	166	15	157	100000 - 130000	170000 - 190000	220000 - 390000	440000 - 1100000
	FS	12	182	13	178	41000 - 54000	66000 - 93000	4200 - 8600	8400 - 14000
	HDS	5	106	5	105	120000 - 220000	120000 - 310000	9400 - 19000	7500 - 18000
	BI	29	582	25	389	52000 - 58000	77000 - 86000	390 - 600	410 - 600
	BS	14	161	12	130	76000 - 85000	77000 - 91000	1700 - 2600	1000 - 2300
	DB	13	149	13	133	94000 - 130000	100000 - 120000	20000 - 36000	16000 - 35000
	HRMF	6	171	6	171	25000 - 45000	25000 - 39000	1400 - 2200	990 - 1500
	WC	14	259	14	221	24000 - 42000	16000 - 23000	160000 - 230000	96000 - 170000
	BS	33	409	37	465	34000 - 44000	19000 - 24000	1200 - 1700	430 - 600
	HDS	27	490	27	455	63000 - 80000	36000 - 45000	3500 - 5900	2000 - 3300
TSS	OGS	15	245	15	228	30000 - 41000	13000 - 19000	1300 - 3200	520 - 1100
	WB	26	671	27	644	29000 - 34000	17000 - 19000	100000 - 210000	49000 - 130000
	BI	43	833	38	613	48000 - 56000	20000 - 24000	410 - 530	140 - 210
	DB	36	457	38	491	56000 - 72000	22000 - 26000	13000 - 18000	5000 - 7400
	PM	4	136	11	196	69000 - 120000	20000 - 29000	770 - 1800	120 - 220
	HRMF	18	416	18	416	42000 - 62000	15000 - 19000	1100 - 1500	360 - 460
	RP	61	1088	63	1038	44000 - 56000	12000 - 13000	51000 - 76000	18000 - 28000
	BR	37	802	34	617	43000 - 52000	9000 - 10000	1000 - 1400	140 - 190
	FS	26	393	25	370	56000 - 72000	8700 - 11000	5100 - 8100	740 - 1200
	HRBF	6	113	6	113	25000 - 35000	3000 - 3900	380 - 810	29 - 58

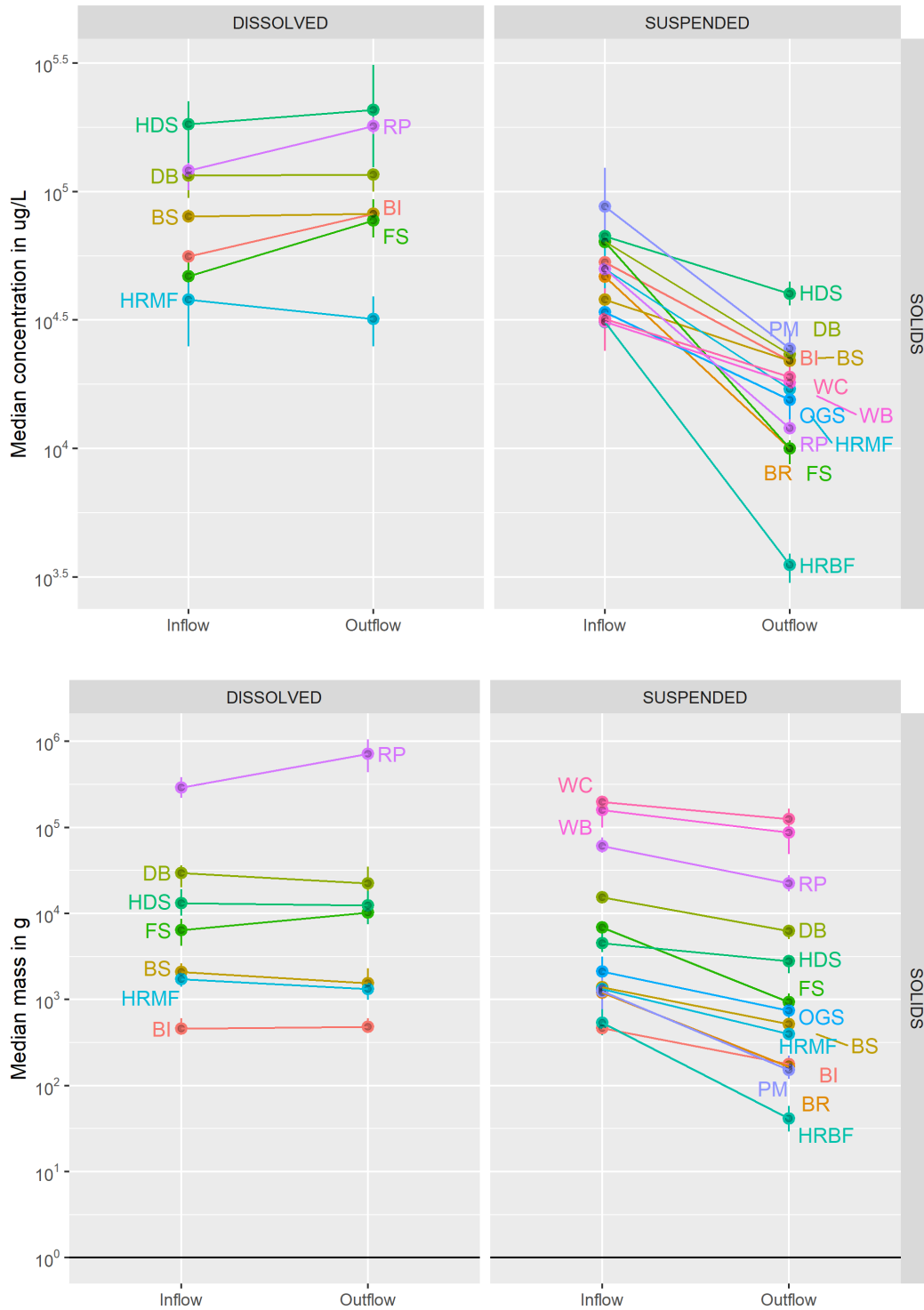


Figure D.1: Bootstrapped median TSS and TDS concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference line representing 1 gram flux at 10^0 .

Table D.2: Observed median concentrations and median mass flux estimates of total nitrogen (TN) and total nitrates + nitrites (NO_x) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Total N	DB	16	191	17	189	1000 - 1200	1000 - 1200	210 - 320	240 - 340
	HDS	5	152	5	138	1600 - 2100	1500 - 1900	560 - 1100	730 - 1400
	BI	9	153	9	131	1100 - 1600	1100 - 1300	11 - 24	16 - 34
	FS	11	177	10	164	1000 - 1300	880 - 1100	340 - 650	220 - 390
	WB	14	1049	14	911	1600 - 1700	1300 - 1400	140000 - 190000	92000 - 130000
	BR	24	373	22	309	1100 - 1300	820 - 1000	39 - 59	16 - 24
	BS	8	146	11	212	910 - 1100	650 - 790	22 - 33	12 - 20
	RP	29	584	29	537	1500 - 1800	1100 - 1200	2100 - 3500	2200 - 3300
	PM	3	121	11	186	490 - 630	900 - 1200	4.2 - 10	6.4 - 12
	FS	23	366	22	343	310 - 360	480 - 590	32 - 60	52 - 75
NOx as N	BR	32	693	29	527	280 - 340	350 - 480	6.6 - 8.5	5.1 - 8.4
	HRMF	8	158	8	157	140 - 230	110 - 180	2.8 - 7.9	1.1 - 4.1
	GR	7	108	11	138	430 - 860	420 - 780	5.4 - 14	1.2 - 3.8
	BS	21	297	21	277	250 - 320	300 - 380	6.4 - 9.9	5.2 - 8.8
	WB	21	1190	21	1009	370 - 420	390 - 460	8400 - 14000	11000 - 19000
	WC	12	223	13	192	270 - 400	170 - 320	1300 - 2200	1000 - 2000
	HDS	13	285	13	252	420 - 570	350 - 480	37 - 73	27 - 55
	DB	22	331	24	337	330 - 460	220 - 290	74 - 110	39 - 58
	BI	42	831	36	602	420 - 480	310 - 370	3.9 - 5.9	1.8 - 3.0
	RP	47	880	49	831	370 - 420	120 - 160	290 - 450	110 - 180

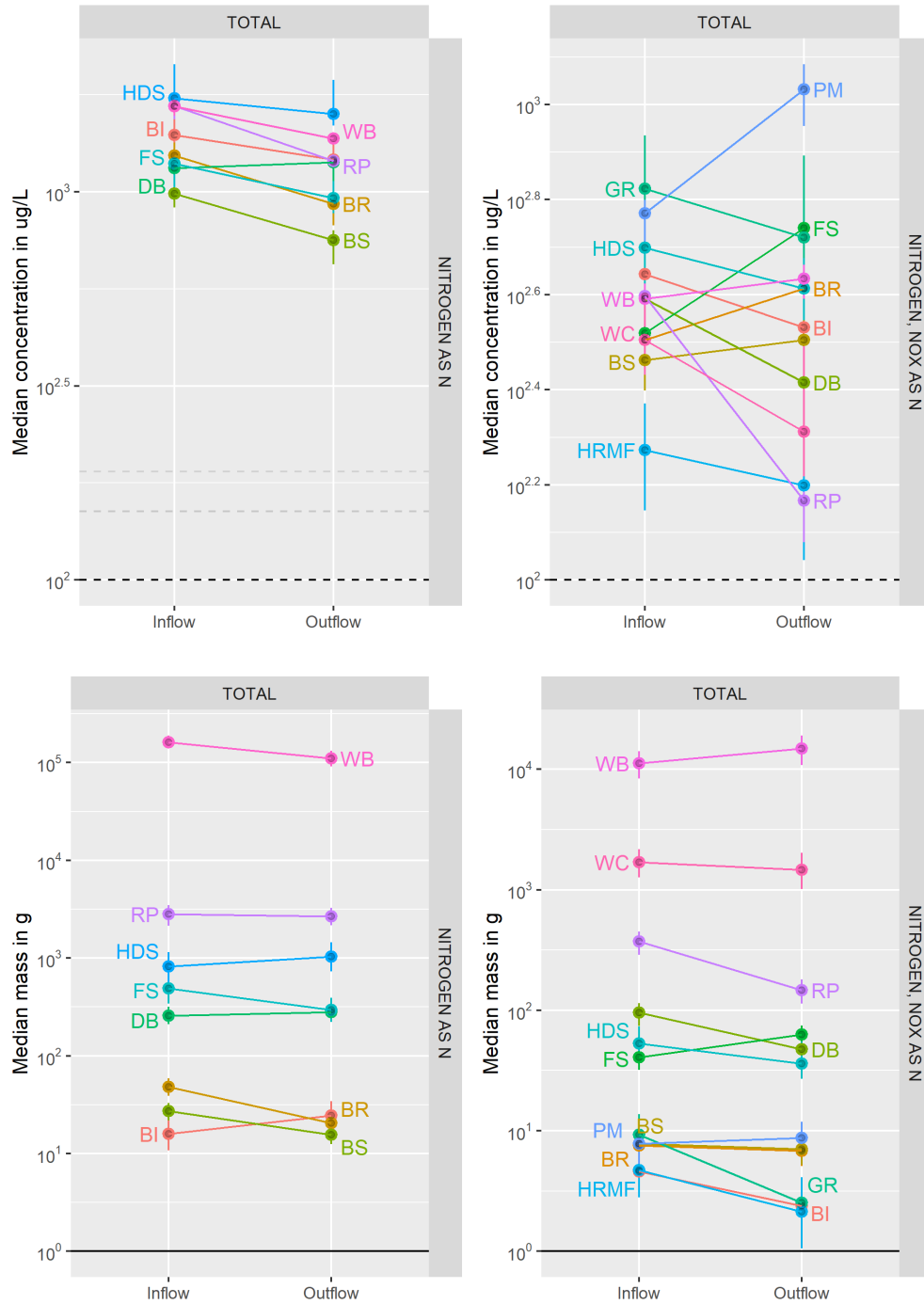


Figure D.2: Bootstrapped median total N and NO_x concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10⁰g).

Table D.3: Observed median concentrations and median mass flux estimates of dissolved and total P, and total reactive P entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Dissolved P	DB	9	115	9	105	80 - 110	90 - 120	13 - 32	13 - 27
	HRMF	9	194	9	194	45 - 80	40 - 60	1.3 - 2.2	1.1 - 1.5
	WB	9	756	8	701	30 - 30	20 - 20	1700 - 2400	1400 - 2000
	HDS	7	126	7	121	60 - 95	50 - 80	3.6 - 8.2	2.7 - 6.5
	RP	19	393	20	392	120 - 140	60 - 80	220 - 350	150 - 250
	GR	7	109	11	138	130 - 310	380 - 460	1.3 - 3.0	1.4 - 2.5
Total P	BI	43	836	38	601	170 - 200	210 - 260	1.4 - 1.7	1.8 - 2.6
	BR	42	812	38	635	180 - 210	200 - 270	4.2 - 5.7	3.1 - 4.4
	BS	28	366	29	348	170 - 220	190 - 230	6.8 - 8.5	4.1 - 6.2
	WC	14	244	14	210	210 - 260	180 - 220	930 - 1500	930 - 1700
	HDS	23	339	23	305	210 - 260	160 - 200	18 - 34	14 - 27
	OGS	8	140	8	141	200 - 470	90 - 300	8 - 23	5.7 - 15
	DB	33	432	35	454	240 - 270	180 - 220	49 - 74	39 - 53
	HRMF	18	346	18	348	100 - 140	70 - 90	2.6 - 3.7	2.1 - 2.9
	PM	3	122	11	190	180 - 280	98 - 120	2.2 - 4.5	0.5 - 1.2
	FS	23	356	22	337	150 - 190	90 - 100	17 - 25	9.1 - 14
Total PO4	WB	26	1289	26	1101	140 - 160	80 - 85	7100 - 10000	3200 - 4600
	HRBF	6	109	6	109	80 - 100	40 - 50	1.2 - 1.9	0.42 - 1.1
	RP	59	1032	60	973	240 - 280	120 - 140	240 - 320	150 - 220
	BR	23	507	22	385	20 - 28	110 - 140	0.54 - 0.76	1.6 - 2.6
	BI	22	217	13	162	3.3 - 6.5	13 - 28	0.039 - 0.082	0.064 - 0.21
	WC	3	121	3	101	10 - 20	30 - 40	89 - 170	220 - 420
	DB	12	121	15	155	23 - 42	20 - 42	5.0 - 25	5.9 - 15
	HRMF	8	120	8	120	6.5 - 9.9	6.0 - 9.8	0.067 - 0.23	0.18 - 0.31
	WB	9	745	10	686	20 - 20	10 - 20	1400 - 2100	1200 - 1800
	RP	13	328	12	272	16 - 30	6.5 - 12	27 - 70	3.8 - 23

Figure D.3: Comparison of bootstrapped median total and dissolved P and total PO4 concentrations (top) and mass flux (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Dotted lines (concentration plots) represent the most commonly reported detection limits. Solid horizontal line (mass plots) is a reference line representing 1 gram.

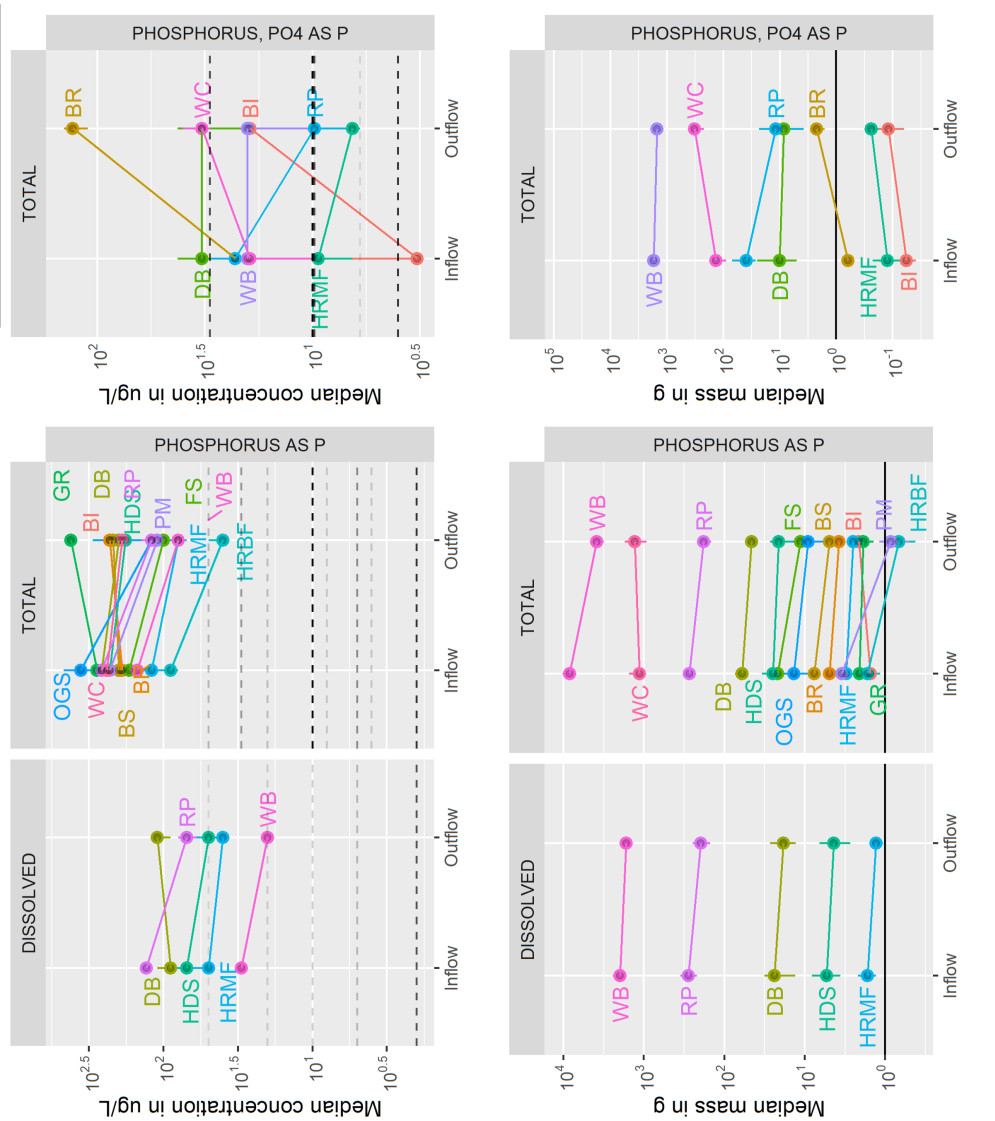


Table D.4: Observed median concentrations and median mass flux estimates of dissolved and total cadmium (Cd) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow n sites	Inflow n obs	Outflow n sites	Outflow n obs	Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
Dissolved Cd	BR	9	253	8	184	0.050 - 0.070	0.10 - 0.12	0.0021 - 0.0032	0.0011 - 0.0022
	BI	29	579	25	390	0.10 - 0.10	0.10 - 0.10	0.0010 - 0.0014	0.00073 - 0.0010
	DB	10	180	10	179	0.43 - 0.50	0.50 - 0.50	0.046 - 0.10	0.036 - 0.083
	FS	9	111	9	106	0.20 - 0.20	0.10 - 0.20	0.0096 - 0.017	0.0083 - 0.015
Total Cd	DB	14	213	14	211	0.50 - 0.50	0.50 - 0.50	0.14 - 0.31	0.084 - 0.15
	HDS	9	139	9	134	0.25 - 0.32	0.22 - 0.26	0.010 - 0.019	0.0094 - 0.017
	BR	12	229	13	221	0.20 - 0.30	0.16 - 0.21	0.0053 - 0.0092	0.0033 - 0.0059
	BS	16	185	14	165	0.26 - 0.42	0.24 - 0.25	0.013 - 0.019	0.0037 - 0.0063
	RP	27	487	28	487	0.50 - 0.50	0.25 - 0.30	0.26 - 0.42	0.21 - 0.37
	FS	15	191	15	188	0.29 - 0.43	0.20 - 0.25	0.026 - 0.043	0.017 - 0.026
	BI	29	582	25	390	0.42 - 0.50	0.20 - 0.24	0.0037 - 0.0046	0.0013 - 0.0019

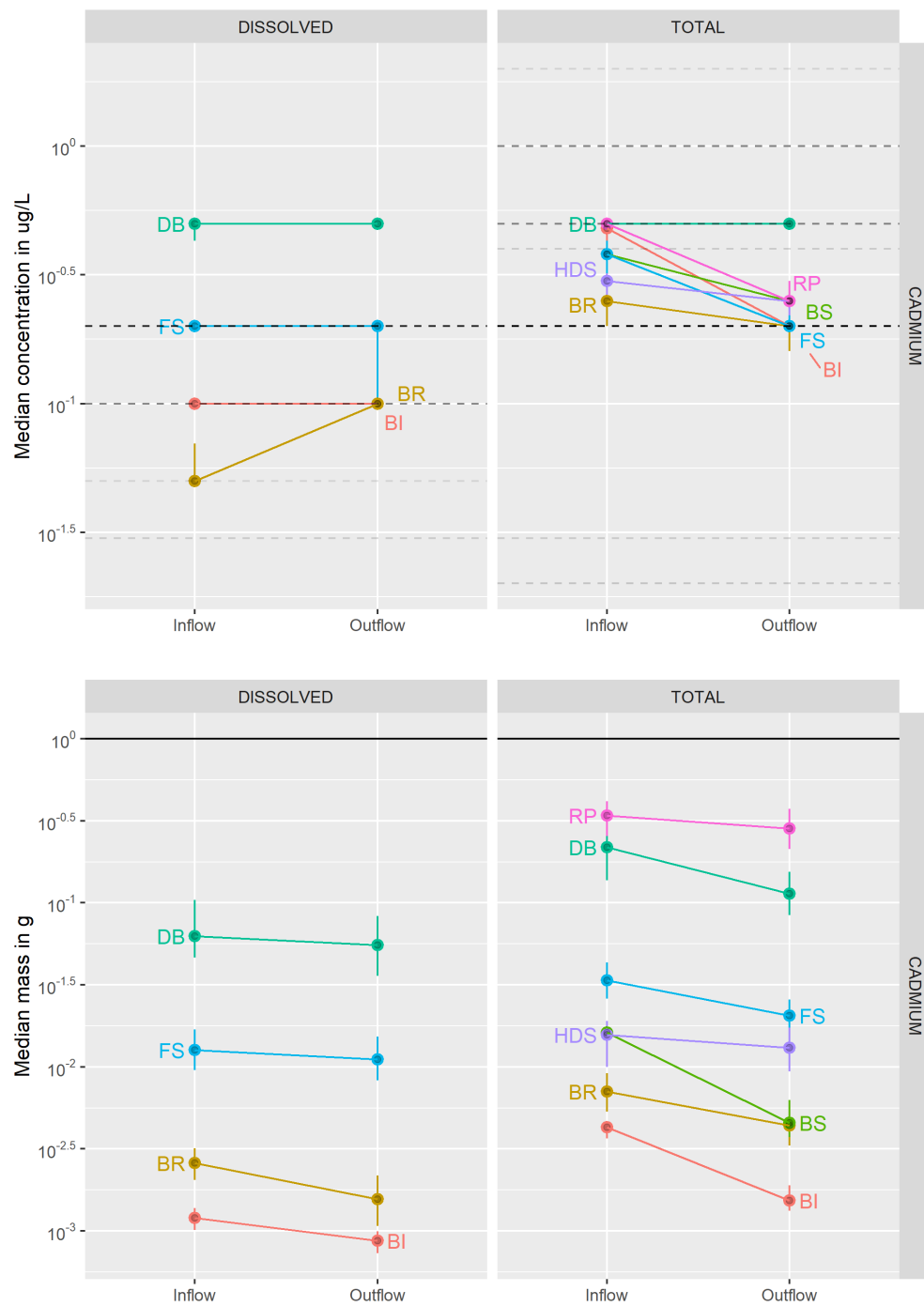


Figure D.4: Bootstrapped median total and dissolved Cd concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10^0 g).

