Stormwater Quantity and Quality: An Assessment of Runoff Probability Estimation Methods and the Effects of Bioretention Soil Mixtures on Aquatic Toxicity

Submitted By

Kelly DeGuzman

A thesis submitted to the Graduate Faculty of

Auburn University

Auburn, Alabama

in partial fulfillment of the requirements

for the Degree of Master of Science

May 4, 2024

Keywords: Stormwater, Runoff, Bioretention cells, Aquatic toxicity,

Master's Committee:

Advisor: Thorsten Knappenberger, Associate Professor, Crop, Soil, and Environmental Sciences

Members: Eve Brantley, Professor, Crop, Soil, and Environmental Sciences

Yaniv Olshansky, Assistant Professor, Crop, Soil, and Environmental Sciences

ABSTRACT

Stormwater runoff occurs when precipitation exceeds the available storage of a watershed. Urbanization changes the natural hydrology of a watershed as cover types, slopes, flow paths, and antecedent moisture conditions (AMC) are altered. Generally, there are two concerns associated with increased stormwater runoff: water quality and water quantity. Stormwater quality and quantity may be controlled through several types of stormwater management practices (SMPs). The goal of low impact development (LID) is to manage stormwater in a way that preserves the pre-development hydrological characteristics of a watershed, including its ability to infiltrate, evaporate, filter, and detain stormwater (Dietz, 2007; Prince George's County, 1999; USDA NRCS, 1986; USEPA, 2000).

The runoff threshold of a catchment is often reported in the literature as a measure used to evaluate the effectiveness of LID techniques and is oftentimes determined using linear regression. However, due to the heteroscedastic nature of precipitation and runoff data, linear regression analyses can result in invalid conclusions, so the use of a binomial regression model was investigated. Precipitation and runoff data collected from five studies were assessed for homoscedasticity and applied to four linear regression models for the evaluation of the linear regression runoff threshold (LRRT) and the associated 95% bootstrapping confidence intervals. Log-transformation corrected the heteroscedasticity of two of the five precipitation and runoff datasets. While mixed-effect linear models accounted for heteroscedasticity, these models often resulted in extremely large confidence intervals. A binomial regression model was created to determine the likelihood of runoff based on precipitation depth. For each catchment, the 10%-90% runoff probability range (p10-p90) is reported to provide the user with a more comprehensive understanding of when a catchment produces runoff than the LRRT. The p10-p90 range reflects

the effects that environmental factors may have on runoff generation. For example, impervious catchments with limited interaction with the vegetation and soil produce a narrow p10-p90 range. Conversely, LID practices encourage the interaction of runoff with environmental factors and result in a wider p10-p90 range. Utilization of the binomial regression methodology presented herein is recommended for evaluation of the likelihood of runoff for each precipitation depth.

Heavy metal concentrations in stormwater discharges are regulated and monitored as these pollutants can be toxic to aquatic communities (USEPA, 2007). The total concentration of a heavy metal does not represent the concentration of a metal ion available to aquatic organisms. For a heavy metal to be bioavailable for uptake through the gill, it must be in a dissolved form. When evaluating the toxicity of stormwater, the bioavailability and speciation of a heavy metal should be considered. The influent and effluent of four bioretention soil mixtures (BSMs) from ten storms were studied for the purpose of examining speciation shifts of stormwater pollutants and investigating potential changes to stormwater toxicity following filtration through BSMs. Further, this study sought to determine which, if any, of the BSMs were more adept at decreasing pollutant bioavailability. Visual MINTEQ 3.1 was used to predict pollutant speciation and the Windward Environmental, LLC Biotic Ligand Model (BLM) (v 3.41.2.45) was used to determine toxic concentrations of heavy metal species. No noticeable speciation shifts were noted within bioretention cell (BRC) effluent for cadmium, copper, lead, and zinc. However, the speciation of chromium effluent was dependent on the initial pollutant concentrations. A multiple-factor analysis (MFA) indicated that the four BSMs do not differ from one another in reducing the BLM-identified toxic limit (relative toxicity) of stormwater effluent for the BLM-selected aquatic organisms. BRCs are most effective at reducing the toxicity of stormwater when the stormwater contains high pollutant concentrations. At low pollutant concentrations, BRCs may increase the toxicity of the effluent stormwater through export of copper.

ACKNOWLEDGEMENTS

I would like to express my appreciation to all those who have been part of this journey. This thesis would not have been possible without the guidance and the help of many individuals who in one way or another have contributed to its completion.

I would like to express my deepest gratitude to my advisor, Dr. Thorsten Knappenberger, who I could not have undertaken this journey without. His guidance, knowledge, and expertise in the topics of my study have been instrumental in the completion of this thesis. As a distance learning student, I am particularly appreciative of our early morning zoom calls and his flexibility and accommodation for long-distance communication. His patience and willingness to explain each detail is truly valued and has not gone unnoticed. I am also grateful to the other members of my committee, Dr. Eve Brantley and Dr. Yaniv Olshansky. Without their dedication and valuable input, this thesis would not have been possible. The conclusions of this thesis would not have been as compelling or possible without the collaboration of my committee.

I extend my gratitude and appreciation to Anand Jayakaran, Eban Bean, Andrew Tirpak, and Zhiyong Fu for providing data from their respective studies for analysis. I would also like to extend gratitude to Tham Hoang and Robert Santore for their provided guidance and support. Additionally, I would like to extend my sincere thanks to my professional colleague Greg Fleischer for his continued encouragement from day one and for always pushing me outside of my comfort zone. I have grown both professionally and academically because of his support.

Without the ongoing support of a few special people in my life, completing my Master of Science would not have seemed attainable. Many thanks go to my family, who made this journey less daunting. I would like to thank my husband Rich who persuaded me not to take the easy road, knowing I would not be satisfied. I am also grateful for my parents for listening to me when things became stressful and encouraging mental health days. I would like to thank my children, Richard and Penelope, who unknowingly have sacrificed time with me during this pursuit. They say it takes a village to raise a child, but it also takes one to complete a thesis. Thank you to my village.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ADP	Antecedent dry periods
AIC	Akaike information criterion
AM0C	Antecedent moisture conditions
ANCOVA	Analysis of covariance
ASI	Antecedent Soil Moisture Index
BLM	Biotic Ligand Model
BMP	Best management practice
BP	Breusch-Pagan
BRC	Bioretention cell
BSM	Bioretention soil mixture
Ca	Calcium
Cd	Cadmium
CFD	Cumulative distribution function
CGP	Concrete grid pavers
Cr	Chromium
Cu	Copper
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
ET	Evapotranspiration
HC5	Hazardous concentration for 5% of the population
HDPE	High-density polyethylene
IWS	Internal water storage
LC50	Lethal concentration for 50% of the population
LID	Low impact development
LRRT	Linear regression runoff threshold
MFA	Multiple-factor analysis

Mg	Magnesium
NO_2	Nitrite
NO_3	Nitrate
NPDES	National Pollutant Discharge Elimination System
p10-p90	10%-90% runoff probability range
Pb	Lead
PC	Porous concrete
PICP	Permeable interlocking concrete pavers
PO_4	Phosphate
RE	Random effects
RT	Runoff threshold
SMP	Stormwater management practice
SSD	Species sensitivity distribution
TD	Topdressing
TKN	Total Kjeldahl nitrogen
TN	Total nitrogen
TP	Total phosphorus
TSS	Total suspended solids
USEPA	United States Environmental Protection Agency
WSDE	Washington State Department of Ecology
WTR	Water treatment residuals
Zn	Zinc

CHAPTER 1 – INTRODUCTION

Stormwater runoff occurs when precipitation exceeds the available storage of a watershed. It is plainly seen that as a landscape is developed, the flow of stormwater is altered. Urbanization affects stream health, channel stability, and water quality. Generally, there are two concerns associated with increased stormwater runoff: water quality and water quantity. As stormwater moves across a surface, it accumulates pollutants such as sediment, nutrients, pathogens, oil and grease, heavy metals, and thermal loadings creating water quality concerns for downstream waters (USEPA, 2003). Additionally, as impervious cover increases, so do the volume and rate of runoff, both of which are water quantity concerns (Dietz, 2007). To mitigate these impacts, implementation of stormwater management practices (SMPs) is required and often enforced by government agencies through permitting programs. Further the design, operation, and maintenance of SMPs is also often supervised by these agencies (NYSDEC, 2015; WSDE, 2019).

While stormwater runoff has been occurring since the dawn of time, as our local and global climates change stormwater has become of increased focus to regulatory agencies and environmental reviews. The United States Environmental Protection Agency (USEPA) recognizes the concern of stormwater quality and quantity and enforces programs aimed to mitigate potential impacts that occur from urbanization. The Clean Water Act Section 303(d) program requires identification of impaired waterbodies utilizing available data, development of water quality assessments and plans, and approaches implemented to restore and protect waterbodies. The EPA also administers of the National Pollutant Discharge Elimination System (NPDES) program which regulates stormwater discharges from construction activities, industrial activities, municipal sources, and more throughout the U.S. The responsibility for implementation of these programs is delegated to each state.

Stormwater quality and quantity may be controlled through distinct types of stormwater management practices (SMPs). Inevitably, the goal of managing the quality of stormwater runoff is to prevent pollutants from reaching receiving waterbodies where they can impact aquatic life. Whereas mitigation of water quantity impacts is aimed at attenuating flow and managing peak discharge rates to prevent downstream impacts associated with erosion and flooding. Many SMPs offer benefits to both quality and quantity concerns. For example, infiltration and evapotranspiration (ET) practices reduce runoff, thereby mitigating discharge volume and reducing pollutant loadings (Davis et al., 2006). SMPs are commonly classified into the following categories by their function: infiltration, ET, filtration, and detention. Further, hybrid or dual systems may be designed to provide more than one function. Popular SMPs within urbanized areas include bioretention systems, underground infiltrations systems, green roofs, sand filters, and permeable pavements, although new practices are continually developed (NYSDEC, 2022).

RESEARCH OBJECTIVES

My research is rooted in stormwater management, focusing on the two main aspects of stormwater - water quality and water quantity. The reason I chose to research stormwater is that as our global landscape continues to urbanize, it loses the ability to reduce stormwater and increases the amount of runoff and pollutants reaching waterbodies. While Federal and State governments implement programs that require incorporation of low impact development (LID) practices to mitigate downstream impacts including erosion, flooding, and increased pollutant loadings, they often fall short. My research aims to refine the techniques that are commonly used for evaluating the effectiveness of low impact development practices.

THESIS STRUCTURE

To address the research objectives, the thesis begins with a literature review (Chapter 2), followed by estimating runoff probability from precipitation data: a binomial regression analysis (Chapter 3), and evaluations of the effect of bioretention soil mixtures on metal speciation and toxicity to aquatic species (Chapter 4).

CHAPTER 2 – LITERATURE REVIEW

WATER QUALITY

Point source pollution is an easily identifiable source of pollution that can be traced to a specific location, typically associated with industrial waste and sewage treatment discharges that emanate from a pipe. Conversely, non-point source pollution is commonly associated with stormwater, where the pollution is derived from diffuse sources, such as urban and agricultural land use. In fact, agricultural discharges are the principal non-point source for water quality impacts on lakes and rivers (USEPA, 2005).

As landscapes urbanize, pollutant types and loadings are altered. Land use is a major contributing factor to the types and quantity of pollutants found in stormwater (NYSDEC, 2015; WSDE, 2019). Agricultural operations regularly discharge sediments, nitrogen, phosphorus, pathogens, metals, salts, pesticides, herbicides, and fungicides and are the leading contributor to water quality concerns in rivers (Munn et al., 2018; USEPA, 2005). Exposed soils from construction activities regularly produce sediment-laden water if not effectively managed. Stormwater discharges from stabilized residential lawns commonly contain nitrogen and phosphorus from fertilizer use and pet wastes, in addition to copper and zinc from roofing materials (Davis et al., 2001; Dietz, 2007; Radovanovic & Bean, 2022). Roadways accumulate oil and grease from automobiles, salts from winter treatments, and sediment from construction vehicles (WSDE, 2019). Roadways are also known sources of heavy metals including copper, zinc, lead, and cadmium from automobile brake pads and tire wear (Camponelli et al., 2010; Davis et al., 2001). Stormwater generated from industrial facilities may contain a variety of pollutants that are relatively dependent upon the nature of operations at the facility; however, much of the aforementioned pollutants can be expected. Meanwhile, under predevelopment conditions, these undeveloped lands do not have similar pollutant sources or loadings. Rather, these natural areas serve as a filter and prevent the introduction of such pollutants into runoff and downstream waters.

The term first flush often refers to rain events during the beginning of the wet season or the first few minutes of a storm when runoff exhibits high pollutant concentrations and warmer water temperatures (Roy-Poirier et al., 2010). The USEPA quantifies the first flush as approximately 1.3 mm (½ inch) of rainfall (USEPA, 2000). Many SMPs are designed with the goal of treating first flush storms to preserve the quality of downstream waters (Morzaria-Luna et al., 2004).

IMPACTS

Sediment-laden stormwater is primarily generated when stormwater runs across surfaces with exposed soil. As water slows, suspended sediment is deposited. This settling process is observed in drainage sumps and settling tanks or, on a larger scale, where deltas and sand bars are created as rivers empty into the ocean. High sediment loadings can impact the longevity of stormwater management practices as they can become clogged, and functionality is decreased. To prevent the need for excessive maintenance associated with sediment deposition, contributing drainage areas should be fully stabilized prior to SMPs being brough online. This problem is well recognized as pretreatment mechanisms are often required components design standards for water quality treatment practices (NYSDEC, 2015; WSDE, 2019). Typical pretreatment devices that promote sediment settling and prevent such impacts include vegetated filter strips, sumps, forebays, or manufactured devices. Lastly, suspended sediment in stormwater is a good indication that other pollutants will be present as soil particles often have pollutants adsorb to their surface (Roy-Poirier et al., 2010).

Nutrient loadings in runoff from agricultural fields are contingent on cropping patterns, fertilizer application rates, and topography (Huang et al., 2017). Just as nitrogen and phosphorus promote crop growth, the increased nutrient loadings can lead to algae blooms within waterbodies. When the excessive amount of algae dies, microbe populations breakdown the dead organisms and deplete the oxygens levels within a waterbody (Davis et al., 2006; Huang et al., 2017). This process is known as eutrophication and is a concern as can result in decreases in an ecosystem's biodiversity (Huang et al., 2017).

Aquatic communities are impacted by the pollutants collected in stormwater runoff and discharged untreated to receiving waterbodies. Species can be impacted by increased suspended sediment, as visibility and light penetration become reduced, and eutrophication and decreased dissolved oxygen supply from increased nutrient loadings. Bioavailable heavy metals can cause developmental impacts or mortality to aquatic species (Bui et al., 2016; Hoang et al., 2004; Hoang & Tong, 2015).

WATER QUANTITY

Urbanization changes the natural hydrology of a watershed as cover types, slopes, flow paths, and antecedent moisture conditions (AMC) are altered. Meandering watercourses, infiltrating soils, and vegetation are replaced with continuous hard surfaces and straight, smooth conveyance systems (USEPA, 2003). The reduction of vegetated surfaces and expansion of impervious areas increases runoff volumes and rates and reduces groundwater recharge as roofs and pavement cannot infiltrate stormwater (Dietz, 2007). Land use affects the volume of water a watershed produces as it relates to the amount of impervious cover and the level of soil disturbance. Commercial and industrial areas generally have more impervious cover than residential lawns (USDA NRCS, 1986).

Prior to the development of an area, the natural landscape retains a prescribed volume of water before discharge at an established rate for a particular rainfall intensity. The storage capacity of a watershed is impacted by vegetation type, soil type, and AMC (Detty & McGuire, 2010; Penna et al., 2011; Scaife et al., 2020; S. Wang et al., 2022). Factors that impact runoff rates include flow lengths, patterns (i.e., sheet flow vs. channelized flow), and slope (USDA NRCS, 1986). Appropriately sized watercourses are formed within natural watersheds and serve to safely carry the generated flows. Flooding is a natural event that occurs when the capacity of a waterway is exceeded during extreme precipitation events.

Not all vegetation equally intercepts and reduces stormwater. While replacing a forest with a parking lot will undoubtedly increase runoff volumes, it should be understood that modifying a forested area into a soccer field will also increase discharge volumes and rates, as forests have greater storage capabilities than manicured lawns (USDA NRCS, 1986). This disparity occurs as the leaves of trees and shrubs intercept raindrops prior to hitting the ground, slowing and absorbing water. Additionally, lawn areas are often disturbed and compacted during construction activities (Radovanovic & Bean, 2022; USDA NRCS, 1986). Further, the AMC of a soil impacts the generation of runoff, which likewise becomes affected by vegetative cover modifications (Detty & McGuire, 2010; Penna et al., 2011; Scaife et al., 2020). Forest floor litter is nature's mulch, serving to retain prevent erosion, lock in soil moisture, and prevent compaction of the underlying soils (USEPA, 1999).

IMPACTS

Urbanized areas often channel storm flows into sewer pipes, resulting in concentrated discharges and decreased time of concentrations, resulting in the primary concerns associated with impacts to water quantity: erosion and flooding (Freni et al., 2010; USEPA, 2003). Erosion is a natural process that is exacerbated by the impacts of urbanization, specifically increased runoff rates. A surface is considered stabilized if it prevents erosion from occurring as stormwater runs over it, such as paved, vegetated, and mulched areas. However, as stormwater moves across a non-stabilized landscape as sheet flow, surface erosion can occur, and within watercourses where channelized flows scour, sediment streambank erosion can occur (USEPA, 2003). The eroded sediments are then transported downstream, where they can impact aquatic life and landscapes.

As a watershed is developed, increased impervious surfaces and traditional conveyance systems (gutter and pipes) decrease the travel time of water to a particular design point (USDA NRCS, 1986). This travel time is commonly referred to as the time of concentration and influences peak discharge rates. While impervious surfaces are extremely efficient at preventing soil erosion, they also decrease the time of concentration of a watershed (Hood et al., 2007; USDA NRCS, 1986). Impervious surfaces also contribute to increased runoff volume due to a watershed's reduced groundwater recharge and storage capacity. Within urbanized areas, greater volumes of water reach waterbodies in less time, and downstream flooding is more likely to occur as the naturally formed watercourse cannot manage received flows (USEPA, 2000).

LOW IMPACT DEVELOPMENT

Prior to the last 25 years, stormwater was predominantly managed using gray infrastructure. Gray infrastructure is a general term that encompasses pipes, gutters, drains, and retention basins and serves to convey and store stormwater for the purpose of flood prevention and pollutant removal. Gray infrastructure tends to manage flows at the “end of pipe” meaning that all stormwater from a watershed is funneled to a single location, like a stormwater pond, where it is treated and detained (Freni et al., 2010; USEPA, 2000). More recently, the focus of stormwater management has shifted to low impact development techniques, commonly referred to as green infrastructure (Prince George’s County, 1999). Low impact development practices differ from gray infrastructure because they aim to manage stormwater near the source. The goal of LID is to manage stormwater in a way that preserves the pre-development hydrological characteristics of a watershed, including its ability to infiltrate, evaporate, filter, and detain stormwater (Dietz, 2007; Prince George’s County, 1999; USDA NRCS, 1986; USEPA, 2000).

As previously noted, urbanization often decreases the time of concentration within a watershed, allowing larger volumes of stormwater to reach receiving waterbodies sooner than it did prior to expansion. These changes can be offset by incorporating LID practices that slow stormwater through ponding and promote slope reduction (USDA NRCS, 1986). Low impact development techniques are often thought of as physical structures; however, site planning practices, such as the preservation of natural areas and the reduction of impervious areas, are equally influential (NYSDEC, 2015; WSDE, 2019). Natural areas that are often targeted for preservation include undisturbed forests, stream corridors, wetlands, and wetland buffers. During the site planning phase of a development project, the design of roadways, sidewalks, driveways, building footprints, and parking lots should identify areas where impervious areas can be reduced. Beyond site planning, physical LID practices may be incorporated into a project design in place of traditional development techniques, such as using permeable pavements in place of conventional impervious asphalt and concrete (Ball & Rankin, 2010). Alternatively, SMPs, such as bioretention cells and underground infiltration systems, can be interspersed throughout a project site (Johnson & Hunt, 2020).

Low impact development mitigates water quality and quantity impacts associated with evolution in urban and suburban land use. Federal and state programs often require incorporating LID practices to alleviate potential downstream impacts, including erosion, flooding, and increased pollutant loadings. LID practices are often evaluated by pollutant removal rates; however, when the water quality of influent stormwater is good, studies generally result in lower removal rates which can be misleading (Davis et al., 2006). It has been suggested that mass removal rates may be a more accurate indicator of practice performance and should be reported when possible (Roy-Poirier et al., 2010). Physical LID practices may be divided into four groups based on function: infiltration, evapotranspiration, filtration, and detention.

INFILTRATION

Infiltration practices aim to provide volume reduction through the capture and storage of stormwater. These systems provide storage capacity and prevent stormwater runoff from discharging downstream. Infiltration practices may be further divided into vegetated infiltration practices, which include stormwater planters, bioretention systems without underdrains, rain gardens, vegetated swales, or underground infiltration practices, such as porous pavement, dry wells, and prefabricated chambers. Both water quality and quantity benefits may be derived from infiltration practices.

EVAPOTRANSPIRATION

Evapotranspiration practices are vegetated surficial mechanisms designed with limited storage capabilities as they typically have an impermeable liner or barrier that prevents the effective infiltration of stormwater. The porosity of the soil media and selected vegetation are the key components of a successful ET practice. These practices are designed with overflow devices that allow excess stormwater to safely discharge to downstream locations. Common ET practices include green roofs, stormwater planters, and rain gardens. Like infiltration practices, ET practices mitigate water quality and quantity impacts, although to a lesser extent due to limited storage capacity.

FILTRATION

The primary function of a filtration practice is to treat stormwater prior to downstream discharge, although these practices may also provide limited rate and volume attenuation potential. These devices use filter media to remove sediment and pollutants. Filter media may contain sand, an engineered soil mixture, or a manufactured product. Sand filters, bioretention systems with underdrains, and manufactured filter devices are well-known filtration practices often in integrated into development designs.

DETENTION

Systems that provide temporary storage prior to discharge to downstream waters are defined as detention practices. These practices provide for management of peak discharges but do not readily reduce the volume of stormwater. Further, limited treatment of the stormwater may be achieved as the detention system allows for sediment and adsorbed pollutants to settle out of suspension prior to discharge. Detention systems include dry and wet basins, blue roofs, detention ponds, underground chambers, and constructed wetlands.

MITIGATIVE IMPACTS

Low impact development practices can be utilized to help achieve pre-development hydrological conditions within residential watersheds (Hood et al., 2007; Tirpak et al., 2021). Low impact development strategies and practices are evaluated by how well they function to mitigate impacts of urbanization. The functionality of a water quality practice is often evaluated based on its ability to reduce pollutant loadings and concentrations in effluent (Davis et al., 2006). Water quantity controls are assessed based on discharge volume and rate reduction. Many LID practices function to provide mitigation to both water quality and quantity concerns. Detention and retention practices mitigate impacts from increased stormwater volume and rate (Freni et al., 2010). Detention practices serve to temporarily store and subsequently discharge stormwater. Retention practices, including infiltration and ET practices, function to reduce discharges of stormwater to receiving waterbodies. Generally, LID practices that retain stormwater, such as porous pavements, bioretention systems, and underground infiltration systems, are more environmentally beneficial as they provide mitigation for both water quality and quantity concerns (Ball & Rankin, 2010; Hunt et al., 2006).

LID practices are commonly evaluated by their ability to mitigate peak flow by either increasing the lag time (delay between peak precipitation and peak discharge) or through reducing peak flow rate. Hood et al. (2007) found that a watershed that incorporated best management practices (BMPs), including bioretention areas, grassed swales, and eliminated standard curb and gutter stormwater collection systems exhibited significantly greater lag times than traditional watersheds. Tirpak et al. (2021) concluded that, in general, permeable pavers were capable of increasing several measures of lag, including centroid lag-to-peak (the time from the centroid of precipitation to the peak discharge), centroid lag (the time from the centroid of precipitation to the centroid of discharge), and peak lag-to-peak (the time from the peak rainfall intensity to the peak discharge). Peak flow mitigation from permeable pavement systems was further reported by Winston et al. (2018). Bioretention systems have also been shown to mitigate peak discharges within urban environments (Davis, 2008; Winston et al., 2016).

Storm size, storm intensity, and storm duration impact the extent to which lag time is affected (Davis, 2008; Hood et al., 2007; Tirpak et al., 2021). Hood et al. (2007) noted that within the LID watershed with smaller, short-duration storms resulted in significantly larger lag times. Consistently, Tirpak et al. (2021) observed decreased benefits of permeable pavement systems during larger, more intense storms preceded with lower antecedent dry periods (ADP). Additional factors that influence the effectiveness of LID practices at mitigating peak discharges include loading ratio (ratio of contributing area to practice area), internal water storage (IWS), and exfiltration rates of the underlying soils (Tirpak et al., 2021; Winston et al., 2018). The success of a bioretention system is typically affected by its design parameters, including ponding storage, surface infiltration rate, IWS capacity, and the exfiltration rate of the underlying soil (Brown & Hunt, 2011; Davis, 2008; Winston et al., 2016). Environmental factors like AMC also affect stormwater discharges where LID practices often perform best when AMCs are low (Davis, 2008; Hood et al., 2007; Tirpak et al., 2021). To further increase lag times, Tirpak et al. (2021) suggested using an orifice to regulate discharges or creating an IWS.

Stormwater management practices implemented at the source can be smaller and promote local infiltration; however, they often require increased maintenance demand (Freni et al., 2010). Alternatively, centralized SMPs installed at the “end-of-pipe” require a larger area for operation and have cheaper maintenance costs (Freni et al., 2010). Stormwater management designs that incorporate a mixture of at-source and centralized practices provide for a compromise between

SMP effectiveness and space constraints. Detention practices are effective at decreasing peak discharges by increasing the duration of a storm; however, detention basins do not reduce storm volume (Ferguson, 1995). Volume controls must be implemented to prevent flooding conditions at downstream obstructions, as infiltration provides a “more complete and reliable” solution (Ferguson & Deak, 1994).

Infiltration practices simultaneously reduce stormwater volume, peak discharges, and pollutants (Ferguson, 1995). Infiltration practices reduce stormwater flows to a drainage system (Freni et al., 2010). Low impact development techniques have been shown to significantly reduce runoff volume in a watershed, attributed to decreased impervious cover and increased ability for infiltration, regardless of storm sizes, storm size and storm durations, and AMC (Hood et al., 2007). Tirpak et al. (2021) found a 78-89% volume reduction during smaller storms with greater ADP and a 43-45% volume reduction during larger storms with less ADP. The reductions were attributed to infiltration into underlying soils and, to a lesser extent, evaporation. Bioretention systems can effectively eliminate flow from small storms, ultimately reducing pollutant loadings to zero (Davis et al., 2006).

