

Leveraging Data for Green Infrastructure Performance Analysis and Prediction

By

Andrew Kurzweil

Thesis

Submitted to Department of Civil and Environmental Engineering
College of Engineering
Villanova University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE, CIVIL ENGINEERING

May 2021



VILLANOVA
UNIVERSITY
College of Engineering

Andrew Kurzweil

Leveraging Data for Green Infrastructure Performance Analysis and Prediction

May 2021

Advisors:

Robert G. Traver, Ph.D., P.E., D. WRE, F.EWRI, F.ASCE ,

Bridget Wadzuk, Ph.D., and

Gerald Zaremba, Ph.D.

Villanova University

College of Engineering

Center for Resilient Water Systems

Department of Civil and Environmental Engineering

800 East Lancaster Avenue

Villanova, Pennsylvania 19085

Copyright © Andrew Kurzweil, May 2021. All Rights Reserved.

Brief portions of this material may be quoted in relevant works. Requests for extended quotation should be addressed to the above listed author or advisors.

The data presented herein were collected as part of a research initiative by Villanova University's Center for Resilient Water Systems (VCRWS) under a grant funded by the Pennsylvania Department of Transportation (PennDOT). These data are to be considered preliminary, and may not be used elsewhere except with explicit written permission from PennDOT or VCRWS. PennDOT and VCRWS make no guarantees as to the integrity of these data, which are provided for informational purposes only and are subject to change upon further review.

Abstract

In the twenty-first century, harnessing the power and abundance of environmental data has the potential to unleash a greater understand of and appreciation for the natural environment and enable better integration of the built environment by dynamically responding to the natural environment's processes. Reliably, accurately, and consistently collecting quality observations for monitoring, analysis, and future design adjustments becomes necessary when high volumes of data are able to be processed rapidly and automatically. In this thesis, I will demonstrate the need for, and benefits of, a data-oriented approach to Green Infrastructure design and analysis using the I-95 Revive project and PennDOT's SMP A site as a case study for robust monitoring networks capable of minute-by-minute measurements over a years-long span to enable better insight to the natural processes occurring in the rain garden and the potential effects on downstream hydraulics.

Acknowledgement

This thesis is a result of my studies at Villanova University and my research as a member of the Center for Resilient Water Systems' PennDOT Hydrology Team. I would like to thank my advisors, Dr. Robert Traver, Dr. Bridget Wadzuk, and Dr. Gerald Zaremba, for their endless support, for sharing their wealth of knowledge and experience, and for entertaining my continuous requests to try new things. Madhat Fares, Danielle Galloway, Shaelynn Heffernan, and Matina Shakya, my fellow graduate students, made field work a breeze and provided many hours of troubleshooting, laughs, bouncing ideas around, and more troubleshooting. The Pennsylvania Department of Transportation and AECOM's partnership were key in inspiring, funding, and critiquing our work, and I am forever grateful for their support. Last, but certainly not least, thank you to my family for providing a rich educational background and my friends for their encouragement and support along the way.

Table of Contents

1	Introduction	1
2	Standardized Data Collection	2
2.1	Background	2
2.2	Site Configuration	3
2.2.1	Measurement: Sensors and Structures	7
2.2.2	Data Collection and Transmission	8
2.3	Challenges	9
2.4	Solutions	10
2.4.1	Temporary Ramp	10
2.4.2	Inlet Flow Measurement	12
2.4.3	Future Work	14
2.5	Conclusions	15
3	Robust Data Storage	17
3.1	Background	17
3.2	VCRWS IDM	17
3.2.1	Metadata Standardization and Data Formatting	17
3.2.2	Query Performance	17
3.3	Next Steps	17
3.3.1	Streaming Data Ingest	17
3.3.2	Programmatic Data Retrieval	17
3.4	Conclusions	17
4	Performance Driven Data Analysis	18
4.1	Background	18
4.2	Data and Modeling	20
4.2.1	Rainfall Events	20
4.2.2	Soil Moisture Events	21
4.3	Key Performance Indicator Definitions	23
4.3.1	Recession Rate	23
4.3.2	Infiltration Drying Rate	24