Table D.5: Observed median concentrations and median mass flux estimates of dissolved and total chromium (Cr) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Dissolved Cr	FS	9	111	9	106	1.0 - 1.0	1.0 - 1.0	0.046 - 0.074	0.052 - 0.081
	BI	29	582	25	390	2.3 - 3.0	2.0 - 2.6	0.023 - 0.029	0.014 - 0.021
	BR	4	193	3	114	0.90 - 1.2	0.50 - 0.60	0.021 - 0.039	0.0061 - 0.014
Total Cr	WC	5	124	5	109	3.0 - 5.0	4.0 - 5.0	20 - 37	22 - 51
	BI	30	601	25	391	5.7 - 6.4	3.8 - 4.8	0.036 - 0.056	0.025 - 0.033
	RP	15	227	14	216	3.4 - 4.9	2.0 - 3.9	3.7 - 8.3	3.3 - 8.0
	FS	9	112	9	106	1.6 - 2.5	1.0 - 1.2	0.12 - 0.18	0.066 - 0.10
	BR	6	164	6	149	2.7 - 4.1	1.0 - 2.2	0.074 - 0.13	0.022 - 0.037

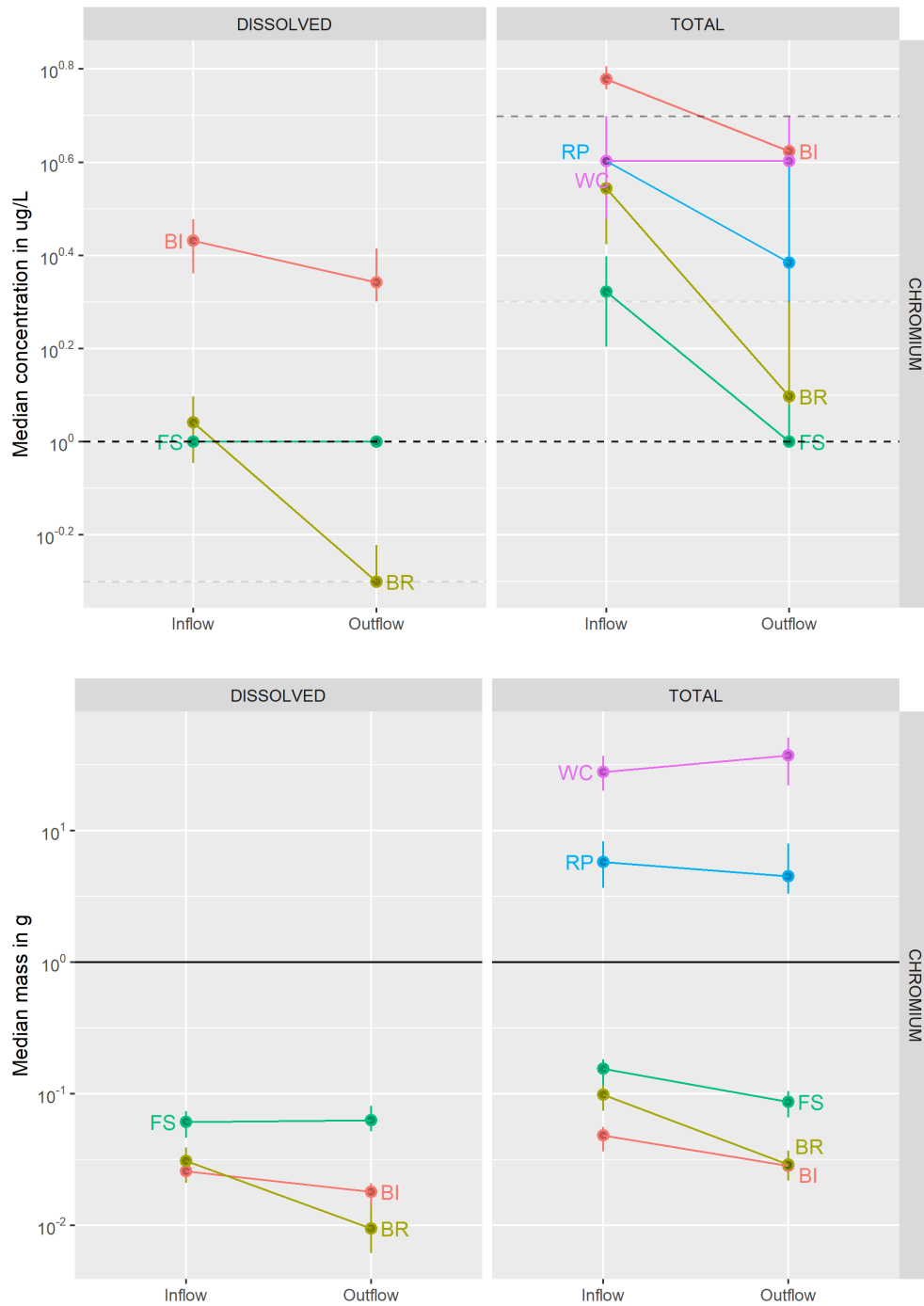


Figure D.5: Bootstrapped median total and dissolved Cr concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10^0 g).

Table D.6: Observed median concentrations and median mass flux estimates of dissolved and total copper (Cu) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Dissolved Cu	HRMF	13	217	13	217	3.6 - 4.5	3.7 - 5.0	0.065 - 0.10	0.10 - 0.13
	HDS	9	123	9	123	7.0 - 9.6	7.1 - 9.9	0.21 - 0.41	0.22 - 0.45
	PM	3	125	5	114	5.2 - 6.1	5.6 - 7.0	0.055 - 0.12	0.034 - 0.073
	BR	13	310	13	216	5.3 - 6.6	5.2 - 7.0	0.18 - 0.28	0.092 - 0.17
	BS	16	174	16	141	5.3 - 7.8	4.9 - 6.7	0.24 - 0.40	0.077 - 0.12
	FS	10	158	10	153	4.1 - 6.0	3.5 - 4.7	0.25 - 0.36	0.21 - 0.33
	DB	10	209	10	202	3.9 - 5.7	3.0 - 4.4	1.1 - 1.6	0.61 - 1.0
	WB	7	119	9	114	3.5 - 4.4	1.9 - 3.6	0.35 - 0.74	0.19 - 0.48
	RP	17	335	15	321	5.0 - 5.6	3.7 - 4.3	2.0 - 3.5	1.2 - 1.9
	BI	32	634	27	431	9.3 - 12	6.1 - 7.4	0.066 - 0.099	0.025 - 0.038
	WC	8	151	8	134	5.0 - 9.6	5.0 - 10	30 - 52	27 - 48
	HDS	14	216	14	211	13 - 16	12 - 14	0.76 - 1.2	0.59 - 1.0
Total Cu	OGS	9	114	9	123	8.0 - 18	7.6 - 14	0.61 - 1.1	0.42 - 0.89
	BS	20	261	20	226	14 - 18	11 - 13	0.57 - 0.95	0.18 - 0.28
	HRMF	15	278	15	278	11 - 14	7.0 - 9.2	0.21 - 0.33	0.17 - 0.25
	BR	23	447	24	415	10 - 12	6.0 - 7.0	0.23 - 0.31	0.12 - 0.15
	FS	19	319	19	308	10 - 11	6.2 - 7.5	0.95 - 1.4	0.53 - 0.73
	DB	17	276	19	288	8.0 - 10	4.5 - 6.0	2.6 - 4.3	1.0 - 1.5
	PM	3	125	11	183	13 - 19	8.0 - 9.0	0.16 - 0.33	0.057 - 0.094
	RP	41	805	41	757	9.5 - 10	5.0 - 5.5	7.6 - 11	3.7 - 6.0
	BI	34	679	28	452	21 - 24	9.2 - 11	0.14 - 0.22	0.043 - 0.062
	WB	12	210	13	192	8.9 - 10	4.9 - 5.0	2.5 - 4.8	1.6 - 3.8

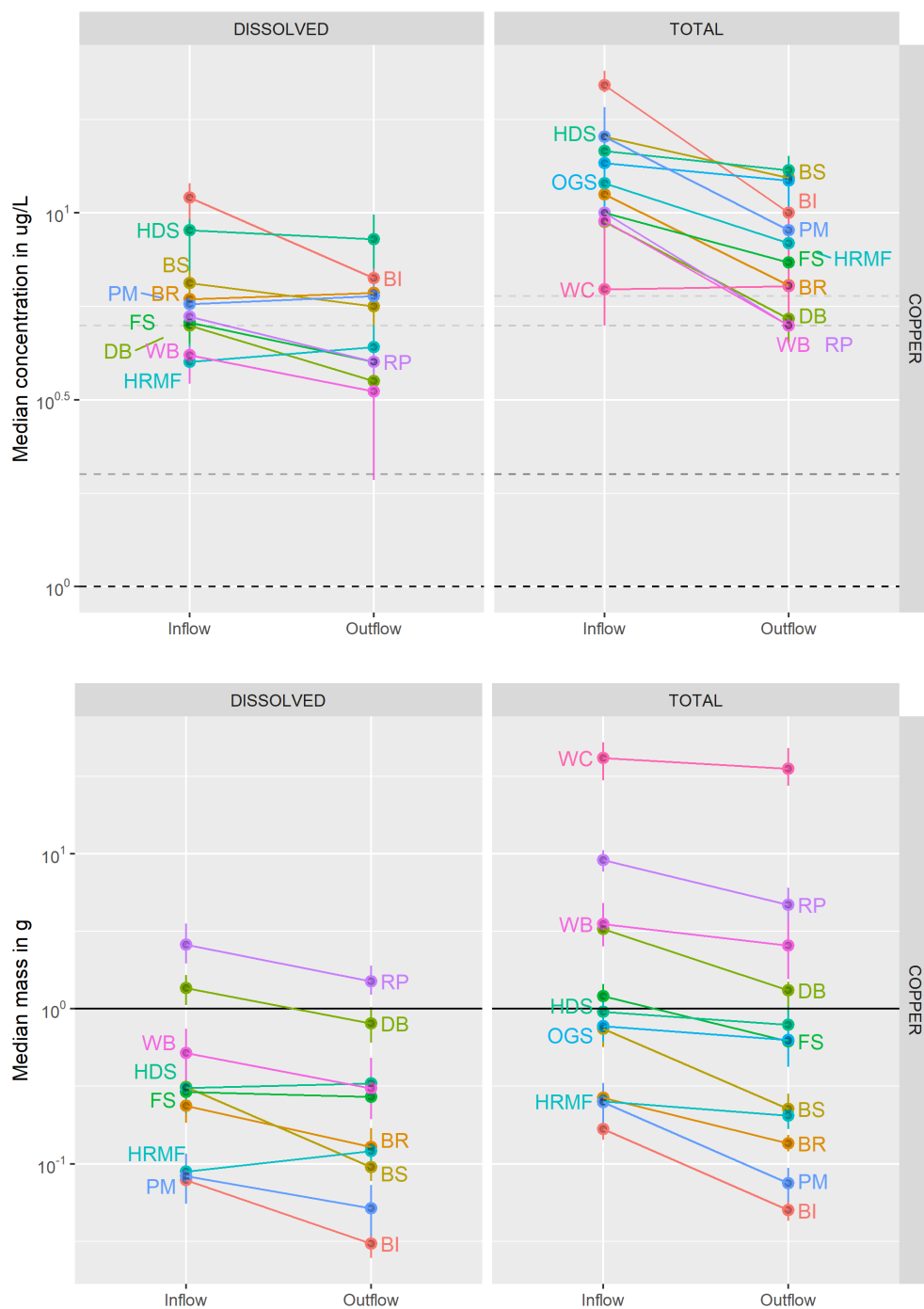


Figure D.6: Bootstrapped median total and dissolved Cu concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10⁰g).

Table D.7: Observed median concentrations and median mass flux estimates of dissolved and total lead (Pb) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Dissolved Pb	BI	29	589	25	402	0.50 - 0.50	0.50 - 0.50	0.0041 - 0.0058	0.0037 - 0.0053
	DB	10	209	10	203	1.0 - 1.1	1.0 - 1.4	0.22 - 0.41	0.16 - 0.30
	FS	9	149	9	144	1.0 - 1.0	1.0 - 1.0	0.053 - 0.084	0.041 - 0.064
	BR	11	265	11	194	0.40 - 0.50	0.25 - 0.50	0.01 - 0.016	0.0033 - 0.0070
	RP	14	193	14	203	1.0 - 1.5	1.0 - 1.0	1.1 - 2.4	0.60 - 1.8
Total Pb	HDS	9	139	9	134	8.3 - 12	6.4 - 9.0	0.36 - 0.77	0.24 - 0.49
	DB	16	272	17	260	5.0 - 8.1	2.5 - 5.1	2.0 - 4.5	1.0 - 1.8
	BS	19	220	19	201	5.8 - 12	4.9 - 7.4	0.29 - 0.5	0.099 - 0.16
	WC	12	204	12	168	5.0 - 8.0	2.7 - 5.0	15 - 26	10 - 19
	BI	32	648	27	434	7.2 - 9.1	3.1 - 4.5	0.055 - 0.090	0.021 - 0.047
	HRMF	7	103	7	103	12 - 22	5.0 - 5.0	0.19 - 0.45	0.11 - 0.28
	RP	42	785	43	767	8.0 - 10	2.5 - 3.0	9.6 - 16	4.8 - 7.0
	BR	22	317	21	289	5.0 - 6.2	1.0 - 2.1	0.15 - 0.27	0.027 - 0.042
	FS	20	309	18	283	9.1 - 12	1.5 - 2.0	0.77 - 1.3	0.16 - 0.27

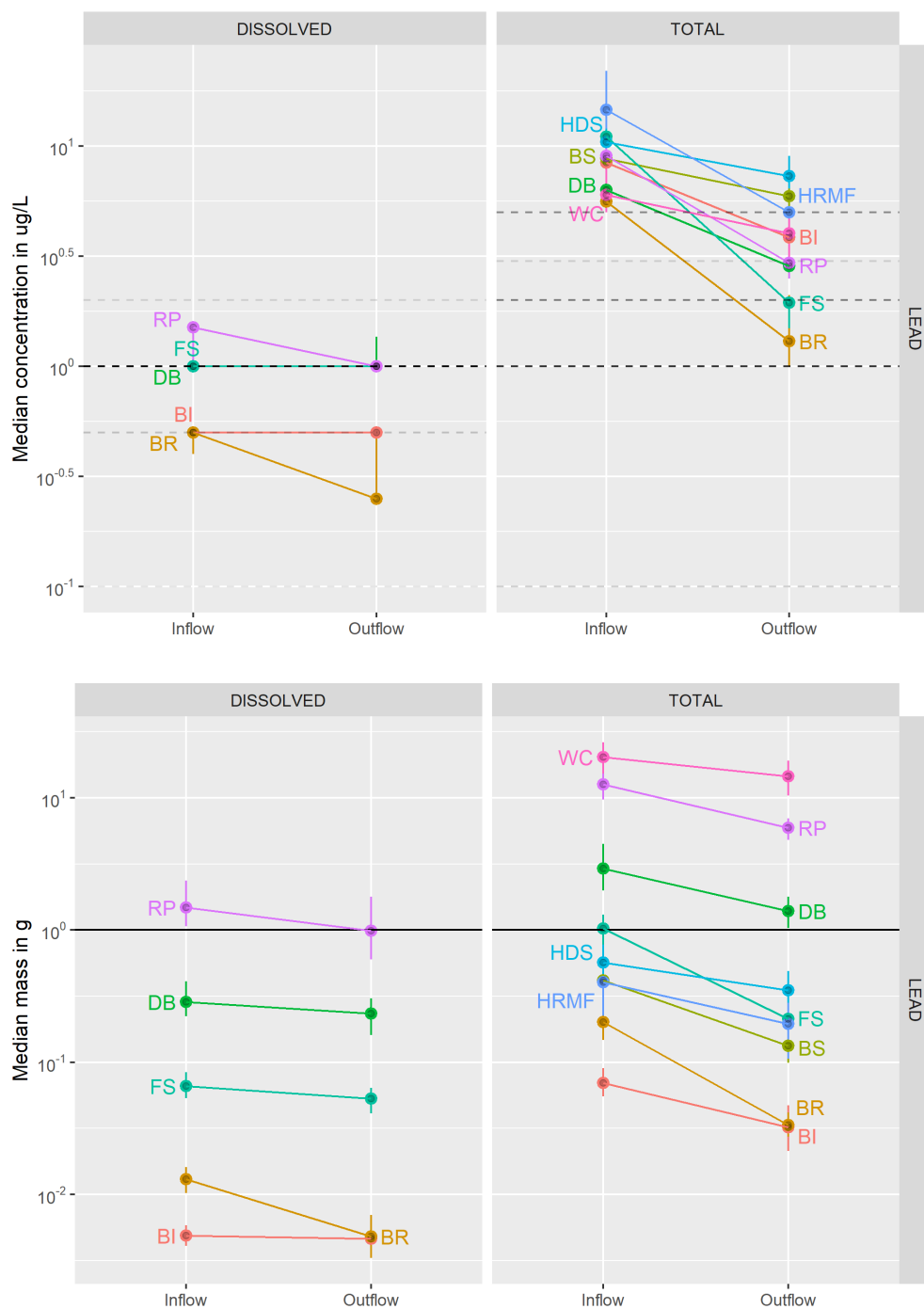


Figure D.7: Bootstrapped median total and dissolved Pb concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10^0 g).

Table D.8: Observed median concentrations and median mass flux estimates of dissolved and total zinc (Zn) entering and exiting BMP technologies. Bold values indicate there is no overlap between the inflow and outflow median ranges.

Analyte	BMP Type	Site and observation count				Median Concentration		Median Event Mass Flux	
		Inflow		Outflow		Inflow range (ug/L)	Outflow range (ug/L)	Inflow range (g)	Outflow range (g)
		n sites	n obs	n sites	n obs				
Dissolved Zn	HDS	9	122	9	123	35 - 47	36 - 52	0.95 - 2.0	1.2 - 2.4
	HRMF	13	225	13	225	15 - 17	16 - 20	0.31 - 0.47	0.48 - 0.66
	DB	10	209	10	203	13 - 17	9.2 - 15	3.2 - 5.8	2.2 - 3.3
	BS	16	174	16	141	28 - 35	17 - 22	1.3 - 2.2	0.25 - 0.40
	RP	20	356	19	335	19 - 24	10 - 13	8.9 - 21	3.5 - 9.2
	BI	32	634	27	431	30 - 39	15 - 20	0.18 - 0.26	0.071 - 0.10
	BR	13	298	13	227	17 - 22	11 - 14	0.62 - 0.98	0.14 - 0.28
	WB	7	119	9	114	21 - 26	7.0 - 9.4	2 - 3.7	0.7 - 1.3
	PM	3	125	5	114	18 - 25	7.0 - 10	0.24 - 0.44	0.051 - 0.096
	FS	10	158	9	143	32 - 51	6.0 - 12	2.5 - 4.5	0.48 - 0.95
Total Zn	OGS	9	115	9	123	54 - 84	67 - 100	3.4 - 9.2	3.1 - 6.6
	HDS	18	269	18	264	71 - 87	56 - 69	3.1 - 5.2	2.3 - 4.0
	WC	10	189	10	167	21 - 30	13 - 20	120 - 180	78 - 130
	HRMF	18	341	18	341	52 - 68	34 - 43	1.2 - 1.6	0.83 - 1.1
	BS	24	307	24	272	52 - 76	35 - 46	2.6 - 3.8	0.62 - 0.96
	DB	19	294	22	316	40 - 58	19 - 27	14 - 25	4.9 - 7.6
	WB	16	247	15	213	52 - 63	20 - 25	14 - 23	4.3 - 14
	RP	49	904	50	852	48 - 51	18 - 20	38 - 54	16 - 24
	BI	35	703	29	480	99 - 120	32 - 40	0.71 - 1.0	0.17 - 0.25
	PM	3	126	11	183	69 - 100	22 - 25	0.87 - 1.9	0.16 - 0.27
	FS	22	361	21	333	63 - 80	14 - 20	8.2 - 12	1.5 - 2.1
	BR	25	491	26	457	55 - 68	12 - 14	1.2 - 1.8	0.24 - 0.32

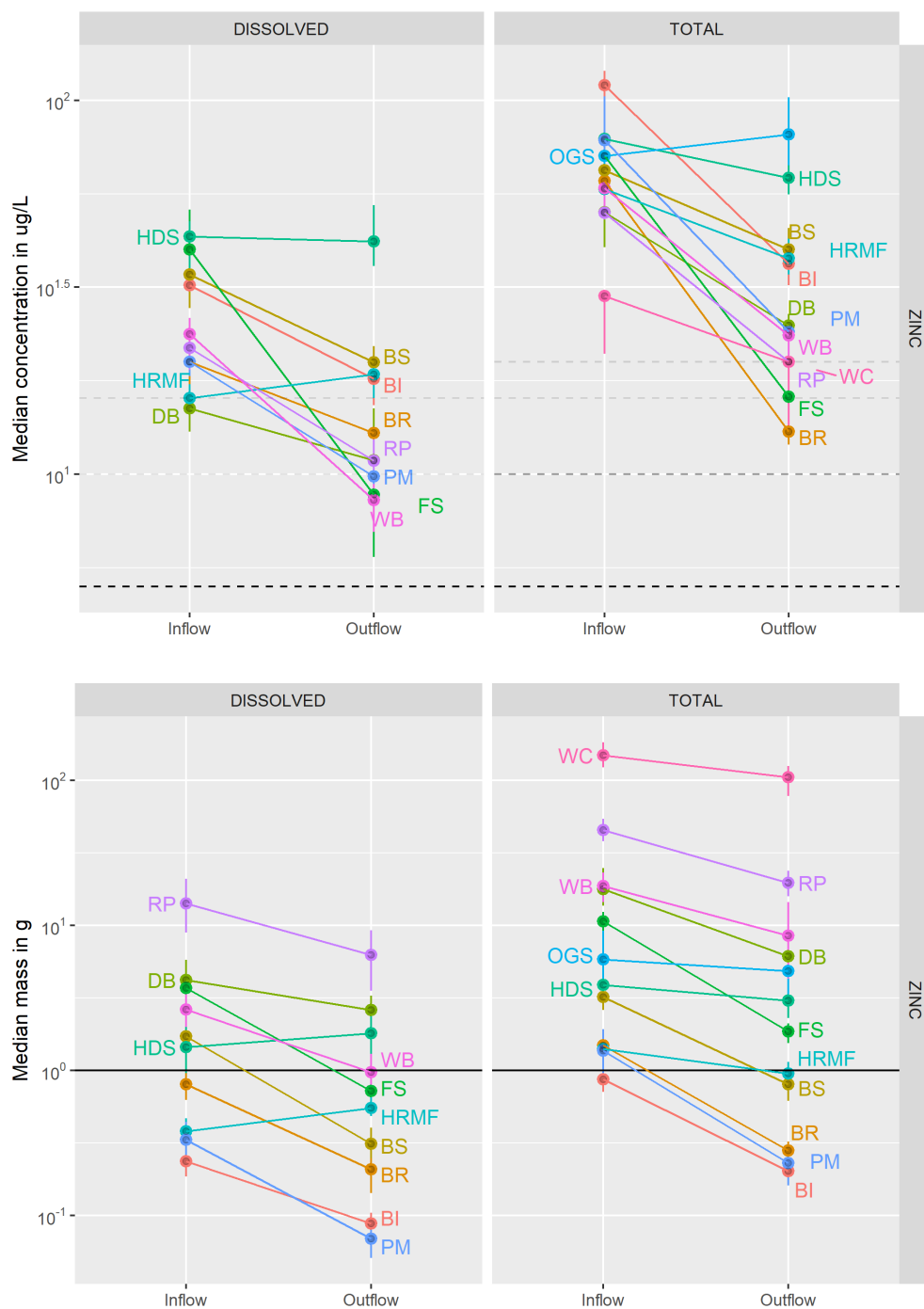


Figure D.8: Bootstrapped median total and dissolved Zn concentrations (top) and mass fluxes (bottom). BMPs with overlapping median inflow and outflow 90th percentile credible interval ranges are labelled to the left of the pairplots (no difference), those without overlap are labelled to the right (statistically different). Reference lines at the most commonly reported concentration detection limits (dotted lines) and 1 gram mass flux (solid line, 10^0 g).