RUNOFF GENERATION

Rainfall-runoff relationships have been long studied with the goal of developing methods that accurately quantify runoff given a particular rainfall event. Several factors impact when a surface may produce runoff, and the development of a one-size-fits-all model is highly unlikely. As stated above, factors often evaluated when determining runoff probability include cover type, slope, rainfall depth, rainfall intensity, and AMC. It is also difficult to apply prediction approaches across small-scale catchments and large-scale watersheds. Rainfall characteristics and AMCs are well-known parameters that influence runoff generation from vegetated catchments (Kohler & Kinsley, 1951). Research often represents the soil water storage of a watershed with an index commonly referred to as the Antecedent Soil Moisture Index (ASI).

Forested areas are very effective at reducing runoff as the tree leaves and forest litter can intercept and store precipitation (USDA NRCS, 1986). In fact, many studies of forested watersheds quantify stormflow from storm events as a combination of overland surface flow and shallow subsurface flow, as opposed to overland runoff alone (Detty & McGuire, 2010; Mosley, 1979; Penna et al., 2011; Scaife et al., 2020; S. Wang et al., 2022). Penna et al. (2011) found that only after exceedance of the soil moisture threshold did the forested hillslopes within a headwater catchment in Italy (the Alps) produce runoff. Detty & McGuire (2010) found a significant linear correlation between stormflow and ASI and precipitation above a particular precipitation threshold. Scaife et al. (2020) found that rainfall intensity did not have a significant impact in humid forested habitats on runoff production and attributed their findings to the highly saturated hydraulic conductivity of the soils, which in turn enables rapid infiltration. Rapid infiltration enables stormwater to discharge to a stream through subsurface means and promotes groundwater recharge. Mosley (1979) found that despite the hydraulic conductivity of the soil, subsurface flow reached the stream through macropores within the soil structure. However, Scaife et al. (2020) determined that ASI and groundwater levels were good indicators of the dominant runoff processes of a watershed (overland flow, sub-surface flow, deeper flow). Saffarpour et al. (2016) determined that the hydrological response of an agricultural catchment was primarily influenced by the ASI and rainfall characteristics. Nonetheless, little to no relationship between runoff and rainfall depth or rainfall intensity was observed when considered alone. Within twelve vegetated catchments, Wang et al. (2022) observed positive relationships between ASI and runoff, although there was no indication of a threshold responses. Despite the absence of an ASI threshold, Wang et al. (2022) did observe that runoff increased linearly after a defined precipitation threshold and that the linear relationship was stronger when P + ASI were cumulatively considered.

Saffarpour et al. (2016) also noted that the seasonal variation of the ASI resulted in seasonal changes to runoff response. Consistently, Penna et al. (2011) observed that during dry conditions, below the soil moisture threshold, precipitation events produced very little runoff and under wet conditions, when the soil moisture threshold was exceeded greater runoff was produced (Mosley, 1979).

Conversely, in less vegetated, urbanized catchments, stormwater runoff is more closely affected by the percentage of impervious cover. The goal of LID practices is to mimic pre-development watershed hydrology similar to that observed from forested watersheds. Environmental factors that impact the performance of LID practices include soil moisture, soil types, soil depth, rock fragments, root count, bedrock outcrops, and season. The greater interaction between runoff and vegetation and soil, the greater benefit may be derived as it supports hydrological processes such as groundwater recharge and evapotranspiration. Low impact development techniques promote increased vegetated cover, increasing the watersheds ability to infiltrate, and thus reduce runoff volumes over traditional development (Hood et al., 2007).

The effectiveness of an LID practice is often evaluated using runoff thresholds. Runoff occurs when the ability for a watershed to intercept, infiltrate, and store precipitation has been exceeded. Many researchers look to determine a runoff threshold that represents the precipitation depth that must occur within a catchment before surficial runoff generation that incorporates the aforementioned factors (Ali et al., 2013). Rainfall characteristics and AMC are well known parameters that influence runoff generation from vegetated catchments, although many methods do not account for frozen condition or snow pack (Kohler & Kinsley, 1951).

There are a variety of methods for determining thresholds. Many studies conduct a basic linear regression analysis, where the precipitation depth is plotted on a scatterplot versus the measured runoff from a storm. The x-intercept of the line of best fit, generated using the least squares method, is then interpreted as the runoff threshold. This method of linear regression has long been used to determine runoff thresholds (Fink and Frasier, 1977; Li and Gong, 2002; Hood et al., 2007; Bean et al., 2007; Ross et al., 2021; Tirpak et al., 2021). Using this methodology, Hood et al. (2007) compared a traditionally developed watershed and an LID watershed and concluded that the traditional watershed had a lower threshold than the LID watershed. The linear regression method was utilized by Bean et al. (2007) to estimate the storage depths, and, ultimately, curve numbers of permeable pavement systems. Tirpak et al. (2021) also established the runoff threshold of permeable interlocking concrete pavers using the linear regression and x-intercept methodology. Detty & McGuire (2010) and Saffarpour et al. (2016) utilized the linear regression technique to evaluate thresholds associated with ASI, runoff, and ASI + runoff. Saffarpour et al. (2016) acknowledged that rainfall depth and intensity alone did not adequately describe runoff generation behavior.

Runoff thresholds are commonly more complex and are dictated by catchment-specific variables. Other relationships explored by the literature include piecewise linear, step or Heaviside, Dirac, and sigmoid functions (Ali et al., 2013; S. Wang et al., 2022). Ali et al., (2013) reviewed much literature which describes non-linear relationships among hydrological inputs and responses. Piecewise linear regression has been commonly implemented to describe the relationship between precipitation and runoff. Scaife et al. (2020) modeled runoff against ASI + runoff using a piecewise regression analysis, also known as a hockey stick model, to determine the runoff threshold. Ross et al. (2021) also recognized that the relationship between precipitation and runoff could have been more consistently described using a piecewise linear regression to identify a break in the relationship between precipitation and runoff within various watersheds.

The runoff threshold is often determined for specific stormwater management practices. For comparison of effectiveness, Winston et al. (2016) calculated the runoff thresholds of bioretention systems overlying silty clay loam and silt loam. Davis (2008) also assessed the division between storm events that produced runoff and those that did not for the purpose of estimating a threshold rainfall intensity of the bioretention system.

AQUATIC TOXICITY FROM HEAVY METALS

The total measured concentration of a metal does not adequately represent the bioavailability of a metal for uptake by an aquatic species. Rather, a measure of the dissolved contents is a better approximation, as only a dissolved metal can pass through the gill and produce toxic effects (Kinerson et al., 1996). The USEPA standards loosely define dissolved metals to be any metal that can pass through a 0.45-micron filter. This definition may still overestimate the bioavailable metal concentration as any particulate metal less than 0.45 microns is considered dissolved. A better means for determining the bioavailability of metals is through consideration of metal speciation, often achieved through the use of a chemical speciation model, discussed further below. This is recognized by the USEPA, which implements a BLM that incorporates water quality parameters to determine speciation and bioavailability of metals, to determine freshwater aquatic life copper criteria (USEPA, 2007).

Toxicity is defined as the accumulation of a metal on a receptor at or greater than a threshold concentration such that a pre-defined lethal or sublethal effect will occur (Santore et al., 2001). Heavy metals can be detrimental to aquatic communities as they can have lethal and sublethal effects (Bui et al., 2016; Hoang et al., 2004; Hoang & Tong, 2015). The extent to which an organism may be affected is dependent on the characteristics of the organism, the pollutant of concern, exposure time and concentration, and the environment, as these factors contribute to the sensitivity of an aquatic species and the bioavailability of heavy metals to accumulate on a biotic ligand, a surface where metal may accumulate and result in toxic effects.

The chemical species of a metal effects bioavailability (USEPA, 2007). Free ions of zinc and copper are considered the most influential on the toxicity (Hoang & Tong, 2015; Magnuson et al., 1979). Nonetheless, they are not the only toxic metal species; in fact, Bui et al. (2016) found that copper hydroxide and carbonate may also contribute to toxicity, and Hoang & Tong (2015) reported aqueous zinc carbonate and zinc bicarbonate are likely bioavailable as well. Metal speciation is affected by water quality characteristics, including temperature, pH, hardness, and dissolved organic carbon (DOC), which in turn affect bioavailability and the overall toxicity of water (Di Toro et al., 2001; Pagenkopf, 1983; Santore et al., 2001). The effects of DOC, hardness, and pH on chemical speciation and aquatic toxicity are well-known to produce a protective effect on aquatic species from heavy metals (Birceanu et al., 2008; Bui et al., 2016; Gillis et al., 2008; Hoang et al., 2004; Hoang & Tong, 2015).

SPECIATION AND TOXICITY MODELS

Using laboratory data, computer models have been developed with the goal of predicting toxicity effects on aquatic species. Several chemical speciation models have been developed and are commonly implemented in research. Visual MINTEQ 3.1 (Gustafsson, 2020) is a chemical equilibrium model that, among other functions, may be used to predict the chemical speciation of inorganic ions and complexes within water samples. Visual MINTEQ also accounts for the effects of dissolved or precipitating solid phases based on water quality parameters. The WHAM (Windermere Humic Aqueous Model) utilizes the assumption that binding at an organism's biotic ligand can be estimated by how metals bind to humic substances. More recently, the WHAM-FTOX and WHAM-FTOXβ models were developed to evaluate the toxic effect of multiple metals on aquatic species (Tipping et al., 2021, 2023). These models assume protons and metal cations are additively related to their occupancies at a biotic ligand. Another program used to determine the bioavailability of metals is bio-met. This tool estimates the bioavailability of metals based on water quality parameters and determines the results from compiled BLMs. A common BLM is provided by Windward Environmental, LLC (v 3.41.2.45). This model utilizes water quality parameters to estimate the speciation of heavy metals and may further provide the concentration at which a metal species is expected to cause a predefined toxic effect (i.e., EC20, LC50, etc.) to a given aquatic species' life stage. Sierra et al. (2017) compared the speciation predictions of Visual MINTEQ, WHAM 7, and bio-met with field-measured conditions and found that the model-based chemical speciation values were slightly higher. Field-measured results for Zn and Cu speciation were most closely related to that predicted by bio-met, yet WHAM7 was effective at determining the speciation of Cu when DOM was incorporated. For Zn, Ni, Cd, and Hg, Sierra et al. (2017) found that WHAM7 and Visual MINTEQ were comparable.

WATER QUALITY IMPACTS

Generally, hardness affects toxicity as calcium ions (Ca^{2+}) and magnesium ions (Mg^{2+}) compete with the heavy metal ions (Cd^{2+}, Cu^{2+}, Pb^{2+}, Zn^{2+}) for binding sites at the biotic ligand resulting in a reduction to toxicity. Hoang & Tong (2015) documented this mitigative property of hardness on zinc toxicity to *Pomacea paludosa*, a freshwater snail. However, Rogevich et al. (2008) described no effect from hardness on copper toxicity to *Pomacea paludosa*. Freshwater mussels also benefitted from moderately hard water, as a decrease in copper sensitivity was observed as hardness increased (Gillis et al., 2008). Positive correlation among hardness and the LC50 of three Vietnamese aquatic species was also reported by Bui et al., (2016). Protective effects of hardness is also documented for nickel on fathead minnows (Hoang et al., 2004) and cadmium in fish species (Kumar & Singh, 2010).

Complexation with DOC also decreases the bioavailability of free metal ions as bound metals are not available for uptake. Hoang & Tong (2015) found the affinity for zinc ions to bind with DOC increased as pH increased, thus decreasing zinc toxicity to *Pomacea paludosa*. Rogevich et al. (2008) has similar findings in that copper toxicity to *Pomacea paludosa* decreased as DOC increased. Freshwater mussels exposed to copper at low concentrations of DOC demonstrated a two-fold increase in EC50 levels and up to a ten-fold increase at higher DOC concentrations (Gillis et al., 2008). Nickel toxicity to fathead minnows exhibited the protective effect of DOC in waters with high alkalinity, hardness, and pH (Hoang et al., 2004).

Generally, studies have found that pH and metal toxicity are inversely related, whereas when pH increases, the toxicity from heavy metals decreases. Hoang & Tong (2015) reported that the concentration of Zn^{2+}, the most bioavailable zinc species, decreased as the pH increase. The findings of X. F. Li et al. (2019) were in agreement, yet found no impacts to Zn toxicity when pH ranged from 8-11. Further, using Visual MINTEQ software, Rogevich et al. (2008) determined that when pH exceeded 7, concentrations of Cu^{2+} in solution decreased. Consistently, Bui et al., (2016) observed increased copper toxicity to aquatic species resulting from low pH. Hoang et al.

(2004) reported the influence of pH on nickel toxicity was dependent on the hardness and alkalinity of the water; where at low hardness and alkalinity, pH displayed no effect on toxicity, yet at higher hardness and alkalinity level the lethal concentration for 50% of the population (LC50) increased with pH.

In addition to water quality parameters, the age of an organism also affects toxicity, where adult species are less susceptible to heavy metals than juveniles. This has been observed in fathead minnows exposed to nickel (Hoang et al., 2004) and in freshwater snails subjected to copper (Rogevich et al., 2008).

Although recent models have been developed (Tipping et al., 2023), less studied is the interaction and competition among multiple metals at the biotic ligand. Birceanu et al. (2008) found that Pb and Cd in combination resulted in a less than additive binding at the biotic ligand in rainbow trout and attributed their findings to the ability of Cd to out compete Pb for binding at the biotic ligand. Nonetheless, they found the overall effects from the metal mixture resulted in ionic disturbances that produced a more than additive toxic effect on the fish. It has been suggested that the BLM should be reevaluated to include assumptions for metal mixtures.

BIORETENTION SYSTEMS

Bioretention systems are a vegetated LID practice that function to retain, filter, and detain stormwater, well-suited for commercial and residential areas within urban environments (USEPA, 1999). Bioretention systems provide at- or near-source treatment and are well adapted for urban areas as they have continually demonstrated efficient removal of common stormwater pollutants and mitigation of peak discharges (Brown & Hunt, 2011; Davis et al., 2003, 2006; Glass & Bissouma, 2005; Johnson & Hunt, 2020, 2020b; H. Li & Davis, 2008; Roy-Poirier et al., 2010;

Sun & Davis, 2007; Winston et al., 2016). Bioretention is a term that is often used to collectively describe rain gardens, vegetated swales, infiltration basins, and more. Stormwater runoff is directed to a bioretention system either via overland sheet flow or discharged through an outfall with the employment of dissipation devices. Bioretention systems are comprised of a ponding zone, mulch layer, vegetation, bioretention soil mixture (BSM), drainage layer, in some cases, an underdrain when underlying soils are sufficiently permeable (USEPA, 2000). Each layer serves a particular function in the treatment of stormwater. Stormwater may exit the system through several outlets: infiltration, evapotranspiration, underdrain, or overflow bypass (Davis et al., 2006; Johnson & Hunt, 2020). Bioretention systems may be designed using volume-based methodologies, where the system is designed to manage frequent small storm events as a function of Darcy's Law, area-based, where the SMP is sized based on the contributing impervious area, or other methods based on the peak runoff mitigation, anticipated loadings, or computer modeling (Roy-Poirier et al., 2010).

BIORETENTION SYSTEM COMPONENTS

PONDING ZONE

Bioretention cells are designed with a concave surface that provides for the ponding of stormwater prior to percolation through the BSM. Following the exceedance of the ponding zone, stormwater may bypass the natural filter and enter a downstream conveyance system. The ponding zone allows for a calculated volume of stormwater to be stored and treated by the practice prior to bypass and typically ranges from 15-30 cm (NYSDEC, 2022; USEPA, 1999; WSDE, 2019). The ponding zone also functions to mitigate peak discharges from urban areas (Davis, 2008; Johnson & Hunt, 2020; Winston et al., 2016). When water levels exceed the design depth for ponding, water is diverted away from the practice via an overflow pipe, manhole, or other mechanism for relief, and diverted flows will discharge directly to receiving waterbodies. Therefore, this zone is important because it slows water and allows sediment to fall out of suspension prior to bypass. Ponded water is typically designed to drain within a prescribed period to prevent issues associated with long-standing water, such as mosquito production.

MULCH LAYER

The mulch layer is one of the most important components in a bioretention system as it provides a multitude of functions. First, mulch protects the underlying soil as it serves to prevent erosion, lock in soil moisture, and prevent BSM compaction. Maintaining soil moisture prevents drying of underlying soils and promotes vegetation health. Preventing compaction of the BSM surface preserves the soil's permeability, a key element to the bioretention system's success. Next, the mulch layer is the first component with filtration capabilities. Following analysis of soil core samples, Davis et al. (2003) attributed the high heavy metal removal efficiencies observed to the mulch layer. The mulch layer has also been credited with significant removal of total Kjeldahl nitrogen (TKN) (Davis et al., 2006). Additionally, petroleum-based products are regularly trapped and broken down by microorganisms found within the mulch layer (USEPA, 1999). When installing the mulch, it should be applied uniformly and should consist of shredded hardwood that has aged for at least six months and, to promote soil and air interaction, should not exceed 8 cm in depth (NYSDEC, 2022; USEPA, 1999; WSDE, 2019). Maintenance or replacement of the mulch layer is required to remove accumulated metals and continued performance (Glass & Bissouma, 2005).

BIORETENTION SOIL MIXTURE

Bioretention soil mixtures are comprised of a combination of sand and organic material and may include a small silt or clay component and serve as the main filtering mechanism. As stormwater percolates through the BSM, stormwater pollutants may be removed as particulate matter (organics and suspended solids) is physically trapped within soil pores or as pollutants are adsorbed to soil particles (USEPA, 1999). Li & Davis (2008) reported the upper 15 cm of the mulch and soil profile is where most of the removal occurred. The selection of the BSM should consider particle size as it influences retention time and pollutant removal rates (Brown & Hunt, 2011). Further, the BSM should be designed to provide for adequate contact time between stormwater and soil particles as it promotes the removal of heavy metals, phosphorus, and hydrocarbons via adsorption (USEPA, 1999). As runoff infiltrates the BSM, adsorption to the finer soil particles, such as silt and clay, is the primary mechanism for phosphorus removal (Davis et al., 2006). Soil texture also impacts heavy metal removal rates, with a BSM consisting of sand, topsoil, and compost exhibiting a lower removal rate when compared to BSMs with higher percentages of fines (silt and clay) (Davis et al., 2003). Although finer soil particles promote pollutant adsorption, they also contribute to poor infiltration capabilities and increased bypass of untreated stormwater (Johnson & Hunt, 2020). Lastly, integrated organic material can affect pH and electrical conductivity of the BSM and can export DOC, nutrients, and copper (Mullane et al., 2015).

UNDERDRAIN

Bioretention cells are commonly equipped with a perforated underdrain pipe, typically located within a stone drainage layer. When storage of the drainage layer is exceeded, stormwater may exit the system through the underdrain. A traditional underdrain prevents long-term water storage and the occurrence of anaerobic conditions within the bioretention system limiting their ability to remove nitrogen as nitrogen gas. Alternatively, some systems are designed with a perched underdrain or an upturned elbow or lack an underdrain completely to allow for IWS (Brown & Hunt, 2011; Davis, 2008; Winston et al., 2016). Incorporating IWS into the bioretention system allows for a reservoir to accumulate filtered stormwater, producing a saturated anaerobic zone that can encourage denitrification (USEPA, 1999). However, the upper limits of the IWS should be sited far enough below the surface to sustain an aerobic zone of 0.3-0.45 m to ensure pollutant trapping (Brown & Hunt, 2011) and prevent freezing in northern climates (Brown & Hunt, 2011). In addition to water quality benefits, the IWS area also promotes volume reduction and pollutant mass removal through infiltration into underlying soils, especially when in situ soils have low permeability (Brown & Hunt, 2011; Davis et al., 2006).

VEGETATION

Vegetation is a key component to the long-term success of a bioretention system as it affects soil structure and porosity and may uptake pollutants from the soil (Davis et al., 2006). Trees and shrubs more effectively uptake and trap pollutants than herbaceous species as they do not die every season (USEPA, 1999). Vegetation also functions to remove water from the system through evapotranspiration, although Dietz & Clausen (2005) found this process to account for only 0.4% of volume reduction and Brown & Hunt (2011) up to 5%. A properly designed planting plan can produce a low maintenance, aesthetically pleasing landscaped feature. Additional benefits of vegetation beyond stormwater treatment, attenuation, and aesthetics, include the provision of soil stabilization to prevent erosions, shade, and noise and wind barriers (Roy-Poirier et al., 2010; USEPA, 1999).

WATER QUALITY IMPACTS

Bioretention systems improve water quality through physical, chemical, and biological processes. Physical processes include flow rate reduction, sedimentation, and filtering. Chemical processes include adsorption and volatilization. Common biological processes are microbial activity, decomposition, and plant uptake (Davis et al., 2003; USEPA, 1999). Mass removal of pollutants is commonly attributed to the reduction of outflow from bioretention systems as outflow concentrations are often observed higher than inflow (Davis et al., 2006; Hunt et al., 2006). While retained stormwater eliminates flow and pollutant loadings, filtered flows are capable of alleviating peak discharge rates and reducing pollutant loadings.

NUTRIENT REMOVAL

Bioretention systems are typically less efficient at removing nutrients than heavy metals (Davis et al., 2003, 2006). Nonetheless, effluent concentrations have been found to be removed to standards that are representative of fair benthic health where tolerant benthic species, such as crayfish and crustaceans, may be present (Brown & Hunt, 2011). Generally, nutrient removal through bioretention systems depends upon initial input concentrations, where higher input concentrations result in higher output concentrations (Brown & Hunt, 2011; Davis et al., 2006; Glass & Bissouma, 2005). After observing these phenomena, Davis et al. (2006) suggested that a lower limit input threshold concentration may exist for certain removal efficiencies to be observed and, further, that very low effluent concentration may be unattainable.

Soil texture affects nutrient removal in bioretention cells (Davis et al., 2006). Brown & Hunt (2011) reported higher nutrient removal rates for sandy clay loam than sandier soils and attributed this difference to the greater hydraulic retention time, which permitted increased interaction time between pollutants and negatively charged clay soil particles. For example, Davis et al. (2006) noted total phosphorus removal increased with depths up to 77-87% within bioretention box studies with limited impacts from storm duration or intensity.

The literature provides mixed results regarding the removal of nitrogen compounds (Knappenberger et al., 2022). Dietz & Clausen (2005) investigated roof runoff treatment through a rain garden and found no significant reductions to nitrate-nitrogen, TKN, or organic nitrogen although the gardens did significantly reduce ammonia-N (NH3-N). However, nitrates do not adsorb to soil particles, so nitrate removal through a BSM is unlikely (Davis et al., 2006). Despite this, Hunt et al. (2006) found that bioretention systems reduced nitrate-nitrogen by 13-75%. Davis et al. (2006) reported TKN removal rates ranging from 74-83% in laboratory box studies and 52-67% in field studies. A significant portion of these removals occurred within the first few inches of the bioretention system and was attributed to the mulch layer. Knappenberger et al. (2022) suggested the compost and bioretention age were contributors to nitrogen and phosphorus export, where less N-N and TP removal occurred as the percent of compost increased and greater removal occurred as the systems aged. Davis also concluded that the TKN removal rates were impacted by stormwater flow rates, where higher rates resulted in lower removal efficiencies. Lastly, systems designed with a conventional underdrain demonstrated mass removal rates of TKN varying from -5 to 45% (Hunt et al., 2006).

Although nitrification and increased nitrate concentrations have been reported from bioretention cells, a significant reduction of total nitrogen (TN) has also been reported in the literature (Davis et al., 2006; Dietz & Clausen, 2005). Systems designed with conventional underdrain systems exhibited mass removal rates of 40% TN (Hunt et al., 2006), while other field studies resulted in the removal of 49-59% TN (Davis et al., 2006). Systems designed to facilitate an anaerobic zone can promote denitrification (USEPA, 1999). However, in a field study in NC conducted by Hunt et al. (2006) to investigate the effect of IWS on the removal rate of nitrogen, no significant difference from the provision of an anaerobic zone was observed.

HEAVY METAL REMOVAL

It is well-documented that bioretention systems are known to be extremely effective at removing heavy metals from stormwater and unlike nutrient removal, many studies have found relative consistency among removal rates during laboratory and field evaluations. (Davis et al., 2003; Hunt et al., 2006; H. Li & Davis, 2008; Sun & Davis, 2007). Davis et al. (2003) studied bioretention systems in both the laboratory and field and found bioretention boxes filled with 2.5 cm of mulch over sandy loam substrate exhibited mass removal rates between 98-99% for copper, lead, and zinc. However, the study's field sites produce more variable results with Cu mass removal rates ranging from 43-97%, Pb removal from 70-95%, and Zn from 64-95% (Davis et al., 2003). The variability among the field-documented removal rates was attributed to the composition of the BSM, where higher rates were reported from the BSM consisting of a higher percentage of soil and fines (silt and clay). In a field study in North Carolina, USA, conducted by Hunt et al. (2006) using organic sandy soil, metal mass removal rates were also close to the laboratory tests conducted by Davis et al. (2003) with 99% Cu removal, 81% Pb removal and 98% zinc removal. Glass & Bissouma (2005) found moderate removal rates from a BSM comprised of 50% sand, 20% organic material, 10% clay, and 20% mulch with reductions of 81% of Cu, 75% of Pb, and 79% of Zn, in addition to 66% Cd removal, 53% Cr removal, 53% Fe removal, 17% Al removal, and 11% As removal. This system could reduce influent concentrations of Pb and Cd, exceeding the EPA drinking water standard to levels below the maximum allowable level.

Several factors, beyond BSM components, may further affect pollutant removal. Similar to nutrient loadings, greater removal efficiencies of zinc, copper, lead, and cadmium have been observed with higher metal loadings within bioretention influent (Sun & Davis, 2007). Further storm intensity and duration also impact removal rates as less intense and longer duration storms slightly improve system performance, yet no significant effects were detected from longer duration storms (Davis et al., 2003). However, as BRCs are designed to bypass flows when a certain level of head is produced, the mass removal rates of the heavy metals is effectively decreased during high-intensity storm flows (Davis et al., 2003). Metals in the dissolved form are more bioavailable for plant and animal uptake than those affixed to particulate matter. Sun & Davis (2007) reported that 0.5 - 3.3% of the heavy metals removed from the system were due to accumulations in grasses, with more bioaccumulation occurring within the roots than the shoots. Davis et al. (2003) recognized that a significant percentage of removal was encountered during the upper portion of the systems profile (i.e., mulch layer and soil media), possibly owing to the filtering of particulate matter within these upper layers. Knappenberger et al. (2022) found no effect on metal removal due to BSM composition or system age.