4.4 Results and Discussion	26
4.4.1 Recession Rate	26
4.4.2 Infiltration Drying Rate	32
4.5 Conclusions	34
5 Conclusion	36
Bibliography	37

List of Figures

2.1	Natural vs Urban water budget (Philadelphia Water Department, 2021).	3
2.2	SMP A site layout.	4
2.3	Check dam at the downstream end of the central gabion blanket section of SMP A. . .	4
2.4	Gabion blanket section showing highly sloped banks.	5
2.5	B1 outlet structure with weir plate covering CSS connection.	5
2.6	Misaligned catchments allow water to bypass and flow down the temporary ramp. . . .	6
2.7	Temporary ramp at the downstream end of SMP A.	6
2.8	Steven's Water Hydraprobe.	7
2.9	0.8-foot OpenChannelFlow flume plan view.	11
2.10	Flume installation at the bottom of construction ramp, 2019.	11
2.11	Area-Velocity measurement comparison results	12
2.12	Low flow measurement setup	13
2.13	Low flow measurement installation	13
2.14	Splashing at N9 low flow installation.	14
2.15	N9 low flow data for 47mm storm event	14
2.16	Ultrasonic depth plus Blue Siren velocity preliminary flow comparison.	15
2.17	Vegetronix SDI-12 Analog Sensor Translator.	16
4.1	Example SWCC.	20
4.2	Distribution of event total rainfall depths.	21
4.3	Distribution of event lengths.	22
4.4	Distribution of event ordinal dates.	23
4.5	Typical soil moisture curve.	25
4.6	Illustration of infiltration drying rate.	26
4.7	Average water temperature vs Recession Rate for SMP A storm event.	27
4.8	Average water temperature vs Recession Rate	28
4.9	Average Relative Humidity vs Recession Rate	29
4.10	Average Barometric Pressure vs Recession Rate	30
4.11	Average Ponding Depth vs Recession Rate	31
4.12	Normal Q-Q plot for ponding recession regression.	31
4.13	Time to Desaturation between end of ponding and 10cm.	32
4.14	Time to Desaturation between 10 and 35cm.	33
4.15	Time to Desaturation between 35 and 60cm.	33

List of Tables

4.1	Ponding recession rate regression results.	30
4.2	Kendall seasonal trend results.	32

Introduction

Modern cities are facing new and quickly emerging threats from increasingly frequent large storms, which wreak havoc on urban streams and tax urban drainage and sewer systems to dangerous levels.

Standardized Data Collection

2.1 Background

Collecting valid data is the first step to any rigorous experiment. The methods used to collect this data must be reliable and easily repeatable for other parties who wish to replicate results. That is to say identical inputs to a measurement system, given the same external conditions, should yield the same results. This is a fundamental premise of the scientific method.

Measuring water in any system, natural or anthropogenic, is not without its challenges. After all, nature obeys the laws of entropy, and energy tends to dissipate towards its lowest potential. The water cycle that forms our climate is driven by solar heat inputs to this energy system. Rainfall is generated from the accumulation of evaporated water in the atmosphere, which then condenses and falls to the ground. No two storms will ever be the same in magnitude, duration, or intensity profile, so constructed solutions must be flexible enough to handle storms ranging from slow, drawn out events, to short, intense events. In urbanized environments, the high percentage of surface area covered in impermeable asphalt or buildings causes increased runoff and reduced evaporation and infiltration (Figure 2.1). Urban environments have long faced challenges when it comes to handling this increased runoff in a neutral manner that avoids disastrous flooding while also minimizing the impact to all parts of the environment. In Philadelphia, combined sewer systems (CSS) were the preferred method of handling large volumes of rainfall for nearly the city's entire history (Akhavan and Jianpeng, 2015), but have a major negative impact on the environment via combined sewer overflows (CSO). Only recently have green initiatives such as Philadelphia's "Green Streets, Green City" expanded, allowing the city to expedite the creation of environmentally friendly means of removing stormwater (Philadelphia Water Department, 2018; Callahan, 2019). These systems allow urban stormwater, which primarily falls on impervious surfaces and leads to increased runoff as compared to natural environments, to be captured and managed largely in place (Heffernan et al., 2016) by a combination of green infrastructure methods including rain gardens, tree trenches, green roofs, and pervious concrete.