E Contaminant concentrations and mass flux reduction estimates

Table E.1: Treatment effects on TSS, in units of OM change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate net removal and red to indicate net export. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux			Difference in Differences			Descriptive
		Order of OM	Order of Magnitude p-value	Reduction range (%)	Order of OM	Order of Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change			
TSS	WC	-0.24	2.4E-05 ***	25 - 55	-0.17	6.4E-02		0.07	-16%	constructed channel wetlands decrease TSS concentrations in effluent, but may have little effect on mass flux		
	BS	-0.22	7.5E-10 ***	30 - 49	-0.44	1.9E-12 ***	52 - 73	-0.22	39%	grass swales reduce TSS concentrations and volumetric reduction effects boost solids removal by 40%		
	HDS	-0.26	1.4E-10 ***	34 - 53	-0.26	1.0E-04 ***	26 - 60	-0.01	1%	hydrodynamic separators reduce TSS concentrations and mass flux at the same rate, indicating a common process: swirl separation		
	OGS	-0.34	4.2E-08 ***	39 - 65	-0.37	1.3E-03 **	30 - 74	-0.03	7%	oil and grit separators reduce TSS concentrations and mass flux at the same rate, indicating a common process: settling tanks with baffles and sediment traps		
	WB	-0.28	1.3E-23 ***	40 - 53	-0.20	4.5E-02 *	2 - 59	0.08	-20%	constructed wetland basins reduce TSS concentrations and mass flux at the same rate, indicating a common process: settling ponds, with probable resuspension and recirculation of captured sediment, resulting in relatively poor net capture		
	BI	-0.37	3.9E-38 ***	51 - 62	-0.45	8.6E-19 ***	55 - 71	-0.08	17%	grass biofilter strips reduce TSS concentrations and mass, volumetric reduction effect boosts net capture by 15-20%		
	DB	-0.38	7.9E-25 ***	52 - 65	-0.44	1.2E-11 ***	51 - 72	-0.05	11%	detention basins reduce TSS concentrations and mass at about the same rate, indicating a common process, likely particle settling or saltation over grass. Possible slight volumetric reduction resulting in very good net capture.		
	PM	-0.54	1.4E-12 ***	60 - 79	-0.88	3.6E-13 ***	78 - 92	-0.34	55%	modular block porous pavement indicates decreased concentration of TSS in effluent, and additional net capture due to volumetric reduction boost of 55%		
	HRMF	-0.52	4.4E-35 ***	64 - 75	-0.47	3.4E-16 ***	56 - 73	0.05	-11%	high rate media filters reduce TSS concentrations and mass through a common process, physical straining		
	RP	-0.59	1.4E-77 ***	71 - 78	-0.53	3.3E-21 ***	62 - 77	0.06	-16%	retention ponds have very good concentration reduction through settling, although net capture may be dampened by resuspension and recirculation of captured particles		
	BR	-0.76	4.1E-100 ***	80 - 85	-0.90	1.1E-63 ***	84 - 90	-0.15	29%	bioretention has very good TSS concentration reduction, and volumetric reduction effects boost net capture by 30%		
	FS	-0.80	1.7E-57 ***	81 - 87	-0.86	3.1E-26 ***	81 - 90	-0.06	12%	sand media filters have excellent TSS concentration reduction via physical straining alone, no additional volumetric reduction		
	HRBF	-0.91	2.8E-26 ***	84 - 90	-1.02	6.3E-12 ***	82 - 95	-0.11	22%	high rate biofiltration has excellent TSS concentration reduction, and volumetric reduction effect boost of 20%		

Table E.2: Treatment effects on TDS, represented by the order of magnitude (OM) change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate net removal and red to indicate net export. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. DID OM estimate color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		Order of Magnitude OM	p-value	Reduction range (%)	Order of Magnitude OM	p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results		
TDS	RP	0.22	2.8E-06 ***	-106 - -34	0.37	2.4E-03 **	-301 - -33	0.15	-40%	retention ponds show concentration increase and mass export of total dissolved solids, indicating that the standing water experiences an evaporative concentration effect		
	FS	0.23	1.8E-07 ***	-105 - -40	0.21	3.0E-02 *	-158 - -3	-0.02	3%			
	HDS	0.03	7.5E-01		0.02	8.3E-01		-0.01	2%	sand media filters show increases in TDS concentration and probable mass flux at the same rate, this indicates the export occurs due to a common process, such as temporary capture and leaching of solutes		
	BI	0.16	2.3E-11 ***	-59 - -30	0.03	5.7E-01		-0.13	26%			
	BS	0.02	5.3E-01		-0.19	6.9E-02		-0.21	39%	hydrodynamic separators have no effect on TDS concentration or mass flux		
	DB	-0.01	8.5E-01		-0.17	1.7E-01		-0.16	31%			
	HRMF	0.00	6.0E-01		-0.13	2.4E-01		-0.13	25%	grass biofilter strips show increased TDS concentrations in effluent, but do not exhibit net export due to a volumetric reduction effect of 25%		

Table E.3: Treatment effects on total nitrogen concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with "**", "***" and "***" respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		Order of Magnitude OM	p-value	Reduction range (%)	Order of Magnitude OM	p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results		
Total N	DB	0.02	4.8E-01		0.02	8.3E-01		0.00	0%	detention basins have little effect on total N concentrations or mass flux		
	HDS	-0.03	3.6E-01		-0.05	6.2E-01		-0.02	5%	hydrodynamic separators have little effect on total N concentrations or mass flux		
	BI	-0.05	2.0E-01		0.18	1.0E-01		0.23	-68%	grass biofilter strips may have little effect on total N concentrations or mass flux, and could possibly export N from fertilizer		
	FS	-0.07	4.9E-02 *	0 - 30	-0.17	7.1E-02		-0.10	20%	sand media filters show small reduction of total N concentrations, but volumetric reduction effect may or may not mediate net capture		
	WB	-0.08	1.6E-15 ***	13 - 21	-0.18	1.1E-03 **	16 - 49	-0.10	21%	wetland basins reduce concentrations of total N, and volumetric reduction effects boost net capture by 20% leading to very effective mass reduction		
	BR	-0.11	6.5E-05 ***	13 - 32	-0.36	2.4E-08 ***	42 - 67	-0.24	43%	bioretention reduces concentrations of total N, and volumetric reduction effects boost net capture by 40-45%		
	BS	-0.13	4.6E-05 ***	15 - 36	-0.10	2.0E-01		0.03	-6%	grass swales reduce concentrations of total N, but may not reduce mass flux		
	RP	-0.15	4.7E-20 ***	24 - 34	-0.05	4.0E-01		0.10	-25%	retention ponds reduce concentrations of total N in effluent, but have little effect on net export, possibly pointing to resuspension of settled materials		

Table E.4: Treatment effects on total nitrate concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences			Descriptive Interpretation of method 1a and 1b results
		Order of Magnitude OM	Reduction range (%)	Reduction range (%)	Order of Magnitude OM	Order of Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change				
NOx as N	PM	0.27	1.5E-07 ***	-125 - -49	-0.02	7.5E-01	-0.28	48%	modular block porous pavement indicates increased concentration of nitrate in effluent, but does not show net export due to volumetric reduction of 45-50%				
	FS	0.20	2.2E-11 ***	-80 - -39	0.14	8.4E-02	-0.06	14%	sand media filters indicate increased concentration of nitrate in effluent, but may or may not show net export				
	BR	0.16	1.3E-05 ***	-71 - -22	-0.07	1.2E-01	-0.23	41%	bioretention indicates increased concentration of nitrate in effluent, but does not export nitrate due to volumetric reduction				
	HRMF	0.00	9.2E-01		-0.20	1.4E-01	-0.20	37%	high rate media filtration may have little effect on nitrate concentrations				
	GR	-0.01	8.6E-01		-0.57	1.5E-03 **	-0.57	73%	green roofs have good net nitrate capture, due to high volumetric reduction effects of 70-75%. Fertilization effects are unknown				
	BS	0.07	9.1E-02		-0.03	9.2E-01	-0.10	20%	grass bioswales show little to no effect on nitrate concentration and net capture, with possible interference from fertilizer additions				
	WB	0.01	7.4E-01		0.03	6.9E-01	0.02	-6%	constructed wetland basins may have little effect on nitrate concentrations and mass flux				
	WC	-0.15	3.1E-02 *	0 - 49	-0.08	3.3E-01	0.06	-16%	constructed wetland channels reduce nitrate concentrations, but may or may not have an effect on mass flux. Net capture and export likely vary widely by site				
	HDS	-0.10	5.9E-03 **	6 - 33	-0.09	3.6E-01	0.01	-2%	hydrodynamic separators show small to moderate reduction of nitrate concentrations, but little net capture				
	DB	-0.16	1.1E-04 ***	17 - 43	-0.29	5.9E-05 ***	-0.13	26%	detention basins have good reduction of nitrate concentrations, and volumetric reduction effects boost mass removal by 25%				
	BI	-0.15	5.8E-07 ***	18 - 37	-0.23	3.3E-06 ***	-0.09	18%	grass biofilter strips reduce nitrate concentrations, and a small volumetric reduction effect boosts mass removal by 15-20%				
	RP	-0.40	6.2E-34 ***	54 - 66	-0.39	3.6E-09 ***	0.01	-3%	retention ponds show very good reduction of nitrate concentration and mass flux at the same rates, indicating a single removal process, likely biological or chemical. Resuspension of settled nitrate particles is unlikely.				

Table E.5: Treatment effects on total phosphorus concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		OM	Order of Magnitude	Reduction range (%)		OM	Order of Magnitude	Reduction range (%)		DID OM estimate	Net DID change	Descriptive
			p-value				p-value					
Total P	GR	0.48	1.8E-13 ***	-333 - -100		0.06	6.8E-01			-0.42	62%	green roofs have increased concentrations of total P in effluent, but large volumetric reduction effects boost capture by 60%, resulting in net zero mass flux. Fertilizer is likely an important factor
	BI	0.17	2.2E-11 ***	-65 - -31		0.14	3.5E-03 **	-73 - -10		-0.03	7%	grass biofilter strips increase concentration of total P in effluent and are net exporters of P. Fertilizer is likely an important factor
	BR	0.07	2.4E-02 *	-33 - 0		-0.13	1.3E-02 *	7 - 40		-0.20	36%	bioretention may slightly increase concentrations of total P in effluent, but volumetric reduction effects boost mass capture by 35%, resulting in net capture, not export
	BS	0.04	2.5E-01			-0.16	6.0E-03 **	9 - 47		-0.20	37%	grass swales show no effects on concentrations of total P, but volumetric reduction effects boost mass capture by 35-40%, resulting in net capture, not zero effect
	WC	-0.07	2.3E-02 *	2 - 26		-0.01	8.0E-01			0.06	-14%	constructed wetland channels may slightly decrease total P, but due to high flow volumes and net volumetric increase, there is little evidence of net capture. Groundwater or site-specific hydrologic features may be in play
	HDS	-0.09	1.1E-02 *	5 - 31		-0.13	1.2E-01			-0.04	8%	hydrodynamic separators reduce concentrations of total P, but may or may not result in net capture
	OGS	-0.20	1.2E-02 *	7 - 58		-0.27	5.6E-02			-0.07	14%	oil and grit separators reduce concentrations of total P, and may result in net capture
	DB	-0.11	2.5E-05 ***	12 - 30		-0.14	1.6E-02 *	5 - 45		-0.04	8%	detention basins reduce total P concentrations and mass flux at the same rate, pointing to a single removal process, possibly salination or biological uptake
	HRMF	-0.18	7.3E-05 ***	18 - 43		-0.10	1.5E-01			0.07	-19%	high rate media filtration reduces total P concentrations in effluent, but may or may not reduce mass flux, possible leaching of captured P material
	PM	-0.29	1.2E-07 ***	35 - 60		-0.69	3.6E-08 ***	64 - 88		-0.39	60%	modular block porous pavement shows reduction of total phosphorus and volumetric reduction effects boost capture by 60%
	FS	-0.28	5.8E-16 ***	39 - 54		-0.33	1.2E-06 ***	36 - 66		-0.05	11%	sand media filters reduce concentration and mass flux of total P at the same rate, pointing to a single removal process, physical straining of particulates
	WB	-0.25	9.4E-38 ***	40 - 50		-0.40	3.4E-10 ***	48 - 69		-0.15	29%	constructed wetland basins reduce concentrations of total P, and volumetric reduction effects boost net capture by 30%, resulting in very effective total P capture
	HRBF	-0.30	1.9E-10 ***	40 - 57		-0.39	4.9E-03 **	28 - 76		-0.08	18%	high rate biofiltration has good total P concentration reduction, and possible biological effects boost net capture by 20%
	RP	-0.30	1.2E-37 ***	46 - 56		-0.26	1.4E-06 ***	30 - 58		0.04	-9%	retention ponds reduce total P concentration and mass flux by the same rate, likely indicating a single removal process, probably settling of particulates

Table E.6: Treatment effects on dissolved and reactive P concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration			Change in Mass Flux			Difference in Differences		Descriptive Interpretation of method 1a and 1b results
		Order of OM	Magnitude p-value	Reduction range (%)	Order of OM	Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change	
Dissolved P	DB	0.00	6.1E-01		-0.09	4.9E-01		-0.09	19%	detention basins likely have little effect on dissolved P concentration and on mass flux
	HRMF	0.00	2.0E-01		-0.11	1.9E-01		-0.10	21%	high rate media filtration likely has little effect on dissolved P concentration and may or may not have a small effect mass flux
	WB	-0.10	1.0E-07 ***	0 - 33	-0.08	1.8E-01		0.02	-4%	constructed wetland basins reduce effluent concentrations of dissolved P, mass flux may or may not decrease
	HDS	-0.08	1.5E-01		-0.10	2.8E-01		-0.02	5%	hydrodynamic separators show little effect on reduction of dissolved P concentration and mass flux
	RP	-0.24	4.7E-12 ***	33 - 50	-0.18	9.6E-03 **	9 - 52	0.06	-14%	retention ponds reduce dissolved P concentrations and mass flux at about the same rate, pointing to a single removal process, probably biological
Total PO4	BR	0.70	2.2E-41 ***	-513 - -300	0.41	1.9E-08 ***	-251 - -86	-0.29	49%	bioretention increases PO4 concentration in effluent, and despite volumetric reduction effects of 50%, it remains a net exporter
	BI	0.30	1.0E-03 **	-250 - -33	0.24	5.1E-02		-0.07	14%	grass biofilter strips increase PO4 concentrations in effluent and are likely net exporters. Fertilizer is likely an important factor
	WC	0.24	6.1E-05 ***	-130 - -33	0.33	6.4E-03 **	-246 - -31	0.09	-22%	constructed wetland channels increase PO4 concentrations in effluent and export large fluxes of PO4. Groundwater or site-specific hydrologic features may increase net discharge
	DB	0.00	7.5E-01		-0.21	1.9E-01		-0.21	38%	detention basins do not reduce PO4 concentration, but may or may not show a volumetric reduction effect that boosts net capture
	HRMF	0.00	4.2E-01		0.22	5.5E-02		0.22	-66%	high rate media filters do not reduce PO4 concentrations, but may or may not affect mass flux
	WB	0.00	2.4E-02 *	0 - 20	-0.05	3.7E-01		-0.05	11%	wetland basins reduce PO4 concentrations somewhat, but do not exhibit net mass capture
	RP	-0.30	1.0E-04 ***	28 - 67	-0.42	3.4E-03 **	26 - 80	-0.12	24%	retention ponds reduce PO4 concentrations and may have a secondary removal process that affects PO4 capture, although volumetric reduction effects are unlikely

Table E.7: Treatment effects on dissolved and total cadmium concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		OM	Order of Magnitude	p-value	Reduction range (%)	OM	Order of Magnitude	p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive
Dissolved Cd	BR	0.08	2.6E-02 *	-82 - 0		-0.24	3.5E-02 *	6 - 65		-0.32	52%	bioretention tends to have higher concentrations of dissolved Cd in effluent due to very low influent concentrations, but volumetric runoff reduction effects boost removal by +50% resulting in net capture, not net export
	BI	0.00	3.5E-02 *	0 - 0		-0.17	2.7E-04 ***	16 - 46		-0.17	33%	grass strips do not change dissolved Cd concentrations in effluent, but volumetric runoff reduction effect boosts removal by +30% resulting in net capture, not zero effect
	DB	0.00	9.1E-01			-0.10	2.6E-01			-0.10	21%	detention basins have little effect on dissolved Cd concentrations or mass flux
	FS	0.00	1.5E-02 *	0 - 50		-0.07	4.1E-01			-0.07	14%	sand filtration systems likely do not reduce concentrations of dissolved Cd, and result in little net capture
Total Cd	DB	0.00	2.9E-02 *	0 - 12		-0.20	2.5E-02 *	3 - 58		-0.19	36%	detention basins may reduce total Cd concentrations a small amount, but volumetric reduction effect boosts total Cd removal by +35% resulting in very good net capture
	HDS	-0.10	6.6E-02			-0.17	1.2E-01			-0.08	16%	hydrodynamic separation has little effect on total Cd concentrations or mass flux
	BR	-0.08	6.2E-02			-0.21	1.1E-02 *	10 - 58		-0.13	26%	bioretention volumetric runoff reduction effect boosts total Cd removal by +25% resulting in net capture, not zero effect despite lower influent concentrations
	BS	-0.17	1.1E-08 ***	20 - 42		-0.53	1.7E-11 ***	60 - 79		-0.36	56%	grass swale volumetric reduction effects boost total Cd removal by 55%, resulting in reduced concentrations and good net capture
	RP	-0.20	7.6E-07 ***	20 - 48		-0.12	9.7E-02			0.08	-21%	retention ponds reduce concentrations of total Cd in effluent, but have little effect on net export, pointing to resuspension of settled materials or redox mobilization
	FS	-0.22	1.9E-05 ***	20 - 56		-0.25	1.8E-03 **	20 - 61		-0.03	6%	sand filtration systems reduce concentrations and mass flux of total Cd by about the same rates, pointing to a physical filtration
	BI	-0.30	6.7E-21 ***	44 - 52		-0.41	7.8E-15 ***	51 - 69		-0.11	23%	grass biofilter strips reduce total Cd concentrations, and volumetric runoff reduction effects boost removal by +20-25%

Table E.8: Treatment effects on dissolved and total chromium concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration			Change in Mass Flux			Difference in Differences		
		Order of OM	Magnitude p-value	Reduction range (%)	Order of OM	Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results
Dissolved Cr	FS	0.00	5.8E-01		0.04	6.4E-01		0.04	-9%	sand media filtration likely has little effect on dissolved Cr concentrations or flux
	BI	-0.03	1.6E-01		-0.17	3.0E-04 ***	16 - 46	-0.14	28%	grass strip biofilters do not reduce dissolved Cr concentrations but volumetric reduction effects boost net capture by 25-30%
	BR	-0.14	1.3E-03 **	0 - 50	-0.62	8.7E-08 ***	60 - 86	-0.48	67%	bioretention reduces concentrations of dissolved Cr, and good volumetric reduction effects boost capture by an additional 65%
Total Cr	WC	0.00	8.3E-01		0.04	6.1E-01		0.04	-9%	constructed channel wetlands likely have little effect on total Cr concentrations and flux
	BI	-0.12	8.1E-06 ***	14 - 34	-0.25	3.0E-07 ***	30 - 55	-0.13	26%	grass strip biofilters reduce total Cr concentrations well and volumetric reduction effects boost capture by 25-30%
	RP	-0.20	1.5E-04 ***	15 - 50	-0.18	1.7E-01		0.02	-4%	retention ponds may reduce concentrations of total Cr in effluent, but have little effect on net export, pointing to resuspension of settled materials
	FS	-0.18	1.2E-04 ***	23 - 50	-0.19	2.4E-02 *	7 - 55	-0.02	3%	sand filtration systems reduce concentrations and mass flux of total Cr by about the same rates, pointing to a common mechanism: physical filtration
	BR	-0.40	2.0E-16 ***	56 - 69	-0.62	3.0E-10 ***	64 - 84	-0.22	40%	bioretention reduces total Cr concentrations and volumetric runoff reduction effect boosts net capture by 40%

Table E.9: Treatment effects on dissolved copper concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration			Change in Mass Flux			Difference in Differences		
		OM	Order of Magnitude	Reduction range (%)	OM	Order of Magnitude	Reduction range (%)	DID OM estimate	Net DID change	Descriptive
Dissolved Cu	HRMF	0.00	1.0E+00		0.17	1.2E-02 *	-104 - -7	0.17	-48%	high rate media filtration likely accumulates Cu in influent stormwater and exports dissolved Cu in effluent
	HDS	-0.03	5.8E-01		-0.02	7.0E-01		0.00	0%	hydrodynamic separation has little effect on dissolved Cu concentrations or mass flux
	PM	-0.02	5.7E-01		-0.24	3.1E-02 *	6 - 64	-0.21	39%	modular block porous pavement may act as a sink for dissolved Cu via volumetric reduction effects, resulting in 40% net capture instead of zero effect
	BR	-0.02	5.3E-01		-0.25	3.0E-03 **	17 - 62	-0.23	41%	bioretention has little effect on dissolved Cu concentrations, but volumetric reduction effects boost capture by 40%, resulting in net capture instead of zero effect
	BS	-0.07	1.8E-01		-0.43	1.1E-07 ***	47 - 73	-0.36	56%	grassed swales has little effect on dissolved Cu concentrations, but volumetric reduction effects boost capture by 55%, resulting in good net capture instead of zero effect
	FS	-0.09	8.0E-02		-0.05	3.8E-01		0.04	-9%	sand media filters has little effect on dissolved Cu concentrations or mass flux
	DB	-0.09	3.0E-02 *	0 - 35	-0.18	2.2E-02 *	8 - 53	-0.09	20%	detention basins may moderately reduce dissolved Cu concentrations, but volumetric reduction effects boost dissolved Cu mass removal by 20%, resulting in good net capture
	WB	-0.18	7.6E-04 ***	15 - 48	-0.27	2.5E-02 *	6 - 67	-0.09	19%	constructed wetland basins reduce dissolved Cu concentrations, but volumetric reduction effects are small and likely vary site by site
	RP	-0.14	2.3E-10 ***	20 - 33	-0.28	2.6E-04 ***	26 - 64	-0.15	29%	retention ponds reduce dissolved Cu concentrations and mass at roughly the same rates, pointing to a common removal mechanism, possibly adsorption or biological uptake. Possible volumetric effects that boost net capture by up to 30%
	BI	-0.17	2.4E-08 ***	22 - 41	-0.35	9.0E-11 ***	43 - 65	-0.18	35%	grass biofilter strips lower dissolved Cu concentration and volumetric reduction boosts net capture by 35%

Table E.10: Treatment effects on total copper concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		Order of OM	Magnitude p-value	Reduction range (%)	Order of OM	Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results		
Total Cu	WC	-0.04	3.4E-01		-0.09	3.9E-01		-0.05	11%	constructed wetland channels likely have little effect on total Cu concentrations or mass flux		
	HDS	-0.04	2.0E-01		-0.05	5.8E-01		-0.01	1%	hydrodynamic separation has little effect on total Cu concentrations and mass flux		
	OGS	-0.09	1.1E-01		-0.14	8.8E-02		-0.05	11%	oil/grit separators and baffle boxes show little effect on total Cu concentrations and mass fluxes		
	BS	-0.15	5.2E-05 ***	17 - 41	-0.58	1.1E-14 ***	63 - 81	-0.43	63%	grass swales reduce total Cu concentrations and large volumetric reduction effects boost removal by 60-65%		
	HRMF	-0.23	9.5E-08 ***	29 - 52	-0.07	2.1E-01		0.16	-46%	high rate media filtration shows good reduction of total Cu concentrations, but no net capture, possibly indicating Cu leaching from filter		
	BR	-0.23	1.6E-17 ***	34 - 48	-0.31	1.3E-08 ***	38 - 62	-0.08	17%	bioretention reduces total Cu concentrations and mass flux at about the same rate, indicating a common removal process. Possible small volumetric runoff reduction effects that boost net capture by up to 20%		
	FS	-0.25	2.6E-14 ***	34 - 52	-0.33	9.6E-07 ***	37 - 65	-0.08	17%	sand media filters reduce total Cu concentrations and mass flux at about the same rate, indicating a common removal process. Possible small volumetric runoff reduction effects that boost net capture by up to 20%		
	DB	-0.29	2.0E-09 ***	35 - 58	-0.41	2.0E-07 ***	45 - 73	-0.12	25%	detention basins reduce total Cu concentrations, small volumetric reduction effects boost net capture by 25%		
	PM	-0.29	9.4E-10 ***	37 - 59	-0.55	2.5E-08 ***	56 - 83	-0.27	46%	modular block porous pavement shows good reduction of total Cu concentration and volumetric reduction effects boost net capture by 45%		
	RP	-0.29	7.7E-37 ***	43 - 53	-0.25	5.5E-06 ***	28 - 57	0.04	-10%	retention ponds reduce total Cu concentration and mass flux by the same rate, likely indicating a common biological or chemical removal mechanism		
	BI	-0.30	9.9E-23 ***	43 - 57	-0.46	9.6E-18 ***	56 - 73	-0.16	31%	grass biofilter strips lower total Cu concentration and volumetric reduction boosts net capture by 30%		
	WB	-0.31	1.2E-15 ***	43 - 58	-0.23	2.1E-02 *	8 - 63	0.08	-19%	constructed wetland basins reduce dissolved Cu concentrations and mass by roughly the same amount but vary site by site		