POLLUTANT LEACHING

Although bioretention systems have the ability to effectively filter stormwater, leaching of DOC, base cations (Ca, Mg, and Na), heavy metals, and nutrients has also been reported (Chahal et al., 2016; Knappenberger et al., 2022; H. Li & Davis, 2008; Mullane et al., 2015). Backfilled soil can be a source of pollutants, and successive storms can mobilize previously deposited pollutants (Hunt et al., 2006). While leachate concentrations have been found to decrease over the duration of an individual storm and following subsequent storms, persistent leaching of DOC, nitrogen, phosphorus, and copper has been observed, although concentrations plateau following the first few storms (Mullane et al., 2015). Increased total suspended solid (TSS) concentrations are often faulted for the increases in the mass of TP in bioretention effluent (Dietz & Clausen, 2005; Hunt et al., 2006). Discharged organic matter in bioretention outflow can result in the depletion of oxygen in downstream waters due to bacteria-led decay processes (Roy-Poirier et al., 2010). Copper leaching is a common consequence of integrating compost, and like other stormwater pollutants, copper concentrations have the tendency to decrease both throughout the duration of a particular storm and following subsequent storms (H. Li & Davis, 2008; Mullane et al., 2015). More specifically, Mullane et al. (2015) found that the majority of leached copper was dissolved and discharged during the first 5 to 6 storms. Of concern, both Mullane et al. (2015) and Chahal et al. (2016) detected leachate concentrations of copper greater than the State of Washington's effluent benchmark concentrations, highlighting the importance of investigating the source of selected backfill materials and understanding the bioavailability of leached pollutants.

WATER QUANTITY IMPACTS

Bioretention systems perform best during small, more frequent events (Johnson & Hunt, 2020; USEPA, 2000) and commonly target effluent volume and peak discharge reduction rates of 33% (Brown & Hunt, 2011; Davis, 2008). Unfortunately, the bioretention systems studied by Davis (2008) exhibited only a 30 to 42% chance of meeting this target. Runoff reduction may be achieved via exfiltration into the in-situ soil or through evapotranspiration. Both processes provide stormwater quality and quantity benefits to downstream water through the mitigation of peak discharges and reduction of first flush storms. Properly designed bioretention cells are capable of providing major reductions in stormwater volumes and discharge rates for 33-50% of storms, even with a liner (Davis, 2008). Larger storms that occur over a long period of time are more likely to be mitigated as precipitation intensity approaches the infiltration capability of the system (Johnson & Hunt, 2020).

The effectiveness of a bioretention system at reducing the volume and rate of stormwater is most notably impacted by the ponding storage and surface infiltration rate, IWS capacity, and the exfiltration rate of the underlying soil; however, AMCs also affect stormwater discharges (Brown & Hunt, 2011; Davis, 2008; Winston et al., 2016). When the ponding depth and infiltration capacity of the surface layer are exceeded, bypass of untreated stormwater occurs. Johnson & Hunt (2020) found that 75-85% of stormwater inflow did not bypass the bioretention systems, and Davis (2008) determined that overflow occurred during approximately 15% of the studied storms. Internal water storage areas allow stormwater to exfiltrate from the bioretention system in between storm events and is primarily important when systems are sited over low-permeability soils (Winston et al., 2016). Therefore, it may be expected that bioretention systems designed with IWS are capable of reducing peak flows to a greater extent than conventional underdrain systems (Davis, 2008) and that larger IWSs produce greater volume reductions (Brown & Hunt, 2011). Exfiltration is a primary outlet for bioretention systems. The degree to which exfiltration provides hydrological benefits is dependent upon the permeability of the underlying soil. BRCs overlying silty clay loam and silt loam have been found to exhibit average exfiltration rates ranging of 1.7 mm/h and 4.3 mm/h, respectively (Winston et al., 2016). Conversely, exfiltration from a bioretention system above sand demonstrated rates ranging from 60–90 mm/h (Brown & Hunt, 2011).

Reducing stormwater runoff is one of the primary goals of LID practices as volume reduction improves downstream water quality and peak discharge rates (Davis, 2008; Hunt et al., 2006). Bioretention systems are often designed to manage 1.3 mm (½ inch) of runoff from the area tributary to the practice, resulting in full attenuation of small precipitation events (Davis, 2008; Winston et al., 2016). While all storms are not able to be fully reduced, Winston et al. (2016) reported volume reduction rates ranging from 36-59% and Johnson & Hunt (2020) achieved 64-90% reduction. Johnson & Hunt (2020) also concluded that the age of the systems did not prevent achievement of the 33% volume reduction goal.

The characteristics of the underlying soils impact the volume retention capabilities of a bioretention system. Nonetheless, bioretention cells of varying permeability were all found to significantly reduced the runoff volume, with a majority of the stormwater exiting via exfiltration and ET (Johnson & Hunt, 2020). While sandy clay loam is more effective at treating stormwater, sandy soil provides for greater volume reduction benefits, approaching 100% (Brown & Hunt, 2011). Sandy clay loam exhibited volume retention rates between 75-87% (Brown & Hunt, 2011). Systems over silty clay loam and silt loam can still achieve 36-59% volume reduction (Winston et al., 2016). Accordingly, high volume-to-discharge ratios (volume out/volume in) are observed from systems sited over clayey soils (Johnson & Hunt, 2020). Yet, even overlying clayey soil volume reductions up to 50% are possible, highlighting the value of ET and exfiltration in the success of these systems (Hunt et al., 2006). Further, flow reduction may vary based on season, increasing in the summer and decreasing in the winter (Hunt et al., 2006).

Bioretention systems provide peak flow reduction and increased lag time between rainfall and runoff peaks, equally important water quantity goals of LID practices (Dietz & Clausen, 2005). However, peak flow mitigation decreases as rainfall intensity increases, especially if ponding is exceeded and bypass occurs (Winston et al., 2016). Contributing factors to outflow duration include ponding storage, IWS, and exfiltration rate (Winston et al., 2016). Davis (2008) found that bioretention cells could delay discharge up to two hours, resulting in reduced and delayed peak flows. Further, low flows were observed from the underdrains from several hours up to several days following the end of a storm, indicating temporary internal storage of stormwater. Prolonged discharges make determining the impact of the bioretention systems on volume reduction difficult, principally when continued discharge overlaps the commencement of the subsequent precipitation event (Brown & Hunt, 2011; Davis, 2008). Even in the absence of exfiltration, internal water storage systems can delay and reduce discharges such that only 1/4 to 1/6 of the inflow is discharged within the first 24 hours, indicating successful management of stormwater flows from bioretention systems (Davis, 2008).

OPERATION AND MAINTENANCE

Many site contractors and homeowners lack the knowledge needed to properly install and maintain these SMPs (Woodward et al., 2008). Education to developers, contractors, and homeowners can be the first line of defense in ensuring that the success of these systems occurs (Morzaria-Luna et al., 2004). Aesthetically pleasing BRCs can be expected to be well-received and maintained by homeowners. Legal instruments, such as deed restrictions, may also be implemented to legally enforce maintenance, although as ownership of a property changes, it is common for these restrictions to be overlooked and SMPs to be unkept or removed. Deed restrictions are only effective if they are enforced, and the enforcement of such at local levels is not always clear or well-perceived by the public (Morzaria-Luna et al., 2004). Working in tandem, education and enforcement can be effective measures to ensure the success of these systems.

As with all LID practices, maintenance is key to the long-term success of a bioretention system. The longevity of a bioretention system depends upon proper installation and provision of the necessary level of maintenance. Routine inspections should be conducted to evaluate vegetation health, erosion, and ponding duration and to assess maintenance requirements (USEPA, 1999). Common maintenance items include proper vegetation establishment, soil drainage capabilities, mulching, neglect, sedimentation, and berm failure (Woodward et al., 2008). Johnson & Hunt (2020) studied three bioretention systems in North Carolina ranging from 8 to 17 years old and concluded that they continued to provide long-term solutions to mitigate stormwater volume and peak discharges. The age of the systems did not affect peak flow mitigation. In fact, older cells facilitated infiltration better than new cells.

Long-term pollutant accumulation in bioretention systems is often estimated, but actual field testing is not well-documented as these practices, which gained popularity in the early 2000s, are only now reaching their design life. Over the years, phosphorus will accumulate within the BSM profile with limited outlets, such as plant uptake (Davis et al., 2006). Further, heavy metal accumulation is also a concern as most metals are deposited within the upper layers of the system, limiting the ability for vegetation with deep root systems to uptake metals (Davis et al., 2003). While the accumulation of metals does not pose a short-term concern to the functionality of a system, excavation and replacement of should be incorporated into a long term maintenance plan for the facility as offsite transport and disposal of potentially contaminated media metals can be costly. Davis et al. (2003) estimated that the continual accumulation of heavy metals will exceed EPA regulatory limits for biosolid applications for lead and zinc after 16 years, cadmium 20 years, and copper 77 years. Even so, USEPA (2000) notes soil replacement should occur between 5 and 10 years following establishment of these systems and total mulch replacement is encouraged every 2-3 years (USEPA, 1999).

CHAPTER 3 – ESTIMATING RUNOFF PROBABILITY FROM PRECIPITATION DATA: A BINOMIAL REGRESSION ANALYSIS

INTRODUCTION

Surface runoff discharges occur across various land uses and under a wide array of circumstances. Within a forested hillslope catchment, little or no runoff may occur during a particular storm (Penna et al., 2011; Saffarpour et al., 2016; S. Wang et al., 2022). Yet, within urban areas, where there is a high percentage of impervious surfaces, surface runoff discharges may occur during small storms. Low impact development is a method utilized in urban and suburban catchments with the goal of preserving the pre-development hydrology. It incorporates a combination of techniques into site design to reduce the development impacts on stormwater water quality and quantity (Dietz, 2007).

Low impact development effectiveness is often evaluated in literature by the difference in runoff thresholds when LID practices are and are not applied (Hood et al., 2007; Tirpak et al., 2021). The runoff threshold represents the precipitation depth that must occur within a catchment before surficial runoff generation. The runoff threshold may further describe the storage depth of a watershed, system, or surface (Bean et al., 2007). Understanding the thresholds for which a catchment produces runoff discharge is important in determining the effectiveness of an LID technique or soil and vegetation management practices that affect surface runoff generation.

Linear regression analyses have long been used to determine runoff thresholds (Fink and Frasier, 1977; Li and Gong, 2002; Hood et al., 2007; Bean et al., 2007; Ross et al., 2021; Tirpak et al., 2021). Yet, literature has also recognized that runoff thresholds are likely more complex, dictated by catchment-specific variables, and may follow non-linear relationships (Ali et al., 2013).

Typical linear regression is performed by plotting the precipitation depth versus the runoff observed for each storm. A line of best fit is generated using the least squares method, and the x-intercept of the generated line is interpreted as the runoff threshold. Hood et al. (2007) utilized this technique to compare the runoff threshold of traditional and LID development sites. They concluded that runoff occurs from the traditional watershed when precipitation exceeds 3.0 mm and from the LID watersheds following 6.0 mm of precipitation. Bean et al. (2007) determined estimated storage depths and curve numbers from linear regression models when evaluating the effectiveness of permeable pavement systems on runoff reduction and water quality. Ross et al. (2021) recognized that the relationship between precipitation and runoff could have been more consistently linear and utilized piecewise linear regression analysis to identify a break in the relationship between precipitation and runoff within various watersheds. Tirpak et al. (2021) established the runoff threshold of permeable interlocking concrete pavers using the linear regression and x-intercept methodology. Overall, linear modeling provides the user with a finite value that, when exceeded, runoff is expected to occur.

Homoscedasticity (constancy of variance) is an important assumption of statistical analysis and is a useful way to check the adequacy of a linear regression model (Crawley, 2015; Zuur et al., 2008). When the assumption of homoscedasticity is not met, the resulting linear model may lead to invalid inferences (Box & Cox, 1964; Breusch & Pagan, 1979). Homoscedasticity can be determined visually by plotting the residuals (the difference between the predicted value of the linear model and the observed value) against the fitted values. Figure **1**a and b provide examples of scatterplots and linear regression analyses of a homoscedastic and a heteroscedastic dataset, representative of real-world observations, respectfully. The linear regression analysis yields a fitted line that is used to describe the data by minimizing the residuals or the distance between the data and the predictor line. The homoscedastic dataset displays constant variance along the fitted line, while the heteroscedastic dataset displays non-constant variance along the fitted line (variance increases as precipitation increases). Figure **1**c and 1d illustrate the residuals against the fitted values of the linear regression for visual interpretation of the residual structure. If the residual plot is without structure, or as Crawley (2015) describes, "like the sky at night" (Figure **1**c), then homoscedasticity is observed, and the assumption is satisfied. Alternatively, if the residual plot is structured, such as the fan-shaped pattern seen in Figure **1**d, it indicates heteroscedasticity (non-constant variance) and a failure of the linear regression assumption (Crawley, 2015; Seber & Lee, 2003; Triola, 2008).

Visual observation of published runoff datasets (Hood et al., 2007; Bean et al., 2007; Tirpak et al., 2021; Radovanovic and Bean, 2022; Wang et al., 2022) indicated that the data might not be homoscedastic and suggested that the commonly used linear regression to determine the runoff threshold may not be the minimal adequate model. The scatterplot of the published studies displayed increased residual size as storm size increased (non-constant variance), like that observed in Figure 1b. The observed fan-shaped pattern implies that the precipitation-discharge datasets may be heteroscedastic, and the linear regression may not accurately express the relationship between precipitation and discharge (Crawley, 2015; Seber & Lee, 2003; Triola, 2008). Beyond visual observation, heteroscedasticity can be tested using the Breusch-Pagan (BP) Test (Breusch & Pagan, 1979), the White Test (White, 1980), or others. Studies determining runoff threshold with linear regression did not indicate if such tests were completed (Fink and Frasier, 1977; Li and Gong, 2002; Hood et al., 2007; Bean et al., 2007; Ross et al., 2021; Tirpak et al., 2021). We acknowledged that many variables may affect processes that lead to runoff generation, such as AMC, rainfall intensity, slope, groundwater table depth, soil type, soil depth, and percent vegetative cover, to name a few (Detty & McGuire, 2010; Kohler & Kinsley, 1951; Saffarpour et al., 2016; S. Wang et al., 2022). Nonetheless, when only rainfall is selected as the independent variable to determine threshold behavior, as commonly performed, the data appears to be heteroscedastic.

Further indication that the linear regression may not be adequate is that discharge was observed in those studies (Hood et al., 2007; Bean et al., 2007; Tirpak et al., 2021; Radovanovic and Bean, 2022; Wang et al., 2022) before the runoff threshold was exceeded. Additionally, four of the five studies encountered storms when the runoff threshold was exceeded, yet no discharge occurred (Hood et al., 2007; Bean et al., 2007; Radovanovic and Bean, 2022; Wang et al., 2022). Bean et al. (2007), too, acknowledged this weakness in the determination of the runoff threshold. Given the above, it is questionable if linear regression is a minimally adequate way to assess the likelihood of discharge. We investigate if the concept of runoff threshold can be improved through the determination of the probability for discharge to occur rather than utilizing a finite value above which discharge is expected to occur.

Binomial regression could offer a better way to analyze the chance of discharge occurring following a particular precipitation depth. Generally, binomial regression describes the probability of success given a particular value of the independent variable, or in our case, the probability for runoff to occur based on the precipitation (Triola, 2008). Binomial regression does not attempt to definitively determine when discharge will occur, rather, it provides a probability for discharge to occur based on precipitation depth.

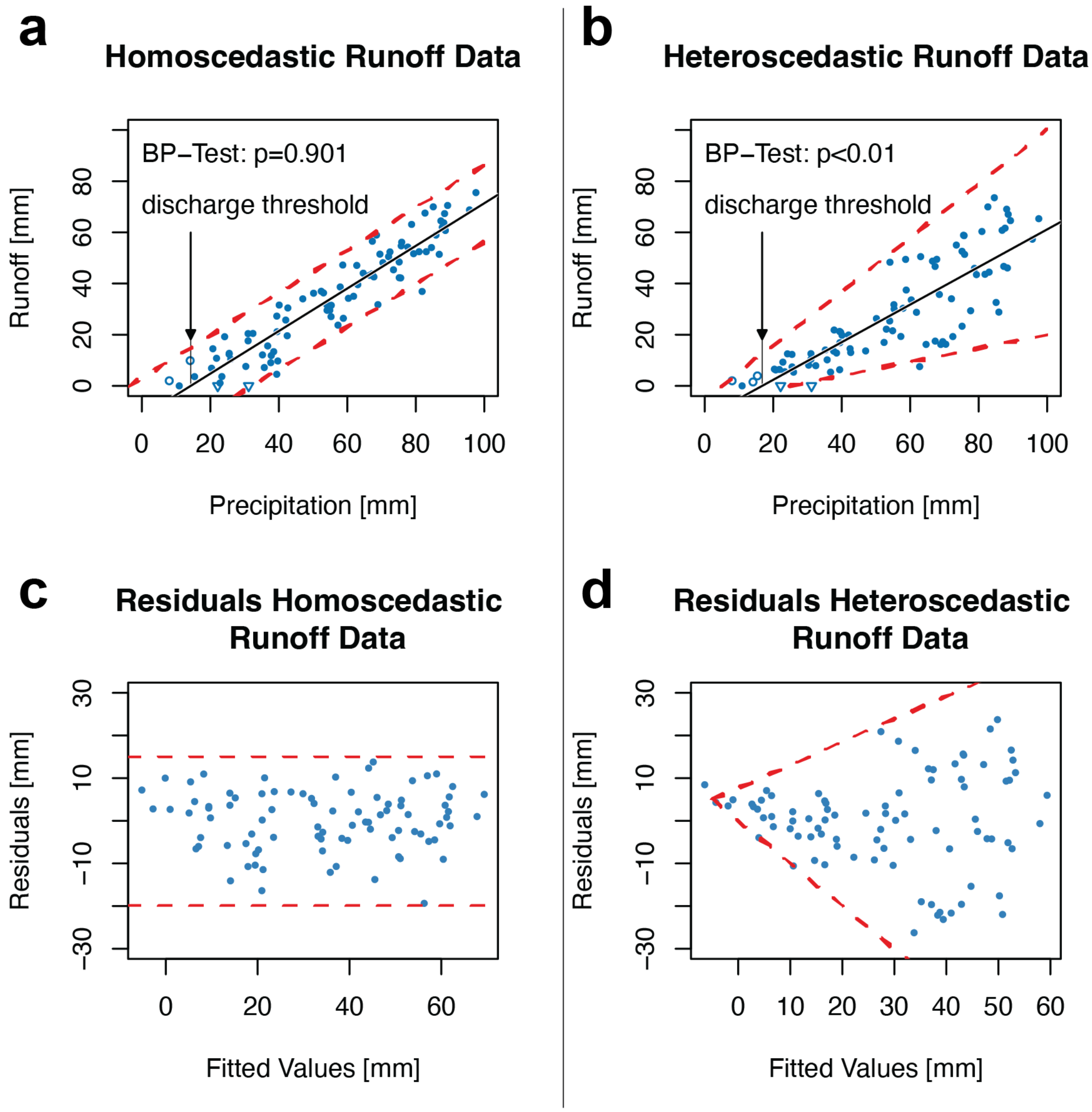


Figure 1: *Homoscedastic and heteroscedastic runoff data. (a) Scatterplot and linear regression analysis of homoscedastic runoff data displaying constant variance. (b) Scatterplot and linear regression analysis of heteroscedastic runoff data displaying non-constant variance. (c) Scatterplot of fitted values vs. residuals plots of homoscedastic runoff data displaying constant variance. (d) Scatterplot of fitted values vs. residuals plots of heteroscedastic runoff data displaying non-constant variance.*

It is understood that the runoff threshold is approximate and may be impacted by several variables, including, but not limited to, storm intensity, soil moisture conditions, and antecedent dry periods. The goal of this study is to provide a more objective approach to the evaluation of catchments and small watershed hydrology. Our objectives are to evaluate binomial regression analysis as an alternative method to linear regression-based runoff thresholds and to determine the likelihood of a watershed producing discharge following a given precipitation amount.

MATERIALS AND METHODS

STUDY SITES

Data for this study were collected from five published articles and were obtained either directly from the authors or extracted from the available tables and figures. These five studies were selected for reporting rainfall and runoff data and to reflect different runoff scenarios. The gathered data include nineteen catchments ranging from 0.03 to 5.5 hectares and varying treatments. Bean et al. (2007) studied permeable pavement systems consisting of concrete grid pavers (CGP), porous concrete (PC), and permeable interlocking concrete pavers (PICP) at four sites in North Carolina (USA) to evaluate their effectiveness in reducing runoff and improving water quality. The permeable pavement systems consisted of 200 mm to 275 mm thick aggregate reservoirs overlying loamy sand to sandy soils. Rainfall and runoff data from the PC and CGP sites were collected over 17 and 26-month periods, respectively. Hood et al. (2007) focused on the comparison of the lag time of traditional and LID sites in Connecticut (USA). The traditional residential watershed was constructed with standard stormwater management systems which utilized curbs and gutters for stormwater collection and conveyance. The LID watershed was a newer construction and incorporated LID practices consisting of bioretention areas, rain gardens, concrete pavers, and grassed swales. Rainfall and runoff data for the watersheds was collected over a 31-month study period. Tirpak et al. (2021) examined the effectiveness of permeable pavements overlying poorly drained, sit loam soils to mitigate peak runoff conditions in Ohio (USA). The retrofit parking lot was equipped with CGPs and an approximately 575 mm aggregate reservoir. Rainfall and runoff data was collected over a 16-month period. Radovanovic and Bean (2022) investigated whether the integration of compost into residential landscapes in Florida (USA) decreased runoff volumes and nutrient loadings. Soils underlying the residential lawns consisted of very sandy soils. Stormwater runoff was measured and collected at catch basins within each defined watershed over a 27-month period. Wang et al. (2022) reviewed rainfall-runoff characteristics of a karst hillslope in southwest China over a two-year period. The karst hillslopes were overlain with approximately 0 to 2 m of calcareous clay to clay-loam and divided into 5 m by 20 m vegetated plots. We refer to the original publications for further details on experimental design, data acquisition, and processing.

DATA PREPARATION

The datasets were analyzed using R (R Core Team, 2024), with the following packages: ggplot2 (Wickham, 2016), nlme (Pinheiro and Bates, 2000, and Pinheiro et al., 2022), lmtest (Zeileis and Hothorn, 2002), multcomp (Hothorn et al., 2008), ResourceSelection (Lele et al., 2023), survey (Lumley, 2010), and reshape2 (Wickham, 2007). Precipitation and runoff data were collected on a per-storm basis. Runoff data of values less than 0.1 mm were considered to be artifacts of the visual data extraction (i.e., from Hood et al., 2007) and hence set to zero. Observations where runoff exceeded precipitation were removed. The precipitation and runoff values were also log-transformed, and to compensate for zero runoff values, a constant of 0.01 mm was added prior to the transformation (Kilmartin and Peterson, 1972). Binomial runoff datasets were generated for each, where runoff values were converted to a binomial response with zero equals no runoff and one equals runoff. A “no runoff” data point was added to each binomial dataset, consisting of zero runoff and zero precipitation, to prevent convergence issues with the binomial modeling.

STATISTICAL ANALYSIS

Runoff data were modeled as continuous and binomial response variables. Runoff as a continuous variable was evaluated with four linear models for each dataset: 1.) analysis of covariance (ANCOVA); 2.) ANCOVA with random effects for variance structure; 3.) log-transformed ANCOVA; and 4.) log-transformed ANCOVA with random effects for variance structure. In the following, we use the term linear models when referring to these four models. Best-fit models were based on the Akaike information criterion (AIC), where the model with the lowest AIC was accepted as the best-fit model (Zuur et al., 2008). For each model, the linear regression runoff threshold (LRRT) and the threshold confidence interval were computed, the Breusch-Pagan test was performed to check if the model residual structure improved (i.e., less fan-shaped), and individual treatments were contrasted to see which were significantly different at the $\alpha=0.05$ level. Treatments were contrasted to determine statistically significant differences. If not significantly different, treatments were combined, and a new model was evaluated based on the AIC. If the AIC did not decrease upon the combination of treatments, the new model was rejected.

ANALYSIS OF COVARIANCE

ANCOVA was used to model the relationship between precipitation and runoff. The parameters derived from an ANCOVA model include slope, intercept, and standard errors. The LRRT was calculated as the x-intercept of the bet-fit regression line. It is understood that treating data as homoscedastic, although it is heteroscedastic, can result in incorrect derived parameters (i.e., the runoff threshold) and flawed conclusions (Breusch & Pagan, 1979). For bootstrapping the 95% confidence interval, the fitted model values from the original data and randomly selected residuals (with replacement) were combined to generate a new dataset from which a new LRRT was computed. This step was repeated 1,000 times. Stormwater data are often log-normally distributed (Kilmartin & Peterson, 1972; Van Buren et al., 1997), and log-transformation of data may rectify the heteroscedasticity of a dataset (Box & Cox, 1964; Zuur et al., 2008). Accordingly, the ANCOVA was performed on the log-transformed data, and the LRRT and the 95% bootstrapping confidence interval were subsequently computed.

ANALYSIS OF COVARIANCE WITH RANDOM EFFECTS (MIXED-EFFECT MODEL)

ANCOVA with random effects was evaluated as a method to address the heterogeneity of the datasets (Zuur et al., 2008). The applied variance structures outlined in Zuur et al. (2008) were employed to both the original and log-transformed datasets. The resultant mixed-effect models with the lowest AIC were determined to be the best fit for each dataset. The LRRT and 95% bootstrapping confidence intervals were determined using the best mixed-effect model for each dataset.

BINOMIAL REGRESSION MODELING

A binomial model was created to determine the likelihood for runoff, the binomial response variable, to occur from a particular watershed as a function of precipitation depth by the methodology outlined in Crawley (2015). The treatments were contrasted as they were for linear modeling. Using the binomial model, the probability for runoff to occur based on precipitation depth can be determined. Next, the binomial runoff probability curves (s-curves) generated by the binomial model for each watershed are used to determine the storm size required for a watershed to have a 10% and 90% chance of producing runoff. We use the 10% and 90% runoff probabilities to determine precipitation depths, which result in a low chance of runoff ($p<0.1$) and a high chance of runoff ($p>0.9$). The goodness of fit of the binomial regression model was evaluated with the Hosmer-Lemeshow test and the likelihood ratio test (Hosmer et al., 2013). The Wald test (Rao &

Scott, 1984) was performed to test for the linear relationship between the independent variable (log-precipitation) and the log-odds of the binomial regression models.

RESULTS AND DISCUSSION

For all reviewed studies, Table 1 presents the LRRT, 95% confidence interval of the LRRT, and BP test statistic for homoscedasticity for the linear models. Further reported in Table 1 is the number of recorded storm events that 1.) have generated runoff before precipitation exceeding the LRRT and 2.) did not generate runoff following exceedance of the LRRT. Table 1 also provides the 10% and 90% runoff probabilities from the binomial regression.