This chapter will discuss best practices pertaining to the remote sensing of green infrastructure conditions developed for SMP A in the GR2 section of PennDOT's I-95 Revive project. This framework has been developed to be reliable, consistent, easy to implement, and expandable to other sites with minimal additional configuration. A primary goal of this research effort is to not only inform better data collection, but also to enable that data to be collected at scale across many sites in a project or across many different projects to allow for more widespread analysis of green infrastructure's performance, maintenance needs, and impact on the natural environment (Burcin et al., 2014). To that end, measurements must be precise and data recording and transmission must be unhindered by randomness or other inconsistency. Myriad challenges exist in this effort, but the focus of this discussion is accurate measurements of flowing water, which exhibits a wide variety of properties

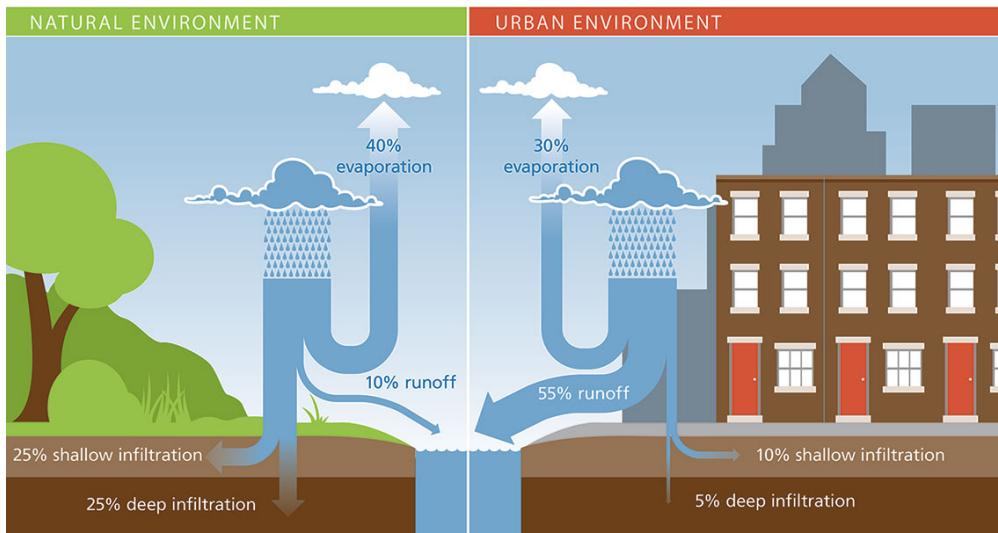


Figure 2.1: Natural vs Urban water budget (Philadelphia Water Department, 2021).

ranging from depth and velocity to random formations of currents and internal feedback in the form of turbulence (Mays, 2010).

The most important aspects of monitoring a GSI system are a reliable data logger with minimal downtime, connected to a central data server (or "cloud") that collects data at a regular interval and redundantly stores that data for further analysis. The data logger must be able to poll all connected sensors at timely intervals such that data measurements are associated with an accurate timestamp. Measurements must be accurate, while their precision can be based on the unit of measure. The ability to continuously monitor GSI conditions enables an understanding of design specifications versus real world performance. While models are useful tools for understanding and managing expectations for GSI, they can only go so far at capturing real world knowledge and the many interactions of GSI conditions that create nearly endless combinations of unique conditions.

2.2 Site Configuration

Situated along the south side of Interstate 95 northbound, SMP A handles inflow from the elevated roadway surface that flow into it from a network of pipes beneath the highway. Most of the directly connected impervious area (DCIA) lies in the northbound section, per LiDAR surveys conducted in 2017. SMP A is a linear bioswale type rain garden split into 3 sections by two check dams placed at roughly the third points of the garden (Figure 2.2).

The upstream portion contains one inlet (N8) that is a 30" diameter reinforced concrete pipe (RCP). This upstream section is planted with local, salt tolerant grasses and small coniferous trees. The surface is mulched, and the cross-sectional geometry is gently sloping. The upstream section



Figure 2.2: SMP A site layout.

terminates with a 45° steel weir plate situated in the center of a plywood check dam that prevents unmeasured flow from transiting to the center section (Figure 2.3).