Table E.11: Treatment effects on dissolved and total lead concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with “*”, “**”, and “***” respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		Descriptive Interpretation of method 1a and 1b results
		Order of OM	Order of Magnitude p-value	Reduction range (%)	Order of OM	Order of Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change			
Dissolved Pb	BI	0.00	8.9E-01		-0.13	1.8E-02 *	5 - 43	-0.13	26%	grass biofilter strips do not reduce concentrations of dissolved lead, a volumetric reduction effect of 25% may result in net capture		
	DB	0.00	9.1E-01		-0.11	2.6E-01		-0.11	22%	detention basins likely have little effect on dissolved Pb concentration and mass flux		
	FS	0.00	2.5E-02 *	0 - 33	-0.10	2.8E-01		-0.10	21%	sand media filters have little or no effect on dissolved Pb concentration and mass flux		
	BR	-0.06	5.5E-02		-0.41	2.5E-05 ***	40 - 75	-0.35	55%	bioretention does not reduce concentrations of dissolved lead, but volumetric reduction effects boosts net capture by 55%		
	RP	-0.07	1.6E-01		-0.17	1.6E-01		-0.10	21%	retention ponds have little effect on dissolved Pb concentrations and mass flux		
	HDS	-0.10	3.9E-02 *	0 - 36	-0.08	3.2E-01		0.02	-4%	hydrodynamic separators may reduce total Pb concentration minimally, but likely do not reduce mass export		
Total Pb	DB	-0.17	1.7E-03 **	0 - 52	-0.35	4.1E-05 ***	34 - 70	-0.19	35%	detention basins reduce total Pb concentrations, and volumetric reduction effects boost net capture by 35%		
	BS	-0.18	5.9E-03 **	11 - 51	-0.47	2.2E-06 ***	48 - 78	-0.29	48%	grass swales reduce total Pb concentrations and volumetric reduction effects boost removal by 50%		
	WC	-0.20	3.3E-03 **	13 - 51	-0.16	7.1E-02		0.04	-9%	wetland channels reduce total Pb concentrations, but may have no effect on mass flux		
	BI	-0.38	1.5E-13 ***	46 - 66	-0.46	1.4E-09 ***	51 - 75	-0.09	18%	grass biofilter strips reduce total Pb concentrations, and a small volumetric reduction effect boosts mass removal by 15-20%		
	HRMF	-0.44	2.4E-08 ***	49 - 76	-0.34	1.6E-02 *	11 - 76	0.10	-24%	high rate media filtration removes concentrations and mass of total Pb at approximately the same rate, likely pointing to a common process. Large variation indicates some filters may leach captured Pb		
	RP	-0.46	5.6E-36 ***	60 - 70	-0.37	1.6E-09 ***	43 - 68	0.09	-24%	retention ponds remove concentrations and mass of total Pb at approximately the same rate, likely pointing to a common process. Wet ponds may resuspend and export previously captured particles		
	BR	-0.60	8.8E-29 ***	67 - 80	-0.75	1.3E-19 ***	75 - 88	-0.15	30%	bioretention reduces total Pb concentrations well, and volumetric reduction effects boost capture by an additional 30%		
	FS	-0.70	3.3E-33 ***	75 - 84	-0.70	2.4E-15 ***	71 - 87	-0.01	1%	sand media filters reduce total Pb concentrations and mass flux at about the same rate, indicating a common removal process		

Table E.12: Treatment effects on dissolved zinc concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with "**", "***", and "****" respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration			Change in Mass Flux			Difference in Differences		
		Order of OM	Magnitude p-value	Reduction range (%)	Order of OM	Magnitude p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results
Dissolved Zn	HDS	0.00	9.5E-01		0.01	7.4E-01		0.01	-1%	hydrodynamic separators likely have no effect on dissolved Zn concentrations or flux
	HRMF	0.05	1.3E-01		0.20	3.8E-03 **	-118 - -16	0.15	-41%	high rate media filtration does not affect dissolved zinc concentrations, but does not export dissolved zinc, likely related to redox effects mobilizing captured Zn material
	DB	-0.12	5.3E-02		-0.21	1.2E-02 *	10 - 58	-0.09	19%	detention basins do not reduce Zn concentration, but show a volumetric reduction effect that boosts net capture by 20%
	BS	-0.24	3.6E-07 ***	28 - 53	-0.64	1.3E-11 ***	66 - 84	-0.40	61%	grassed swales reduce dissolved Zn concentrations, and volumetric reduction effects boost capture by an additional 60%, resulting in very good net capture
	RP	-0.25	7.5E-09 ***	32 - 53	-0.31	4.8E-04 ***	27 - 68	-0.06	14%	retention ponds reduce dissolved Zn concentrations and mass flux at the same rate, pointing to a single unknown sink mechanism that is very effective at mass removal
	BI	-0.31	2.4E-18 ***	43 - 58	-0.49	1.5E-16 ***	58 - 75	-0.18	34%	grass biofilter strips reduce dissolved Zn concentrations, and volumetric reduction effects boost net capture by 35%
	BR	-0.35	5.1E-14 ***	45 - 64	-0.60	1.2E-12 ***	62 - 83	-0.25	44%	bioretention reduces dissolved Zn concentrations, and volumetric reduction effects boost net capture by 45%
	WB	-0.45	2.6E-16 ***	56 - 72	-0.49	8.7E-06 ***	45 - 81	-0.03	8%	wetland basins reduce dissolved Zn concentrations and mass at the same rate, indicating a single removal process, possibly redox or biological mechanisms
	PM	-0.49	1.1E-15 ***	56 - 75	-0.68	1.2E-09 ***	65 - 88	-0.19	35%	porous pavement reduces dissolved Zn concentrations and volumetric reduction effects boost net capture by 35%
	FS	-0.74	5.9E-20 ***	74 - 86	-0.67	1.5E-13 ***	69 - 86	0.07	-17%	sand media filters reduce concentration and mass flux of dissolved Zn at the same rate, pointing to a very effective single removal mechanism, possibly adsorption

Table E.13: Treatment effects on total zinc concentration and mass flux, represented by the order of magnitude change between inflow and outflow. Color gradient for OM change in concentration and mass flux uses green to indicate decrease and red to indicate increase. Estimates with p-values less than 0.05, 0.01, and 0.001 are indicated with **, ***, and **** respectively. For OM change estimates with p values less than 0.05, the converted equivalent percent reduction range is reported based on the confidence interval around the mean OM change value. Difference in differences (DID OM estimate) is a simple difference between mass flux and concentration OM change. Color gradient uses yellow to indicate mass flux treatment performance worse than concentration estimate, blue to indicate better.

Analyte	BMP Type	Change in Concentration				Change in Mass Flux				Difference in Differences		
		Order of Magnitude OM	Reduction range (%)	Order of Magnitude OM	Reduction range (%)	Order of Magnitude OM	p-value	Reduction range (%)	DID OM estimate	Net DID change	Descriptive Interpretation of method 1a and 1b results	
Total Zn	OGS	0.02	8.2E-01	-0.07	5 - 29	-0.07	5.4E-01	64 - 81	-0.08	18%	oil and grit separators likely have little effect on total Zn concentration and mass flux	
	HDS	-0.08	6.6E-03 **	-0.10	5 - 29	-0.10	1.7E-01	54 - 77	-0.01	3%	hydrodynamic separators reduce concentrations of total Zn, but likely do not result in net capture	
	WC	-0.18	6.1E-04 ***	-0.13	16 - 47	-0.13	1.5E-01	31 - 74	0.05	-11%	wetland channels reduce total Zn concentrations, but may have no effect on mass flux or relatively poor net capture	
	HRMF	-0.21	5.6E-08 ***	-0.13	27 - 48	-0.13	7.7E-02	70 - 82	0.09	-22%	high rate media filters reduce total zinc concentrations, but may not result in net capture because of interactions with dissolved Zn	
	BS	-0.24	1.1E-08 ***	-0.58	31 - 52	-0.58	2.6E-14 ***	76 - 88	-0.34	55%	grass swales reduce total Zn concentration well and volumetric reduction effects boost net capture by 55%	
	DB	-0.35	2.1E-12 ***	-0.48	44 - 64	-0.48	5.6E-10 ***	79 - 87	-0.13	27%	detention basins reduce total Zn concentration and volumetric reduction effects boost net capture by 25%	
	WB	-0.43	2.1E-26 ***	-0.37	57 - 69	-0.37	7.1E-04 ***	47 - 67	0.06	-15%	wetland basins reduce total Zn concentrations and mass flux at the same rate, pointing to a single removal process	
	RP	-0.45	1.0E-79 ***	-0.38	61 - 68	-0.38	8.1E-12 ***	70 - 82	0.07	-17%	retention ponds reduce total Zn concentrations and mass flux at the same rate, pointing to a single removal process, possible settling and recirculation of settled particles	
	BI	-0.48	4.9E-49 ***	-0.63	62 - 71	-0.63	1.7E-27 ***	78 - 91	-0.14	28%	grass biofilter strips show good reduction of total Zn, and volumetric reduction effects boost net capture by 25-30%	
	PM	-0.56	9.0E-26 ***	-0.86	66 - 78	-0.86	1.8E-13 ***	76 - 88	-0.30	50%	porous pavement reduces total Zn concentrations very well and volumetric reduction effects boost net capture by 50%	
	FS	-0.68	3.4E-51 ***	-0.76	75 - 83	-0.76	1.1E-23 ***	79 - 87	-0.08	17%	sand media filters reduce concentrations of total Zn, with good net capture and possible volumetric effects boosting net capture slightly	
	BR	-0.68	1.4E-82 ***	-0.79	76 - 82	-0.79	2.6E-41 ***	79 - 87	-0.11	22%	bioretention reduces concentrations of total Zn, and volumetric effects boost net capture by 20-25%	

F Shifts in dissolved fractional ratio

Figure F.1: Empirical cumulative distribution functions for solids in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for twelve BMP types. Bootstrapped inflow and outflow ranges are marked with vertical lines. Shaded ranges represent the shortest half of the data distribution.

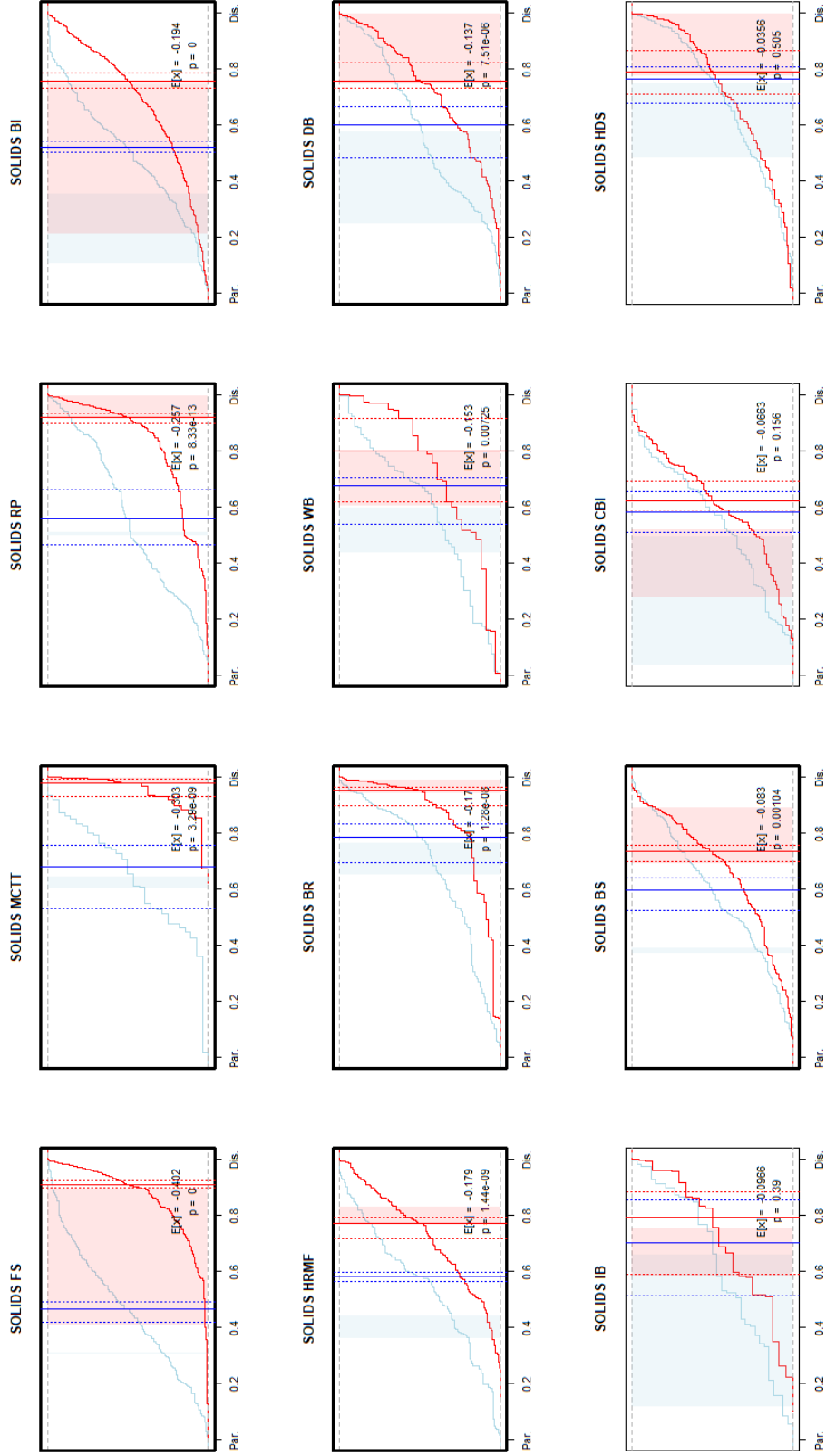


Figure F.2: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved nutrients in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for twelve BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.

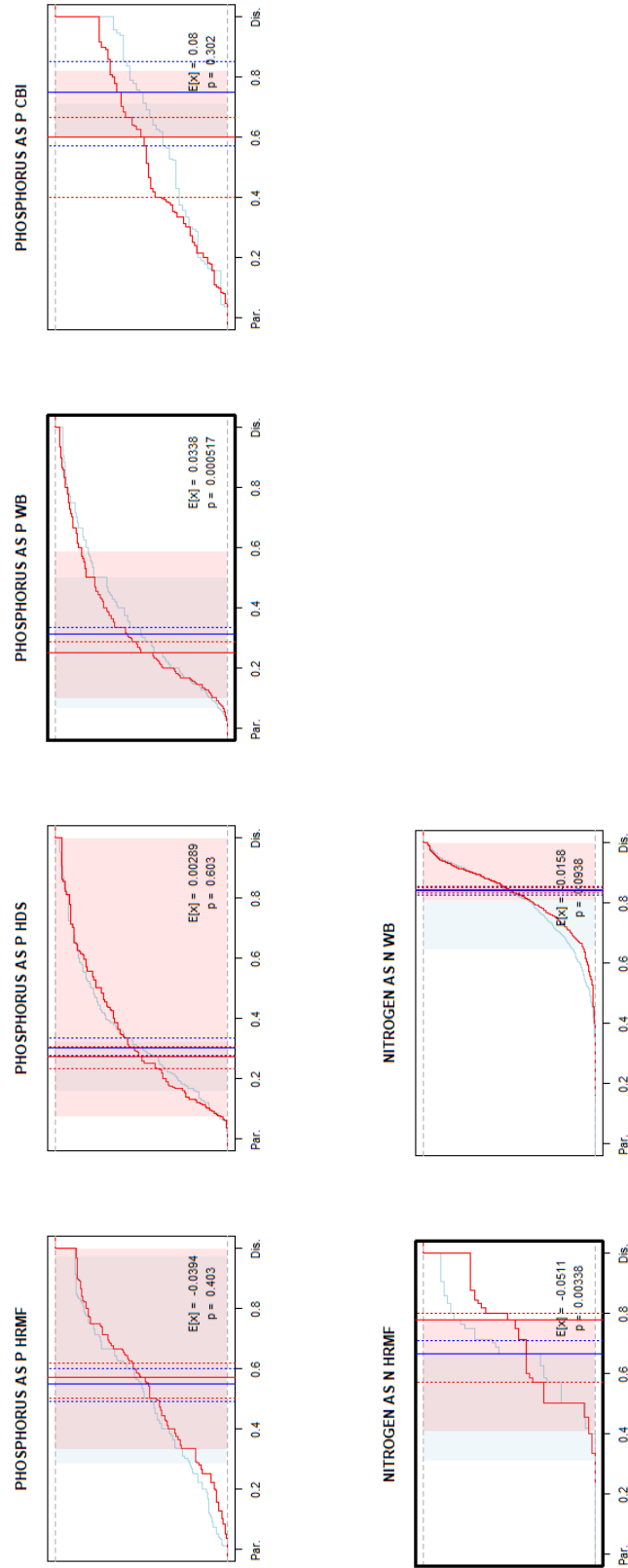


Figure F.3: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved cadmium in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for twelve BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.

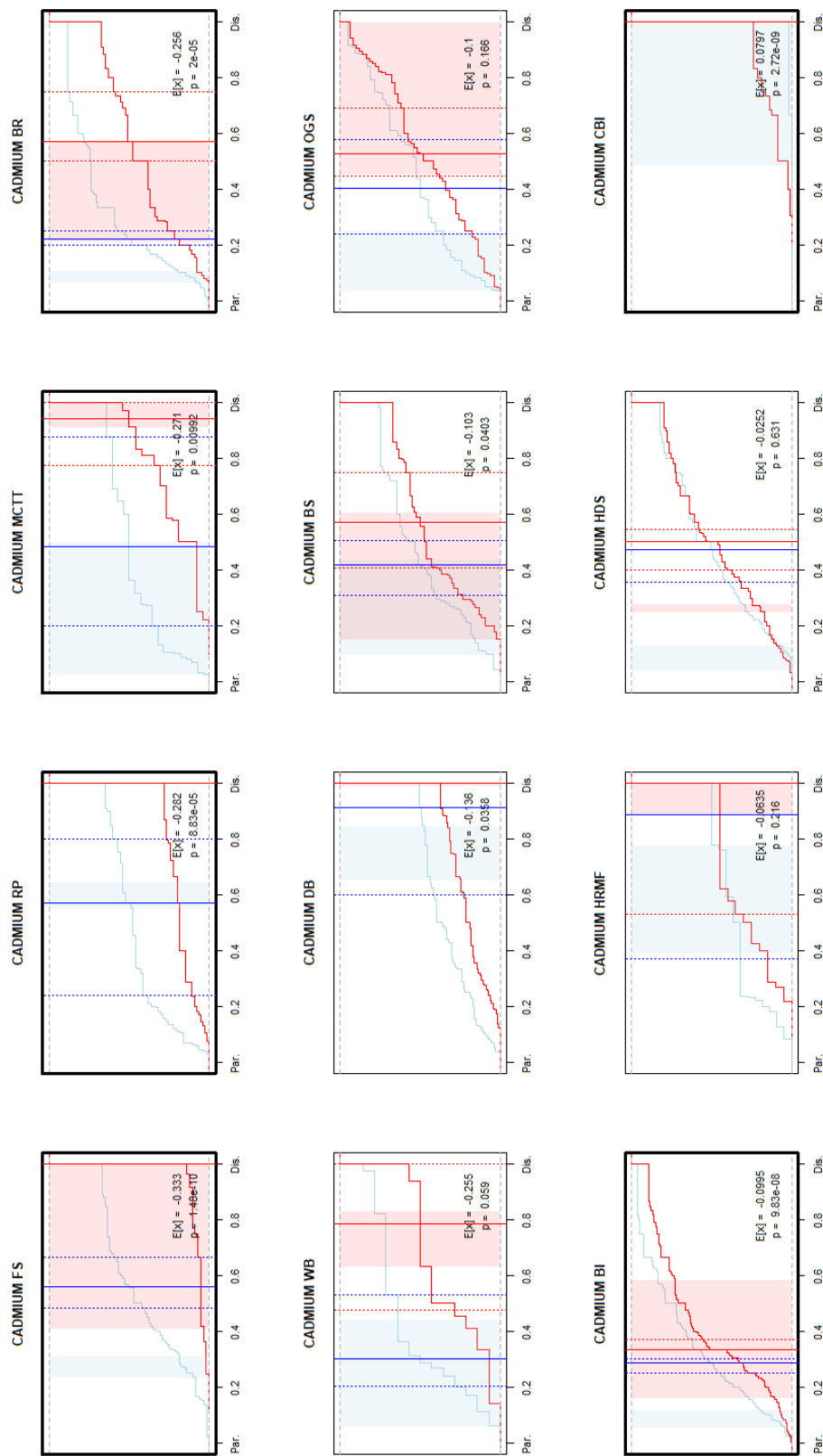


Figure F.4: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved chromium in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for twelve BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.

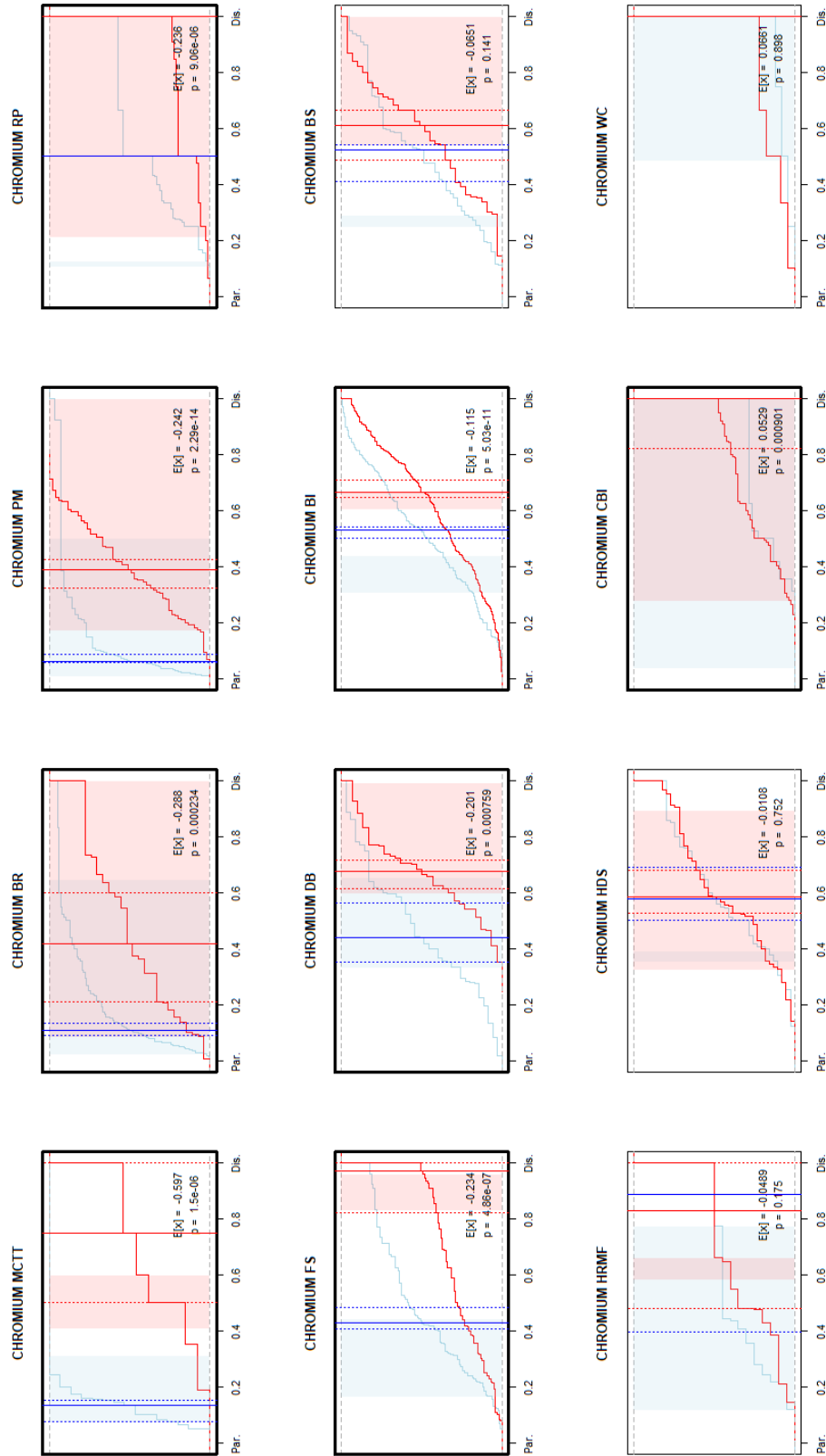


Figure F.5: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved copper in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for eight BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.

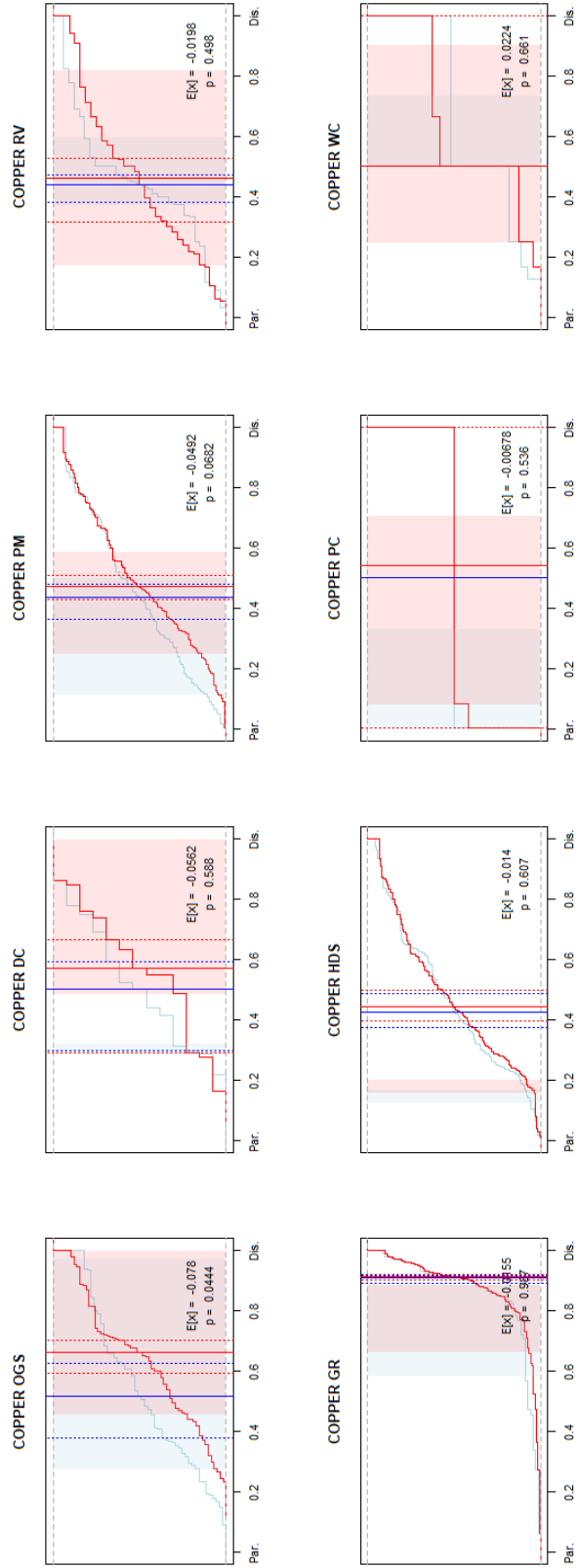


Figure F.6: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved lead in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for five BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.