LINEAR MODELS

The BP test (Table 1) revealed that the original datasets from the five reviewed studies were heteroscedastic and, therefore, violated the assumption of linear regression. When using a linear model to evaluate data, the BP test results indicate that heteroscedasticity may be corrected through log-transformation. Log-transformation remedied the heteroscedasticity of the dataset from two of the five studies (Bean et al., 2007; Radovanovic & Bean, 2022) but failed to do so in the remaining (Hood et al., 2007; Tirpak et al., 2021; Wang et al., 2022) (**Error! Reference source not found.**). The BP test was consistently greater than 0.05 for the mixed-effect modeling, indicating constant variances for all normal and log-transformed datasets, which was to be expected. Observations of runoff for precipitation depths lower than the LRRT and observations of no runoff following exceedance of the LRRT indicate that the linear model does not adequately represent the observed data. Regardless of the log-transformation, runoff occurring before the LRRT and not occurring following exceedance continue to be observed for all studies.

ANCOVA of the heteroscedastic datasets resulted in large, and at times negative, confidence intervals, specifically for the Hood et al. (2007) and Radovanovic and Bean (2022) datasets (**Error! Reference source not found.**). Negative runoff thresholds, which imply that runoff occurs without precipitation, are illogical and indicative that the model does not adequately describe the data. Although no negative log-transformed confidence intervals were observed, the confidence intervals remain large. Additionally, the confidence intervals of the mixed-effect models are large and more frequently become negative (Bean et al., 2007; Hood et al., 2007; Radovanovic & Bean, 2022; S. Wang et al., 2022).

Our study implies that the heteroscedasticity of the precipitation and runoff datasets is best remedied by implementing a mixed-effect model when determining an LRRT. However, negative values of the bootstrapped 95% confidence intervals and runoff occurring before the mixed-effect modeling LRRTs and not occurring following exceedance continued are indications that, although the mixed-effect modeling accounts for the heteroscedasticity, it does not adequately describe the data. Many other factors influence the runoff threshold and perhaps are not best represented in a finite value.

Table 1: *Linear regression runoff thresholds and 95% bootstrapping (n=1,000) confidence intervals for linear regression modeling methods and the 10% and 90% binomial model runoff probabilities.*

Article	Treatment	ANCOVA			ANCOVA with Random Effects			Log-Transformed ANCOVA			Log-Transformed ANCOVA with Random Effects			Binomial Model	
		RT [CI] [mm]	# Below	# Above	RT [CI] [mm]	# Below	# Above	RT [CI] [mm]	# Below	# Above	RT [CI] [mm]	# Below	# Above	p10 [mm]	p90 [mm]
Hood et al. (2007)	Control	0.8 [-0.8; 2.2]	0	2	1.0 [-0.9; 2.0]	0	2	6.9 [4.3; 10.2]	33	0	6.3 [3.6; 9.7]	29	0	0.4	2.3
	LID	5.8 [3.8; 7.4]	0	26	1.6 [-181.6; 195.3]	0	56	34.5 [25.0; 55.6]	32	2	60.0 [35.5; 157.4]	39	0	5.9	30.1
	Traditional	3 [2.3; 3.7]	7	4	2.2 [1.0; 2.7]	2	6	8.3 [6.5; 10.8]	34	0	7.3 [5.3; 9.7]	31	0	0.9	4.5
	BP test	**0.00**			0.93			**0.00**			0.73				
Bean et al. (2007)	CGP	23.4 [23.1; 23.7]	1	10	13.6 [7.9; 15.3]	0	33	57.6 [45.3; 81.9]	6	1	88.9 [59.0; 176.5]	8	0	23.3	80.1
	PC	11.4 [10.8; 12.3]	4	0	2.5 [-76.2; 16.2]	0	3	20.0 [13.4; 32.9]	7	0	20.0 [12.9; 31.2]	7	0	3.3	11.4
	BP-test	**0.00**			0.83			0.29			0.87				
Tirpak et al. (2021)	Modeled - Asphalt	0.9 [0.6; 1.2]	0	0	0.9 [0.5; 1.1]	0	0	1.5 [0.1; 3.5]	0	0	1.3 [0.1; 4.2]	0	0	0.2	0.7
	Retrofit - PICP	5.5 [5.2; 5.9]	11	0	3.9 [3.2; 4.3]	5	5	9.1 [7.2; 11.5]	18	0	5.8 [4.8; 9.9]	12	0	1.9	6.5
	BP test	**0.00**			0.90			**0.00**			0.32				
Radovanovic and Bean (2022)	Compost, no till, no TD	4.5 [3.2; 5.8]	1	0	1.3 [-1.5; 3.4]	1	2	10.1 [4.7; 20.8]	5	0	9.7 [4.0; 21.3]	5	0	0.9	5.8
	Compost, till, no TD	7.0 [6.3; 7.8]	4	0	0.5 [-1.3; 2.3]	0	0	4.3 [0.3; 11.0]	2	0	4.2 [0.2; 11.7]	2	0	0.1	0.6
	Compost, no till, no TD	-24.4 [-32.2; -18.2]	0	4	1.4 [-0.8; 3.1]	0	3	11.2 [6.4; 21.6]	3	0	10.9 [5.5; 22.0]	3	0	2.1	13.4
	Compost, till, TD	2.3 [-1.3; 5.0]	0	1	2.1 [-6.5; 6.3]	0	1	14.3 [8.9; 26.3]	2	0	14.4 [9.2; 23.5]	2	0	2.1	13.4
	Compost, no till, TD	10.3 [9.5; 11.0]	2	0	2.2 [-1.1; 4.9]	0	2	14.4 [9.5; 23.6]	4	0	14.4 [9.2; 23.5]	4	0	2.1	13.4
	No compost, no till, TD	8.2 [7.6; 8.7]	4	0	1.3 [-0.2; 2.5]	0	2	8.1 [4.5; 13.3]	4	0	6.2 [2.4; 11.2]	2	0	0.9	5.8
	No compost, no till, TD	-6.9 [-16.0; -0.9]	0	3	1.7 [-3.3; 4.3]	0	3	13.1 [7.7; 25.1]	2	0	13.6 [8.7; 23.6]	2	0	2.1	13.4
	No compost, no till, no TD	8.0 [7.0;9.0]	2	2	2.3 [-1.2; 4.8]	0	4	16.4 [10.8; 32.1]	5	1	15.9 [10.0; 30.5]	5	1	2.1	13.4
	BP test	**0.04**			0.52			0.25			0.99				
Wang et al. (2022)	Walnut	16.3 [12.9; 19.3]	16	26	2.5 [-51.7; 13.6]	0	52	121.6 [73.2; 356.5]	57	0	178.9 [86.4; 1350.2]	57	0	3.5	78.3
	Corn	17.5 [11.2; 21.9]	8	37	5.4 [-167.8; 24.9]	0	68	196.3 [103.6; 781.3]	41	0	593.1 [176.1; 45,488.9]	41	0	3.5	78.3
	Shrubs	15.6 [3.7; 22.6]	7	20	5.8 [-279.8; 144.1]	0	51	101.3 [65.8; 203.7]	49	0	151.6 [81.1; 515.2]	52	0	3.5	78.3
	Grass	17.7 [14.5; 20.1]	19	28	3.2 [-40.8; 13.1]	0	54	121.4 [73.4; 381.8]	57	0	185.1 [90.9; 1,575.1]	57	0	6.1	134.3
	BP test	**0.00**			0.69			**0.00**			0.38				

Note: For each model, the number of runoff events that occurred before the respective runoff threshold is reported as "# Below" and the number of discharge events that did not occur following exceedance of the runoff threshold is reported as "# Above". **Bold** = BP test results in p <0.05 (heteroscedastic)
Abbreviations: ANCOVA, Analysis of covariance; BP, Breusch-Pagan; CGP, Concrete Grid Pavers; CI, 95% Confidence Interval; LID, low impact development; p10, 10% runoff probability; p90, 90% runoff probability; PC, Porous Concrete; PICP, Permeable Interlocking Concrete Pavers; RT, Runoff threshold; TD, Topdressing.

BINOMIAL REGRESSION MODELING

Goodness of fit tests were performed for all binomial regression models. The Wald tests, and the likelihood ratio tests were significant ($p<0.01$) for all binomial regression models. The Hosmer-Lemeshow tests to evaluate differences between objected and fitted data was not significant for any binomial regression model ($p>0.05$). The binomial runoff probability curves (s-curves) for Hood et al. (2007), Bean et al. (2007), and Tirpak et al. (2021) and the LRRTs are presented in Figure 2. The binomial regression resulted in a larger 10%-90% runoff probability range (p10-p90) for catchments that incorporated LID practices than those that did not. For example, the modeled data from Tirpak et al. (2021) resulted in a p10-p90 of 0.2–0.7 mm, while the parking lot retrofit with PICP has a p10-p90 of 1.9–6.5 mm (Table 1). Steep slopes of the s-curve, as seen for the control, traditional, and modeled watersheds in Figure 2, indicate a smaller p10-p90, whereas gradual slopes observed for catchments with less impervious surface or LID practices indicate a larger p10-p90. This disparity is best explained by the impact of environmental factors within a catchment or watershed.

The LRRTs from the four linear models for each reviewed study were compared to the binomial runoff probability curve. Figure 3 presents a box plot summarizing the probabilities for runoff to occur at the LRRT for each linear method. The ANCOVA LRRTs, established with heteroscedastic datasets, correspond with a wide 25 to 75 percentile range of runoff probabilities, from approximately 30-80%, with a median of approximately 65% (i.e., when precipitation reaches the LRRT, there is a 65% chance the catchment will produce runoff) (Figure 3). The wide 30-80% range indicates that the traditionally used ANCOVA LRRT does not consistently correspond with a particular chance of runoff production and is not a reliable measure.

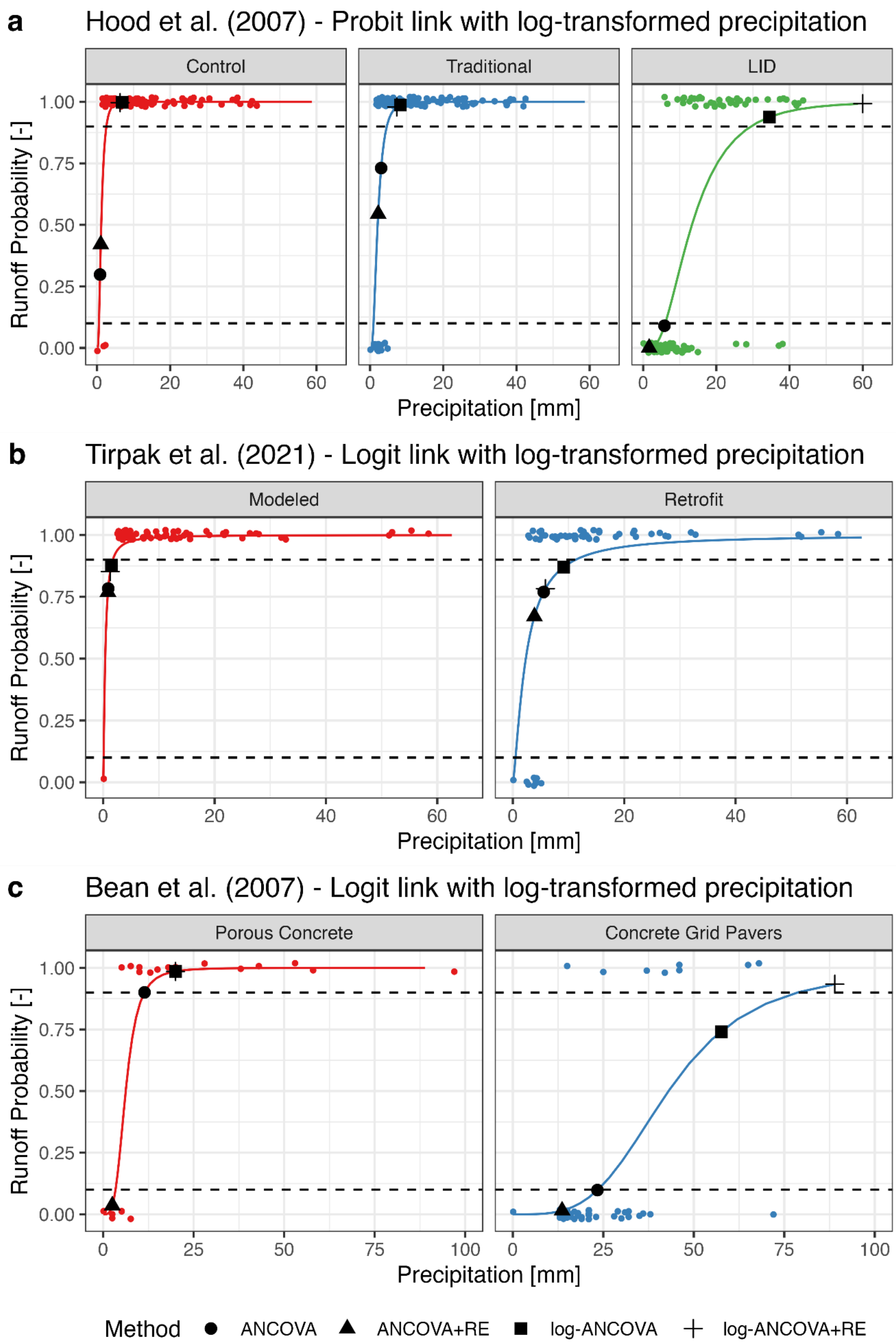


Figure 2: *Binomial Models estimating the probability of runoff based on precipitation depths with linear regression model runoff thresholds. (a) Hood et al. (2007) (b) Tirpak et al. (2021) (c) Bean et al. (2007). Abbreviations: ANCOVA, Analysis of Covariance; LID, low impact development; RE, random effects.*

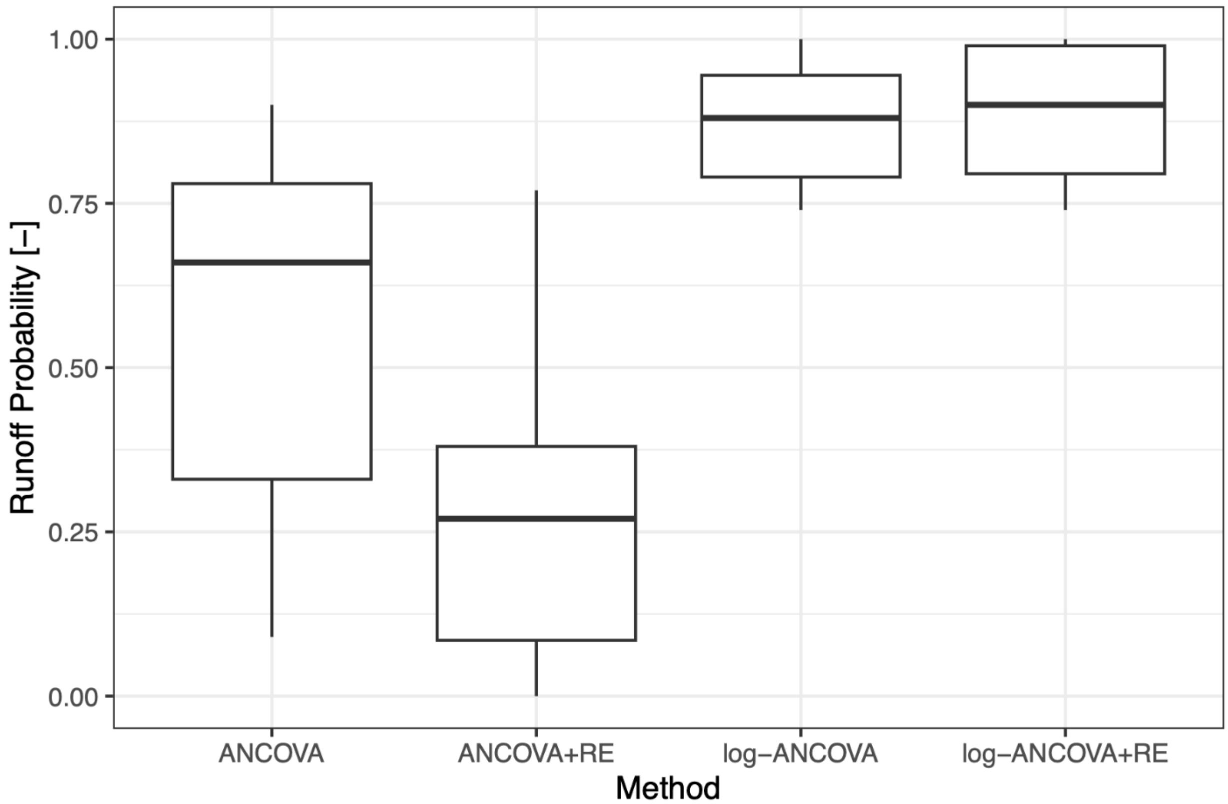


Figure 3: *Box and whisker plot summarizing the probabilities for runoff to occur at the linear regression runoff threshold (LRRT) for each linear method. The boxes encompass the 25-75 percentile range, and the whiskers extend to the minimum and maximum values of the LRRT runoff probabilities. Abbreviations: ANCOVA, Analysis of Covariance; RE, random effects.*

The ANCOVA with random effects model (mixed-effect model), while addressing heteroscedasticity of runoff datasets, results in a median LRRT corresponding to a 25% runoff probability, much lower than the ANCOVA method. The mixed-effect model results in a smaller 25 to 75 percentile range of LRRT corresponding runoff probability of approximately 15-40% (Figure 3), indicating a more consistent model performance across the datasets. Figure 3 illustrates the convergence of the log-transformed ANCOVA and mixed-effect models to a higher runoff probability with median LRRTs around 90%. The log-transformed models both resulted in narrower ranges of corresponding runoff probabilities than the models that utilized the non-log-transformed data (Figure 3).

The effects of LID practices on runoff generation are variable. There are a number of factors that impact the performance of LID practices and runoff generation, including storm size, precipitation intensity, antecedent dry periods, underlying soil types, soil depth, rock fragments, root count, bedrock outcrops, growing season, sediment loadings, and traffic load. Hood et al. (2007) noted that LID watersheds exhibited a higher runoff threshold than traditional development. However, the effects of the LID practices were most notably observed during small, short-duration storms with low antecedent moisture conditions. Tirpak et al. (2021) also observed that benefits from LID diminished as storm size or intensity increased and antecedent dry periods decreased. Considering the variability of the environmental factors that influence runoff generation, we believe the p10-p90 provides the user with a more comprehensive understanding of when runoff may be generated than the LRRT.

Stormwater that interacts with a vegetated surface or underlying substrate allows environmental factors to influence runoff generation. Bean et al. (2007) concluded that the permeable pavement systems with a storage layer increased runoff reduction. Tirpak et al. (2021) attributed the effects of the PICP, inclusive of a sub-base, to the stormwater's ability to infiltrate into underlying soils. The observed variability of runoff generation from LID practices is better reflected in the p10-p90 range than it is with the LRRT. Conversely, runoff thresholds for highly impervious watersheds, where influence from environmental factors is considerably less, have small p10-p90 ranges and may be sufficiently represented by an LRRT.

Presentation of runoff threshold as a range, as opposed to a finite LRRT, provides a better understanding of how environmental factors influence a catchment. Large p10-p90 ranges indicate high variability and more environmental interaction, whereas small ranges represent less variability and less interaction. Provision of this range provides the user with the ability to understand the precipitation depths that have a low chance of producing runoff ($p<0.1$), a high chance of producing runoff ($p>0.9$), or any value between. The range also allows for a review of both best- and worst-case runoff scenarios. It is recommended that the selected runoff probability be based on the potential downstream impacts. For example, should downstream impacts have the potential to result in flooding homes or significant streambank erosion, a user may choose to implement the 10% runoff probability value into the project design.

COMPARISON OF FINDINGS

Binomial regression concurs with the general conclusions of the reviewed studies. Our findings indicate that the use of the LRRT may be sufficient to predict the amount of precipitation required to produce runoff for highly impervious catchments but is deficient in doing so for less impervious catchments or those that implement LID practices. Contrasting the watersheds concurred with the observed significant differences reported by the authors. We agree with the studies that within catchments that implement LID practices, there is an observed level of variability attributable to environmental factors.

Hood et al. (2007) concluded that runoff depth from the LID watershed was significantly less and more variable than control and traditional watersheds. This variability is not reflected in the use of an LRRT. The mixed-effect model corrected for heteroscedasticity yet resulted in a lower LRRT for the LID watershed than the traditional watershed, indicating that the mixed-effect model does not amply describe the data. Figure 2a shows the influence of the LID practices on the runoff generation with a more gradual s-curve and wider p10-p90 range than the more impervious control and traditional catchments.

Bean et al. (2007) found CGPs reduced runoff more than PC. Binomial regression concurred as the CGP paver site has a higher 10% runoff probability (23.3 mm) than PC (3.3 mm). The difference in the performance of the two porous pavement systems is clearly illustrated by the shape of the s-curve (Figure 2c) and it can be assumed that CGPs allow for more interaction with the underlying soil than the PC.

Tirpak et al. (2021) observed statistically significant runoff reductions and increased runoff thresholds with permeable interlocking concrete pavers (PICP) compared to asphalt. The steep slope of the modeled parking lot s-curve (Figure 2b) and narrow p10-p90 range are indicative of the impervious nature of the catchment. Whereas the retrofit PICP watershed increases the interaction between the runoff and soil and, therefore, results in a gradual slope of the binomial curve and a wider p10-p90 range.

CONCLUSIONS

Runoff thresholds computed from linear regression can be flawed as the precipitation-runoff datasets might be heteroscedastic. For two of the five studies reviewed, heteroscedasticity was remedied via log-transformation of the data. However, it is better to correct for heteroscedasticity by implementing a mixed-effect model with applied variance structures. Confidence intervals for the linear models were relatively large, and at times negative. Runoff occurring before the LRRTs and not occurring following exceedance is a further indication that these models may not adequately describe the data.

The best method to analyze precipitation and runoff data in the five evaluated studies was to model the likelihood of runoff with a binomial regression model that allows calculating the likelihood of runoff for each precipitation depth. The 10%-90% runoff probability range acknowledges the variability of LID practices and associated environmental factors while also accurately representing the limited variability of traditional development techniques. The more gradual slope of the binomial curve is attributed to pervious surfaces, LID practices, and environmental interaction and results in a larger runoff p10-p90 range. We offer the p10-p90 range as a more accurate representation of runoff thresholds as it accounts for the potential variability of a catchment. Further, the binomial regression presented herein concurs with the general conclusions of the original studies.

CHAPTER 4 – EFFECTS OF BIORETENTION SOIL MIXTURES ON METAL SPECIATION AND TOXICITY TO AQUATIC COMMUNITIES

INTRODUCTION

Unmitigated urbanization will increase the quantity and decrease the quality of stormwater runoff (USEPA, 2000). Many heavy metals are introduced into urbanized stormwater as a byproduct of accumulated tire wear and brake dust. It is well known that high pollutant loadings cause poor water quality conditions within receiving waterbodies and are problematic for aquatic plant and wildlife species. Water quality impacts caused by new development can be mitigated through the implementation of properly designed LID practices, such as BRCs. Bioretention cells are a common LID practice consisting of a shallow landscaped depression, commonly used to reduce and filter stormwater runoff from impervious areas and typically contain (from top to bottom) a ponding zone, mulch layer, vegetation, BSM, and an aggregate layer with an underdrain where native soils have low permeability (USEPA, 2000). BSM generally comprises a combination of compost to provide organic matter and nutrients to promote phyto and microbial remediation and a mineral aggregate to ensure stormwater infiltration through the BSM (Roy-Poirier et al., 2010).

The ability of BRCs and various BSMs to effectively filter stormwater pollutants, including heavy metals, has been well-investigated (Davis et al., 2003; Dietz & Clausen, 2005; Glass & Bissouma, 2005; Knappenberger et al., 2022; H. Li & Davis, 2008; Sun & Davis, 2007). Bioretention cells can remove heavy metals by filtering particulates and through the adsorption of pollutants to soil particles. Reported removal rates of nutrients from bioretention cells are variable and are likely due to the variable compositions and depths of BSMs (Chahal et al., 2016; Davis et al., 2006; Hunt et al., 2006; Mullane et al., 2015). Less variability of removal rates was observed for heavy metals, with most studies concluding that BRCs are highly effective at removing lead, copper, and zinc (Davis et al., 2003; Hunt et al., 2006; H. Li & Davis, 2008; Sun & Davis, 2007; USEPA, 2000). Davis et al. (2003) observed that a significant percentage of heavy metal removal occurred within the upper portion of the BSM profile and, upon evaluation of BRC core samples, attributed this high efficiency to the mulch layer. Knappenberger et al. (2022) examined metal removals by BSMs of varying composition and found that all reduced total and dissolved copper, lead, and zinc.

In contrast, several studies have found that BRCs increased pollutant loadings with peak concentrations of constituents leached from BRCs occurring at the beginning of a storm and generally decreasing as the storm progresses and following successive storms (Chahal et al., 2016; Mullane et al., 2015). Bioretention cells with an organic component in the BSM can leach nitrate, TKN, phosphorus, and total and dissolved copper into stormwater effluent where most copper is leach during the first 3-6 storms following establishment, and the majority of nitrate and TKN are leached during the first three storms (Chahal et al., 2016; Hunt et al., 2006; Mullane et al., 2015).

Heavy metal concentrations in stormwater discharges are regulated and monitored as these pollutants can be toxic to aquatic communities (USEPA, 2007). The total concentration of a heavy metal does not represent the concentration of a metal ion available to aquatic organisms. A heavy metal must be in a dissolved form to be bioavailable for uptake through the gill. The USEPA recognizes that dissolved metal concentrations are better estimates for toxicity to aquatic organisms than total metal concentrations (Kinerson et al., 1996). Toxicity is the metal accumulation on a receptor at or greater than a threshold concentration, so a predefined lethal or sublethal effect will occur (Santore et al., 2001). Dissolved metals better estimate toxicity because as water passes through a species' gill, dissolved metals may be adsorbed to a biotic ligand, a gill surface where metal may accumulate, resulting in toxic effects. It is further recognized that dissolved metal concentrations are only an approximation as USEPA standards loosely define dissolved metals to be any metal that can pass through a 0.45-micron filter. This definition allows all particulate-bound metals less than 0.45 microns to be assumed dissolved and ultimately considered bioavailable.