Figure 2.3: Check dam at the downstream end of the central gabion blanket section of SMP A.

The center section is lined on the south side by an impermeable fabric to prevent infiltration from damaging neighboring structures' foundations. The entire section is lined with gabion basket devices, which help prevent erosion and cut down on weed growth. There is a single 18" RCP inlet (N9) in this section and the cross section has more pronounced slopes (Figure 2.4). The gabion blanket section is similarly terminated at the downstream end by a 45° steel weir plate flush with a check dam to retain as much water as possible.

The downstream portion of the garden contains one more 18" RCP inlet (N10), as well as two concrete overflow structures connected to Philadelphia's CSS known as B1 and B2. These structures



Figure 2.4: Gabion blanket section showing highly sloped banks.

have PT devices attached to the outside of the structure for measuring the ponded water level. Inside the structures are 22.5° steel weir plates inside covering the CSS connection (Figure 2.5), with another pair of PT devices for measuring the outflow from the system, again using the standard weir equation.



Figure 2.5: B1 outlet structure with weir plate covering CSS connection.

Due to SMP A's location at the border between the completed GR2 section of the project and the upcoming GR1 section to the south, there is a temporary construction ramp at the downstream end of the basin that leads from the elevated roadway to grade level adjacent to Frankford Ave. The ramp is not intended to convey inflow, but the grade of the highway, misalignment of catchments (Figure 2.6) on the highway, and lack of barrier at the top of the ramp mean that a significant amount of flow enters the basin after flowing down the ramp. In June 2019, approximately 60

sandbags were added to an existing asphalt berm on the ramp, along with an H-flume with a PT sensor to capture the hydraulics of the ramp.



Figure 2.6: Misaligned catchments allow water to bypass and flow down the temporary ramp.



Figure 2.7: Temporary ramp at the downstream end of SMP A.

2.2.1 Measurement: Sensors and Structures

To aid in repeatable, accurate, and timely measurements, sensors are deployed at a variety of locations throughout SMP A. Pressure transducers in a variety of locations monitor both ponding level (outside the B1 and B2 outlet structures) and the depth behind weir plates (inside B1 and B2, and behind two check dams separating garden sections). The flow rate over weir plates can be calculated using the standard weir equation (USBR, 2001).

Inlet flow at N8, N9, and N10 was originally measured by a single BlueSiren Dual Wave Doppler Area-Velocity (AV) sensor at each inlet. These sensors are fixed to expandable steel bands that secure the apparatus inside the end of the RCP by means of a screw jack. Located at the downstream end of the AV sensor is a PT that measures flow depth, and outputs a 0-5V analog signal proportionate to the observed depth. The voltage response must be calibrated with several known depths to establish a valid conversion equation (see section 2.3). At the upstream end of the AV sensor, the dual wave doppler measures flow velocity and outputs an 8-bit serial signal, with the number of doppler pulses reflected by the water corresponding to the velocity in millimeters per second. Both parts of the AV sensor require 12V direct current (DC) power supply, which is managed by the CR6 data logger.

Soil moisture is measured at 2 locations using sensors at 10cm, 35cm, and 60cm depths, with an additional, redundant sensor at 35cm for quality assurance purposes. Stevens Water Hydraprobe sensors (Figure 2.8) are used to measure soil moisture level, conductivity, resistivity, temperature, plus real and imaginary dielectric permittivity. Hydraprobes use 4 stainless steel prongs to measure the soil's parameters and "take into account the energy storage and energy loss across the soil area using a 50MHz radio frequency wave" (Stevens Water, 2021). The sensors are located in the middle of the downstream basin between outlet structures B1 and B2, as well as in the upstream basin approximately 2 meters from the check dam.



Figure 2.8: Stevens Water Hydraprobe (Stevens Water, 2021).

Additional climate sensors include air temperature and relative humidity, solar radiation, wind speed and direction, and barometric pressure. These sensors are located atop a 60-foot telephone pole adjacent to the I-95 bridge over Frankford Avenue that extends roughly 40 feet above the highway surface for measurements uninterrupted by traffic or highway structures. These sensors are collectively referred to as the "Weather Station," and are collectively a separate network from the main garden.