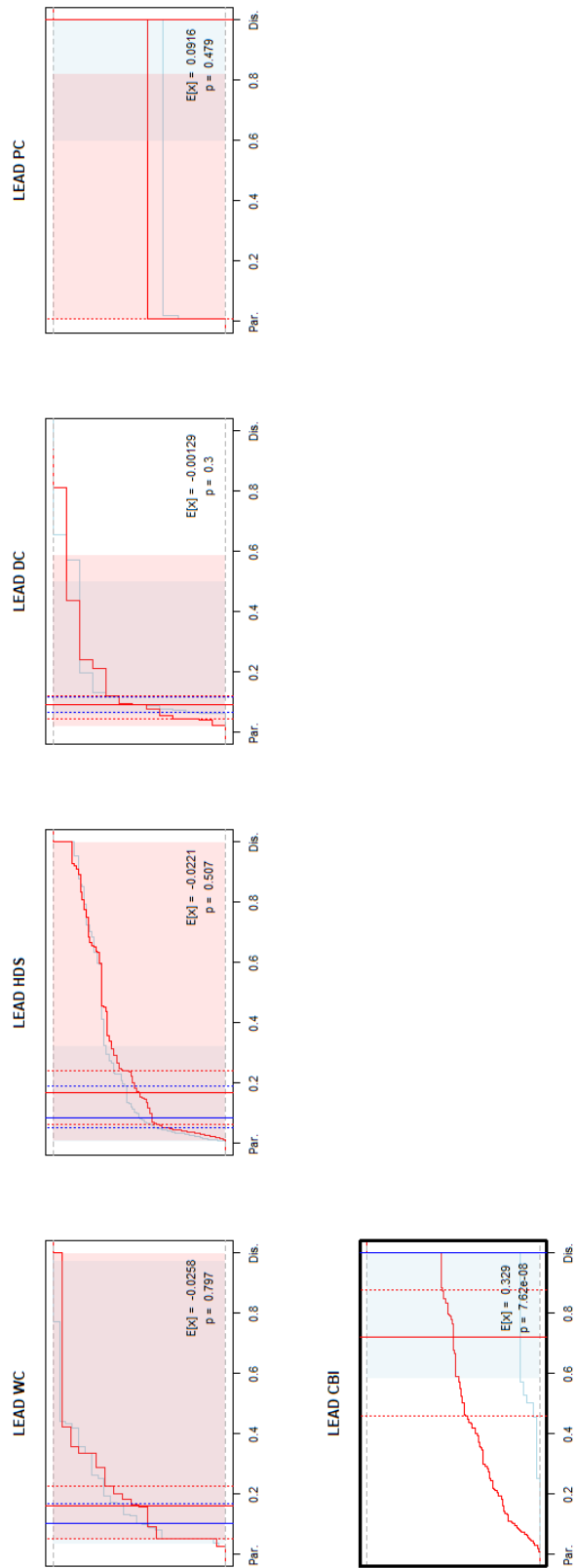
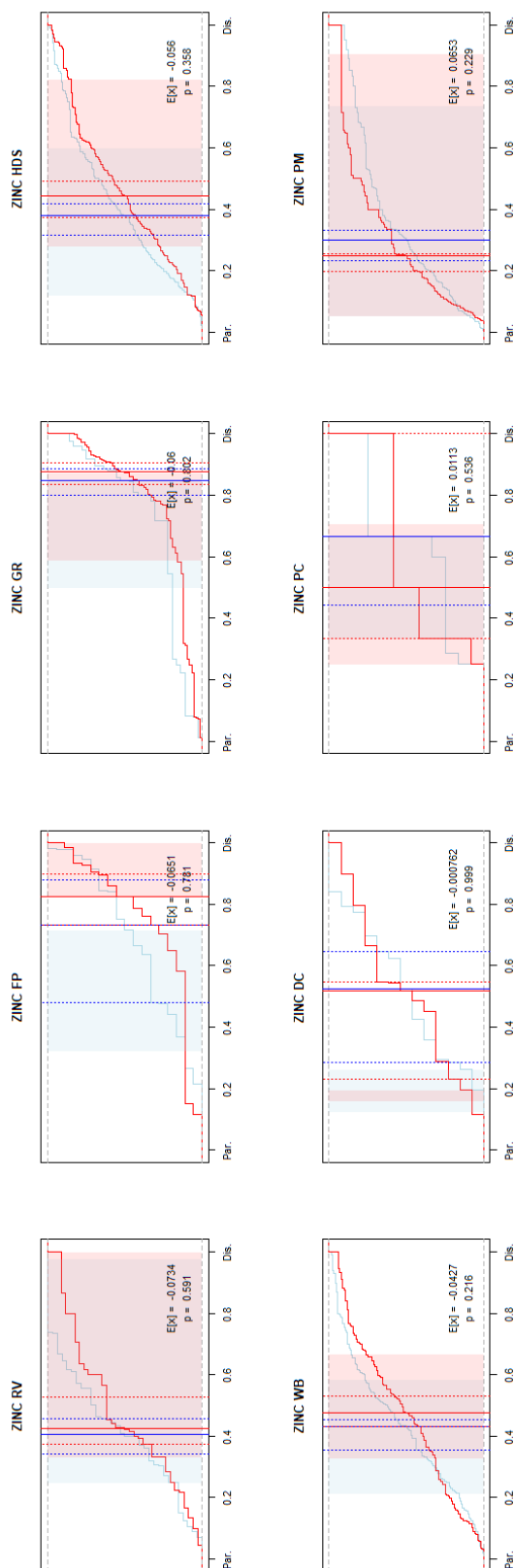


Figure F.7: Empirical cumulative distribution functions for the dissolved fractional ratio of particulate and dissolved zinc in untreated stormwater runoff (inflow, in blue) and treated effluent (outflow, in red) for eight BMP types. Left side represents 100% particulate fraction, right side represents 100% dissolved fraction as percent of total contaminant load. Median bootstrapped inflow and outflow ranges are marked with vertical lines, shaded areas represent the shortest half of the dataset.



G Causal effects in order of magnitude change

G.1 BMP type effects on hydrologic performance

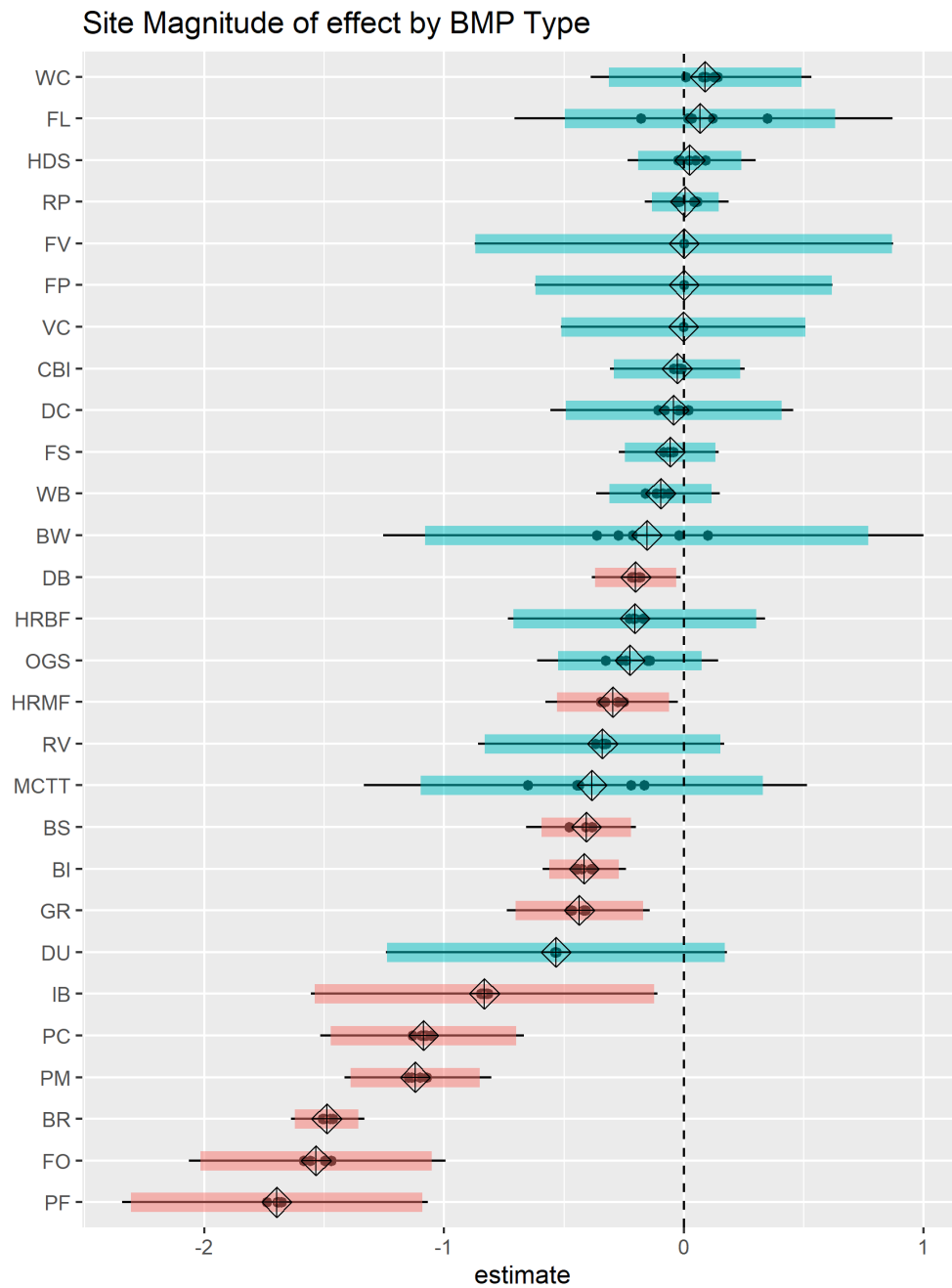


Figure G.1: Site effect size estimates for each BMP type (units of OM change). Imputed estimates (small black points) are pooled (black diamond) with color-coded pooled standard error range. Colors indicate BMP types that reduce effluent (red, to the left of the dotted line) or have no effect (blue, crossing dotted line).

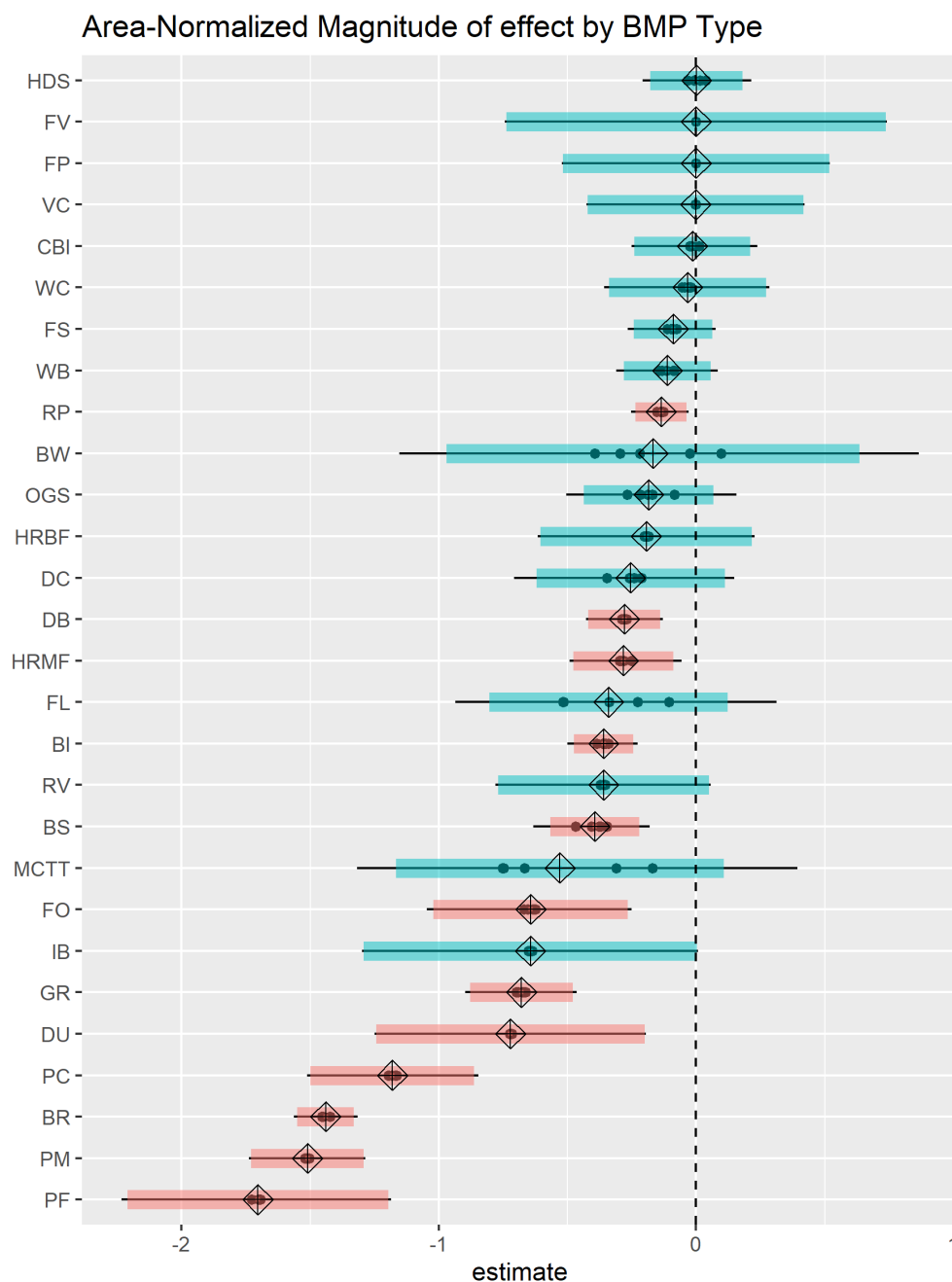


Figure G.2: Area-normalized site effect size estimates for each BMP type feature sets (units of OM change). Imputed estimates (small black points) are pooled (black diamond) with color-coded pooled standard error range. Colors indicate BMP types that reduce effluent (red, to the left of the dotted line) or have no effect (blue, crossing dotted line).

G.2 Individual effects by structural feature

	features	n sites	n obs	estimate	std.error	t statistic	df	p.value
reference condition	liner	150	3899	-0.08	0.20	-0.39	3993	0.70
	unamended	215	5086					
	no surface ponding	85	2270					
	surface drainage	225	5917					
	unplanted	115	2576					
	enclosed	319	8996					
	pointflow	249	6899					
	overflow path	292	7892					
	channel shape	125	2233					
treatment conditions	no liner	220	6112	-0.39	0.20	-1.98	6325	0.05 *
	media amended	153	4887	-0.32	0.14	-2.25	1617	0.02 *
	seive or screen	2	38	0.15	0.61	0.25	1560	0.80 --
	intermittent ponding	129	4514	0.04	0.14	0.29	3714	0.77 --
	permanent ponding	156	3227	0.62	0.16	3.98	5329	0.00 ***
	underdrain	91	2644	0.13	0.14	0.94	5722	0.35
	elbow or IWZ	37	1128	-0.44	0.17	-2.53	4236	0.01 *
	internal baffles	17	322	-0.27	0.25	-1.09	3051	0.28
	grass or sedum	236	5963	-0.07	0.14	-0.52	2268	0.61
	shrubs and trees	19	1472	-0.84	0.21	-3.90	9737	0.00 ***
	open to air	51	1015	-0.30	0.19	-1.53	3768	0.12
	sheetflow	121	3112	-0.15	0.12	-1.25	8571	0.21
	bypass routing	78	2119	0.25	0.11	2.24	2258	0.03 *
	basin shape	245	7778	-0.27	0.11	-2.57	2663	0.01 *

Table G.1: Pooled Model Set 2 effect estimates and std errors for each structural feature, units are in orders of magnitude. T-statistic and degrees of freedom for pooled calculation are also reported. Estimates that are statistically significant are marked with ‘***’ ($p < 0.01$) or ‘*’ ($p < 0.05$) depending on their level of significance, poor estimates are marked with ‘—’ in red based on unusually large Z-statistics ($|x| > 5$), which may indicate that the parameters are near the edge of the range and the random effects are near zero for the sites represented.

G.3 Covariate balance after matching

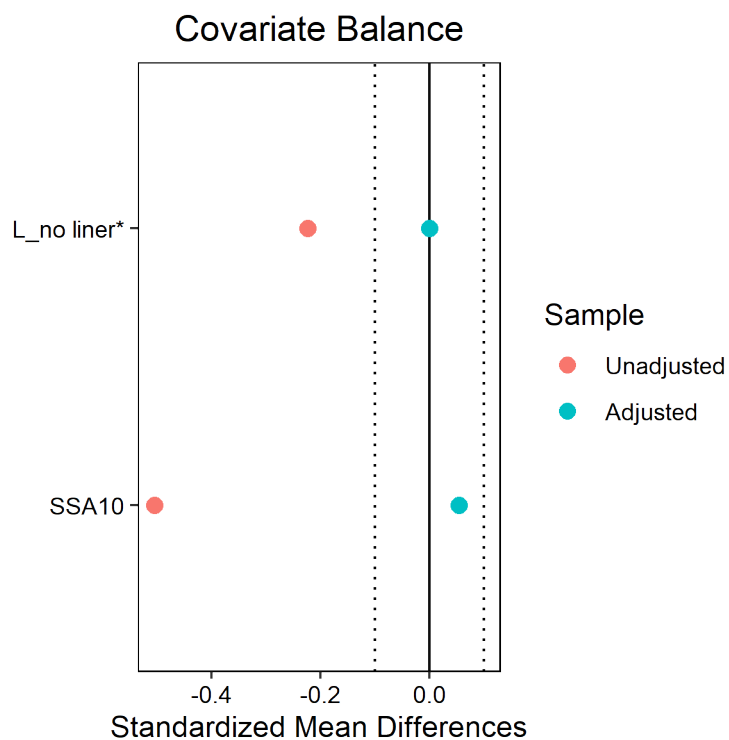


Figure G.3: Love plot indicating standardized mean difference for liner (L, no liner) and site surface area (SSA 10) covariates before (unadjusted in red) and after (adjusted in blue) matching. An absolute value of standardized mean difference below 0.1 is considered a relatively well-matched feature (blue dots should be between black dotted lines).

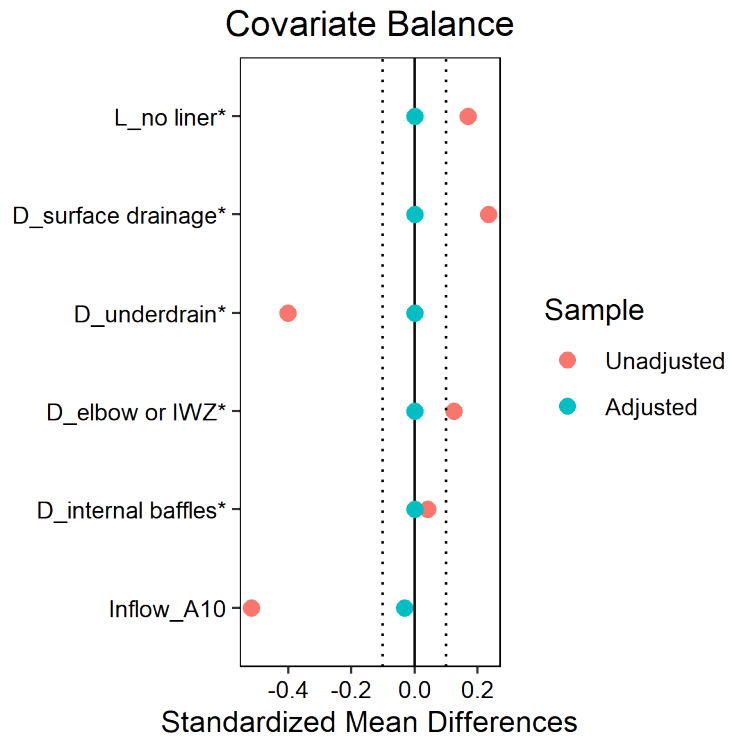


Figure G.4: Love plot indicating standardized mean difference for liner and site surface area covariates before (unadjusted) and after (adjusted) matching.

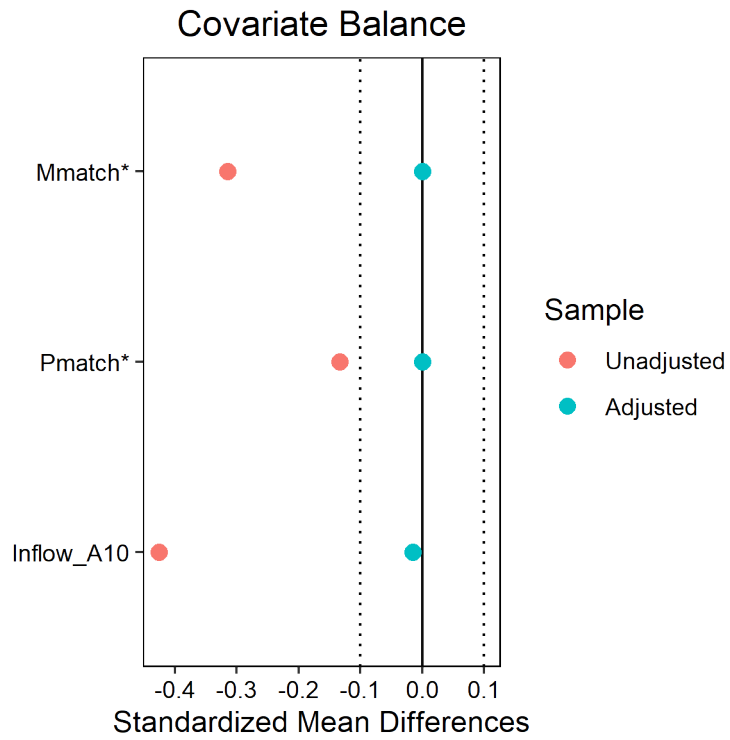
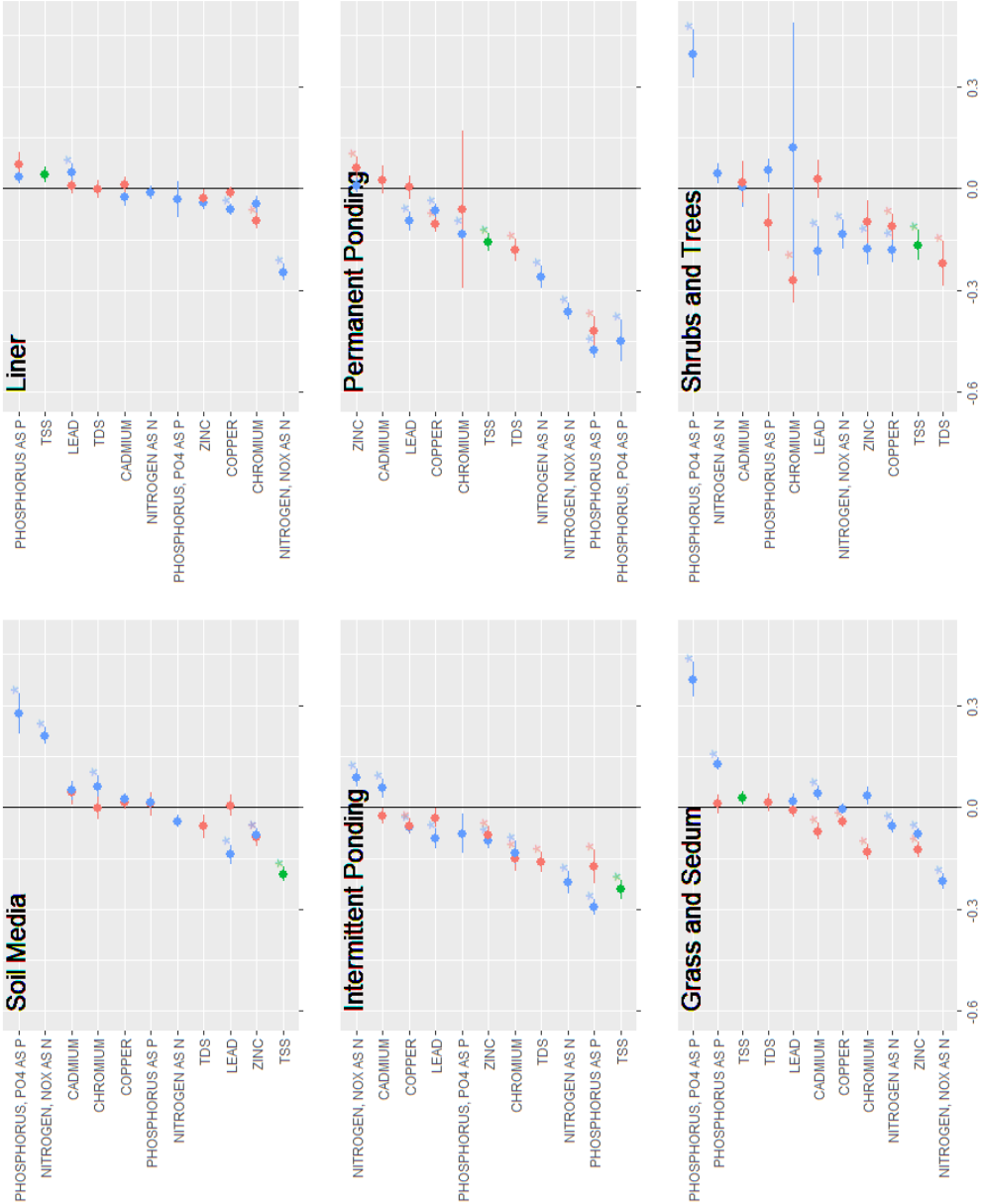


Figure G.5: Love plot indicating standardized mean difference for media, ponding, and site surface area covariates before (unadjusted) and after (adjusted) matching.