In addition to distinguishing between total and dissolved metal concentrations, the speciation of a metal should be considered when evaluating toxicity. Free ions of zinc and copper are the most bioavailable species for each metal and, therefore, are considered the most influential on toxicity (Hoang & Tong, 2015; Magnuson et al., 1979). However, water quality characteristics, including temperature, pH, hardness, and DOC, affect metal speciation, bioavailability, and overall toxicity of stormwater (Di Toro et al., 2001; Pagenkopf, 1983; Santore et al., 2001). Specifically, hardness affects toxicity as calcium ions (Ca^{2+}) and magnesium ions (Mg^{2+}) compete with the free heavy metal ions (Cd^{2+}, Cu^{2+}, Pb^{2+}, Zn^{2+}) for binding sites at the biotic ligand and, as such, as hardness increases, toxicity generally decreases (Bui et al., 2016; Hoang & Tong, 2015; Kumar & Singh, 2010). Complexation with DOC also decreases the bioavailability and toxicity of free metal ions. These relationships are called the "protective effect" of the water quality parameters on metal toxicity.

The effects of water quality parameters (i.e., pH, hardness, alkalinity) on the toxicity of heavy metals to aquatic organisms have been well studied (Bui et al., 2016; Hoang & Tong, 2015; Rogevich et al., 2008). Further, the efficiency of bioretention systems at reducing heavy metals and nutrients has been investigated since the inception of these systems (Davis et al., 2003, 2006; Hunt et al., 2006; H. Li & Davis, 2008). However, the reviewed literature lacks the assessment of how bioretention systems alter water quality and how that alteration may affect the toxicity of heavy metals to an aquatic community. The objectives of this study are to evaluate:

1. speciation shifts of stormwater pollutants following filtration through BSMs.
2. changes in toxicity of stormwater following filtration through BSMs using a BLM.
3. determine which BSMs, if any, were more adept at decreasing pollutant bioavailability and toxicity.

MATERIALS AND METHODS

EXPERIMENTAL DESIGN

STORMWATER COLLECTION AND DISTRIBUTION

The study site was on the Washington State University Puyallup Research and Extension Center campus in Puyallup, Washington, USA (Knappenberger et al., 2022). Stormwater runoff from a 1,675 square meter impervious catchment on the campus, consisting primarily of roofs and pavement, was collected and diverted to a single catch basin. Stormwater runoff was directed from the catch basin to an 11.370 m^3 cistern for temporary storage and pollutant dosing as initial testing of the study area revealed that the stormwater had relatively low concentrations of metals, nutrients, and total suspended solids (TSS). Utilizing weirs, the dosed stormwater from the cistern reservoir was equally routed into 16 mesocosms during storms between 2013 and 2015.

Of the eleven storms (S1 to S11) investigated by Knappenberger et al. (2022), ten were selected for this study based on the completeness of water quality data collection and mesocosm construction, excluding storm S1 from our analysis. Phase 1 of the study (October 1, 2012 – May 31, 2013) encompassed S2, S3, and S4, Phase 2 (October 1, 2013 – May 31, 2014) S5, S6, S7, and

S8 and Phase 3 (October 1, 2014 – May 31, 2015) S9, S10, and S11. Due to naturally low pollutant concentrations observed in storms before Phase 1, a mixture of common stormwater constituents was added to the collected runoff. Target concentrations were chosen to reflect the typical potency of stormwater found in western Washington State and are summarized in **Error! Reference source not found.**. For use as a baseline for evaluating pollutant removal or leaching from the BRCs when influent stormwater contained low concentrations of constituents, Storm S2 was not dosed. The copper concentration of S2 represents only the 4th percentile of the NPDES Phase I permittee copper samples collected within western Washington State between 2007 and 2013 (Knappenberger et al., 2022). In each phase, the amount of added stormwater pollutants increased. Supplemented pollutants included phosphorus (KH_2PO4), nitrate (KNO_3), ammonium (NH_4OH), cadmium ($CdCl_2$), chromium (K_2CrO_4), copper ($CuSO_4$), lead ($PbNO_3$), and zinc ($ZnCl_2$). The stormwater was agitated within the cistern during storm events to prevent particulate-bound pollutants from settling.

Table 2 - Target pollutant concentrations from the dosing regimen by phase.

Pollutant	Target Concentrations			
	unit	Phase 1	Phase 2	Phase 3
TSS	[mg/L]	75.0	100.0	200.0
Hardness	-	No target	No target	No target
Total Cd	[µg/L]	0.3	0.3	0.3
Total Cr	[µg/L]	6.0	10.0	10.0
Total Cu	[µg/L]	20.0	25.0	22.4
Total Pb	[µg/L]	2.0	2.0	5.0
Total Zn	[µg/L]	150.0	150.0	610.0
TN	[mg/L]	1.0	2.0	2.0
NO_3+NO_2	[mg/L]	0.3	0.8	1.0
Ammonia	[mg/L]	0.4	0.5	1.0
TP	[mg/L]	0.25	0.3	0.364

Cd: Cadmium; Cr: chromium; Cu: Copper; NO3: nitrate; NO2: Nitrite; Pb: lead, PO4: Phosphate; TN: Total nitrogen; TP: Total phosphorus; TSS: Total suspended solids; Zn: Zinc

BIORETENTION MESOCOSM CONSTRUCTION

The 16 mesocosms were constructed of 152 cm diameter, 132 cm tall high-density polyethylene (HDPE) cylinders. Each mesocosm comprised 31 cm of an aggregate sand layer with a 5 cm diameter perforated underdrain pipe. Over the coarse sand layer, 61 cm of bioretention media was packed. Four bioretention media mixtures containing varying percentages of sand, compost, water treatment residuals (WTR), and shredded bark, were utilized, summarized in **Error! Reference source not found.**. We maintained a consistent nomenclature for BSM type with a previous study by Knappenberger et al. (2022); the amount of compost in the mix identifies the mix. Therefore, the Washington State Department of Ecology (WSDE) recommended BSM containing a 40% compost to 60% sand by volume mix (WSDE, 2019) was named Mix40. Each mesocosm was planted with the same vegetative species, including *Deschampsia cespitosa*, *Deschampsia cespitosa* 'Northern Lights', and *Cornus sericea* 'kelseyi' and replicated four times.

Table 3: Composition of bioretention soil mixtures.

Treatment	Sand	Compost	WTR	Shredded Bark
	[%]	[%]	[%]	[%]
Control				
Mix15	60	15	15	10
Mix20	80	20	-	-
Mix30	60	30	10	-
Mix40	60	40	-	-

SAMPLE COLLECTION AND ANALYSIS

Influent and effluent flow-weighted composite water quality samples were collected from ten storms. Influent stormwater samples were collected from a weir bypassing the mesocosms during each storm. Mesocosm effluent stormwater samples were collected from the perforated underdrain pipe using ISCO portable samplers (ISCO 6712, Teledyne Isco, Lincoln, NE USA) and set on ice to cool the stormwater immediately. Up to 40 flow-weighted composite water quality sample aliquots were collected throughout each storm for a total 400 mL sample to ensure a representative event mean concentration of each stormwater pollutant. The flow was monitored using tipping-bucket flowmeters (Model TB1L, Hydrological Services, Sydney, Australia).

Influent and effluent samples were analyzed for a suite of stormwater constituents, including chemical oxygen demand, DOC, ammonia, nitrite + nitrate, TKN, total phosphorus, ortho phosphorous, total and dissolved cadmium, copper, zinc, lead, and chromium, and common water quality parameter including calcium, magnesium, and pH. These analyses were performed according to the following methods: EPA 410.4, EPA 9060, EPA 350.1M, EPA 353.2, EPA 351.2, EPA 365.2, EPA 365.2, EPA 200.8, EPA 6010C and EPA 6010B by a commercial lab.

CHEMICAL SPECIATION MODELING

Visual MINTEQ 3.1 (Gustafsson, 2020) was used to predict the chemical speciation of the metals within sampled stormwater based on field-measured water quality parameters. Measured input parameters included pH, calcium, magnesium, ammonia, nitrate, phosphate, DOC, and total cadmium, copper, zinc, lead, and chromium. Atmospheric CO_2 was incorporated into the model to account for the interaction between the stormwater and the atmosphere. The chloride concentration (Cl^-), a common stormwater constituent, was estimated to maintain the charge balance between anions and cations. The Gaussian Model was used to model the complexation of the influent and effluent stormwater constituents with dissolved organic matter (DOM) (Grimm et al., 1991).

TOXICITY MODELING

The interaction between metals and the biotic ligand, a surface where metal may accumulate and result in toxic effects on aquatic organisms, is a function of the water quality parameters and speciation of metals (Di Toro et al., 2001; Pagenkopf, 1983; Santore et al., 2001). A BLM (Windward Environmental, LLC, 2019, v 3.41.2.45) was used to estimate the speciation of metals within the BRC influent and effluent based on field-measured dissolved metals and water quality parameters. The BLM estimated the full site chemistry based on measured temperature, pH, DOC, and hardness levels. Washington State's median ion ratio was used to estimate the concentrations of the major cations (Ca^{2+}, Mg^{2+}, Na^{+}, and K^{+}) and anions (SO_4^{2-} and Cl^{-}), allowing the model to run with only temperature, pH, DOC, and hardness data. To improve the model, the BLM-estimated values for Ca and Mg cations were replaced with measured values as these data were readily available. The BLM further provided the concentration at which a given stormwater pollutant is expected to cause a predefined toxic effect (i.e., EC20, LC50, etc.) to a given aquatic organism's life stage based upon the general chemistry of the stormwater (Table S1, supporting information).

The BLM output was analyzed in R (Kassambara & Mundt, 2020; Knappenberger, 2017; Lê et al., 2008; R Core Team, 2024; Thorley & Schwarz, 2018; Wickham, 2007, 2016; Wickham et al., 2019; Wickham & Bryan, 2023) to evaluate the relative toxicity of the various bioretention effluents. For this study, the relative toxicity of a water sample is defined as the proportion of the measured concentration to the BLM-identified toxic limit for a given species and endpoint using the equation below:

$$\text{Relative Toxicity} = \frac{\text{Predicted Concentration}}{\text{BLM Toxic Concentration}} \qquad \text{Eq. 1}$$

A multiple-factor analysis (MFA) was used to reduce the dimensionality of the large quantitative dataset. The MFA provides a simplified visual representation of our data by creating new dimensions compromised of linear combinations of the initial variables. The number of dimensions required to describe the data adequately was determined where the eigenvalues were greater than one with a minimum of two dimensions. A cluster analysis was further performed to identify similarity groupings within the dataset (Kassambara, 2017).

A species sensitivity distribution (SSD) is a statistical distribution that describes the sensitivity of various species to a particular pollutant of concern and is commonly used to perform ecological risk assessments and develop regulatory water quality standards. Using the toxic endpoints for an aquatic community, a fitted cumulative distribution function (CFD) curve may be constructed, representing the statistical best-fit distributions for the data. The CDF curves are then utilized to estimate the hazardous concentration for a particular population percentage to be affected or predict the percentage of an aquatic community that may be affected at a given concentration. Species sensitivity distributions were constructed using the BLM-identified toxic concentrations for free copper (Cu^{2+}) for the aquatic organisms for which the BLM model predicted acute toxicities. Because the toxic concentrations of free copper depend on the water quality parameters, an SSD was generated for each BRC during each storm, and the predicted hazardous concentration for 5% (HC_5) of the aquatic community was determined. Based on measured free copper concentrations, the percentage of affected species in the aquatic community was further estimated to gauge the effectiveness of each BSM at reducing toxicity.

STATISTICAL ANALYSIS

The Tukey test was used to establish whether the mean percentages of affected species were statistically different among the four BSMs. Statistical analyses were performed using R (Hothorn et al., 2008; Mendiburu, 2023), where differences were significant at $\alpha <= 0.05$.

RESULTS AND DISCUSSION

SPECIATION SHIFTS

The MINTEQ-predicted chemical speciation for calcium, magnesium, DOM, cadmium, chromium, copper, lead, and zinc were reviewed for general trends among stormwater influent and effluent, the four varieties of BSMs, and the various dosing levels (Figures S1 through S9). Knappenberger et al. (2022) previously concluded that all of the bioretention mesocosms subject of this study reduced total and dissolved copper, lead, and zinc for all dosed storms. Yet, while the concentration of the heavy metals decreased following filtration through all BSMs and across all storms, no noticeable speciation shifts were noted.

Cadmium speciation in stormwater influent and effluent primarily consisted of Cd^{2+} and Cd-DOM complex, where free cadmium comprised 72%-93% of cadmium species and Cd-DOM complex 6%-27%. Bioretention influent was predicted to contain lead speciated as primarily Pb-DOM (77%-91%) and Pb^{2+} (7%-16%). Influent samples during storms S7 and S8 also contained observable proportions of PbCO3 (aq), $PbOH^{+}$, and $PbHCO_3^{+}$. Effluent lead species were consistent with those predicted in the influent. Using MINTEQ to investigate the cumulative effect of Cd and Pb on rainbow trout, Birceanu et al. (2008) also found that free Cd and Pb ions were prevalent predicted species, although DOM was not input to the model.

The influent stormwater contained copper predominately speciated as Cu-DOM (59%-83%), Cu^{2+} (13%-27%), $CuCO_3$(aq) (0%-18%), and $CuOH^+$ (0%-6%). CuCO3 (aq) was not predicted in effluent samples at any significant level during several storms. Copper speciation among effluent samples is relatively comparable to influent, although slightly higher percentages of Cu-DOM was observed (70%-93%). Similar to the results of this study, Bui et al. (2016) found that the majority of dissolved copper was bound with DOM in water samples from the Dongnai and Mekong Rivers in Vietnam.

Consistent with the findings of Hoang & Tong (2015), we found Zn^{2+} was the dominant species of zinc in the stormwater samples, accounting for up to 90% of zinc species influent and effluent samples, followed by Zn-DOM which ranged from up to 20% of zinc species in influent and 38% in effluent samples. Erten-Unal et al., (1998) predicted the speciation of heavy metals in solution derived from metal salts and at varying pH levels. At pH levels similar to that observed, our study agreed that Zn and Cd were predominantly speciated as the free divalent cations. However, our results differed in the predicted lead speciation, which forms a stable complex with DOM. This difference is attributed to the lack of DOC integration in the modeling of the previous study.

Although heavy metal concentrations for all dosed storms decreased following filtration, the un-dosed storm (S2) resulted in the export of copper and chromium from the BRC. Specifically, increased concentrations of Cu-DOM, Cu^{2+}, and $CuOH^+$ were observed. The export of copper from BRCs when the influent stormwater was relatively clean was also documented by Chahal et al. (2016) and Mullane et al. (2015) and is often attributed to the integration of compost. Chromium was predicted in the forms of $Cr(OH)_3$ (aq), $CrOH^{2+,}$ and $Cr(OH)_2{}^+$ with a very minor proportion as Cr-DOM.

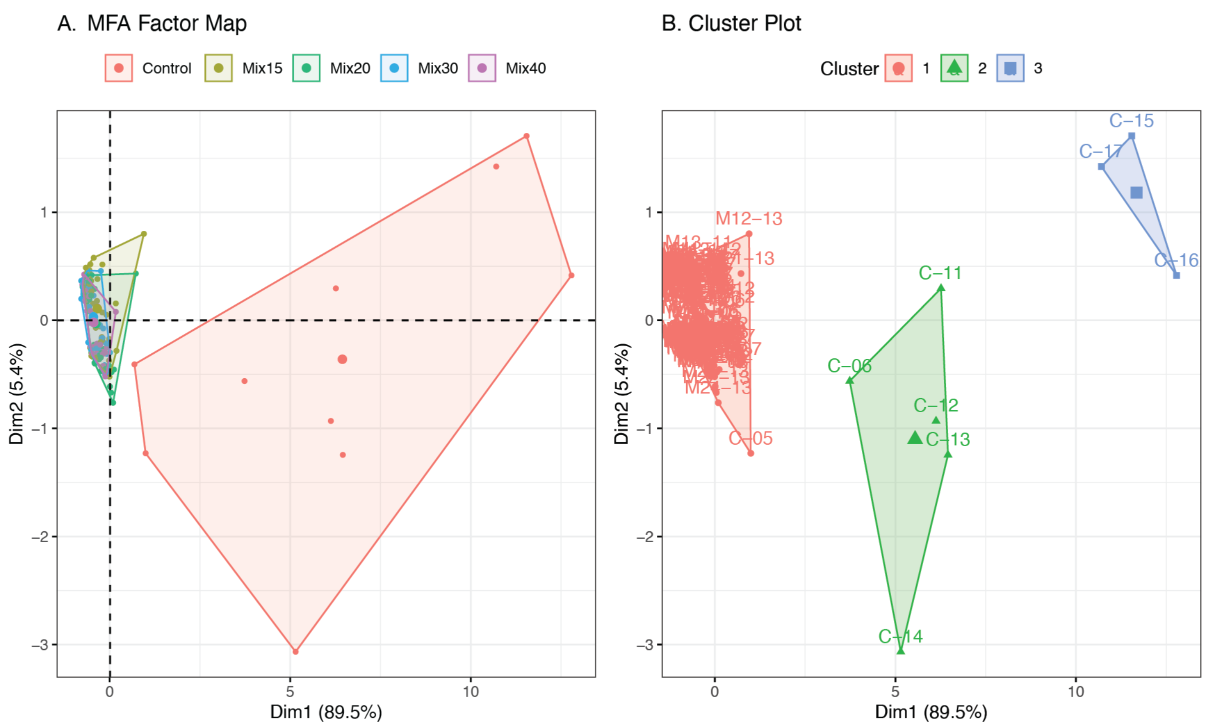


Figure 4: (A) Multiple factor analysis (MFA) of the relative toxicity of the control (influent) and bioretention effluent for each storm plotted along dimensions 1 and 2 grouped by the control and bioretention soil mixtures (BSM). (B) Cluster plot of the relative toxicity of the control (influent) and bioretention effluent for each storm from each mesocosm plotted along dimensions 1 and 2.

VARIATION AMONG SOIL MIXTURES

The MFA was conducted using the relative toxicities of the BLM-identified cadmium, copper, lead, and zinc species. The MFA cumulatively described 94.9% of the variance of the data, summarized along dimensions 1 and 2 (**Figure 4**A and **Figure 4**B). In ascending order, the variables that predominately contribute to dimension 1 are the relative toxicities of zinc, cadmium, and lead. The variables that predominately contributed to dimension 2 are the relative toxicities of cadmium species for various aquatic organisms.

Stormwater effluent from the four BSMs and control (influent) are represented in **Figure 4**A. The MFA revealed that the BSMs are distinctly different from the control with respect to dimensions 1 and 2. Overlap of the BSM polygons shown in **Figure 4**A indicates that the relative toxicity of the stormwater effluent does not substantially differ among mixtures. This contradicts the reviewed literature, which found variation among heavy metal and nutrient removal efficiencies of differing soil texture (Brown & Hunt, 2011; Davis et al., 2003; Johnson & Hunt, 2020b). The lack of variation is consistent with the previous study by Knappenberger et al. (2022), which found no effect on heavy metal removal due to the BSM's composition or system age.

The observed distinction between the control (influent) and BSM effluent implies that the toxicity of the stormwater changes as it passes through a BRC. A cluster analysis of the water samples resulted in three distinct groupings (**Figure 4**B). The clusters are best described as cluster 1: un-dosed storm (S2) influent and all BRC effluent; cluster 2: low to medium-dosed storms (S3, S4, S5, S6, S7, and S8) influent; and cluster 3: highly dosed storms (S9, S10, and S11) influent. This clustering implies that the BRCs can reduce the toxicity of stormwater to levels associated with low pollutant concentrations. The un-dosed storm (S2) control (C-05), although included in cluster 1, slightly differentiates itself from the BRC effluent cluster negatively along dimension 2 (y-axis), which is predominately controlled by the relative cadmium toxicities. The mean values for the measured water quality parameters of the three clusters revealed increased Ca^{2+}, Mg^{2+}, and DOM concentrations in clusters 2 and 3, the BRC effluents, consistent with the chemical speciation findings.

To evaluate the effectiveness of the BSMs at reducing toxicity to the aquatic community, SSDs for each mesocosm and each storm were constructed utilizing BLM-estimated acute toxicity concentrations of free copper. Using the SSDs, the HC_5 and percentage of species affected for each BSM (replicated four times) were extracted and are summarized in **Figure 5** and **Figure 6**, respectively. Although WTRs have been found to remove Cd^{2+}, Cu^{2+}, Pb^{2+}, and Zn^{2+} (Duan & Fedler, 2022) and reduce the bioaccessibility of Cu, Pb, and Zn (C. Wang et al., 2012), the presence of 15% and 10% of WTRs in Mix15 and Mix30, respectively, did not significantly differentiate these mixes from the others through speciation shifts or toxicity reduction. In fact, Mix15 was the least effective mix at reducing toxicity during most storms.

Of the four studied BSMs, Mix40 (Washington State's recommended bioretention media) yielded the most significant reduction of affected species during most storms, implying that most benefit is derived from increased quantities of organic material within the BSM. However, caution should be taken as copper has also been found to leach from compost in relatively young BRCs, and a balance between the protective effect of DOM and increased copper concentrations should be considered (Chahal et al., 2016; H. Li & Davis, 2008; Mullane et al., 2015).

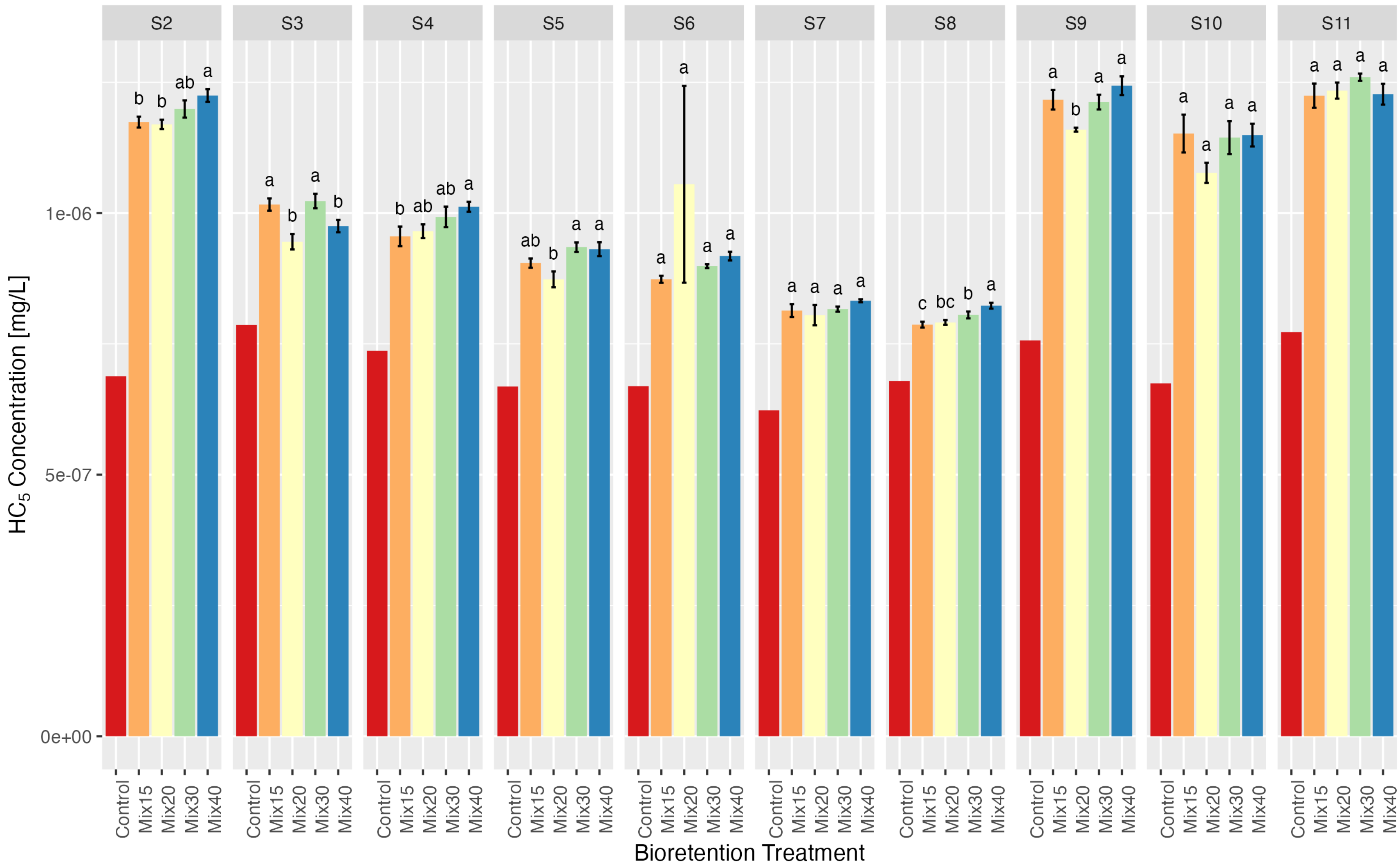


Figure 5: Bar plot of the hazardous concentrations for 5% of the aquatic community to be affected for each of the ten studied storms for the control and the mean of the four replicates for each bioretention soil mixture. Means followed by the same letter for a variable are not significantly different at $\alpha = 0.05$.