2.2.2 Data Collection and Transmission

The entire sensor network is controlled and monitored by three Campbell Scientific CR6 data logger devices. These are located on poles attached to the B2 outlet, at the base of the weather station pole, and on a small post in the ground next to inlet N8. The data loggers are inside protective boxes and have solar panels mounted adjacent to provide power and charge the 12V battery also located in each box. The CR6 runs a program written in Campbell Scientific's CRBasic programming language, which is a procedural language similar to the BASIC family of languages. The program defines variables, storage tables, polling and logging frequencies, and monitor's the CR6's health and battery status.

Communications at SMP A are achieved via Campbell Scientific's Pakbus protocol, which allows multiple CR6 loggers at one site to be connected via external Wi-Fi modules (NL240) or RF devices (RF407). In early 2019, the existing RF407 radio devices were swapped out for NL240 Wi-Fi devices which improved range, reliability, data link speeds, and security. Campbell Scientific LoggerNet software is used to remotely manage the data loggers, collect data every 3 hours, and identify issues that require on site maintenance.

Remote connections are achieved via a Verizon 4G-LTE cellular modem which provides internet access for remote connections and download of data. The modem is paired with the CR6 at the Weather Station, as this setup has the fewest instruments attached, and therefore the lowest power requirements. Remote access is necessary for downloading data at a regular interval and monitoring the network's health.

There are four types of measurement performed by the CR6 data loggers: analog voltage difference, serial, pulse count, and digital. Each have their advantages and disadvantages.

Analog voltage difference sensors, such as the BlueSiren AV depth reading, convert a 5V or 12V DC supply into a 0-5V response, which is read by the data logging device. This method requires a calibration equation to translate the response voltage into meaningful units for the measured variable. The reading is performed by the CR6 device and is almost always separated from the point of measurement (sensor) by a significant length of wire. Calibration is performed by applying a known reading to the sensor (depth of water in the case of a PT sensor), and recording the response voltage at several different readings. A linear relationship (slope and intercept) are then calculated from the points collected, and this equation is written directly into the CR6 program. Due to electrical resistance of copper wire varying with temperature (Eargle, 2002), the resistance of a copper wire increases by 15% over a 40° Celsius range when warmed by direct sunlight. This is amplified by the long length of wire running between the sensors and data logger - nearly the entire 300 foot length of SMP A in some cases. Increased resistance is not compensated for by the CR6, so measurements made at temperatures different from the calibration temperature are subject to error.

Serial sensor send pulses of electrical signals at a given voltage and baud rate. The CR6 data logger reads these serial values by opening a serial buffer on the port a serial sensor is connected to, and can respond to a wide variety of serial protocols and refresh rates. The Blue Siren sensors used at SMP A communicate velocity via UART - TTL at a baud rate of 4800. Serial communications are less prone to error or disturbance introduced by the transmitting wires, but have limits on how far a valid signal will transmit. Additionally, only one serial sensor can be active on any given data logger port at once, so some inlet AV sensors are connected to switched 12V power outputs on the CR6 to allow multiple sensors to share a single physical port. This complicates both the physical wiring of the sensors at the site, as well as the code required to handle toggling the power state in coordination with taking measurements. Additionally, it is unknown what effect this power cycling of the AV sensors has on their performance.

Pulse count communications are used primarily by rain gauges. Measurement occurs when a specific event happens, such as the tip of a tipping rain gauge. The sensor briefly connects the measurement port to ground, creating an electrical pulse that is detected by the CR6. This means pulse count sensors are among the simplest in terms of wiring, as they require only a measurement and ground lead.

Finally, digital sensors, including the CS451 PT, Hydraprobes, pyranometer, and thermometer, use the SDI-12 protocol which allows up to 62 addressed sensors on a single communications port. The SDI-12 protocol allows the CR6 to poll sensors for up to date data on demand, and read the results back nearly instantaneously. The digital transmission of data eliminates the need to run separate communications cables for each sensor, and consolidates the ports used on a CR6 to just one for most applications. Sensors are able to relay several measurements to the CR6 in an array of values without concerns surrounding wire length or environmental conditions.