G.4 Structural treatment effects on contaminant concentrations

Figure G.6: Comparison of individual structural treatment effects on solids, nutrients and trace metals, units are OM change, pink indicates dissolved fraction, green represents suspended fraction, blue represents total.



References

References

- 2nd Nature, LLC. Lake Tahoe BMP monitoring evaluation process: Synthesis of existing research (final report). Technical report, USFS Lake Tahoe Basin Management Unit, 2006.
- L. M. Ahiablame, B. A. Engel, and I. Chaubey. Effectiveness of low impact development practices: Literature review and suggestions for future research. *Water Air Soil Pollut.*, 223(7):4253–4273, Sept. 2012. doi: <https://doi.org/10.1007/s11270-012-1189-2>.
- R. Amick and E. Burgess. Exfiltration in sanitary sewer systems. Technical Report EPA/600/R-01/034 (NTIS PB2003-103053), U.S. Environmental Protection Agency, 2000. URL https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=99137&Lab=NRMRL.
- J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press, Jan. 2009. URL <https://press.princeton.edu/books/paperback/9780691120355/mostly-harmless-econometrics>.
- A. Arguez, I. Durre, S. Applequist, M. Squires, R. Vose, X. Yin, and R. Bilotta. U.S. Climate Normals Product Suite (1981-2010), 2010. URL <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00822>.
- A. Askarizadeh, M. A. Rippey, T. D. Fletcher, D. L. Feldman, J. Peng, P. Bowler, A. S. Mehring, B. K. Winfrey, J. A. Vrugt, A. AghaKouchak, S. C. Jiang, B. F. Sanders, L. A. Levin, S. Taylor, and S. B. Grant. From rain tanks to catchments: Use of low-impact development to address hydrologic symptoms of the urban stream syndrome. *Env. Sci. Tech.*, 49(19):11264–11280, Oct. 2015. doi: <http://dx.doi.org/10.1021/acs.est.5b01635>.
- K. R. Ayub, N. A. Zakaria, R. Abdullah, and R. Ramli. Water balance: Case study of a constructed wetland as part of the bio-ecological drainage system (BIOECODS). *Water Science and Technology*, 62(8):1931–1936, 2010. doi: <http://dx.doi.org/10.2166/wst.2010.473>.
- C. Balderas Guzman, R. Wang, O. Muellerklein, M. Smith, and C. G. Eger. Comparing stormwater quality and watershed typologies across the United States: A machine learning approach. *Water Res.*, 216:118283, June 2022. doi: <http://dx.doi.org/10.1016/j.watres.2022.118283>.
- T. Ballesterio, R. Roseen, and J. Houle. UNH Stormwater Center 2012 Biennial Report. Technical report, University of New Hampshire, Durham, NH, 2012. URL <https://scholars.unh.edu/stormwater/37>.
- D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.*, 68(3), Apr. 2013. doi: <http://dx.doi.org/10.1016/j.jml.2012.11.001>.
- M. Bateman, E. H. Livingston, and J. Cox. Overview of Urban Retrofit Opportunities in Florida. In *National Conference on Retrofit Opportunities for Water Resource Protection in Urban Environments: Proceedings, Chicago, IL, February 9-12, 1998*, page 166. US

- EPA National Risk Management Research Laboratory, 1999. URL <https://p2infohouse.org/ref/41/40288.pdf>.
- E. Z. Bean, W. F. Hunt, and D. A. Bidelsbach. Evaluation of four permeable pavement sites in eastern North Carolina for runoff reduction and water quality impacts. *J. Irrig. Drain. Eng.*, 133(6):583–592, 2007. URL [https://ascelibrary.org/doi/abs/10.1061/\(ASCE\)0733-9437\(2007\)133:6\(583\)](https://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9437(2007)133:6(583)).
- R. Berghage, A. Jarrett, D. Beattie, K. Kelley, S. Husain, F. Rezai, B. Long, A. Negassi, R. Cameron, and W. Hunt. Quantifying evaporation and transpirational water losses from green roofs and green roof media capacity for neutralizing acid rain. Technical report, Pennsylvania State University and North Carolina State University, 2007. URL <https://decentralizedwater.waterrf.org/documents/04-dec-10sg/04-dec-10sg.pdf>.
- R. D. Berghage, D. Beattie, A. R. Jarrett, C. Thuring, F. Razaei, and others. Green roofs for stormwater runoff control. Technical Report EPA/600/R-09/026, National Risk Management Research Laboratory, US EPA, and the Pennsylvania State University, Mar. 2009. URL https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NRMRL&dirEntryId=205444.
- A. S. Bhaskar and C. Welty. Water balances along an urban-to-rural gradient of Metropolitan Baltimore, 2001–2009. *Environ. Eng. Geosci.*, 18(1):37–50, Feb. 2012. doi: <https://doi.org/10.2113/gseegeosci.18.1.37>.
- W. F. Bleam. *Soil and Environmental Chemistry*. Academic Press, Nov. 2016.
- B. M. Bolker, M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.*, 24(3):127–135, Mar. 2009. doi: <http://dx.doi.org/10.1016/j.tree.2008.10.008>.
- P. M. Bradley, K. M. Romanok, K. L. Smalling, J. R. Masoner, D. W. Kolpin, and S. E. Gordon. Predicted aquatic exposure effects from a national urban stormwater study. *Environ. Sci.: Water Res. Technol.*, Mar. 2023. doi: <https://doi.org/10.1039/D2EW00933A>.
- K. Bratieres, T. D. Fletcher, A. Deletic, and Y. Zinger. Nutrient and sediment removal by stormwater biofilters: a large-scale design optimisation study. *Water Res.*, 42(14):3930–3940, Aug. 2008. doi: <http://doi.org/10.1016/j.watres.2008.06.009>.
- M. E. Brooks, K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Maechler, and B. M. Bolker. *glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling*, 2017.
- H. Brown. *Next generation infrastructure: Principles for post-industrial public works*. Island Press/Center for Resource Economics, 2014.
- R. A. Brown and M. Borst. Quantifying evaporation in a permeable pavement system. *Hydrol. Process.*, 29(9):2100–2111, Apr. 2015. doi: <https://onlinelibrary.wiley.com/doi/10.1002/hyp.10359>.

- R. A. Brown, W. F. Hunt, III, and W. R. Skaggs. Long-term modeling of bioretention hydrology with DRAINMOD. Technical report, Water Resources Research Institute of the University of North Carolina, 2011. URL <https://repository.lib.ncsu.edu/bits/tream/handle/1840.4/8155/NC-WRRI-415.pdf?sequence=1>.
- N. Buccola and G. Spolek. A pilot-scale evaluation of greenroof runoff retention, detention, and quality. *Water Air Soil Pollut.*, 216(1-4):83–92, Mar. 2011. doi: <https://doi.org/10.1007/s11270-010-0516-8>.
- E. Buckingham. Studies on the movement of soil moisture. Technical report, US Bureau of Soils Government Printing Office. AGRIS library website, 1907. URL <https://agris.fao.org/agris-search/search.do?recordID=US201300008520>.
- C. A. Burton, T. M. Hoefen, G. S. Plumlee, K. L. Baumberger, A. R. Backlin, E. Gallegos, and R. N. Fisher. Trace elements in stormflow, ash, and burned soil following the 2009 Station Fire in southern California. *PLoS One*, 11(5):e0153372, May 2016. doi: <http://dx.doi.org/10.1371/journal.pone.0153372>.
- M. Butler. Ambitious new plan aims to reinvent homelessness response in Ithaca. Ithaca Voice website, Apr. 2023. URL <https://ithacavoice.org/2023/04/ambitious-new-plan-aims-to-reinvent-homelessness-response-in-ithaca/>. Accessed: 2023-4-19.
- BWE, Inc. Cabrillo Heights neighborhood park. BWE Structural and Civil Engineering Firm website, Apr. 2021. URL <https://bwesd.com/project/cabrillo-heights-neighborhood-park/>. Accessed: 2023-4-15.
- P. V. Caldwell, M. J. Vepraskas, R. W. Skaggs, and J. D. Gregory. Simulating the water budgets of natural Carolina bay wetlands. *Wetlands*, 27(4):1112–1123, Dec. 2007. doi: [https://doi.org/10.1672/0277-5212\(2007\)27\[1112:STWBON\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2007)27[1112:STWBON]2.0.CO;2).
- CALTRANS Division of Environmental Analysis. Roadside Vegetated Treatment Sites (RTVS) Study Final Report. Technical Report CTSW-RT-03-028, California Department of Transportation, Nov. 2003. URL <https://dot.ca.gov/-/media/dot-media/programs/environmental-analysis/documents/env/ctsw-rt-03-028-a11y.pdf>.
- CALTRANS Division of Environmental Analysis. BMP Retrofit Pilot Program Final Report. Technical Report CTSW-RT-01-050, California Department of Transportation, Jan. 2004. URL <https://dot.ca.gov/-/media/dot-media/programs/environmental-analysis/documents/env/ctsw-rt-01-050-a11y.pdf>.
- D. D. Carpenter and P. Kaluvakolanu. Effect of roof surface type on storm-water runoff from full-scale roofs in a temperate climate. *J. Irrig. Drain. Eng.*, 137(3):161–169, Mar. 2011. doi: [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000185](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000185).
- T. B. Carson, D. E. Marasco, P. J. Culligan, and W. R. McGillis. Hydrological performance of extensive green roofs in New York City: observations and multi-year modeling of three full-scale systems. *Environ. Res. Lett.*, 8(2):024036, June 2013. doi: <https://doi.org/10.1088/1748-9326/8/2/024036>.
- T. L. Carter and T. C. Rasmussen. Use of green roofs for ultra-urban stream restoration in

- the Georgia Piedmont (USA). In *2005 Georgia Water Resources Conference Proceedings*. Georgia Institute of Technology, 2005. URL <https://smartech.gatech.edu/handle/1853/47483>.
- T. L. Carter and T. C. Rasmussen. Hydrologic behavior of vegetated roofs. *J. Am. Water Resour. Assoc.*, 42(5):1261–1274, June 2007. doi: <https://doi.org/10.1111/j.1752-1688.2006.tb05299.x>.
- F. K. S. Chan, J. A. Griffiths, D. Higgitt, S. Xu, F. Zhu, Y.-T. Tang, Y. Xu, and C. R. Thorne. “Sponge City” in China—A breakthrough of planning and flood risk management in the urban context. *Land Use Policy*, 76:772–778, 2018. doi: 10.1016/j.landusepol.2018.03.005.
- J. Choi and J. W. Harvey. Quantifying time-varying ground-water discharge and recharge in wetlands of the northern Florida everglades. *Wetlands*, 20(3):500–511, Sept. 2000. doi: [https://doi.org/10.1672/0277-5212\(2000\)020<0500:QTGDAR>2.0.CO;2](https://doi.org/10.1672/0277-5212(2000)020<0500:QTGDAR>2.0.CO;2).
- City of San Diego. 43rd Street and Logan Avenue Project. <https://www.sandiego.gov/stormwater/pilot-projects/43rdlogan>, 2023. URL <https://www.sandiego.gov/stormwater/pilot-projects/43rdlogan>.
- City of Tacoma, Washington. Section S8F - Stormwater Treatment Best Management Practice Evaluation, Phase I Municipal Stormwater NPDES Permit, Quality Assurance Project Plan, Appendix A. Technical Report Permit No.: WAR04-4003, Section S8F - Stormwater Treatment Best Management Practice Evaluation, Oct. 2012. URL <https://cms.cityoftacoma.org/enviro/surfacewater/NPDESAnnual/AppA.pdf>.
- City of Tacoma, Washington. Section S8F - Stormwater Treatment Best Management Practice Evaluation, Final Report. Technical report, Section S8F - Stormwater Treatment Best Management Practice Evaluation, Mar. 2015. URL <https://cms.cityoftacoma.org/enviro/surfacewater/NPDESAnnual/S8FReport.pdf>.
- J. Clary, R. Pitt, and B. Steets. Pathogens in urban stormwater systems. *UWRRRC Technical Committee Report, Environmental and Water Resources Institute, American Society of Civil Engineers*, 2014. URL https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/7330444.
- J. Clary, R. Pitt, B. Steets, and C. Olson. Pathogens in Urban Stormwater Systems: Where Are We Now? *Journal of Sustainable Water in the Built Environment*, 8(1):02521004, Feb. 2022. doi: <https://doi.org/10.1061/JSWBAY.0000969>.
- CNT. The value of green infrastructure: A guide to recognizing its economic, environmental and social benefits. Technical report, Center for Neighborhood Technology and American Rivers, 2010.
- Contech ES. NJCAT technology verification: High efficiency continuous deflective separator (CDS). Technical report, New Jersey Corporation for Advance Technology program, Aug. 2012. URL <http://www.njcat.org/uploads/newDocs/NJCATECHNOLOGYVERIFICATIONMSBCDSFINAL81012.pdf>.

- Contech ES. Inline CDS Standard Detail, May 2014. URL <https://www.conteches.com/Portals/0/Documents/Standard%20Details/CDS5640-10-C-DTL.pdf?ver=2018-08-03-112220-720%27>.
- Contech ES. Stormceptor General Specifications (Various sizes). In *Contech Engineering Solutions Technical Guides*, Online Technical Guides. Contech Engineering Solutions website, 2023a. URL <https://www.conteches.com/technical-guides/search?filter=13I9882Q80>.
- Contech ES. Vortechs Specification and standard details. In *Contech Engineering Solutions Technical Guides*, Online Technical Guides. Contech Engineering Solutions website, 2023b. URL <https://www.conteches.com/technical-guides/search?filter=0JFIA3M8SD>.
- S. R. Corsi, S. R. Greb, R. T. Bannerman, and R. E. Pitt. Evaluation of the multi-chambered treatment train, a retrofit water-quality management device. Technical report, US Geological Survey, 1999. URL <https://pubs.usgs.gov/of/1999/0270/report.pdf>.
- I. T. Cousins, J. H. Johansson, M. E. Salter, B. Sha, and M. Scheringer. Outside the Safe Operating Space of a New Planetary Boundary for Per- and Polyfluoroalkyl Substances (PFAS). *Environ. Sci. Technol.*, 56(16):11172–11179, Aug. 2022. doi: <http://dx.doi.org/10.1021/acs.est.2c02765>.
- CPWJ. Infiltration BMP: Permeable pavement studies allow urban drainage to update its design criteria manual. *Colorado Public Works Journal*, 8(5):21–33, 2012. URL https://www.cityoffortmorgan.com/DocumentCenter/View/3073/2012CPW_Pavers.
- D. T. Crisp. Input and output of minerals for an area of Pennine Moorland: The importance of precipitation, drainage, peat erosion and animals. *J. Appl. Ecol.*, 3(2):327–348, 1966. doi: <https://doi.org/10.2307/2401256>.
- J. Czemieli Berndtsson. Green roof performance towards management of runoff water quantity and quality: A review. *Ecol. Eng.*, 36(4):351–360, Apr. 2010. doi: <https://doi.org/10.1016/j.ecoleng.2009.12.014>.
- L. K. Dally, D. P. Lettenmaier, S. J. Burges, and M. M. Benjamin. Operation of Detention Facilities for Urban Stream Quality Enhancement. Technical Report Water Resources Series Technical Report No. 79, University of Washington Department of Civil and Environmental Engineering, 1983.
- W. L. Daniels, A. R. Cummings, M. Schmidt, N. Fomchenko, G. K. Speiran, M. Focazio, and G. M. Fitch. Evaluation of methods to calculate a wetlands water balance. Technical report, Virginia Center for Transportation Innovation and Research, Aug. 2000. URL <https://vtechworks.lib.vt.edu/handle/10919/46718>.
- A. P. Davis, R. G. Traver, W. F. Hunt, R. Lee, R. A. Brown, and J. M. Olszewski. Hydrologic Performance of Bioretention Storm-Water Control Measures. *J. Hydrol. Eng.*, 17(5):604–614, May 2012. URL [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000467](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000467).
- M. R. C. De Sousa, F. A. Montalto, and S. Spatarì. Using life cycle assessment to evaluate

- green and grey combined sewer overflow control strategies. *J. Ind. Ecol.*, 16(6):901–913, Dec. 2012. doi: <https://doi.org/10.1111/j.1530-9290.2012.00534.x>.
- A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM - algorithm plus discussions on the paper. *J. Royal Stat. Soc. B*, 1977. doi: <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x>.
- M. Denchak. Green infrastructure: How to manage water in a sustainable way. Natural Resources Defense Council website, 2022. URL <https://www.nrdc.org/stories/green-infrastructure-how-manage-water-sustainable-way>. Accessed: 2023-4-5.
- C. Denich and A. Bradford. Estimation of evapotranspiration from bioretention areas using weighing lysimeters. *J. Hydrol. Eng.*, 15(6):522–530, June 2010. doi: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000134](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000134).
- J. Drake, A. Bradford, and T. Van Seters. Hydrologic performance of three partial-infiltration permeable pavements in a cold climate over low permeability soil. *J. Hydrol. Eng.*, 19(9): 04014016, Sept. 2014. doi: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000943](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000943).
- J. I. Drever. *The Geochemistry of Natural Waters: Surface and Groundwater Environments*. Prentice Hall, 1997. URL https://play.google.com/store/books/details?id=mbYP_AQAAIAAJ.
- C. T. Driscoll, C. G. Eger, D. G. Chandler, C. I. Davidson, B. K. Roodsari, C. D. Flynn, K. F. Lambert, N. D. Bettez, and P. M. Groffman. Green infrastructure: Lessons from science and practice. Technical report, Science Policy Exchange, 2015.
- F. L. Duley and C. E. Domingo. Effect of grass on intake of water. *Historical Research Bulletins of the Nebraska Agricultural Experiment Station (1913-1993)*, 159, Apr. 1949. URL <https://digitalcommons.unl.edu/ardhistrb/59/>.
- F. L. Duley and L. L. Kelly. Effect of soil type, slope, and surface conditions on intake of water. *Historical Research Bulletins of the Nebraska Agricultural Experiment Station (1913-1993)*, 112, May 1939. URL <https://digitalcommons.unl.edu/ardhistrb/66/>.
- C. G. Eger. *Nutrient retention in roadside retrofit rain gardens*. MS Thesis, The Ohio State University, 2012. URL https://rave.ohiolink.edu/etdc/view?acc_num=osu1343837490.
- C. G. Eger, D. G. Chandler, B. Kasaei Roodsari, C. I. Davidson, and C. T. Driscoll. Water budget triangle: A new conceptual framework for comparison of green and gray infrastructure. In *ASCE International Conference on Sustainable Infrastructure 2014*, pages 1010–1017. ICSI (ASCE), 2014. doi: <https://doi.org/10.1061/9780784478745.095>.
- C. G. Eger, D. G. Chandler, and C. T. Driscoll. Hydrologic processes that govern stormwater infrastructure behaviour. *Hydrol. Process.*, 31(25):4492–4506, Dec. 2017. doi: <https://doi.org/10.1002/hyp.11353>.
- J. B. Ellis, D. M. Revitt, P. Lister, C. Willgress, and A. Buckley. Experimental studies of sewer exfiltration. *Water Sci. Technol.*, 47(4):61–67, 2003. URL <https://www.ncbi.nlm.nih.gov/pubmed/12666802>.

- C. H. Emerson. *Evaluation of the additive effects of stormwater detention basins at the watershed scale*. MS Thesis, Drexel University, 2003.
- Ergas Sarina J. and Fassman-Beck Elizabeth. Pathogens and Fecal Indicators in Stormwater. *Journal of Sustainable Water in the Built Environment*, 9(3):01023001, Aug. 2023. doi: <https://doi.org/10.1061/JSWBAY.SWENG-514>.
- ETV, US EPA, NSFI. ETV Report: Stormwater source area treatment device: The stormwater management StormFilter using ZPG filter media. Technical Report 04/17/WQPC-WWF, The Environmental Technology Verification program, July 2004. URL <https://nemallc.com/Resources/Documents/BMP%20Performance/stormfilter%20bmp%20review%20epa%202004.pdf>.
- ETV, US EPA, NSFI. ETV Joint Verification Statement: Terre Kleen 09. Technical Report 06/29/WQPC-WWF, The Environmental Technology Verification program, July 2008. URL <https://archive.epa.gov/nrmrl/archive-etv/web/pdf/600etv08028s.pdf>.
- E. Fassman-Beck, E. Voyde, R. Simcock, and Y. S. Hong. 4 Living roofs in 3 locations: Does configuration affect runoff mitigation? *J. Hydrol.*, 490:11–20, May 2013. doi: <https://doi.org/10.1016/j.jhydrol.2013.03.004>.
- E. Fassman-Beck, W. Hunt, R. Berghage, D. Carpenter, T. Kurtz, V. Stovin, and B. Wadzuk. Curve number and runoff coefficients for extensive living roofs. *J. Hydrol. Eng.*, 21(3): 04015073, Mar. 2016. doi: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001318](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001318).
- R. Field, A. Tafuri, J. Y. Lin, and S. L. Yu. Multi-Chambered Treatment Train (MCTT) For Treating Stormwater Runoff From Highly Polluted Source Areas. URL https://www.researchgate.net/publication/266042238_Multi-Chambered_Treatment_Train_MCTT_For_Treating_Stormwater_Runoff_From_Highly_Polluted_Source_Areas.
- R. Fioretti, A. Palla, L. G. Lanza, and P. Principi. Green roof energy and water related performance in the Mediterranean climate. *Build. Environ.*, 45(8):1890–1904, Aug. 2010. doi: <https://doi.org/10.1016/j.buildenv.2010.03.001>.
- T. D. Fletcher, W. Shuster, W. F. Hunt, R. Ashley, D. Butler, S. Arthur, S. Trowsdale, S. Barraud, A. Semadeni-Davies, J.-L. Bertrand-Krajewski, P. S. Mikkelsen, G. Rivard, M. Uhl, D. Dagenais, and M. Viklander. SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage. *Urban Water J.*, 12(7):525–542, Oct. 2015. doi: <https://doi.org/10.1080/1573062X.2014.916314>.
- R. Fox. 2022 EPA nutrient reduction memorandum, Mar. 2022. URL <https://www.epa.gov/nutrient-policy-data/2022-epa-nutrient-reduction-memorandum>.
- Geosyntec Consultants and Wright Water Engineers. International stormwater best management practices (BMP) database technical summary: Volume reduction. Technical report, Water Environment Research Foundation, Federal Highway Administration, Environment and Water Resources Institute (ASCE), Jan. 2011.
- K. L. Getter, D. B. Rowe, and J. A. Andresen. Quantifying the effect of slope on extensive

- green roof stormwater retention. *Ecol. Eng.*, 31(4):225–231, Dec. 2007. doi: <https://doi.org/10.1016/j.ecoleng.2007.06.004>.
- C. C. Glass and ETEC. Green roof water quality and quantity monitoring. Technical report, American Society of Landscape Architects, 2007. URL <https://www.ce.washington.edu/sites/cee/files/pdfs/research/hydrology/water-resources/WRS079.pdf>.
- P. Göbel, P. Starke, A. Voss, and W. Coldewey. Field measurements of evapotranspiration rates on seven pervious concrete pavement systems. In *Novatech 2013-8th International Conference on planning and technologies for sustainable management of water in the city*. hal.science, 2013. URL <https://hal.science/hal-03303532>.
- Google. Google Earth locations of BMPs as recorded in the International Stormwater BMP Database, 2023a. URL earth.google.com.
- Google. Google Maps locations of BMPs as recorded in the International Stormwater BMP Database, 2023b. URL maps.google.com.
- A. Graceson, M. Hare, J. Monaghan, and N. Hall. The water retention capabilities of growing media for green roofs. *Ecol. Eng.*, 61:328–334, Dec. 2013. doi: <https://doi.org/10.1016/j.ecoleng.2013.09.030>.
- B. G. Gregoire and J. C. Clausen. Effect of a modular extensive green roof on stormwater runoff and water quality. *Ecol. Eng.*, 37(6):963–969, June 2011. doi: <https://doi.org/10.1016/j.ecoleng.2011.02.004>.
- M. Guizani. Stormwater harvesting in Saudi Arabia: a multipurpose water management alternative. *Water Resour. Manage.*, 30(5):1819–1833, Mar. 2016. doi: <https://doi.org/10.1007/s11269-016-1255-4>.
- N. E. Hamilton and M. Ferry. ggtern: Ternary diagrams using ggplot2. *Journal of Statistical Software, Code Snippets*, 87(3):1–17, 2018. doi: [10.18637/jss.v087.c03](https://doi.org/10.18637/jss.v087.c03).
- H. H. Harper. Navy Canal stormwater facility performance efficiency evaluation. Technical report, Environmental Research and Design, Inc, Feb. 2010a. URL <http://www.erd.org/ERD%20Publications/NAVY%20CANAL%20STORMWATER%20FACILITY%20PERF%20EFFICIENCY%20EVAL-FEB%202010.pdf>.
- H. H. Harper. Elder Creek stormwater facility performance efficiency evaluation. Technical report, Environmental Research and Design, Inc, 2010b. URL <http://www.erd.org/ERD%20Publications/ELDER%20CREEK-FINAL%20REPORT-SEPT%202010.pdf>.
- H. H. Harper. Evaluation of the performance efficiency of the Poppleton Creek wet detention pond. Technical report, Environmental Research and Design, Inc, Mar. 2010c. URL <http://www.erd.org/ERD%20Publications/EVAL%20OF%20PERF%20EFFICIENCY%20OF%20POPPLETON%20CRK%20WET%20DETENTION%20POND-MARCH%202010.pdf>.
- H. H. Harper. Performance efficiency evaluation of the Club II Regional Stormwater Facility (RSF). Technical report, Environmental Research and Design, Inc, May 2011. URL <http://www.erd.org/ERD%20Publications/CLUB%20II-FINAL%20REPORT-REV%20MAY%202011-1.pdf>.