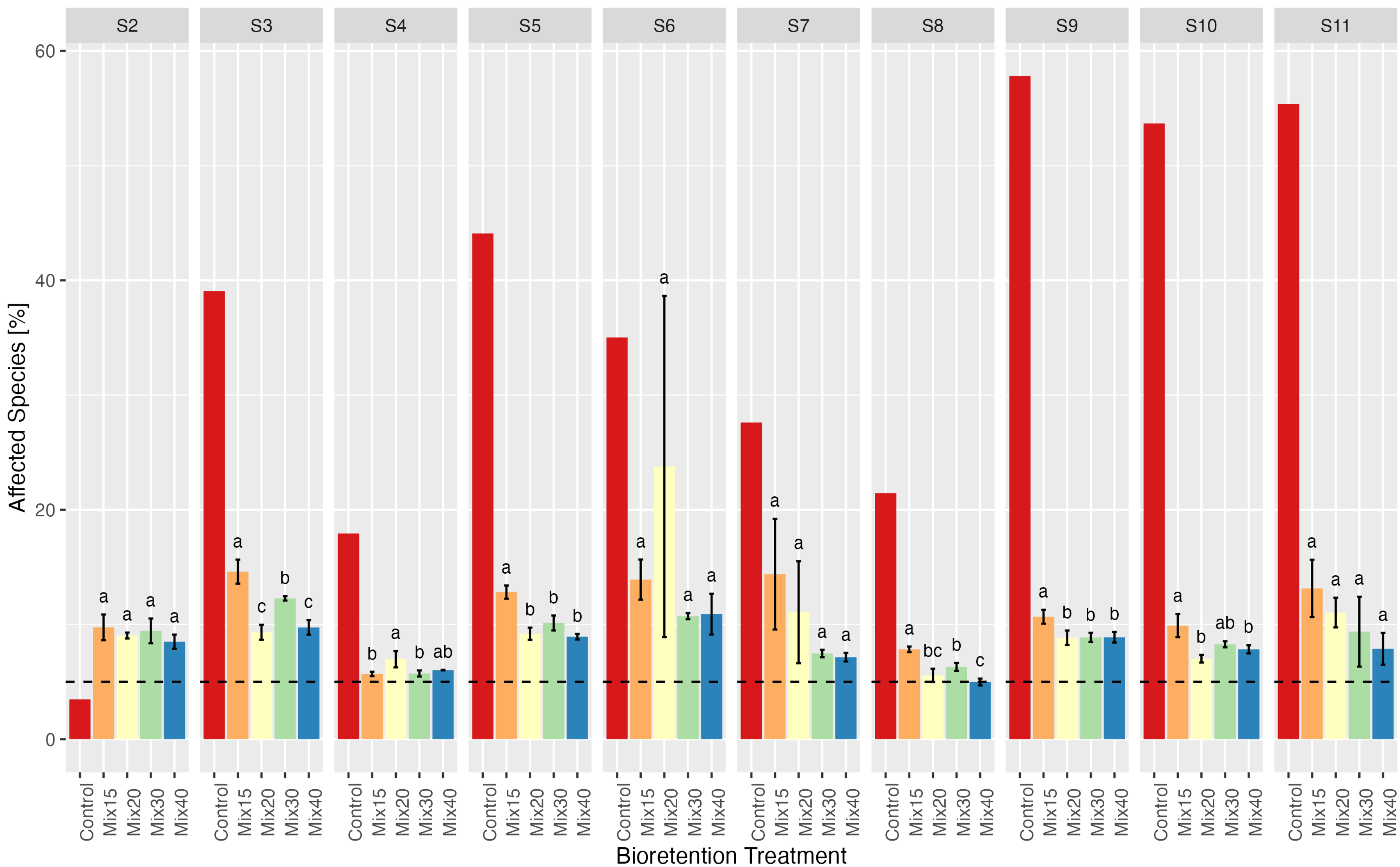


Figure 6: Bar plot of the percentage of the aquatic community anticipated to be affected by the control and the mean effluents of the four replicates for each bioretention soil mixture for each of the ten studied storms. Means followed by the same letter for a variable are not significantly different at α = 0.05.

EFFECTS OF WATER QUALITY ON TOXICITY

Filtration through the four BSMs increased the concentration of Ca^{2+}, Mg^{2+}, and DOC in all four BSMs during all storms. Bui et al. (2016) found a positive correlation between hardness (Ca^{2+} and Mg^{2+}) and the lethal concentration (LC50) of copper for three aquatic organisms. Hoang & Tong (2015) describe the same protective effect from hardness on the toxicity of zinc in freshwater snails and Kumar & Singh (2010) for cadmium in fish. This protective effect occurs as Ca^{2+} and Mg^{2+} compete with heavy metal ions for binding sites at the biotic ligand, consequently preventing accumulation at this sensitive receptor. Complexation with DOM decreases the bioavailability of free metal ions as bound metals are not available for uptake. Heavy metal complexation with DOM has been documented to reduce aquatic toxicity (Bui et al., 2016; Gillis et al., 2008; Hoang et al., 2004; Hoang & Tong, 2015; Rogevich et al., 2008). Due to the mitigating effects described by the literature, we expected to observe a benefit to aquatic species from filtration through the BSMs.

In agreement with the literature, our findings support the concept of the protective effect. We observed that as Ca^{2+}, Mg^{2+}, and DOC concentrations increased following filtration through the BSM, the Cu^{2+} HC_5 concentrations simultaneously increased. Additionally, following filtration through BRCs, the percentage of species affected by the concentration of Cu^{2+} in stormwater samples decreased for all storms except S2 (**Figure 6**), where free copper ions were exported. These observations are indicative of improvements in the toxicity of stormwater from BRCs. Storm S2 is further discussed below regarding the effect of dosing on toxicity.

Consistent with the findings of Bui et al. (2016), the dominant copper species is predicted to be complexed with DOM in all mixes and during all storms and not present as a free ion. Lead speciation also noted that lead was predominantly bound to DOM. Nonetheless, the chemical speciation model predicts that DOM is abundantly present in stormwater influent and effluent, a significant percentage of which is not complexed with heavy metals. While Ca^{2+} and Mg^{2+} may compete with the heavy metals to bind to DOM, the majority speciation of copper and lead with DOM suggests that the increased concentration of Ca^{2+} and Mg^{2+} in effluent does not compete with either heavy metal to the extent that would decrease the availability of DOM for binding.

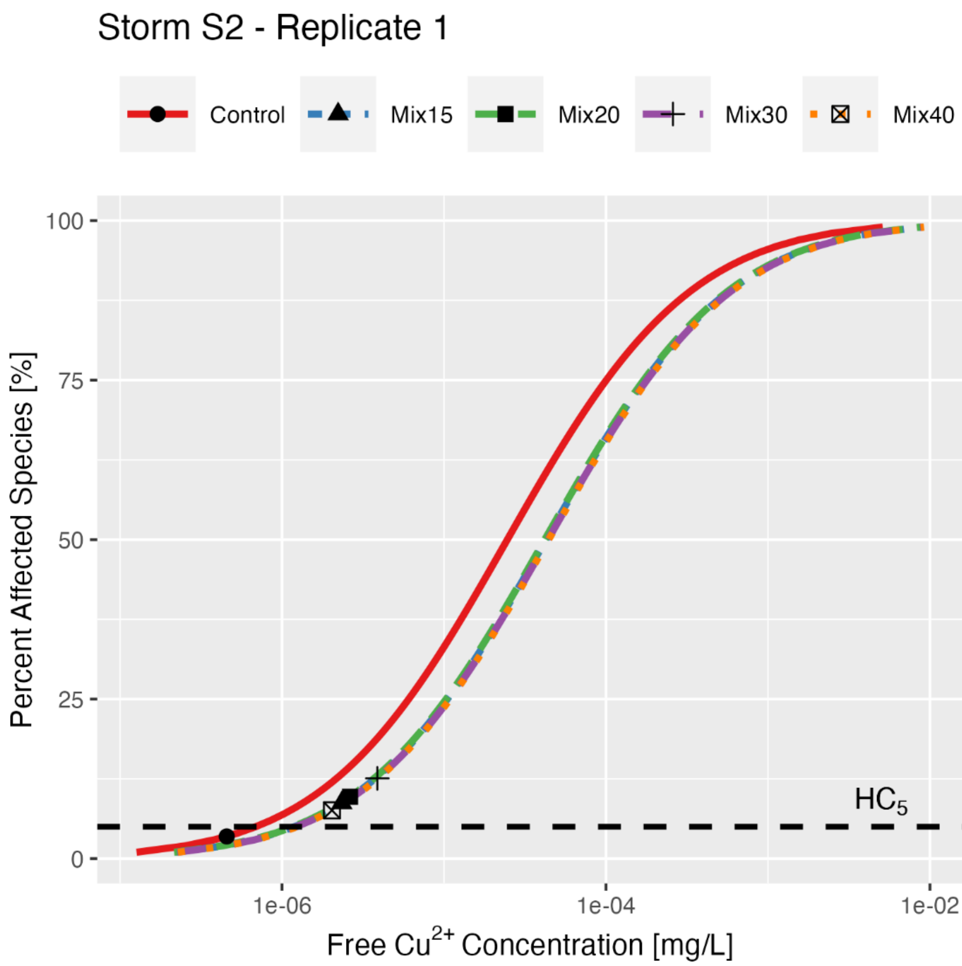


Figure 7: Species sensitivity distribution of the percentage of aquatic species affected by the concentration of free Cu^{2+} ions in the bioretention cells influent and effluent for each bioretention soil mixture and estimated free Cu^{2+} ion concentration.

EFFECTS OF DOSING ON TOXICITY

Generally, the BRCs are most effective at reducing the concentration of the toxic heavy metals in the stormwater during highly dosed storms S9 through S11. This is consistent with Sun & Davis

(2007), who reported that BRCs were most efficient at removing zinc, copper, lead, and cadmium when higher influent concentrations were encountered.

The BRC influent and effluent were compared through their relationships between the concentration of Cu^{2+} and the subsequent effects on the BLM aquatic community. **Figure 7** depicts the SSD of the control (influent) stormwater and the four BSMs for S2. All storms exhibited a rightward shift of the CDF curve, indicating that as stormwater is filtered through the BRC, a higher concentration of free copper is required to impact the aquatic community. This shift is attributable to the water quality changes discussed above and may be used to quantify the protective effect of those changes. The CDF curve shift demonstrates how the increased water quality parameters in filtered stormwater result in the free Cu^{2+} HC_5 increases observed for all BSMs in **Figure 5**.

Recall that the undosed storm (S2) effluent yielded the export of free Cu^{2+} from the BRCs. Measured free Cu^{2+} concentrations for each BSM are depicted on the CDF curves in **Figure 7**. Although the CDF curves demonstrate improved free Cu^{2+} HC_5 concentrations following filtration of S2, the exported concentrations of copper indicate a greater percentage of species affected by the stormwater effluent, corroborating the observed increase in **Figure 6**. The percentage of species affected increased from 3.5% before entering the BRCs to 8.5 - 9.4% affected by the effluent, a greater than two-fold increase attributed to the export of copper from the BRCs. Our study shows that the increased copper concentration affects the toxicity of the stormwater more than the protective effects of water quality can offset.

Nonetheless, there appears to be a benefit to toxicity from filtration through a BRC due to increased water quality parameters, even if Cu concentrations are not reduced. If a pollutant concentration were to increase following filtration, the BRC may still be able to reduce toxicity through the effects of hardness and DOC. Assessing pollutant concentrations alone does not reveal all the benefits of filtration through a BRC.

CONCLUSIONS

Our study propagated several generalized conclusions regarding the toxicity of stormwater as it moves through a bioretention soil media, including:

1. Overall cadmium, copper, lead, and zinc concentrations decreased following filtration through all BSMs and across all dosed storms, yet no noticeable speciation shifts were noted within the effluent.
2. The MFA indicates that the four BSMs are similar in terms of relative toxicity reduction to the aquatic community.
3. Bioretention cells reduce the relative toxicity of stormwater to levels associated with low pollutant concentrations.
4. Mix40 produced the greatest decrease in the percentage of species affected during most storms, implying that most benefit is derived from increased quantities of organic material within the BSM.
5. There may be a benefit to toxicity when running water through a BRC as calcium, magnesium, and dissolved organic carbon increases, even if the pollutant concentration is not reduced. Assessing pollutant concentrations alone does not reveal all the benefits of filtration through a BRC.

While generalized conclusions may be drawn from our research, there is still much to be examined in how bioretention changes the toxicity of stormwater. This study was unable to assess pollutant loadings due to the available data. Consequently, a mass balance could not be evaluated and conclusions regarding pollutant export over the life of the bioretention systems could not be drawn. We suggest that additional research be conducted to investigate the source of the leached copper and chromium.

It should be noted that the standard laboratory test methodology defines dissolved metals as everything that passes a 0.45-micron filter. While this methodology conservatively protects aquatic environments by estimating a higher concentration of dissolved metals than what may actually be present, the laboratory procedure also weakens our study because a metal must be truly dissolved to be bioavailable. Hamid et al. (2023) recently investigated the separation of particulate and dissolved phosphorus and found that 41.8% of particulate matter was smaller than 0.45 micron resulting in incorrectly defining bound phosphorus as dissolved and bioavailable. This definition also results in the inability to determine whether decreases in metal concentrations result from the physical filtering of metal-bound-particulate matter through the BSM or from the adsorption to the BSM, a subject that should be examined in future studies. As our study investigates the tendencies for the toxicity of stormwater to change as it is filtered through a BRC and does not aim to quantify toxicity, our results are valid.

CHAPTER 5 – CONCLUSION

Low impact development techniques are utilized in urban and suburban catchments for the purpose of preserving the pre-development hydrology. LID practices are assessed for effectiveness by various methods dependent on their desired function (i.e., infiltration, filtration, etc.) Through the duration of our studies, we examined runoff thresholds, a measurement commonly used to assess the effectiveness of LID practices, and a methodology for determining the effectiveness of BSM at reducing the toxicity of stormwater to an aquatic community.

Runoff thresholds computed from four linear regression models were determined to not adequately describe the precipitation and runoff data. Using the BP test it was revealed that the original datasets from all five studies were heteroscedastic and, therefore, violated an assumption of linear regression. Such heteroscedasticity may be corrected via log-transformation of the data. However, when using a linear model to assess the runoff threshold of a catchment, it is better to correct for heteroscedasticity by implementing a mixed-effect model with applied variance structures. Other indicators that the linear regression models did not adequately describe the datasets included large confidence intervals, indicating poor representation of the data by the model, and negative confidence intervals, implying that runoff occurred prior to precipitation, which is not possible. Further, there were consistent observations where runoff occurred before precipitation exceeded the runoff threshold and occurrences of no runoff following exceedance of the runoff threshold.

Catchments that incorporated LID practices resulted in higher and wider 10-90% runoff probability ranges. Higher ranges indicate that more water is stored within the watershed before discharge occurs and wider ranges indicate greater variability. These differences between traditional and LID were attributed to the fact that LID practices promote interaction between the stormwater and vegetation and the underlying soil. These conclusions are consistent with the literature in that LID practices increase runoff thresholds and that there is variability in the effectiveness of LID practices. This variability is not well captured by the finite value that linear regression analyses provide. Alternatively, presentation of runoff threshold as a range instead of a finite value, provides the user with a better understanding of how environmental factors may influence a catchment, where large p10-p90 ranges indicate high variability and more environmental interaction. The 10%-90% runoff probability range has a more accurate representation of runoff thresholds because it acknowledges the variability of runoff from LID practices, while also accurately representing the limited variability of traditional developments with smaller and more narrow ranges.

Next, the chemical speciation of heavy metals in bioretention filtered stormwater and the subsequent effects on toxicity to an aquatic community were reviewed. Heavy metal concentrations for all dosed storms from our study decreased following filtration; however, the chemical speciation data revealed no changes among the stormwater influent and effluent, the four BSMs, or the various dosing levels. Nonetheless, the MFA revealed that the bioretention cell effluents are distinctly different from the control, implying that the toxicity of the stormwater changes as it passes through a BRC, although the relative toxicity of the stormwater effluent does not substantially differ among the four tested BSMs. The cluster analysis indicated that the BRCs can reduce the toxicity to levels associated with low pollutant concentrations, like that observed in the undosed stormwater. While the BSMs were not significantly different from one another, Mix40, which is Washington State's recommended bioretention media, yielded the greatest reduction of affected species during most storms. This mix contained the highest percentage of compost at 40%, implying that benefit may be derived from larger quantities of organic material in mixes. However, caution should be taken as copper has also been found to leach from compost.

Additionally, all mixes and all storms had exported Ca^{2+}, Mg^{2+}, and DOC. Due to the protective effects of the water quality parameters noted in the literature, an improvement to toxicity following filtration was expected and, in fact, following filtration the HC_5 concentrations simultaneously increased. Filtration through a BRC revealed reduced percentages of species affected by free Cu^{2+} in all storms, except S2, where free copper was exported. Comparison of influent and effluent SSDs indicates that as stormwater is filtered through a BRC, a higher concentration of free Cu^{2+} is required to impact the aquatic community, attributed to the protective effect of the increased Ca^{2+}, Mg^{2+}, and DOC. However, due to the export of free Cu^{2+} during S2, the undosed storm, results in a greater percentage of species affected by the stormwater effluent. We concluded that the increased copper concentration affects the toxicity of the stormwater more than the protective effects of the exported water quality parameters can offset. Nonetheless, the increased HC_5 concentrations indicate there is still a benefit to toxicity from filtration through a BRC due to increased water quality parameters.

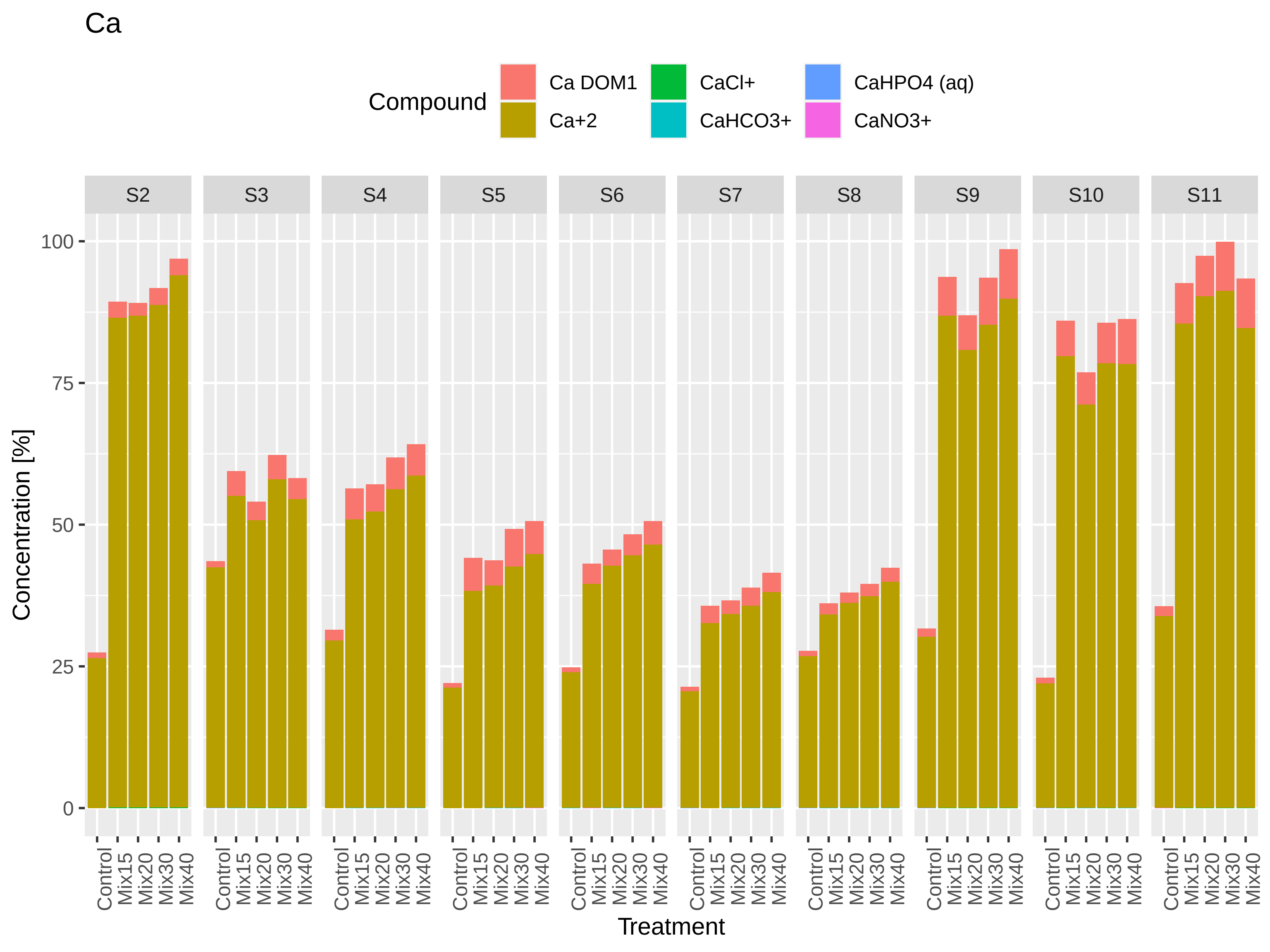
A model that does not adequately describe the data will be deficient in estimation and can lead to flawed conclusions. Further, measurements that do not sufficiently account for the proportions of a constituent which are both present and available will result in overestimation in toxicity. Using accurate methods and measurements is essential to assessing the true effectiveness of a LID practice.

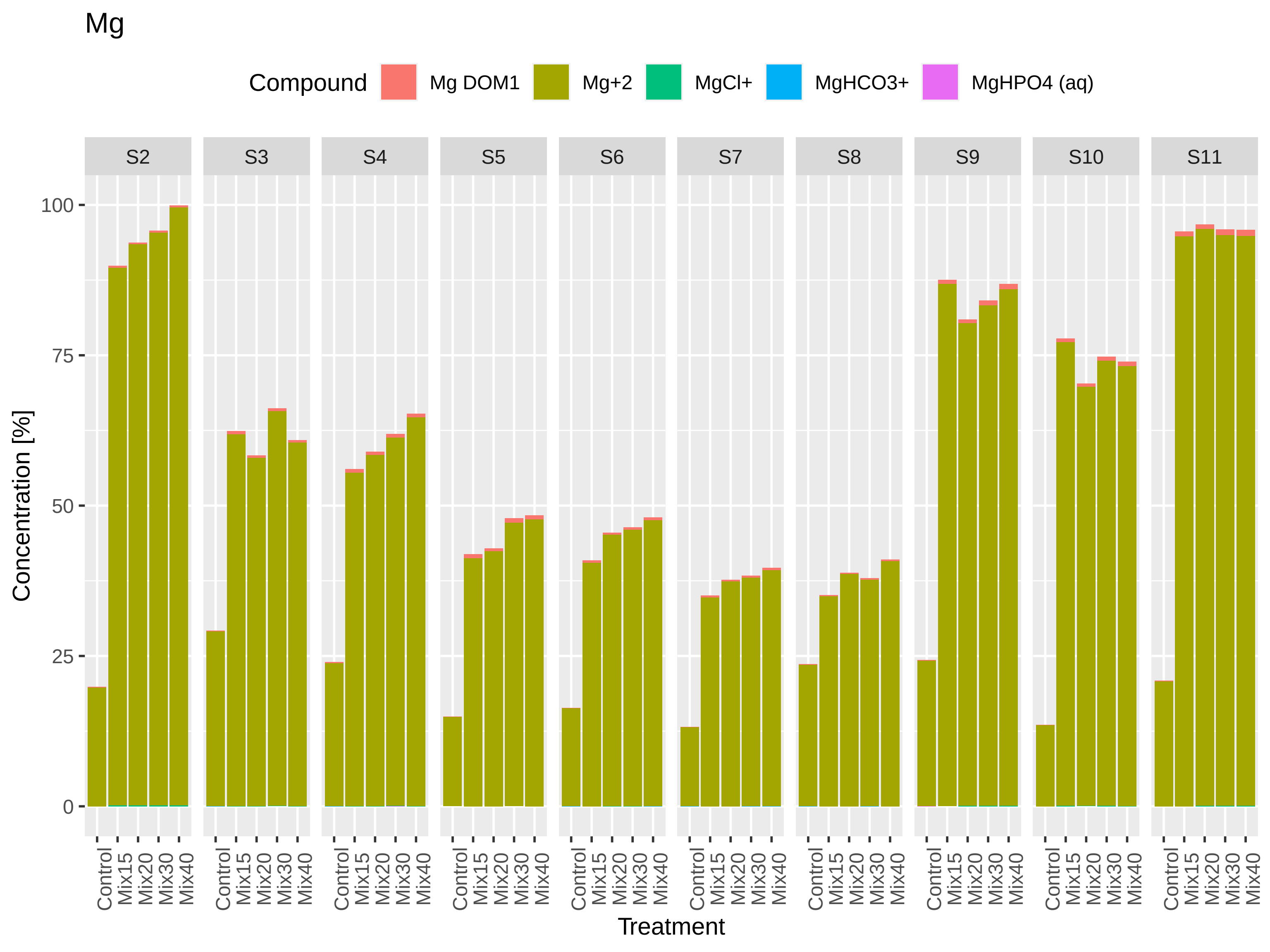
APPENDIX A – SUPPLEMENTAL DATA

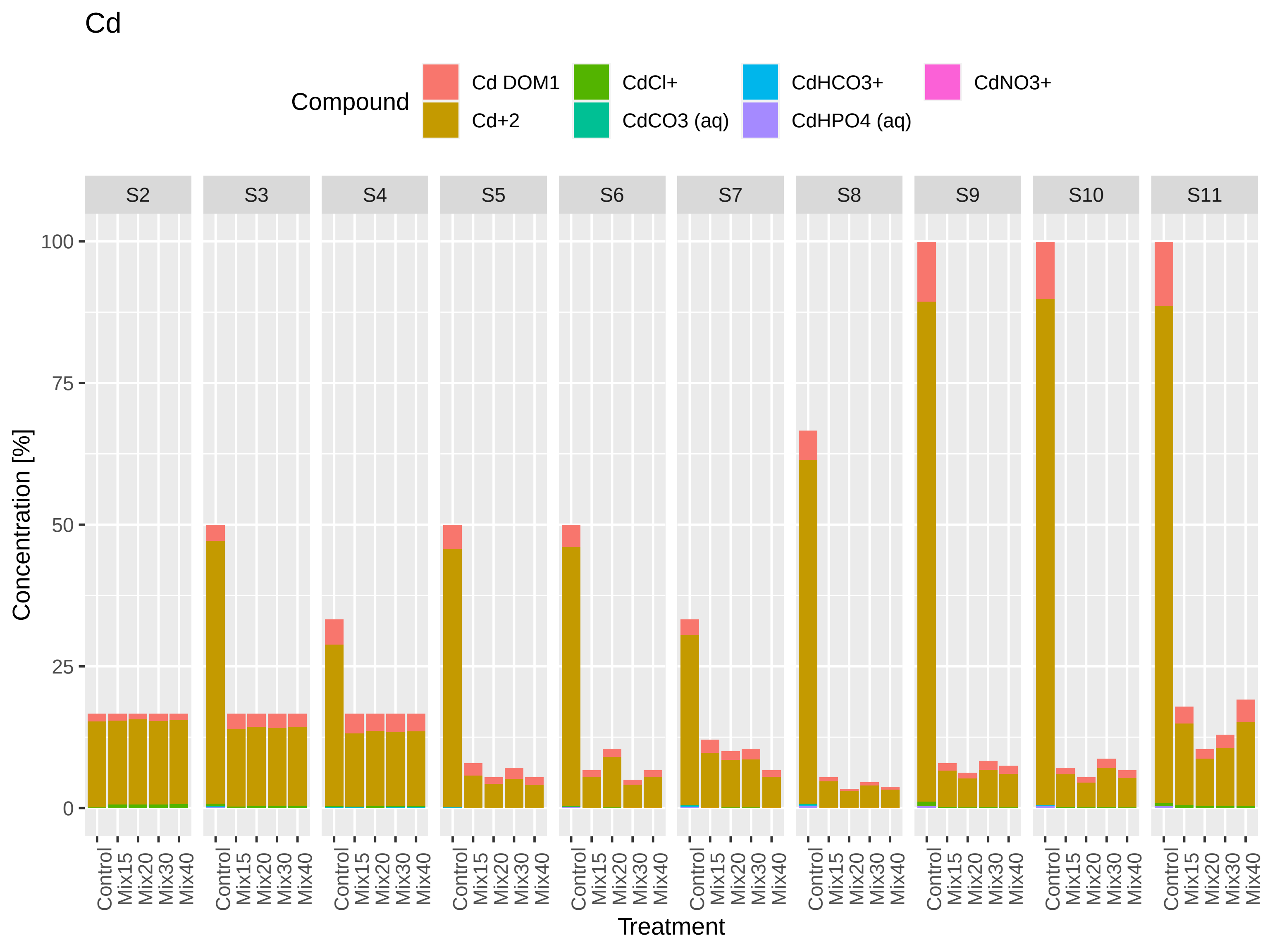
Table S1: Summary of Biotic Ligand Model Toxicity Parameters