2.3 Challenges

Water flowing through pipes or across surfaces is affected by a variety of factors: slope, roughness, temperature, and geometry. The slightest imperfections in the surface or inconsistencies in the roughness can introduce turbulence, a non-uniform flow regime not well suited to measurement by sensors best suited for steady, uniform conditions. The act of falling from highway level to ground level, where measurement takes place, means that water has a high amount of kinetic energy and is thus more likely to splash around, disturbing calmer water and imparting some of its energy. Turbulence is best handled by some form of flow straightener, calming device, or impoundment behind a flow-control device such as a weir plate that uses temporarily stored water to remove energy from incoming water. At lower flow rates, this can also be accomplished through the use of high permeability foam, which has the added benefit of filtering out larger debris particles washed in from the highway surface.

Attempts to characterize the "average" storm year have been successful (Albright and Schramm, 2018), and provide a means of estimating rainfall over a given period of interest. This rainfall can be used as the input for a system model, such as was created by Elizabeth Calt, MSCE '18 (Calt, 2018). The Environmental Protection Agency's (EPA) Stormwater Management Model (SWMM) created by Calt estimates runoff, storage, and overflow volumes and was used to estimate the expected values for the sensors measuring flow through the 3 RCP inlets during both historical rainfall records, and 2 design storms (2 year, 24 hour and 10 year, 24 hour).

Historical data from inlet N8, N9, and N10 show that the depth portion of the AV sensors is frequently negative, which is indicative of improper calibration or sensor malfunction. Despite sensors regularly being calibrated carefully, and returning valid measurements immediately after calibration, measurement drift continues to be an issue. Little of the data in the period of study is valid, with many negative values occurring during storm events, so the best estimates of inflow rates are from the aforementioned SWMM model developed by Elizabeth Calt. In addition to negative inflow rates being invalid, the model suggests that flow rates lower than the Blue Siren AV sensors are capable of measuring are possible, if not frequent depending on the shape and size of storm hydrographs at SMP A.

Finally, the calibration procedure for existing Blue Siren AV sensors is both tricky to perform and subjective based on human error. After removal from the steel band holding it inside the pipe, the sensor is submerged in one of two five-gallon buckets with water measured by hand to the nearest millimeter. Simultaneously, the data logger program must be updated to read out the raw voltage received, and these values are recorded in an Excel spreadsheet. After taking measurements at several depths, typically ranging from 25-150mm, the hand measured depths and corresponding voltages are transformed into a simple linear relationship (slope and intercept), which is then updated and enabled in the data logger's program.

2.4 Solutions

2.4.1 Temporary Ramp

The addition of the 0.8-foot H-flume in June 2019 at the intersection of the temporary ramp and top of slope near the B1 outlet structure was driven by a desire for better understanding of hydraulic inputs to the basin that were bypassing storm drains on the roadway surface and instead flowing down the temporary ramp. The H-flume is made of reinforced fiberglass and was sourced from OpenChannelFlow in Boise, ID, and is sized based on preliminary output from Elizabeth Calt's SWMM model using both design storms and historical rainfall data collected at the site through early 2019. It has a 3 foot approach section, and two cylindrical measurement wells on either side of the pour point which are 150mm deep (Figure 2.9). Using the known offset of these wells, the data

recorded by a CS451 PT can be used along with the calibration curve supplied by OpenChannelFlow to establish the flow rates observed:

$$F = -0.00707921 + 0.04898248 d^{0.5} + 21.70307374 d^{1.5} + 365.9132927 d^{2.5} \quad (2.1)$$

where F is flow in liters per second and d is depth of flow in meters, which gives this flume a range of 0.0085 - 12.94 L/s Flumes have long been an agriculture industry standard for measuring flows into and out of fields and ponds, and they are well equipped to handle a large range of flow rates. While the sandbag wall has been tampered with on several occasions, it largely remains intact and able to direct water through the flume (Figure 2.10). Only minor leakage was observed during a fall 2020 simulated runoff test (SRT) when a high flow rate was pumped into the site.