- H. H. Harper, D. M. Baker, and C. Harper. Performance evaluation of dry detention stormwater management systems. In *Proceedings of the 6th Biennial Stormwater Research and Watershed Management Conference*, pages 162–178. Southwest Florida Water Management District, 1999a.
- H. H. Harper, J. L. Herr, D. Baker, and E. H. Livingston. Performance evaluation of dry detention stormwater management systems. In *6th Biennial Stormwater Research and Watershed Management Conference*, page 162. Penn State University Citeseer website, 1999b. URL <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b51b097b0d198c978b25357b66c32cf3c226a654#page=170>.
- H. H. Harper, D. M. Baker, and C. Harper. Effectiveness of stormwater treatment systems in the Florida Keys. Technical report, Environmental Research and Design, Inc, Dec. 2002. URL <http://www.erd.org/ERD%20Publications/EFFECTIVENESS%20OF%20SW%20TREATMENT%20SYSTEMS%20IN%20THE%20FLORIDA%20KEYS-2002.pdf>.
- H. H. Harper, D. M. Baker, and C. Harper. Performance evaluation of the Palm Bay Basin 7 wet detention pond. Technical Report Final Report, Environmental Research and Design, Inc, May 2003. URL <http://www.erd.org/ERD%20Publications/PERF%20EVAL%20OF%20THE%20PALM%20BAY%20BASIN%207%20WET%20DETENTION%20POND-MAY%2003.pdf>.
- H. H. Harper, R. R. Villapando, and Y. Zhao. Phosphorus Removal in Lake Okeechobee tributaries using cds and baffle box units. In *Proceedings of the 40th Annual Water Resources Conference*, 2004. URL <http://www.erd.org/ERD%20Publications/OKEE-PAPE R.AWRA%20JOURNAL%202004.pdf>.
- P. Hartigan, M. Kelly, and R. Hart. Daily wet pond balance spreadsheet (Excel spreadsheet calculator). City of Austin, TX Watershed website, <https://www.austintexas.gov/department/stormwater-management>, Mar. 2009. URL https://www.austintexas.gov/sites/default/files/files/Watershed/swtreat_wetpond_balance.xls. Accessed: 2023-4-13.
- A. M. Hathaway, W. F. Hunt, and G. D. Jennings. A field study of green roof hydrologic and water quality performance. *Transactions of the ASABE*, 51(1):37–44, 2008. URL <https://elibrary.asabe.org/abstract.asp?aid=24225>.
- D. K. Hein, E. Strecker, A. Poresky, R. Roseen, and M. Venner. Permeable shoulders with stone reservoirs. Technical Report NCHRP 25-25 Task 82, Transportation Research Board, Oct. 2013. URL [https://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP25-25\(82\)_FR.pdf](https://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP25-25(82)_FR.pdf).
- A. Heiss. Meld regression output from multiple imputations with tidyverse. Andrew Heiss’ professional website, Mar. 2018. URL <https://www.andrewheiss.com/blog/2018/03/07/amelia-tidy-melding/>. Accessed: 2023-4-16.
- H. F. Hemond. Biogeochemistry of Thoreau’s bog, Concord, Massachusetts. *Ecol. Monogr.*, 50(4):507–526, Feb. 1980. doi: <https://onlinelibrary.wiley.com/doi/10.2307/1942655>.
- M. A. Hernán and S. L. Taubman. Does obesity shorten life? The importance of well-defined

- interventions to answer causal questions. *Int. J. Obes.*, 32 Suppl 3:S8–14, Aug. 2008. doi: <http://dx.doi.org/10.1038/ijo.2008.82>.
- A. Hess, B. Wadzuk, and A. Welker. Evapotranspiration and infiltration in rain garden systems. In *World Environmental and Water Resources Congress 2015*, pages 261–270. ASCE library website, 2015. doi: <https://doi.org/10.1061/9780784479162.025>.
- A. Hess, B. Wadzuk, and A. Welker. Evapotranspiration in rain gardens using weighing lysimeters. *J. Irrig. Drain. Eng.*, 143(6):04017004, June 2017. doi: [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001157](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001157).
- A. J. Hess. *Monitoring of evapotranspiration and infiltration in rain garden designs*. PhD thesis, Villanova University, 2014. URL https://www1.villanova.edu/content/dam/villanova/engineering/vcase/vusp/AHess_Thesis.pdf.
- D. L. Hey, K. R. Barrett, and C. Biegen. The hydrology of four experimental constructed marshes. *Ecol. Eng.*, 3(4):319–343, Dec. 1994. doi: [https://doi.org/10.1016/0925-8574\(94\)00006-9](https://doi.org/10.1016/0925-8574(94)00006-9).
- M. W. Heymans and I. Eekhout. Chapter 9: Rubin’s Rules. In *Applied missing data analysis with SPSS and (R)Studio*. bookdown.org, Jan. 2019. URL <https://bookdown.org/mwheymans/bookmi/rubins-rules.html>.
- J. J. Hickman, Wadzuk Bridget, and Traver Robert. Evaluating the Role of Evapotranspiration in the Hydrology of a Bioinfiltration Basin Using a Weighing Lysimeter. *World Environmental and Water Resources Congress 2011*, pages 3601–3609, 2011. doi: [https://doi.org/10.1061/41173\(414\)377](https://doi.org/10.1061/41173(414)377).
- J. M. Hickman. *Evaluating the role of evapotranspiration in the hydrology of bioinfiltration and bioretention basins using weighing lysimeters*. PhD thesis, Villanova University, 2011. URL <https://www1.villanova.edu/content/dam/villanova/engineering/vcase/vusp/Hickmen-Thesis%202011.pdf>.
- H. V. Hilprecht. *In the temple of Bêl at Nippur: A lecture delivered before German court and university circles*. University of Pennsylvania, 1904. URL <https://diglib.uibk.ac.at/download/pdf/1439971?name=In%20the%20Temple%20of%20B%C3%AA1%20at%20Nippur>.
- D. Hirschman, K. Collins, and T. Schueler. The Runoff Reduction Method. *Center for Watershed Protection & Chesapeake Stormwater Network*, page 1, 2008. URL <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=514081825b698154e6013559f7e0939df4cd7749>.
- D. E. Ho, K. Imai, G. King, and E. A. Stuart. MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8):1–28, 2011. doi: [10.18637/jss.v042.i08](https://doi.org/10.18637/jss.v042.i08).
- L. Hoffman, G. Loosevelt, and R. Berghage. Green roof thermal and stormwater management performance: The Gratz building case study, New York City. Technical report, NYSERDA, New York City, Pratt Center, 2010.

- J. R. Hofmann and P. A. Hofmann. Darcy's Law and Structural Explanation in Hydrology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1992: 23–35, 1992. URL <http://www.jstor.org/stable/192741>.
- G. G. Hollands. Regional analysis of the creation and restoration of kettle and pothole wetlands. *Wetland Creation and Restoration: The Status of the Science*. EPA/600/3-89/038, pages 287–304, 1989. URL https://books.google.com/books?hl=en&lr=&id=gsULSTfeMScC&oi=fnd&pg=PA287&dq=hollands+1989+regional+analysis+of+the+creation+and+restoration+of+kettle+and+pothole+wetlands&ots=lTlra9cE1C&sig=m8dm_rLEt2cm8QMeh0xx0sCdriE.
- J. Honaker and G. King. What to do about missing values in time-series cross-section data. *Am. J. Pol. Sci.*, 54(2):561–581, Apr. 2010. doi: <https://doi.org/10.1111/j.1540-5907.2010.00447.x>.
- J. Honaker, G. King, M. Blackwell, and M. M. Blackwell. Package ‘Amelia’. *Version*. *View Article*, 2010. URL <https://cran.microsoft.com/snapshot/2016-08-05/web/packages/Amelia/Amelia.pdf>.
- J. Honaker, G. King, and M. Blackwell. Amelia II: A Program for Missing Data. *J. Stat. Softw.*, 45:1–47, Dec. 2011. URL <https://www.jstatsoft.org/index.php/jss/article/view/v045i07/0>.
- J. Honaker, G. King, and M. Blackwell. Introduction to Multiple Imputation. *Amelia package vignette*, Nov. 2022. URL <https://cran.r-project.org/web/packages/Amelia/vignettes/intro-mi.html>.
- J. A. Horwath and R. T. Bannerman. Use of a Stormwater Filtration Device for Reducing Contaminants in Runoff from a Parking Lot in Madison, Wisconsin, 2005-07. Technical Report 0092-05-17, US Geological Survey, Wisconsin Department of Transportation, Sept. 2008. URL <https://wisconsin.gov/documents2/research/05-17stormwaterfiltration-f1.pdf>.
- J. A. Horwath and R. T. Bannerman. Parking Lot Runoff Quality and Treatment Efficiency of a Stormwater-filtration Device, Madison, Wisconsin, 2005-07. Technical Report 2009–5196, US Department of the Interior, US Geological Survey, 2010. URL https://pubs.usgs.gov/sir/2009/5196/pdf/sir2009-5196_web.pdf.
- M. Hu and T. Shealy. Overcoming Status Quo Bias for Resilient Stormwater Infrastructure: Empirical Evidence in Neurocognition and Decision-Making. *J. Manage. Eng.*, 36(4): 04020017, July 2020. doi: [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000771](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000771).
- C. E. Hughes, P. Binning, and G. R. Willgoose. Characterisation of the hydrology of an estuarine wetland. *J. Hydrol.*, 211(1):34–49, Nov. 1998. doi: [https://doi.org/10.1016/S0022-1694\(98\)00194-2](https://doi.org/10.1016/S0022-1694(98)00194-2).
- C. F. Hussain, J. Brand, J. S. Gulliver, and P. Weiss. Water quality performance of dry detention ponds with under-drains. Technical Report MN/RC-2006-43, University of Minnesota, Minnesota Department of Transportation, 2005. URL <https://www.lrrb.org/pdf/200643.pdf>.

- D. Hutchinson, P. Abrams, R. Retzlaff, and T. Liptan. Stormwater monitoring two ecoroofs in Portland, Oregon, USA. Penn State University CiteSeer website, 2003. URL <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5e3b8d9f6ac34230bb9e4cf860ab40630106228c>. Accessed: 2023-4-13.
- HydroLogic Solutions. StormChamber design manual. Technical report, NDS Stormchambers.
- M. Jajarmizadeh, S. Harun, and M. Salarpour. A review on theoretical consideration and types of models in hydrology. *Journal of Environmental Science and Technology*, 5(5): 249–261, 2012. doi: <http://dx.doi.org/10.3923/jest.2012.249.261>.
- J. C. Jakobsen, C. Gluud, J. Wetterslev, and P. Winkel. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med. Res. Methodol.*, 17(1):162, Dec. 2017. doi: <http://dx.doi.org/10.1186/s12874-017-0442-1>.
- M. Jamshidian and M. Mata. Advances in Analysis of Mean and Covariance Structure when Data are Incomplete. In S.-Y. Lee, editor, *Handbook of Latent Variable and Related Models*, pages 21–44. Elsevier, Jan. 2007. doi: <https://doi.org/10.1016/B978-044452044-9/50005-7>.
- Jensen Precast. StormVault Standard Detail, 2020. URL <https://www.jensenprecast.com/water-resources/wp-content/uploads/2020/05/JPHV-5000.pdf>.
- J. A. Jones, I. F. Creed, K. L. Hatcher, R. J. Warren, M. B. Adams, M. H. Benson, E. Boose, W. A. Brown, J. L. Campbell, A. Covich, D. W. Clow, C. N. Dahm, K. Elder, C. R. Ford, N. B. Grimm, D. L. Henshaw, K. L. Larson, E. S. Miles, K. M. Miles, S. D. Sebestyen, A. T. Spargo, A. B. Stone, J. M. Vose, and M. W. Williams. Ecosystem Processes and Human Influences Regulate Streamflow Response to Climate Change at Long-Term Ecological Research Sites. *Bioscience*, 62(4):390–404, Apr. 2012. URL <https://academic.oup.com/bioscience/article-lookup/doi/10.1525/bio.2012.62.4.10>.
- S. E. Jørgensen and W. J. Mitsch. *Ecological engineering: An introduction to ecotechnology*. John Wiley & Sons, 1989.
- S. S. Kaushal and K. T. Belt. The urban watershed continuum: Evolving spatial and temporal dimensions. *Urban Ecosyst.*, 15(2):409–435, June 2012. doi: <https://doi.org/10.1007/s11252-012-0226-7>.
- KCI Technologies. Pollutant removal efficienct of self-converted dry detention ponds. Technical Report Pioneer Grant #11460, Baltimore County, MD Watershed Management and Monitoring Section, Department of Environmental Protection and Sustainability, 2015. URL https://www.chesapeakebay.net/channel_files/23273/attach_b_self_converted_pollutant_removal_study_report_final_12.7.15.pdf.
- K. Kelderman, A. Phillips, T. Pelton, E. Schaeffer, P. MacGillis-Falcon, and C. Bernhardt. The Clean Water Act at 50: Promises Half Kept at the Half-Century Mark. *Environmental Integrity Project: Washington, DC, USA*, 2022.

- E. M. P. Knight, W. F. Hunt, and R. J. Winston. Side-by-side evaluation of four level spreader-vegetated filter strips and a swale in eastern North Carolina. *J. Soil Water Conserv.*, 68(1):60–72, Jan. 2013. doi: <https://doi.org/10.2489/jswc.68.1.60>.
- P. F. Kosmerl. *Water balance of retrofit, right-of-way rain gardens*. MS Thesis, The Ohio State University, 2012. URL https://rave.ohiolink.edu/etdc/view?acc_num=osu1337347745.
- H. A. Lenhart and W. F. Hunt. Evaluating Four Storm-Water Performance Metrics with a North Carolina Coastal Plain Storm-Water Wetland. *J. Environ. Eng.*, 137(2):155–162, Feb. 2011. URL [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000307](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000307).
- R. M. Lent, P. K. Weiskel, F. P. Lyford, and D. S. Armstrong. Hydrologic indices for nontidal wetlands. *Wetlands*, 17(1):19–30, Mar. 1997. doi: <https://doi.org/10.1007/BF03160715>.
- H. Li, L. J. Sharkey, W. F. Hunt, and A. P. Davis. Mitigation of impervious surface hydrology using bioretention in North Carolina and Maryland. *J. Hydrol. Eng.*, 14(4):407–415, Apr. 2009. doi: [https://doi.org/10.1061/\(ASCE\)1084-0699\(2009\)14:4\(407\)](https://doi.org/10.1061/(ASCE)1084-0699(2009)14:4(407)).
- H. Lin, D. Warner, and D. Jacob. Application to Washington State Department of Ecology Water Quality Program for general use designation – Pretreatment applications & conditional use designation – oil treatment Of the Continuous Deflective Separation (CDS) technology. Technical Report (Revised), Contech Stormwater Solutions, June 2007. URL https://swbmpvwrrc.wp.prod.es.cloud.vt.edu/wp-content/uploads/2017/11/HD_Application_CDS_Item-3.pdf.
- D. Line. Final report : Evaluating BMPs for treating stormwater and wastewater from NC-DOT’s highways, industrial facilities, and borrow pits. Technical Report FHWA/NC/2006-05, CTE/NC DOT Joint Environmental Research Program, NC DOT, US DOT, Feb. 2006. URL <https://digital.ncdcr.gov/digital/collection/p16062coll19/id/140391>.
- K. Liu and J. Minor. Performance evaluation of an extensive green roof. *Presentation at Green Rooftops for Sustainable Communities, Washington DC*, pages 1–11, 2005. URL https://www.academia.edu/download/7265468/Performance_evaluation_of_an_extensive_green_roof.pdf.
- Low Impact Development Center. Projects. LID Center website, 2023. URL <https://lowimpactdevelopment.org/projects/>. Accessed: 2023-4-7.
- W. C. Lucas and D. J. Sample. Reducing combined sewer overflows by using outlet controls for Green Stormwater Infrastructure: Case study in Richmond, Virginia. *J. Hydrol.*, 520: 473–488, Jan. 2015. doi: <https://doi.org/10.1016/j.jhydrol.2014.10.029>.
- S. Luell. Evaluating the impact of bioretention cell size and swale design in treating highway bridge deck runoff. Master’s thesis, North Carolina State University, 2011. URL <http://www.lib.ncsu.edu/resolver/1840.16/6921>.
- D. Lüdecke, M. S. Ben-Shachar, I. Patil, P. Waggoner, and D. Makowski. performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, 6(60):3139, 2021. doi: 10.21105/joss.03139.

- P. Madley-Dowd, R. Hughes, K. Tilling, and J. Heron. The proportion of missing data should not be used to guide decisions on multiple imputation. *J. Clin. Epidemiol.*, 110: 63–73, June 2019. doi: <http://dx.doi.org/10.1016/j.jclinepi.2019.02.016>.
- M. A. Mansournia, M. A. Hernán, and S. Greenland. Matched designs and causal diagrams. *Int. J. Epidemiol.*, 42(3):860–869, June 2013. doi: <http://dx.doi.org/10.1093/ije/dyt083>.
- F. Markowetz and R. Spang. Inferring cellular networks—a review. *BMC Bioinformatics*, 8 Suppl 6(Suppl 6):S5, Sept. 2007. doi: <http://dx.doi.org/10.1186/1471-2105-8-S6-S5>.
- W. D. Martin and N. B. Kaye. Hydrologic characterization of undrained porous pavements. *J. Hydrol. Eng.*, 19(6):1069–1079, June 2014. doi: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000873](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000873).
- T. Matthews, A. Y. Lo, and J. A. Byrne. Reconceptualizing green infrastructure for climate change adaptation: Barriers to adoption and drivers for uptake by spatial planners. *Landsc. Urban Plan.*, 138:155–163, June 2015. doi: <https://doi.org/10.1016/j.landurbplan.2015.02.010>.
- R. McElreath. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. CRC Press, Mar. 2020. URL https://play.google.com/store/books/details?id=6H_WDwAAQBAJ.
- Melbourne Water. Introduction to WSUD. Government of Victoria, Australia, Melbourne Water website, 2023. URL <https://www.melbournewater.com.au/building-and-works/stormwater-management/introduction-wsud>. Accessed: 2023-4-7.
- J. Mentens, D. Raes, and M. Hermy. Green roofs as a tool for solving the rainwater runoff problem in the urbanized 21st century? *Landsc. Urban Plan.*, 77(3):217–226, Aug. 2006. doi: <https://doi.org/10.1016/j.landurbplan.2005.02.010>.
- J. Messamer. An evaluation of hydraulic retention time on BMP water quality performance. Master’s thesis, Colorado State University, Fort Collins, Colorado, 2011. URL https://mountainscholar.org/bitstream/handle/10217/52127/Messamer_colostate_0053N_10734.pdf?sequence=1&isAllowed=y.
- G. Millar, B. Yu, and T. Gardner. Rainfall catch efficiency for domestic water supply. In *28th International Hydrology and water resources Symposium, Wollongong, NSW, Australia, Institution of Engineers, Australia*. researchgate.net, 2003. URL https://www.researchgate.net/profile/Grant-Millar-2/publication/279205924_Rainfall_Catch_Efficiency_for_Domestic_Water_Supply/links/5590d70908aed6ec4bf6764d/Rainfall-Catch-Efficiency-for-Domestic-Water-Supply.pdf.
- W. J. Mitsch. What is ecological engineering? *Ecol. Eng.*, 45:5–12, Aug. 2012. doi: <https://doi.org/10.1016/j.ecoleng.2012.04.013>.
- W. J. Mitsch, L. Zhang, E. Waletzko, and B. Bernal. Validation of the ecosystem services of created wetlands: Two decades of plant succession, nutrient retention, and carbon sequestration in experimental riverine marshes. *Ecol. Eng.*, 72:11–24, Nov. 2014. doi: <https://doi.org/10.1016/j.ecoleng.2014.09.108>.

- A. Moran, B. Hunt, and J. Smith. Hydrologic and water quality performance from green roofs in Goldsboro and Raleigh, North Carolina. Technical report, U.S. Department of Energy Office of Scientific and Technical Information, July 2005. URL <https://www.osti.gov/etdeweb/biblio/20861933>.
- S. Morgan, S. Celik, and W. Retzlaff. Green roof storm-water runoff quantity and quality. *J. Environ. Eng.*, 139(4):471–478, Apr. 2013. doi: [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000589](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000589).
- M. F. Müller and M. C. Levy. Complementary vantage points: Integrating hydrology and economics for sociohydrologic knowledge generation. *Water Resour. Res.*, 55(4):2549–2571, Apr. 2019. doi: <https://onlinelibrary.wiley.com/doi/10.1029/2019WR024786>.
- R. Nawaz, A. McDonald, and S. Postoyko. Hydrological performance of a full-scale extensive green roof located in a temperate climate. *Ecol. Eng.*, 82:66–80, Sept. 2015. doi: <https://doi.org/10.1016/j.ecoleng.2014.11.061>.
- NERR. Implementing credits and incentives for innovative stormwater management. Technical Report NOAA grant number NA09NOS4190153 Final Report, North Carolina State University and Ohio Department of Natural Resources for the National Estuarine Research Reserve System Science Collaborative, 2016.
- M. E. Newcomer, J. J. Gurdak, L. S. Sklar, and L. Nanus. Urban recharge beneath low impact development and effects of climate variability and change. *Water Resour. Res.*, 50(2):1716–1734, Feb. 2014. doi: <http://doi.org/10.1002/2013WR014282>.
- M. K. Nungesser and M. J. Chimney. A hydrologic assessment of the Everglades Nutrient Removal Project, a subtropical constructed wetland in South Florida (USA). *Ecol. Eng.*, 27(4):331–344, Oct. 2006. doi: <https://doi.org/10.1016/j.ecoleng.2006.08.007>.
- NYS DEC. Construction stormwater toolbox. NY State Department of Environmental Conservation website, 2023. URL <https://www.dec.ny.gov/chemical/8694.html>. Accessed: 2023-4-13.
- Orange County Public Works. Amendment No. 1 to Clean Water Act Section 401: La Pata Avenue Extension Project Certification Number 12C-056. California Regional Water Quality Control Board, San Diego Region, Oct. 2014. URL https://www.waterboards.ca.gov/rwqcb9/water_issues/programs/401_certification/docs/projects/12C/12C-056_Amendment_1.pdf.
- M. R. Owen, R. T. Pavlowsky, and A. D. Mulling. Final report for water quality monitoring and load reduction evaluation from detention basin retrofits in Springfield, Missouri. Technical Report OEWRI EDR-15-004, The Ozarks Environmental and Water Resources Institute, 2015. URL <https://bearworks.missouristate.edu/cgi/viewcontent.cgi?article=1009&context=oewri-tech>.
- PA DEP. *Pennsylvania Stormwater Best Management Practices Manual*. Pennsylvania Department of Environmental Protection, Bureau of Watershed Management, Dec. 2006. URL <https://www.depgreenport.state.pa.us/elibrary/GetFolder?FolderID=4673>.