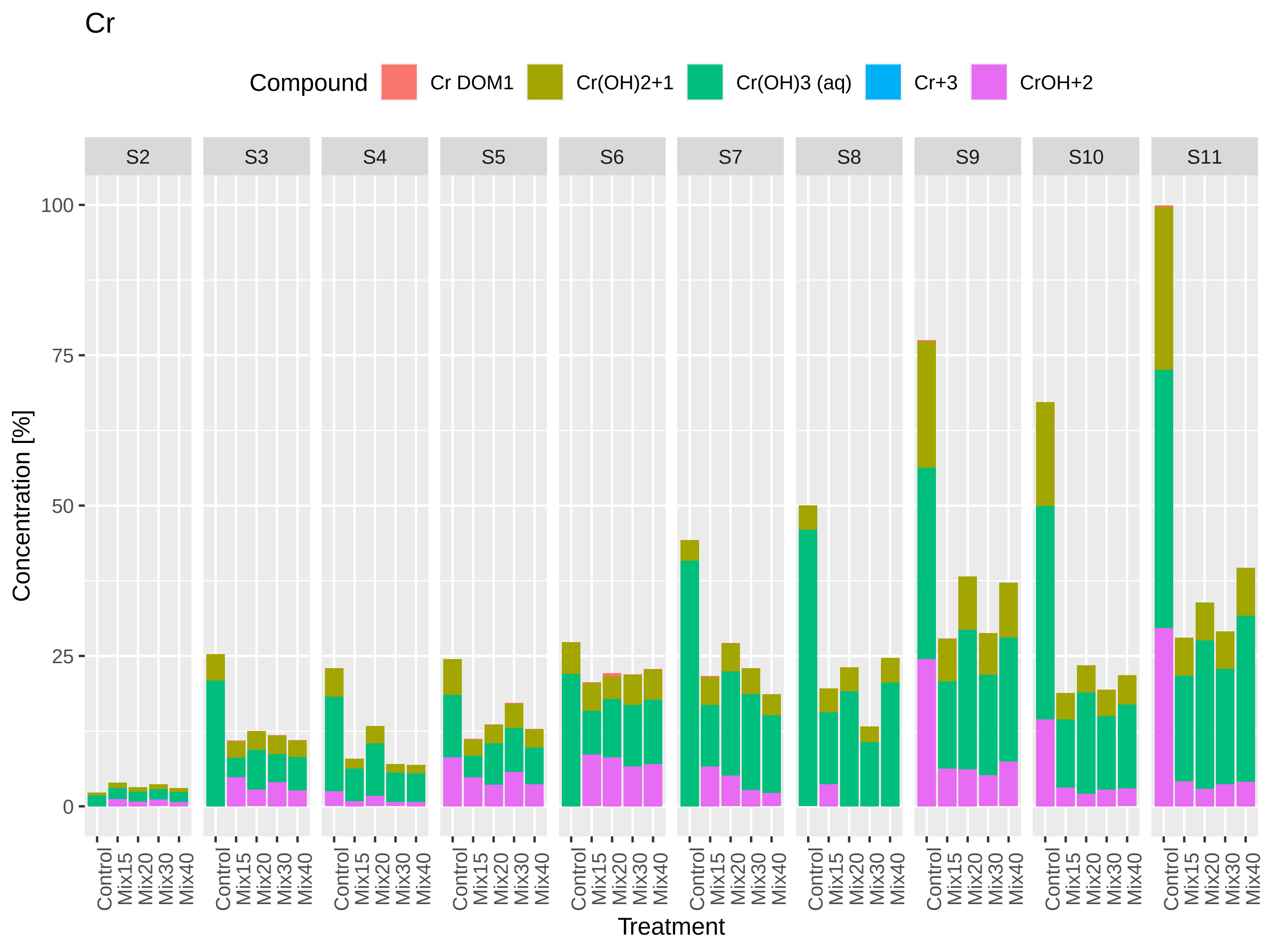
Metal	Organism	Test Type	Lifestage	Endpoint	Quantifier	References	Misc
Cd	Ceriodaphnia dubia						Organism name interpretted from the file name ("Cd_Ceriodaphnia_dubia_04-03-26.dat")
Cd	Oncorhynchus mykiss						Organism name interpretted from the file name ("Cd_Rainbow_Trout_04-03-26.dat")
Cd	Pimephales promelas						Organism name interpretted from the file name ("Cd_Fathead_Minnow_04-03-26.dat")
Cu	Acipenser transmontanus	Acute	Columbia River, 26 dph, 0.08 g, 2.5 cm; Kootenai River, 38 dph, 0.07g, 2.4 cm; Juvenile, 40 dph		LC50	Little et al. 2012; Vardy et al. 2013	SMEA calcuated by geomean
Cu	Acipenser transmontanus	Chronic	Larva, 1 dph, 12.7 mm, 8.57 mg dry weight	growth-dry weight	EC20	Wang et al. 2014	SMEA calcuated by geomean
Cu	Ceriodaphnia dubia	Acute	Neonate (<24 hr)		EC50;LC50	Gensemer et al. 2002; Hyne et al. 2005; Naddy et al. 2002; Naddy et al. 2003; Van Genderen et al. 2007	SMEA calculated by median
Cu	Ceriodaphnia dubia	Acute	Neonate (<24 hr)	Survival	EC50	Larry walker Associates, 2013	
Cu	Ceriodaphnia dubia	Chronic	Neonate (<24 hr)	Reproduction - # of youn	EC20	Wang et al. 2011a	SMEA calcuated by geomean
Cu	Chirinomus tentans	Acute	Larva, 1st instar	death and immobility	EC50	Gauss et al. 1985	SMEA calcuated by geomean
Cu	Daphnia magna	Acute				From the "Cu_Daphnia_Magna_06-10-07.DAT parameter file	
Cu	Daphnia magna	Acute	Neonate (<24 hr)	death and immobility	LC50;EC50	Al-Reasi et al. 2012; Fulton and Meyer 2014; Ryan et al. 2009; Villavicencio et al. 2011	SMEA calcuated by geomean
Cu	Daphnia magna	Chronic	Neonate (<24 hr)	Reproduction	MATC	De Schamphelaere and Janssen 2004	SMEA calculated by median
Cu	Daphnia pulex	Acute				From the "Cu_Daphnia_Pulex_06-10-07.DAT parameter file	
Cu	Daphnia pulex	Acute	Neonate (<24 hr)		LC50	Van Genderan et al. 2007	SMEA calcuated by geomean
Cu	Daphnia pulex	Chronic	Neonate (<24 hr)	Survival	EC20	Winner 1985	SMEA calcuated by geomean
Cu	Daphnia pulicaria	Acute			LC50	Lind et al. Manuscript 1978	SMEA calcuated by geomean
Cu	Lampsilis fasciola	Acute	Glochidia		EC50	Gillis et al. 2008	SMEA calculated by median
Cu	Lampsilis siliquoidea	Acute	Glochidia		LC50	Wang et al. 2007a; 2007c; Wang et al. 2007c	SMEA calcuated by geomean
Cu	Lampsilis siliquoidea	Acute	Juvenile		LC50	Wang et al. 2009	SMEA calculated by median
Cu	Lepomis macrochirus	Acute	4.2 cm, 0.59g		LC50	Inglis and Davis 1972	SMEA calcuated by geomean
Cu	Oncorhynchus tshawytscha	Acute	swim-up, 0.36-0.45 g		LC50	Welsh et al. 2000	SMEA calcuated by geomean
Cu	Oncorhynchus mykiss	Acute				From the "Cu_Rainbow_Trout_06-10-07.DAT parameter file	
Cu	Oncorhynchus mykiss	Acute	swim-up, 0.25 g		LC50	Cacela et al. 1996	SMEA calcuated by geomean
Cu	Pimephales promelas	Acute				From the "Cu_Fathead_Minnow_06-10-07.DAT parameter file	
Cu	Pimephales promelas	Acute	Larva, 1.7d; <24hr, 0.68 mg		LC50	Lind et al. Manuscript 1978; Van Genderen et al. 2007; Welsh et al. 1993	SMEA calculated by median
Cu	Pimephales promelas	Chronic	Larva, <24 hr	Biomass	EC50	Besser et al. 2001; 2005	SMEA calcuated by geomean
Cu	Utterbackia imbecillis	Acute	Juvenile		LC50	Keller unpublished 2000 memo	SMEA calcuated by geomean
Cu	Villosa iris	Chronic	3 mo. old, 2.0 mm	Growth-weight	EC20	Wang et al. 2011a	SMEA calcuated by geomean
Pb	Ceriodaphnia dubia						Organism name interpretted from the file name ("Pb_Ceriodaphnia_dubia_2015-02-18.dat")
Pb	Daphnia magna						Organism name interpretted from the file name ("Pb_Daphnia_Magna_2015-02-18.dat")
Pb	Oncorhynchus mykiss						Organism name interpretted from the file name ("Pb_Rainbow_Trout_2015-02-18.dat")
Pb	Pimephales promelas						Organism name interpretted from the file name ("Pb_Fathead_Minnow_2015-02-18.dat")
Zn	Ceriodaphnia dubia						Organism name interpretted from the file name ("Pb_Ceriodaphnia_dubia_2015-10-27.dat")

Cd; cadmium, Cu; copper, dph; days post hatch, EC20; 20% maximal effective concentration, EC50; half maximal effective concentration, LC50; lethal concentration MATC; maximum allowable concentration, Pb; lead, SMEA; species mean effect accumulation, Zn; zinc

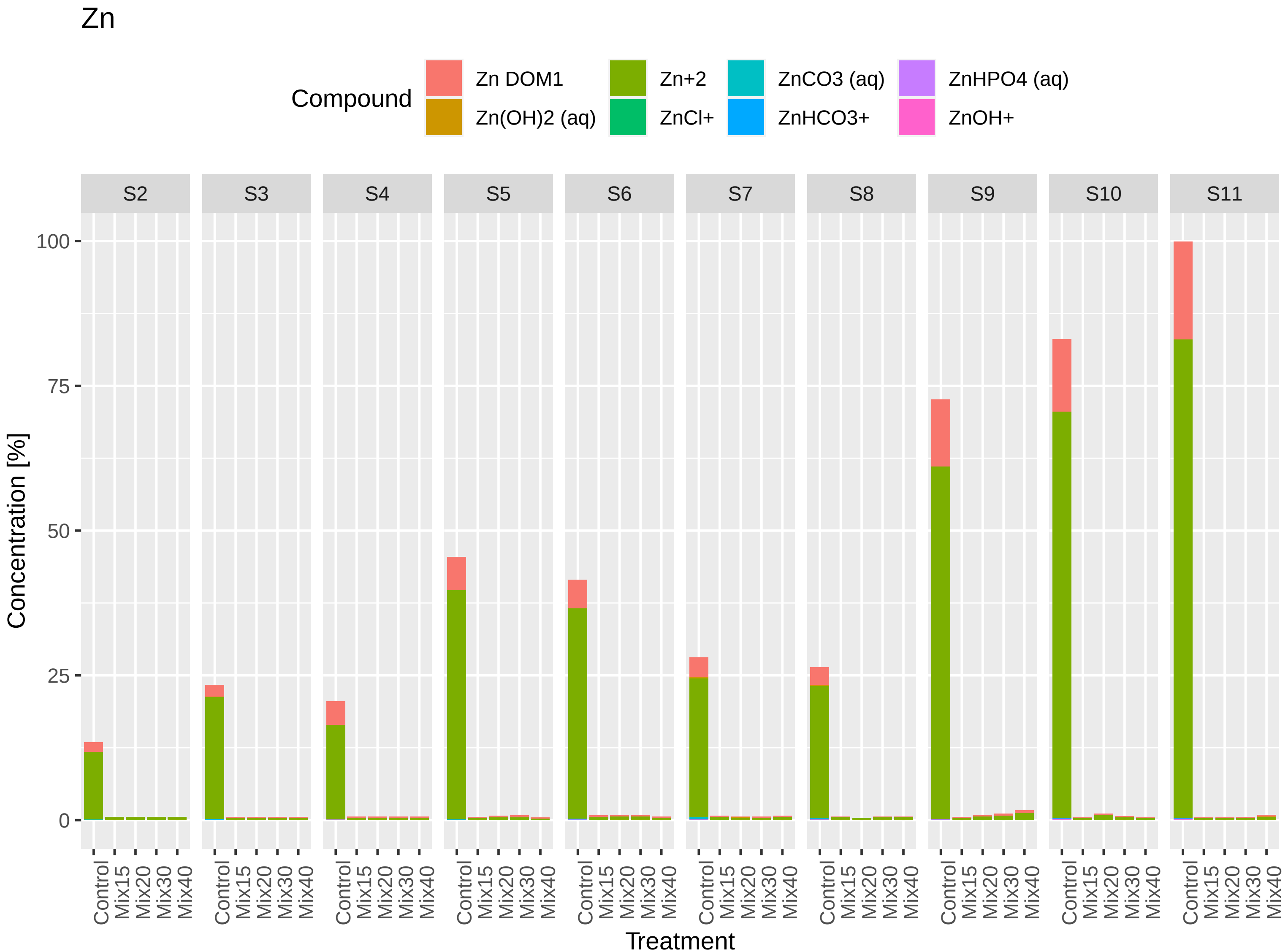

Ca
Compound
Ca DOM1
Ca+2
CaCl+
CaHCO3+
CaHPO4 (aq)
CaNO3+
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
100
75
50
25
0
Concentration [%]
Control
Mix15
Mix20
Mix30
Mix40
Treatment


Mg
Compound
Mg DOM1
Mg+2
MgCl+
MgHCO3+
MgHPO4 (aq)
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment

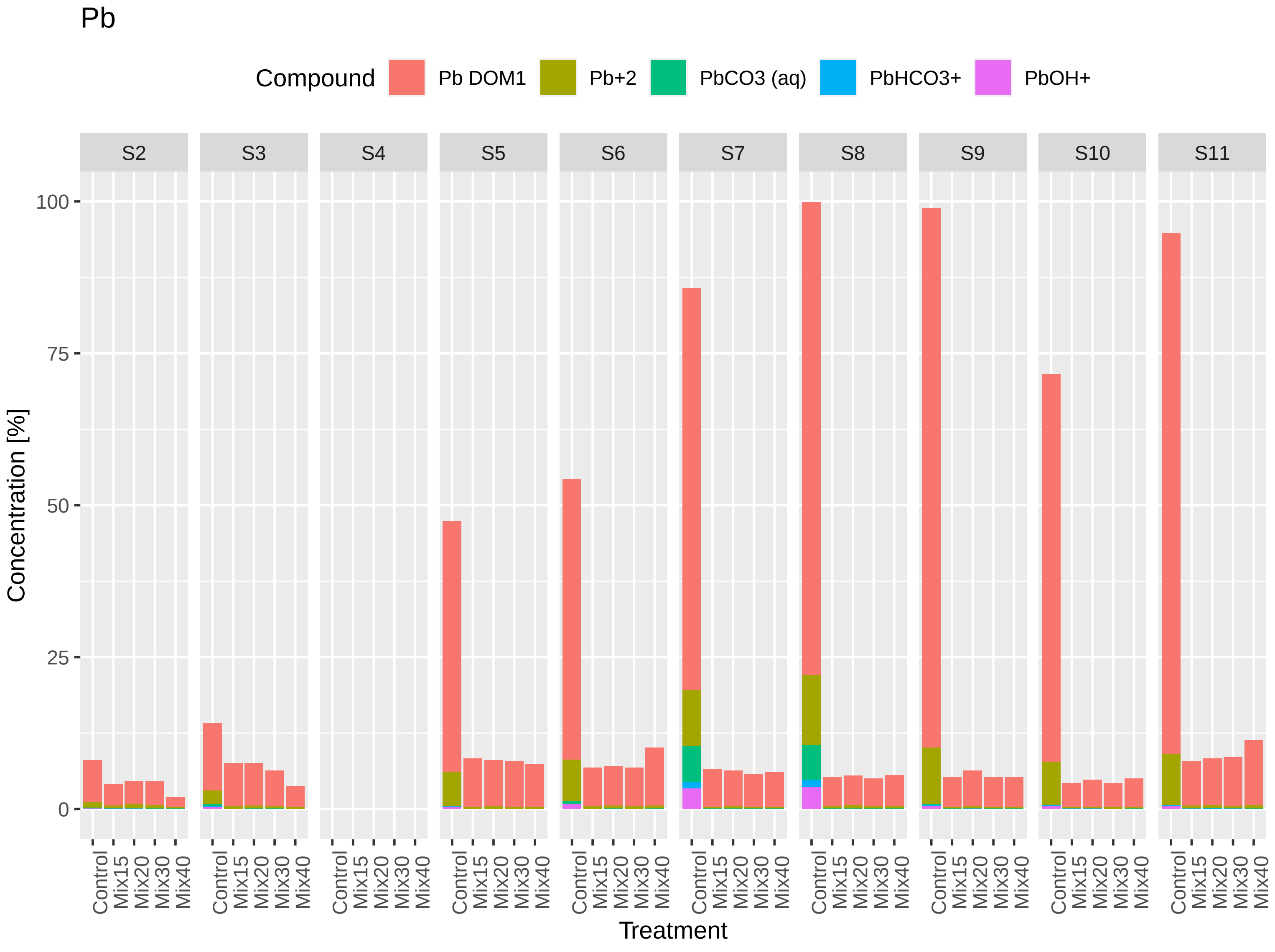

Cd
Compound
Cd DOM1
Cd+2
CdCl+
CdCO3 (aq)
CdHCO3+
CdHPO4 (aq)
CdNO3+
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment

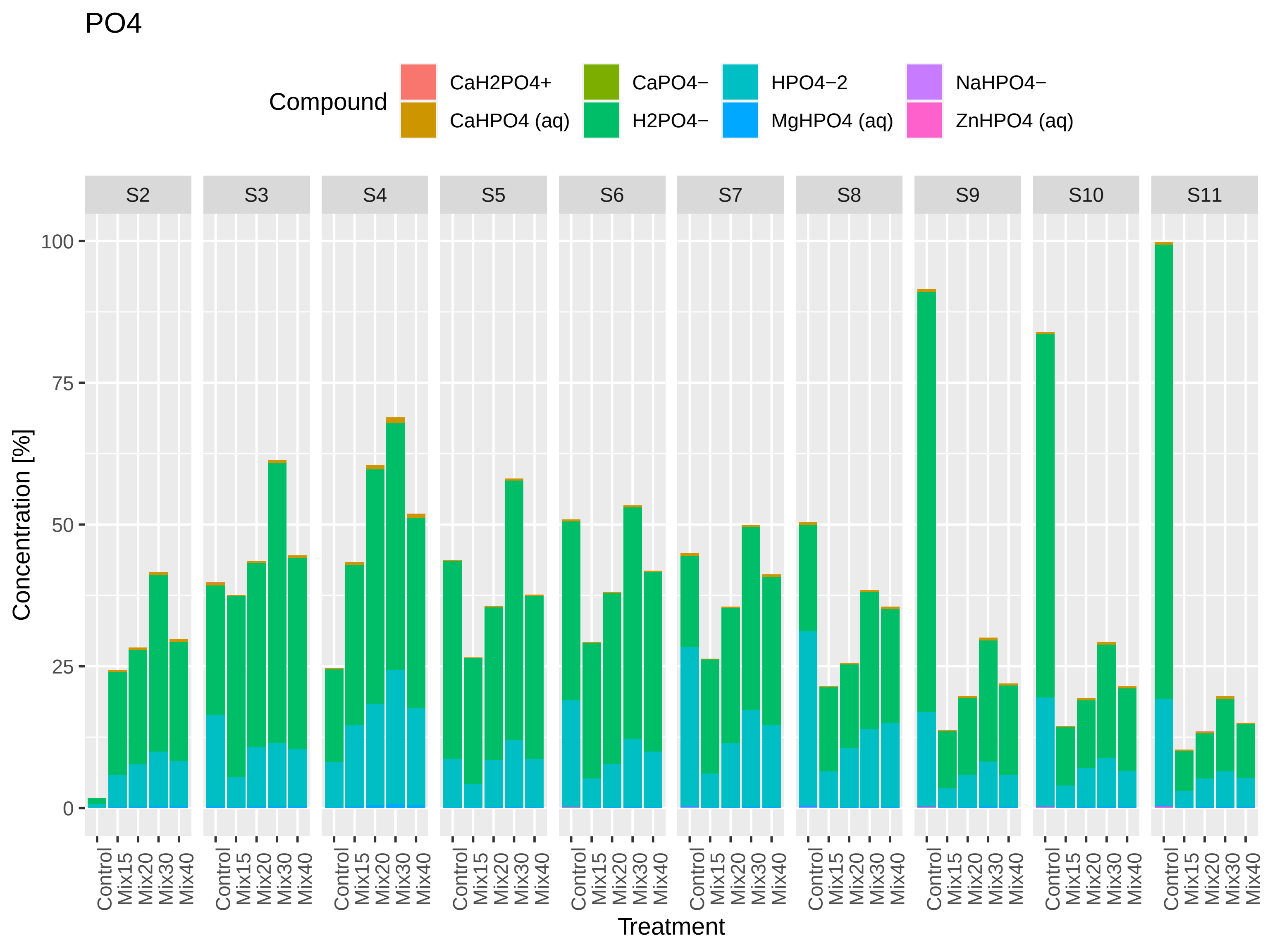


Cr
Compound
Cr DOM1
Cr(OH)2+1
Cr(OH)3 (aq)
Cr+3
CrOH+2
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment

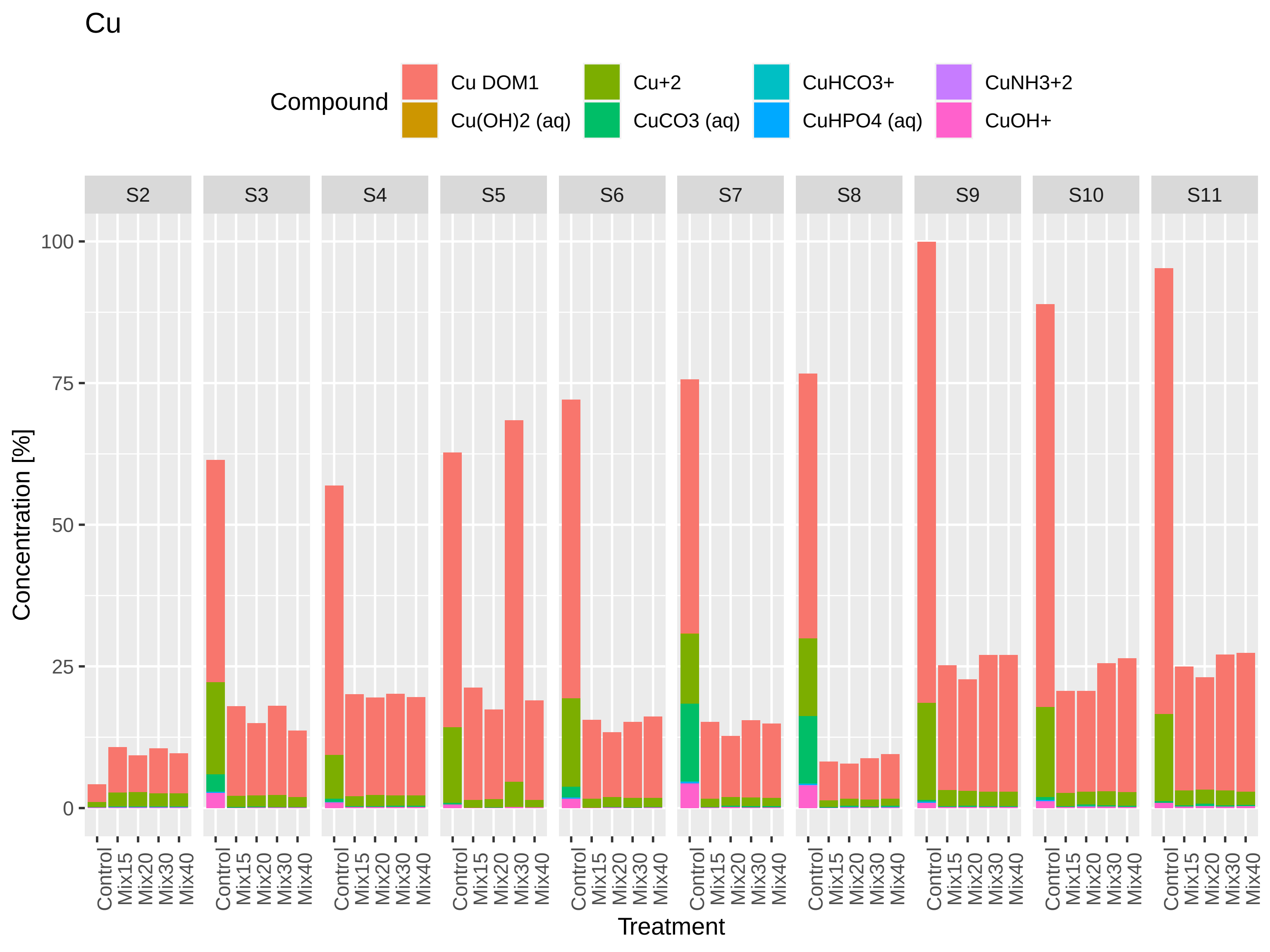


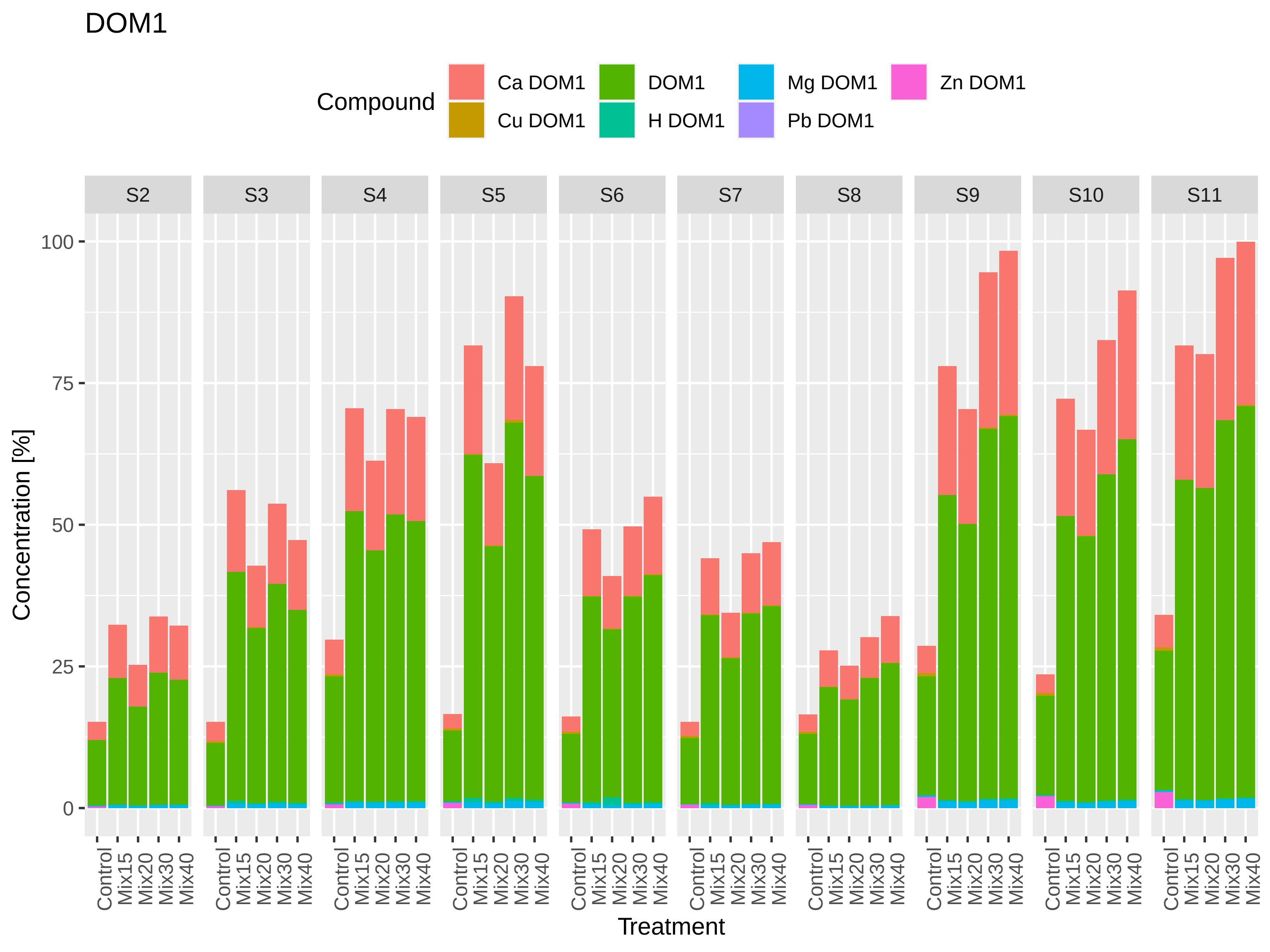
Zn
Compound
Zn DOM1
Zn(OH)2 (aq)
Zn+2
ZnCl+
ZnCO3 (aq)
ZnHCO3+
ZnHPO4 (aq)
ZnOH+
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment


Pb
Compound
Pb DOM1
Pb+2
PbCO3 (aq)
PbHCO3+
PbOH+
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment



PO4
Compound
CaH2PO4+
CaHPO4 (aq)
CaPO4−
H2PO4−
HPO4−2
MgHPO4 (aq)
NaHPO4−
ZnHPO4 (aq)
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
0
25
50
75
100
Control
Mix15
Mix20
Mix30
Mix40
Treatment


Cu
Compound
Cu DOM1
Cu(OH)2 (aq)
Cu+2
CuCO3 (aq)
CuHCO3+
CuHPO4 (aq)
CuNH3+2
CuOH+
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
Concentration [%]
100
75
50
25
0
Control
Mix15
Mix20
Mix30
Mix40
Treatment


DOM1
Compound
Ca DOM1
Cu DOM1
DOM1
H DOM1
Mg DOM1
Pb DOM1
Zn DOM1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
100
75
50
25
0
Concentration [%]
Control
Mix15
Mix20
Mix30
Mix40
Treatment

Bibliography

Ali, G., Oswald, C. J., Spence, C., Cammeraat, E. L. H., McGuire, K. J., Meixner, T., & Reaney, S. M. (2013). Towards a unified threshold-based hydrological theory: Necessary components and recurring challenges. *Hydrological Processes*, *27*(2), 313–318. https://doi.org/10.1002/hyp.9560

Ball, J. E., & Rankin, K. (2010). The hydrological performance of a permeable pavement. *Urban Water Journal*, *7*(2), 79–90. https://doi.org/10.1080/15730620902969773

Bean, E. Z., Hunt, W. F., & Bidelspach, D. A. (2007). Evaluation of Four Permeable Pavement Sites in Eastern North Carolina for Runoff Reduction and Water Quality Impacts. *Journal of Irrigation and Drainage Engineering*, *133*(6), 583–592. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:6(583)

Birceanu, O., Chowdhury, M., Gillis, P., Mcgeer, J., Wood, C., & Wilkie, M. (2008). Modes of metal toxicity and impaired branchial ionoregulation in rainbow trout exposed to mixtures of Pb and Cd in soft water. *Aquatic Toxicology*, *89*(4), 222–231. https://doi.org/10.1016/j.aquatox.2008.07.007

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society.*, *26*(2), 211–252.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287–1294. https://doi.org/10.2307/1911963

Brown, R. A., & Hunt, W. F. (2011). Underdrain Configuration to Enhance Bioretention Exfiltration to Reduce Pollutant Loads. *Journal of Environmental Engineering*, *137*(11), 1082–1091. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000437

Bui, T.-K. L., Do-Hong, L. C., Dao, T.-S., & Hoang, T. C. (2016). Copper toxicity and the influence of water quality of Dongnai River and Mekong River waters on copper bioavailability and toxicity to three tropical species. *Chemosphere*, *144*, 872–878. https://doi.org/10.1016/j.chemosphere.2015.09.058

Camponelli, K., Lev, S., Snodgrass, J., Landa, E., & Casey, R. (2010). Chemical fractionation of Cu and Zn in stormwater, roadway dust and stormwater pond sediments. *Environmental Pollution*, *158*(6), 2143–2149. https://doi.org/10.1016/j.envpol.2010.02.024

Chahal, M. K., Shi, Z., & Flury, M. (2016). Nutrient leaching and copper speciation in compost-amended bioretention systems. *Science of The Total Environment*, *556*, 302–309. https://doi.org/10.1016/j.scitotenv.2016.02.125

Crawley, M. J. (2015). *Statistics: An Introduction Using R* (Second). Wiley.

Davis, A. (2008). Field Performance of Bioretention: Hydrology Impacts. *Journal of Hydrologic Engineering*, *13*(2), 90–95. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:2(90)

Davis, A., Shokouhian, M., & Ni, S. (2001). Loading estimates of lead, copper, cadmium, and zinc in urban runoff from specific sources. *Chemosphere*, *44*(5), 997–1009. https://doi.org/10.1016/S0045-6535(00)00561-0

Davis, A., Shokouhian, M., Sharma, H., & Minami, C. (2006). Water Quality Improvement through Bioretention Media: Nitrogen and Phosphorus Removal. *Water Environment Research*, *78*(3), 284–293. https://doi.org/10.2175/106143005X94376

Davis, A., Shokouhian, M., Sharma, H., Minami, C., & Winogradoff, D. (2003). Water Quality Improvement through Bioretention: Lead, Copper, and Zinc Removal. *Water Environment Research*, *75*(1), 73–82. https://doi.org/10.2175/106143003X140854

Detty, J. M., & McGuire, K. J. (2010). Threshold changes in storm runoff generation at a till-mantled headwater catchment. *Water Resources Research*, *46*(7). https://doi.org/10.1029/2009WR008102

Di Toro, D. M., Allen, H. E., Bergman, H. L., Meyer, J. S., Paquin, P. R., & Santore, R. C. (2001). Biotic ligand model of the acute toxicity of metals. 1. Technical Basis. *Environmental Toxicology and Chemistry*, *20*(10), 2383–2396. https://doi.org/10.1002/etc.5620201034

Dietz, M. E. (2007). Low Impact Development Practices: A Review of Current Research and Recommendations for Future Directions. *Water, Air, and Soil Pollution*, *186*(1–4), 351–363. https://doi.org/10.1007/s11270-007-9484-z

Dietz, M. E., & Clausen, J. C. (2005). A Field Evaluation of Rain Garden Flow and Pollutant Treatment. *Water, Air, and Soil Pollution*, *167*(1–4), 123–138. https://doi.org/10.1007/s11270-005-8266-8

Duan, R., & Fedler, C. B. (2022). Competitive adsorption of Cu2+, Pb2+, Cd2+, and Zn2+ onto water treatment residuals: Implications for mobility in stormwater bioretention systems. *Water Science and Technology*, *86*(5), 878–893. https://doi.org/10.2166/wst.2022.258

Erten-Unal, M., Wixson, B., Gale, N., & Pitt, J. (1998). Evaluation of toxicity, bioavailability and speciation of lead, zinc and cadmium in mine/mill wastewaters. *Chemical Speciation and Bioavailability*, *10*(2), 37–46. https://doi.org/10.3184/095422998782775826

Eshghizadeh, M., Talebi, A., & Dastorani, M. (2018). Thresholds of land cover to control runoff and soil loss. *Hydrological Sciences Journal*, *63*(9), 1424–1434. https://doi.org/10.1080/02626667.2018.1511905

Ferguson, B. (1995). Storm-Water Infiltration for Peak-Flow Control. *Journal of Irrigation and Drainage Engineering*, *6*, 369–469. https://doi.org/10.1061/(ASCE)0733-9437(1995)121:6(463)

Ferguson, B., & Deak, T. (1994). Role of Urban Storm-Flow Volume in Local Drainage Problems. *Journal of Water Resources Planning and Management*, *120*(4), 405–572. https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(523)

Fink, D. H., & Frasier, G. W. (1977). Evaluating Weathering Characteristics of Water-harvesting Catchments from Rainfall-runoff Analyses. *Soil Science Society of America Journal*, *41*(3), 618–622. https://doi.org/10.2136/sssaj1977.03615995004100030040x

Freni, G., Mannina, G., & Viviani, G. (2010). Urban Storm-Water Quality Management: Centralized versus Source Control. *Journal of Water Resources Planning and Management*, *136*(2), 268–278. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:2(268)

Gillis, P. L., Mitchell, R. J., Schwalb, A. N., McNichols, K. A., Mackie, G. L., Wood, C. M., & Ackerman, J. D. (2008). Sensitivity of the glochidia (larvae) of freshwater mussels to copper: Assessing the effect of water hardness and dissolved organic carbon on the sensitivity of endangered species. *Aquatic Toxicology*, *88*(2), 137–145. https://doi.org/10.1016/j.aquatox.2008.04.003

Glass, C., & Bissouma, S. (2005). Evaluation of a parking lot bioretention cell for removal of stormwater pollutants. *WIT Transactions on Ecology and the Environment*, *81*.

Grimm, D. M., Azarraga, L. V., Carreira, L. A., & Susetyo, W. (1991). Continuous multiligand distribution model used to predict the stability constant of copper(II) metal complexation with humic material from fluorescence quenching data. *Environmental Science & Technology*, *25*(8), 1427–1431. https://doi.org/10.1021/es00020a010

Gustafsson, J. P. (2020). *Visual MINTEQ* (3.1) [Visual Basic 2012; Windows]. KTH.

Hamid, A., Wilson, A., Gladfelter, M., & Knappenberger, T. (2023). Long-term missing role of small colloids and nanoparticles on the loading and speciation of phosphorus in catfish aquaculture ponds in west Alabama. *Chemosphere*, *340*, 139906. https://doi.org/10.1016/j.chemosphere.2023.139906

Hoang, T. C., Tomasso, J. R., & Klaine, S. J. (2004). Influence of Water Quality and Age on Nickel Toxicity to Fathead Minnows (Pimephales promelas). *Environmental Toxicology and Chemistry*, *23*(1), 86. https://doi.org/10.1897/03-11

Hoang, T. C., & Tong, X. (2015). Influence of water quality on zinc toxicity to the Florida apple snail (*Pomacea paludosa*) and sensitivity of freshwater snails to zinc. *Environmental Toxicology and Chemistry*, *34*(3), 545–553. https://doi.org/10.1002/etc.2827

Hood, M. J., Clausen, J. C., & Warner, G. S. (2007). Comparison of Stormwater Lag Times for Low Impact and Traditional Residential Development1. *JAWRA Journal of the American Water Resources Association*, *43*(4), 1036–1046. https://doi.org/10.1111/j.1752-1688.2007.00085.x

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (Vol. 398). John Wiley & Sons.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, *50*(3), 346–363.

Huang, J., Xu, C., Ridoutt, B., Wang, X., & Ren, P. (2017). Nitrogen and phosphorus losses and eutrophication potential associated with fertilizer application to cropland in China.

Journal of Cleaner Production, *159*, 171–179. https://doi.org/10.1016/j.jclepro.2017.05.008

Hunt, W. F., Jarrett, A. R., Smith, J. T., & Sharkey, L. J. (2006). Evaluating Bioretention Hydrology and Nutrient Removal at Three Field Sites in North Carolina. *Journal of Irrigation and Drainage Engineering*, *132*(6), 600–608. https://doi.org/10.1061/(ASCE)0733-9437(2006)132:6(600)

Johnson, J. P., & Hunt, W. F. (2020a). *Field Assessment of the Hydrologic Mitigation Performance of Three Aging Bioretention Cells*. https://doi.org/10.1061/JSWBAY.0000925

Johnson, J. P., & Hunt, W. F. (2020b). Field Assessment of the Hydrologic Mitigation Performance of Three Aging Bioretention Cells. *Journal of Sustainable Water in the Built Environment*, *6*(4), 04020017. https://doi.org/10.1061/JSWBAY.0000925

Kassambara, A. (2017). *Practical Guide to Principal Component Methods in R*. STHDA.

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. https://CRAN.R-project.org/package=factoextra

Kilmartin, R. F., & Peterson, J. R. (1972). Rainfall-runoff regression with logarithmic transforms and zeros in the data. *Water Resources Research*, *8*(4), 1096–1099. https://doi.org/10.1029/WR008i004p01096

Kinerson, R., Mattice, J., & Stine, J. (1996). *The Metals Translator: Guidance For Calculating A Total Recoverable Permit Limit From A Dissolved Criterion*. USEPA. https://www3.epa.gov/npdes/pubs/metals_translator.pdf

Knappenberger, T. (2017). *LID: Washington State University Low Impact Development (LID)*.

Knappenberger, T., Jayakaran, A. D., & Stark, J. D. (2022). A Bayesian modeling framework to predict stormwater pollutant reduction in bioretention media. *Ecological Engineering*, *178*, 106582. https://doi.org/10.1016/j.ecoleng.2022.106582

Kohler, M. A., & Kinsley, R. K. (1951). *Predicting the Runoff From Storm Rainfall* (Research Paper No. 34). U.S. Department of Commerce.

Kumar, P., & Singh, A. (2010). *Cadmium toxicity in fish: An overview*.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Lele, S. R., Keim, J. L., & Solymos, P. (2023). *ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data*. https://CRAN.R-project.org/package=ResourceSelection

Li, H., & Davis, A. P. (2008). Heavy Metal Capture and Accumulation in Bioretention Media. *Environmental Science & Technology*, *42*(14), 5247–5253. https://doi.org/10.1021/es702681j

Li, X. F., Wang, P. F., Feng, C. L., Liu, D. Q., Chen, J. K., & Wu, F. C. (2019). Acute Toxicity and Hazardous Concentrations of Zinc to Native Freshwater Organisms Under Different pH Values in China. *Bulletin of Environmental Contamination and Toxicology*, *103*(1), 120–126. https://doi.org/10.1007/s00128-018-2441-2

Li, X.-Y., & Gong, J.-D. (2002). Compacted microcatchments with local earth materials for rainwater harvesting in the semiarid region of China. *Journal of Hydrology*, *257*(1–4), 134–144. https://doi.org/10.1016/S0022-1694(01)00550-9

Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R: A Guide to Analysis Using R*. John Wiley and Sons.

Magnuson, V. R., Harriss, D. K., Sun, M. S., & Taylor, D. K. (1979). *Chemical Modeling in Aqueous Systems: Speciation, Sorption, Solubility, and Kinetics* (E. A. Jenne, Ed.; Vol. 93). AMERICAN CHEMICAL SOCIETY. https://doi.org/10.1021/bk-1979-0093

Mendiburu, F. de. (2023). *agricolae: Statistical Procedures for Agricultural Research*. https://CRAN.R-project.org/package=agricolae

Morzaria-Luna, H. N., Schaepe, K. S., Cutforth, L. B., & Veltman, R. L. (2004). Implementation of Bioretention Systems: A Wisconsin Case Study1. *JAWRA Journal of the American Water Resources Association*, *40*(4), 1053–1061. https://doi.org/10.1111/j.1752-1688.2004.tb01066.x

Mosley, M. P. (1979). Streamflow generation in a forested watershed, New Zealand. *Water Resources Research*, *15*(4), 795–806. https://doi.org/10.1029/WR015i004p00795

Mullane, J. M., Flury, M., Iqbal, H., Freeze, P. M., Hinman, C., Cogger, C. G., & Shi, Z. (2015). Intermittent rainstorms cause pulses of nitrogen, phosphorus, and copper in leachate from compost in bioretention systems. *Science of The Total Environment*, *537*, 294–303. https://doi.org/10.1016/j.scitotenv.2015.07.157

Munn, M. D., Frey, J. W., Tesoriero, A. J., Black, R. W., Duff, J., Lee, K., Maret, T., Mebane, C., Waite, I., & Zelt, R. (2018). *Understanding the Influence of Nutrients on Stream Ecosystems in Agricultural Landscapes* (Circular 1437; Circular, p. 92).

NYSDEC. (2022). Draft Stormwater Management Design Manual. *New York State Department Of Environmental Conservation*. https://extapps.dec.ny.gov/fs/docs/pdf/stormwaterdesignmanual.pdf

NYSDEC, N. Y. S. D. of E. C. (2015). *New York State Stormwater Management Design Manual* (p. 778). https://extapps.dec.ny.gov/fs/docs/pdf/stormwaterdesignmanual2015.pdf

Pagenkopf, G. K. (1983). Gill surface interaction model for trace-metal toxicity to fishes: Role of complexation, pH, and water hardness. *Environmental Science & Technology*, *17*(6), 342–347. https://doi.org/10.1021/es00112a007

Penna, D., Tromp-van Meerveld, H. J., Gobbi, A., Borga, M., & Dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrology and Earth System Sciences*, *15*(3), 689–702. https://doi.org/10.5194/hess-15-689-2011

Pinheiro, J., Bates, D., & R Core Team. (2022). *nlme: Linear and Nonlinear Mixed Effects Models*. https://CRAN.R-project.org/package=nlme

Prince George's County. (1999). *Low-Impact Development Design Strategies An Integrated Design Approach*. Prince George's County, MD Department of Environmental Resources.

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radovanovic, J., & Bean, E. Z. (2022). Runoff Quantity and Quality from Amended Urban Residential Lawns. *J. Sustainable Water Built Environ.*, *8*(4), 04022010-1-04022010–04022019. https://doi.org/10.1061/JSWBAY.0000995

Rao, J. N., & Scott, A. J. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *The Annals of Statistics*, 46–60.

Rogevich, E. C., Hoang, T. C., & Rand, G. M. (2008). The Effects of Water Quality and Age on the Acute Toxicity of Copper to the Florida Apple Snail, Pomacea paludosa. *Archives of Environmental Contamination and Toxicology*, *54*(4), 690–696. https://doi.org/10.1007/s00244-007-9106-1

Ross, C. A., Ali, G. A., Spence, C., & Courchesne, F. (2021). Evaluating the Ubiquity of Thresholds in Rainfall-Runoff Response Across Contrasting Environments. *Water Resources Research*, *57*(1), e2020WR027498. https://doi.org/10.1029/2020WR027498

Roy-Poirier, A., Champagne, P., & Filion, Y. (2010). Review of Bioretention System Research and Design: Past, Present, and Future. *Journal of Environmental Engineering*, *136*(9), 878–889. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000227

Saffarpour, S., Western, A. W., Adams, R., & McDonnell, J. J. (2016). Multiple runoff processes and multiple thresholds control agricultural runoff generation. *Hydrology and Earth System Sciences*, *20*(11), 4525–4545. https://doi.org/10.5194/hess-20-4525-2016

Santore, R. C., Di Toro, D. M., Paquin, P. R., Allen, H. E., & Meyer, J. S. (2001). Biotic ligand model of the acute toxicity of metals. 2. Application to acute copper toxicity in freshwater fish and Daphnia. *Environmental Toxicology and Chemistry*, *20*(10), 2397–2402. https://doi.org/10.1002/etc.5620201035

Scaife, C. I., Singh, N. K., Emanuel, R. E., Miniat, C. F., & Band, L. E. (2020). Non-linear quickflow response as indicators of runoff generation mechanisms. *Hydrological Processes*, *34*(13), 2949–2964. https://doi.org/10.1002/hyp.13780

Seber, G. A. F., & Lee, A. J. (2003). *Linear Regression Analysis* (1st ed.). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780471722199

Sierra, J., Roig, N., Papiol, G. G., Perez-Gallego, E., & Schuhmacher, M. (2017). Prediction of the bioavailability of potentially toxic elements in freshwaters. Comparison between speciation models and passive samplers. *Science of The Total Environment*, *605*, 211–218. https://doi.org/10.1016/j.scitotenv.2017.06.136

Sun, X., & Davis, A. P. (2007). Heavy metal fates in laboratory bioretention systems. *Chemosphere*, *66*(9), 1601–1609. https://doi.org/10.1016/j.chemosphere.2006.08.013

Thorley, J., & Schwarz, C. (2018). ssdtools: An R package to fit Species Sensitivity Distributions. *Journal of Open Source Software*, *3*(31), 1082. https://doi.org/10.21105/joss.01082

Tipping, E., Lofts, S., & Keller, W. (2021). The use of WHAM-FTOX, parameterized with laboratory data, to simulate zooplankton species richness in acid- and metal-contaminated lakes. *Aquatic Toxicity*, Article 105708. https://doi.org/10.1016/j.aquatox.2020.105708

Tipping, E., Lofts, S., & Stockdale, A. (2023). WHAM-FTOXβ – An aquatic toxicity model based on intrinsic metal toxic potency and intrinsic species sensitivity. *Aquatic Toxicity*, *258*, Article 106503. https://doi.org/10.1016/j.aquatox.2023.106503

Tirpak, R. A., Winston, R. J., Feliciano, M., Dorsey, J. D., & Epps, T. H. (2021a). Impacts of permeable interlocking concrete pavement on the runoff hydrograph: Volume reduction, peak flow mitigation, and extension of lag times. *Hydrological Processes*, *35*(4), e14167. https://doi.org/10.1002/hyp.14167

Tirpak, R. A., Winston, R. J., Feliciano, M., Dorsey, J. D., & Epps, T. H. (2021b). Impacts of permeable interlocking concrete pavement on the runoff hydrograph: Volume reduction, peak flow mitigation, and extension of lag times. *Hydrological Processes*, *35*(4), e14167. https://doi.org/10.1002/hyp.14167

Triola, M. F. (2008). *Essentials of Statistics* (Third). Pearson Education Inc.

USDA NRCS. (1986). *Urban Hydrology for Small Watersheds, Technical Release 55*.

USEPA. (1999). Storm Water Technology Fact Sheet Bioretention. *EPA 832-F-99-012*.

USEPA. (2000). Low Impact Development (LID) A Literature Review. *EPA -841-B-00-005*.

USEPA. (2003). Protecting Water Quality from Urban Runoff. *EPA 841-F-03-003*. https://www3.epa.gov/npdes/pubs/nps_urban-facts_final.pdf

USEPA. (2005). Protecting Water Quality from Agricultural Runoff. *EPA 841-F-05-001*. https://nepis.epa.gov/Exe/ZyPDF.cgi/P10039OH.PDF?Dockey=P10039OH.PDF

USEPA. (2007). Aquatic Life Ambient Freshwater Quality Criteria—Copper 2007 Revision. *EPA-822-R-01-001*. https://www.epa.gov/sites/default/files/2019-02/documents/al-freshwater-copper-2007-revision.pdf

Van Buren, M. A., Watt, W. E., & Marsalek, J. (1997). Application of the log-normal and normal distributions to stormwater quality parameters. *Water Research*, *31*(1), 95–104. https://doi.org/10.1016/S0043-1354(96)00246-1

Wang, C., Zhao, Y., & Pei, Y. (2012). Investigation on reusing water treatment residuals to remedy soil contaminated with with multiple metals in Baiyin, China. *Journal of Hazardous Materials*, *237*, 240–246. https://doi.org/10.1016/j.jhazmat.2012.08.034

Wang, S., Yan, Y., Fu, Z., & Chen, H. (2022). Rainfall-runoff characteristics and their threshold behaviors on a karst hillslope in a peak-cluster depression region. *Journal of Hydrology*, *605*, 127370. https://doi.org/10.1016/j.jhydrol.2021.127370

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838. https://doi.org/10.2307/1912934

Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, *21*(12), 1–20.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2023). *readxl: Read Excel Files*. https://CRAN.R-project.org/package=readxl

Windward Environmental, LLC. (2019). *Biotic Ligand Model—Freshwater & Marine* (3.41.2.45) [Computer software]. Windward Environmental, LLC.

Winston, R. J., Dorsey, J. D., & Hunt, W. F. (2016). Quantifying volume reduction and peak flow mitigation for three bioretention cells in clay soils in northeast Ohio. *Science of The Total Environment*, *553*, 83–95. https://doi.org/10.1016/j.scitotenv.2016.02.081

Winston, R. J., Dorsey, J. D., Smolek, A., & Hunt, W. F. (2018). Hydrologic Performance of Four Permeable Pavement Systems Constructed over Low-Permeability Soils in Northeast Ohio. *Journal of Hydrologic Engineering*, *23*(4), 04018007. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001627

Woodward, M., Hunt, W. F., & Hartup, W. (2008). Lessons Learned: The North Carolina Backyard Rain Garden Program. *Low Impact Development for Urban Ecosystem and Habitat Protection*. https://doi.org/10.1061/41009(333)104

WSDE, 2019. (2019). *Stormwater Management Manual for Western Washington* (Technical Report 19-10-021). Washington State Department of Ecology - Water Quality Program. https://apps.ecology.wa.gov/publications/SummaryPages/1910021.html

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2008). *Mixed Effects Models and Extensions in Ecology with R*. DOI 10.1007/978-0-387-87458-6