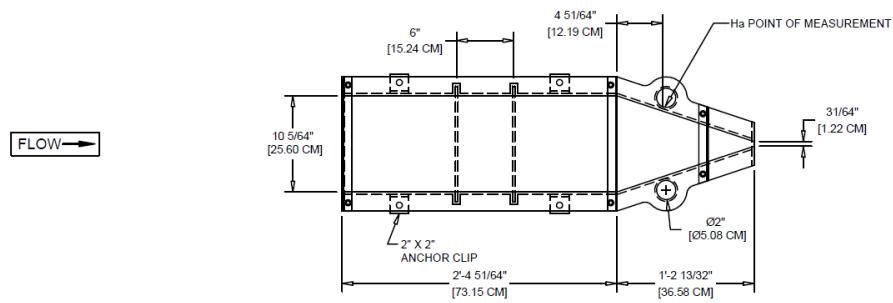


Figure 2.9: 0.8-foot OpenChannelFlow flume plan view.



Figure 2.10: Flume installation at the bottom of construction ramp, 2019.

2.4.2 Inlet Flow Measurement

To estimate the AV sensors' range, a suite of tests were performed using a lab simulation of field conditions. Sensors were attached to an identical steel band and placed inside an 18" corrugated plastic pipe. The roughness of the pipe is different than that of the RCP installed in SMP A. However, given the turbulence introduced by the flow falling from highway level to ground level a short distance upstream of the measurement point, roughness does not control the flow profile (Mays, 2010).

The lower bound for Blue Siren AV sensors was determined in Villanova's water resources lab by calibrating a sensor using the same procedures used in the field, and then recording the sensor output and flume flow rate simultaneously. Analysis showed (Figure 2.11) that the AV sensors are not valid below 1 inch depth and 1.25 inch/second velocity, which equates to 0.1 Liters/second in an 18 inch pipe (N9, N10) or 0.14 Liters/second in an 30 inch pipe (N8). A second Blue Siren system consisting of an Ultrasonic depth sensor mounted to the top of the pipe combined with a "microvelocity" sensor that has a much lower profile than the current Dual Wave doppler AV sensor was tested in an identical fashion and found to have lower velocity accuracy, but consistently better depth accuracy. Figure 2.11 compares the two Area-Velocity measurement methods, with the insert at right showing both systems measuring significantly lower than the reference flow rate below

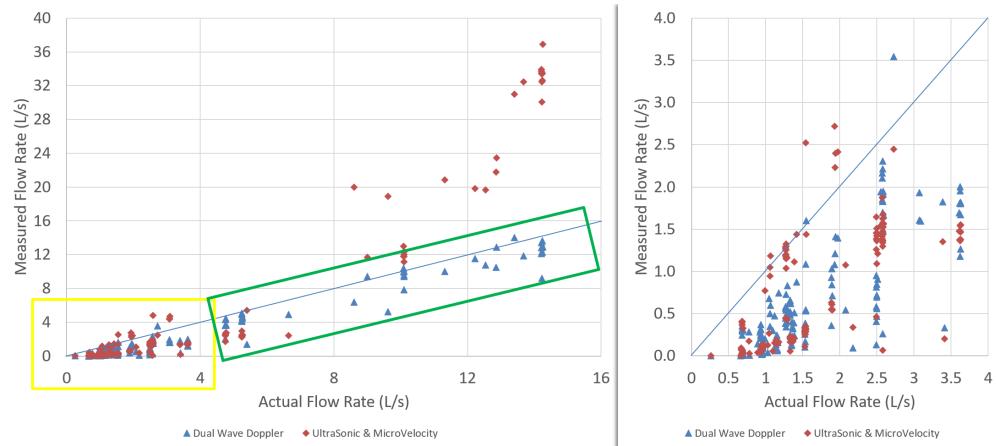


Figure 2.11: Area-Velocity measurement comparison results

In order to capture lower flow rates, a restricting structure such as a weir plate or flume is necessary, so that the flow is forced through a narrower opening where it can be measured more precisely. The solution chosen for implementation is a weir plate sized to fit inside a 4 inch PVC pipe (Figure 2.12), which will sit at the outlet of a catch basin small enough to fit under the pour point of the RCP inlets (Figure 2.13). This solution minimizes the disturbance to existing infrastructure and involves the least standing water between storms, which is a concern for biological safety during hot summer months.