- A. Palla, J. J. Sansalone, I. Gnecco, and L. G. Lanza. Storm water infiltration in a monitored green roof for hydrologic restoration. *Water Sci. Technol.*, 64(3):766–773, 2011. doi: <http://doi.org/10.2166/wst.2011.171>.
- R. Pitt, S. Clark, and J. Voorhees. Water removal in bioretention devices by evapotranspiration processes and related issues affecting performance. Penn State University Citeseer website, 2007. URL <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2de47d0670c7c98934ffcdfeee4f7cc6f54a4dcb>. Accessed: 2023-4-13.
- N. L. Poff and J. K. H. Zimmerman. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshw. Biol.*, 55(1):194–205, Jan. 2010. doi: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2427.2009.02272.x>.
- V. Pons, E. M. H. Abdalla, F. Tscheikner-Gratl, K. Alfredsen, E. Sivertsen, J.-L. Bertrand-Krajewski, and T. M. Muthanna. Practice makes the model: A critical review of stormwater green infrastructure modelling practice. *Water Research*, 236:119958, 2023. ISSN 0043-1354. doi: <https://doi.org/10.1016/j.watres.2023.119958>.
- C. J. Pratt, J. D. G. Mantle, and P. A. Schofield. UK research into the performance of permeable pavement, reservoir structures in controlling stormwater discharge quantity and quality. *Water Sci. Technol.*, 32(1):63–69, Jan. 1995. doi: [https://doi.org/10.1016/0273-1223\(95\)00539-Y](https://doi.org/10.1016/0273-1223(95)00539-Y).
- M. J. Prokop. Determining the effectiveness of the Villanova bio-infiltration traffic island in infiltrating annual runoff. Master’s thesis, Villanova University, 2003. URL <https://www1.villanova.edu/content/dam/villanova/engineering/vcase/vusp/Prokop-Thesis.pdf>.
- A. Reese and J. Parker. Green infrastructure sizing criteria development, part 2. Stormwater Solutions website, Aug. 2014a. URL <https://www.stormh2o.com/bmps/article/13012565/green-infrastructure-sizing-criteria-development-part-2>. Accessed: 2023-4-5.
- A. Reese and J. Parker. Green infrastructure sizing criteria development, part 1. Stormwater Solutions website, July 2014b. URL <https://www.stormh2o.com/bmps/article/13012140/green-infrastructure-sizing-criteria-development>. Accessed: 2023-4-5.
- D. D. Reible, C. N. Haas, J. H. Pardue, and W. J. Walsh. Toxic and contaminant concerns generated by hurricane Katrina. *The Bridge*, 36(1), June 2006. URL <https://www.nae.edu/7623/ToxicandContaminantConcernsGeneratedbyHurricaneKatrina>.
- K. C. Rice and J. D. Jastram. Rising air and stream-water temperatures in Chesapeake Bay region, USA. *Clim. Change*, 128(1):127–138, Jan. 2015. doi: <https://doi.org/10.1007/s10584-014-1295-9>.
- J. Rieckermann, V. Bares, O. Kracht, D. Braun, and W. Gujer. Estimating sewer leakage from continuous tracer experiments. *Water Res.*, 41(9):1960–1972, May 2007. URL <http://dx.doi.org/10.1016/j.watres.2007.01.024>.

- B. Riley, L. Hoffman, and R. Berghage. Green roof thermal and stormwater management performance: The Gratz Building case study, New York City. Technical Report 09-05, NYSERDA, Apr. 2009.
- Y.-N. Rim. *Analyzing runoff dynamics of paved soil surface using weighable lysimeters*. PhD thesis, Berlin, Technische Universität Berlin, Diss., 2011, 2011. URL https://scholar.archive.org/work/kqkoc437dvhy7br5ytq3obm37m/access/wayback/https://depositonce.tu-berlin.de/bitstream/11303/3123/1/Dokument_41.pdf.
- M. Rodell, H. K. Beaudoin, T. S. L'Ecuyer, W. S. Olson, J. S. Famiglietti, P. R. Houser, R. Adler, M. G. Bosilovich, C. A. Clayson, D. Chambers, E. Clark, E. J. Fetzer, X. Gao, G. Gu, K. Hilburn, G. J. Huffman, D. P. Lettenmaier, W. T. Liu, F. R. Robertson, C. A. Schlosser, J. Sheffield, and E. F. Wood. The Observed State of the Water Cycle in the Early Twenty-First Century. *J. Clim.*, 28(21):8289–8318, Nov. 2015. doi: <https://doi.org/10.1175/JCLI-D-14-00555.1>.
- D. J. Rosa, J. C. Clausen, and M. E. Dietz. Calibration and verification of SWMM for low impact development. *J. Am. Water Resour. Assoc.*, 51(3):746–757, June 2015. doi: <https://doi.org/10.1111/jawr.12272>.
- L. A. Rossman and US EPA. Modeling low impact development alternatives with SWMM. *J. Water Manag. Model.*, 2010. doi: <https://doi.org/10.14796/JWMM.R236-11>.
- A. Roy-Poirier, P. Champagne, and Y. Filion. Review of bioretention system research and design: Past, present, and future. *J. Environ. Eng.*, 136(9):878–889, Sept. 2010. doi: [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000227](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000227).
- J. Runge. Causal Inference, Causal Discovery, and Machine Learning with Jakob Runge, Dec. 2022. URL <https://www.youtube.com/watch?v=R5JMeEy9koA>.
- B. Rushton. Broadway Outfall Stormwater Retrofit Project. Technical Report DEP Contract Number WM793, FL Department of Environmental Protection, Southwest Florida Water Management District, June 2006. URL https://www.swfwmd.state.fl.us/sites/default/files/medias/documents/broadway_outfall_retrofit.pdf.
- M. Rutsch, J. Rieckermann, and P. Krebs. Quantification of sewer leakage: a review. *Water Sci. Technol.*, 54(6-7):135–144, 2006. doi: <http://dx.doi.org/10.2166/wst.2006.616>.
- M. Rutsch, J. Rieckermann, J. Cullmann, J. B. Ellis, J. Vollertsen, and P. Krebs. Towards a better understanding of sewer exfiltration. *Water Res.*, 42(10-11):2385–2394, May 2008. doi: <http://dx.doi.org/10.1016/j.watres.2008.01.019>.
- Save the Rain. OnCenter green roof. Government of Onondaga County, NY, USA, Save the Rain website, Oct. 2020. URL <https://savetherain.us/projects/oncenter-green-roof/>. Accessed: 2023-4-7.
- D. A. Schlea. *Retention and management of stormwater runoff with rain gardens and rainwater harvesting systems*. MS Thesis, The Ohio State University, 2011. URL https://rave.ohiolink.edu/etdc/view?acc_num=osu1306853271.

- W. R. Selbig and N. Balster. USGS scientific investigations report 2010-5077: Evaluation of Turf-Grass and Prairie-Vegetated Rain Gardens in a Clay and Sand Soil, Madison, Wisconsin, Water Years 2004-08. US Geological Survey, 2010. URL <https://pubs.usgs.gov/sir/2010/5077/pdf/sir20105077.pdf>. Accessed: 2023-4-23.
- A. Selvakumar, R. Field, E. Burgess, and R. Amick. Exfiltration in sanitary sewer systems in the US. *Urban Water J.*, 1(3):227–234, Sept. 2004. doi: <https://doi.org/10.1080/15730620410001732017>.
- R. Shakya and L. Ahiablame. A synthesis of social and economic benefits linked to green infrastructure. *Water*, 13(24):3651, Dec. 2021. doi: <https://doi.org/10.3390/w13243651>.
- L. J. Sharkey. 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, 2006. URL <http://www.lib.ncsu.edu/resolver/1840.16/2062>.
- Y. Shevah. Chapter 3: Impact of persistent droughts on the quality of the Middle East water resources. In S. Ahuja, editor, *Separation science and technology*, volume 11, pages 51–84. Academic Press, Jan. 2019. doi: <https://doi.org/10.1016/B978-0-12-815730-5.00003-X>.
- A. Shukla, S. Shukla, and M. D. Annable. Using nocturnal water level fluctuations for estimating seepage from stormwater detention systems. *Hydrol. Process.*, 29(26):5465–5476, Dec. 2015. doi: <https://doi.org/10.1002/hyp.10600>.
- M. Simon. EPA’s Summary Report of the Collaborative Green Infrastructure Pilot Project for the Middle Blue River in Kansas City, MO. Technical Report EPA/600/R-16/085, National Risk Management Research Laboratory, Water Supply and Water Management Division, US EPA, Aug. 2016. URL <https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100R8DR.pdf>.
- D. W. Stanley. An evaluation of Pollutant Removal by a Demonstration Urban Stormwater Detention Pond. Technical Report APES 94-07, Albemarle-Pamlico Estuarine Study, NC Department of Environment, Health and Natural Resources, US EPA National Estuary Program, Nov. 1994. URL <https://files.nc.gov/apnep/documents/files/apes/9407ocr.pdf>.
- P. Starke, P. Göbel, and W. G. Coldewey. Urban evaporation rates for water-permeable pavements. *Water Sci. Technol.*, 62(5):1161–1169, 2010. doi: <https://doi.org/10.2166/wst.2010.390>.
- P. Starke, P. Göbel, and W. G. Coldewey. Effects on evaporation rates from different water-permeable pavement designs. *Water Sci. Technol.*, 63(11):2619–2627, 2011. doi: <https://doi.org/10.2166/wst.2011.168>.
- J. Steffen, M. Jensen, C. A. Pomeroy, and S. J. Burian. Water supply and stormwater management benefits of residential rainwater harvesting in U.S. cities. *J. Am. Water Resour. Assoc.*, 49(4):810–824, Aug. 2013. doi: <https://onlinelibrary.wiley.com/doi/10.1111/jawr.12038>.

- V. Stovin. The potential of green roofs to manage Urban Stormwater. *Water Environ. J.*, 24(3):192–199, Sept. 2010. doi: <https://doi.org/10.1111/j.1747-6593.2009.00174.x>.
- V. Stovin, G. Vesuviano, and H. Kasmin. The hydrological performance of a green roof test bed under UK climatic conditions. *J. Hydrol.*, 414-415:148–161, Jan. 2012. doi: <https://doi.org/10.1016/j.jhydrol.2011.10.022>.
- V. Stovin, S. Poë, and C. Berretta. A modelling study of long term green roof retention performance. *J. Environ. Manage.*, 131:206–215, Dec. 2013. doi: <http://doi.org/10.1016/j.jenvman.2013.09.026>.
- V. Stovin, G. Vesuviano, and S. De-Ville. Defining green roof detention performance. *Urban Water J.*, 14(6):574–588, July 2017. doi: <https://doi.org/10.1080/1573062X.2015.1049279>.
- K. R. Strauch, D. L. Rus, and K. E. Holm. Water balance monitoring for two bioretention gardens in Omaha, Nebraska, 2011–14. Technical Report 2015-5188, U.S. Geological Survey, 2016. doi: <https://doi.org/10.3133/sir20155188>.
- W. H. Strosnider, D. R. Hitchcock, M. K. Burke, and A. J. Lewitus. Predicting hydrology in wetlands designed for coastal stormwater management. In *2007 ASAE Annual Meeting*, page 1. elibrary.asabe.org, 2007. URL <https://elibrary.asabe.org/abstract.asp?aid=23079>.
- K. Teague, B. Rushton, and D. Hunneycutt. Stormwater runoff treatment by a filtration system and wet pond in Tampa, Florida. *Tampa, FL: Southwest Florida Water Management District*, 2005.
- A. Teemusk and Ü. Mander. Rainwater runoff quantity and quality performance from a greenroof: The effects of short-term events. *Ecol. Eng.*, 30(3):271–277, July 2007. doi: <https://doi.org/10.1016/j.ecoleng.2007.01.009>.
- G. Tellessen and V. Allen. Field Verification of Performance of an Engineered Phosphorus Removal Media Deployed in a Flow Based Stormwater Filter. Technical report, Contech Engineered Solutions.
- Terre Hill Stormwater Systems. NJCAT technology verification: Terre Kleen hydrodynamic separator. Technical report, New Jersey Corporation for Advance Technology program, Jan. 2017. URL <http://www.njcat.org/uploads/newDocs/TerreKleenNJCATVerificationReportFinal.pdf>.
- J. Textor, B. van der Zander, M. S. Gilthorpe, M. Liškiewicz, and G. T. Ellison. Robust causal inference using directed acyclic graphs: the r package 'dagitty'. *International Journal of Epidemiology*, 45(6):1887–1894, 2016. doi: 10.1093/ije/dyw341.
- C. R. Thorne, E. C. Lawson, C. Ozawa, S. L. Hamlin, and L. A. Smith. Overcoming uncertainty and barriers to adoption of Blue-Green Infrastructure for urban flood risk management. *J. Flood Risk Manag.*, 11(S2):S960–S972, Oct. 2015. doi: <https://doi.org/10.1111/jfr3.12218>.
- Toronto Regional Conservation Authority. Evaluation of an Extensive Green Roof, York

- University, Toronto, Ontario. Sustainable Technologies website, July 2006. URL https://sustainabletechnologies.ca/app/uploads/2013/03/GR_york_fullreport.pdf.
- R. G. Traver and P. A. DeBarry. Designing for Infiltration — A Perspective. In *World Water & Environmental Resources Congress 2003*, Reston, VA, June 2003. American Society of Civil Engineers. doi: [https://doi.org/10.1061/40685\(2003\)325](https://doi.org/10.1061/40685(2003)325).
- UNH Stormwater Center. Overcoming barriers to green infrastructure. Technical Report Green Infrastructure for New Hampshire Coastal Communities Factsheet, University of New Hampshire Stormwater Center, 2014.
- US Census Bureau. Tiger/Line Population of Metropolitan Statistical Areas, 2022a. URL <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-g-eodatabase-file.html>.
- US Census Bureau. American Community Survey, 2021, 2022b. URL <https://www.census.gov/programs-surveys/acs/data.html>.
- US Clivar. Causal inference in Earth system sciences (presentation by Jakob Runge), Apr. 2021. URL https://www.youtube.com/watch?v=wJ_AkNELm6Q.
- US EPA. Green infrastructure opportunities and barriers in the greater Los Angeles region. Technical Report EPA 833-R-13-001, US Environmental Protection Agency, 2012 Green Infrastructure Technical Assistance Program, Council for Watershed Health, Aug. 2013.
- US EPA. Green infrastructure modeling toolkit. US Environmental Protection Agency Water Research website, May 2014. URL <https://www.epa.gov/water-research/green-infrastructure-modeling-toolkit>. Accessed: 2023-4-16.
- US EPA. Green Infrastructure. US Environmental Protection Agency, Office of Water website, 2023. URL <https://www.epa.gov/green-infrastructure>. Accessed: 2023-4-7.
- US EPA Office of Water. Ecoregional nutrient criteria for Rivers and streams. Technical report, US Environmental Protection Agency, Mar. 2013. URL <https://www.epa.gov/nutrient-policy-data/ecoregional-nutrient-criteria-rivers-and-streams>.
- VA Stormwater Management Program. Virginia runoff reduction method, version 3.0. Virginia Department of Environmental Quality Water Permitting Division, Stormwater BMP Clearinghouse website, Apr. 2016. URL <https://swbmp.vwrrc.vt.edu/vrrm/>. Accessed: 2023-4-7.
- N. A. Valinski and D. G. Chandler. Infiltration performance of engineered surfaces commonly used for distributed stormwater management. *J. Environ. Manage.*, 160:297–305, Sept. 2015. doi: <http://dx.doi.org/10.1016/j.jenvman.2015.06.032>.
- S. van Buuren. *Flexible Imputation of Missing Data. Second Edition*. CRC Press, Boca Raton, FL., 2018.
- S. van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate Imputation by Chained Equations in R, 2011.

- T. Van Seters, L. Rocha, D. Smith, and G. MacMillan. Evaluation of green roofs for runoff retention, runoff quality, and leachability. *Water Qual. Res. J. Can.*, 44(1):33–47, Feb. 2009. doi: <https://doi.org/10.2166/wqrj.2009.005>.
- E. Vanuytrecht, C. Van Mechelen, K. Van Meerbeek, P. Willems, M. Hermy, and D. Raes. Runoff and vegetation stress of green roofs under different climate change scenarios. *Landsc. Urban Plan.*, 122:68–77, Feb. 2014. doi: <https://doi.org/10.1016/j.landurbplan.2013.11.001>.
- N. D. VanWoert, D. B. Rowe, J. A. Andresen, C. L. Rugh, R. T. Fernandez, and L. Xiao. Green roof stormwater retention: effects of roof surface, slope, and media depth. *J. Environ. Qual.*, 34(3):1036–1044, May 2005. doi: <http://dx.doi.org/10.2134/jeq2004.0364>.
- G. Vico, R. Revelli, and A. Porporato. Ecohydrology of street trees: design and irrigation requirements for sustainable water use. *Ecohydrol.*, 7(2):508–523, Apr. 2014. doi: <https://onlinelibrary.wiley.com/doi/10.1002/eco.1369>.
- E. L. Villarreal and L. Bengtsson. Response of a Sedum green-roof to individual rain events. *Ecol. Eng.*, 25(1):1–7, July 2005. doi: <https://doi.org/10.1016/j.ecoleng.2004.11.008>.
- D. Vineyard, W. W. Ingwersen, T. R. Hawkins, X. Xue, B. Demeke, and W. Shuster. Comparing green and grey infrastructure using life cycle cost and environmental impact: A rain garden case study in Cincinnati, OH. *JAWRA Journal of the American Water Resources Association*, 51(5):1342–1360, Oct. 2015. doi: <https://doi.org/10.1111/1752-1688.12320>.
- E. Voyde, E. Fassman, and R. Simcock. Hydrology of an extensive living roof under subtropical climate conditions in Auckland, New Zealand. *J. Hydrol.*, 394(3):384–395, Nov. 2010. doi: <https://doi.org/10.1016/j.jhydrol.2010.09.013>.
- B. M. Wadzuk, D. Schneider, M. Feller, and R. G. Traver. Evapotranspiration from a green-roof storm-water control measure. *J. Irrig. Drain. Eng.*, 139(12):995–1003, 2013. doi: [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000643](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000643).
- B. M. Wadzuk, J. M. Hickman, Jr, and R. G. Traver. Understanding the role of evapotranspiration in bioretention: Mesocosm study. *Journal of Sustainable Water in the Built Environment*, 1(2):04014002, 2015. doi: <https://doi.org/10.1061/JSWBAY.0000794>.
- M. Walch. Test-Driving Stormwater BMPs. *Water Environ. Technol.*, 20(4):59–62, 2008. URL https://deldot.gov/Programs/stormwater/pdfs/permit_docs/Phase_I_Report_2008/Appendix_G.pdf.
- C. J. Walsh, A. H. Roy, J. W. Feminella, P. D. Cottingham, P. M. Groffman, and R. P. Morgan. The urban stream syndrome: Current knowledge and the search for a cure. *J. North Am. Benthol. Soc.*, 24(3):706–723, Sept. 2005. doi: <https://doi.org/10.1899/04-028.1>.
- M. P. Wanielista, Y. A. Yousef, L. M. Van DeGraaff, and S. H. Rehmann-Koo. Best management practices - enhanced erosion and sediment control using Swale blocks. Final report. Technical Report FL/DOT/BMR-85-293, University of Central Florida, FL Department

- of Transportation, Sept. 1986. URL <https://stormwater.ucf.edu/research/FILES/Wanielista%20-%20best%20management%20practices.pdf>.
- B. Wardynski, W. Hunt, and R. Brown. Low impact development stormwater educational workshop, Part 6: DRAINMOD: Introduction to the ideal bioretention model. Ohio EPA Municipal Storm Water Training Series, 2011.
- B. J. Wardynski and W. F. Hunt. Are bioretention cells being installed per design standards in North Carolina? A field study. *J. Environ. Eng.*, 138(12):1210–1217, Dec. 2012. doi: [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0000575](https://doi.org/10.1061/(ASCE)EE.1943-7870.0000575).
- WEF and ASCE. *Urban Runoff Quality Management*. American Society of Civil Engineers publications, Jan. 1998.
- C. T. Welborn and J. E. Veenhuis. Effects of runoff controls on the quantity and quality of urban runoff at two locations in Austin, Texas. Technical Report 87-4004, US Geological Survey, 1987. URL <https://pubs.usgs.gov/wri/1987/4004/report.pdf>.
- R. J. Winston, W. F. Hunt, S. G. Kennedy, and J. D. Wright. Evaluation of Permeable Friction Course (PFC), Roadside Filter Strips, Dry Swales, and Wetland Swales for Treatment of Highway Stormwater Runoff. Technical Report FHWA/NC/2007-21, North Carolina Department of Transportation, Jan. 2011.
- R. J. Winston, J. D. Dorsey, and W. F. Hunt. Quantifying volume reduction and peak flow mitigation for three bioretention cells in clay soils in northeast Ohio. *Sci. Total Environ.*, 553:83–95, May 2016. doi: <http://dx.doi.org/10.1016/j.scitotenv.2016.02.081>.
- C. Worstell. Green infrastructure in the state of New Jersey: Statutory and regulatory barriers to green infrastructure implementation. *New Jersey Future, Trenton, NJ*, 2013.
- WRF. International Stormwater BMP Database, version 2019-12-29. Water Research Foundation Consortium with Environmental and Water Resources Institute (ASCE), Federal Highway Administration, US Environmental Protection Agency, Dec. 2019. URL <https://bmpdatabase.org/>. Accessed: 2021-3-20.
- WRF, EWRI, US EPA, US DOT, Geosyntec Consultants, Wright Water Engineers. BMP mapping tool. <https://bmpdatabase.org/bmp-mapping-tool>, 2023. URL <https://bmpdatabase.org/bmp-mapping-tool>. Accessed: 2023-4-19.
- S. L. Yu, G. Michael Fitch, and T. Andrew Earles. Constructed Wetlands for Stormwater Management. Technical report, Virginia Transportation Research Council (VA DOT), 1998. URL https://www.virginiadot.org/vtrc/main/online_reports/pdf/99-r14.pdf.
- Z. Zahmatkesh, S. J. Burian, M. Karamouz, H. Tavakol-Davani, and E. Goharian. Low-impact development practices to mitigate climate change effects on urban stormwater runoff: Case study of New York City. *J. Irrig. Drain. Eng.*, 141(1):04014043, 2015. doi: [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0000770](https://doi.org/10.1061/(ASCE)IR.1943-4774.0000770).
- X. Zhang and M. Hu. Effectiveness of rainwater harvesting in runoff volume reduction in a

planned industrial park, China. *Water Resour. Manage.*, 28(3):671–682, Feb. 2014. doi:
<https://doi.org/10.1007/s11269-013-0507-9>.

Vita

Caitlin Eger

Data Scientist

NH Board of Engineers EIT License 06618 (2015)

Dynamic engineer experienced in causal inference and statistical analysis, with a diverse range of data modeling skills across various industries.

Enthusiastic team leader and project manager.

Experience

2020 – Current

Data Scientist, RTR Technologies, LLC

Marine Corps aviation supply chain support, with a focus on supply forecasting, data specification and model design. Agile software development of predictive models and dashboard displays.

2015 - 2020

NSF Graduate Research Trainee, SU EMPOWER Program

Interdisciplinary research centered on scientific publication, decision support and public presentations. Versatile in hydrology, ecology and environmental engineering with field instrumentation, data collection and coding experience. Skilled grant and proposal writer, securing over \$300,000 for training and research.

2013 - 2015

Staff Scientist, Syracuse Center of Excellence

Public outreach and small grants administration for research on urban stormwater management. Grant proposal pitching and writing to support modernization of public water infrastructure, environmental sustainability and green building design.

Education

2015 – Current (Spring 2023)

Ph.D. Civil & Environmental Engineering
Syracuse University

2009 - 2012

M.S. Ecological Engineering
The Ohio State University

2004 - 2008

B.S. Environmental Science & Chemistry
Juniata College

About Me

- Causal inference geek
- R and Python API designer
- Friendly team collaborator
- Technical communicator
- Adaptive learner

Tools

RStudio, VSCode, SSMS, Docker,
Azure DevOps, Azure Data Studio
LaTeX

Latest Publication

C. Balderas Guzman, R. Wang, O. Muellerklein, M. Smith, **C.G. Eger**. Comparing stormwater quality and watershed typologies across the United States: A machine learning approach, *Water Research*, Volume 216, 2022.

cgeger@syr.edu
[linkedin.com/in/cgeger](https://www.linkedin.com/in/cgeger)
[@Ecologentsia](#)