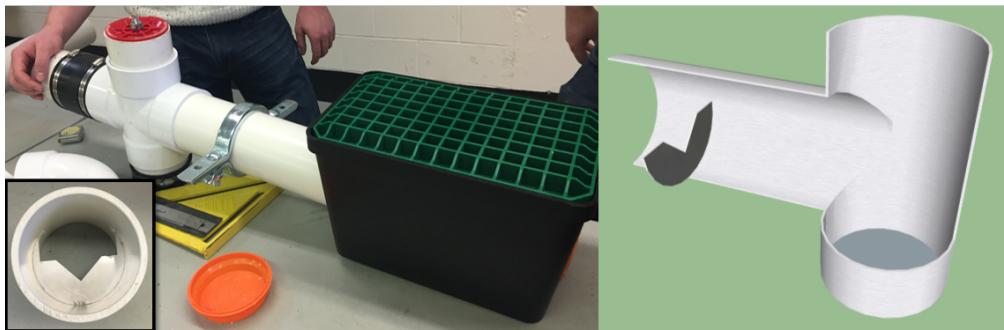


Figure 2.12: (Left) Low flow measurement setup prototype being constructed in the lab. (Right) Cutaway section of the device outlet.



Figure 2.13: Low flow measurement installation

Since installation during summer 2020, the low flow measurement systems at N8, N9, and N10 have captured at least a portion of the flow of each event. However, several issues remain: the capture efficiency of the inlet (green plastic grate) is lower than anticipated, as a significant percent of flow splashes off and is not captured in the system (Figure 2.14). Despite this, the inlets show a clear response to rainfall depths as seen in Figure 2.15. This data, collected at N9 during a 47mm storm event show a total volume of $3.2m^3$, which equates to just 3.7mm of rainfall over the approximate drainage area of $882m^2$ at N9. While this could be explained by changes to the drainage area since the lidar survey was performed in 2017, bypass due to catch basin design, or poor capture efficiency at the pour point into the garden, it is evident that the system needs further tuning.



Figure 2.14: Splashing at N9 low flow installation.

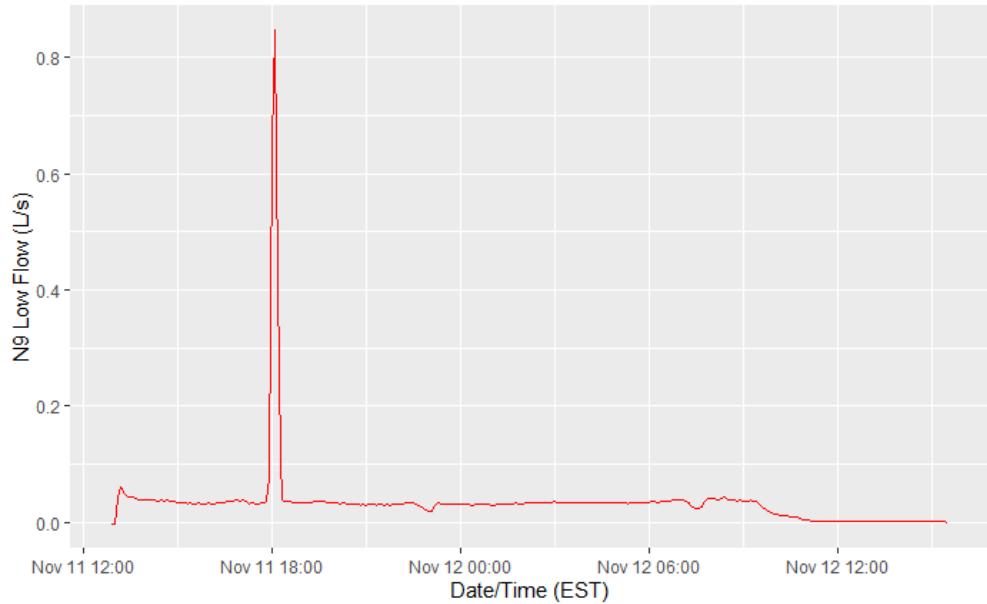


Figure 2.15: N9 low flow data for 47mm storm event

2.4.3 Future Work

Investigation into the combination of Blue Siren AV sensor for velocity readings and Massa Ultrasonic depth sensor is ongoing. Preliminary lab results (Figure 2.16) indicate that this system will be

sufficiently accurate in combination with the aforementioned low flow system due to their overlap in flow range, shown in orange.

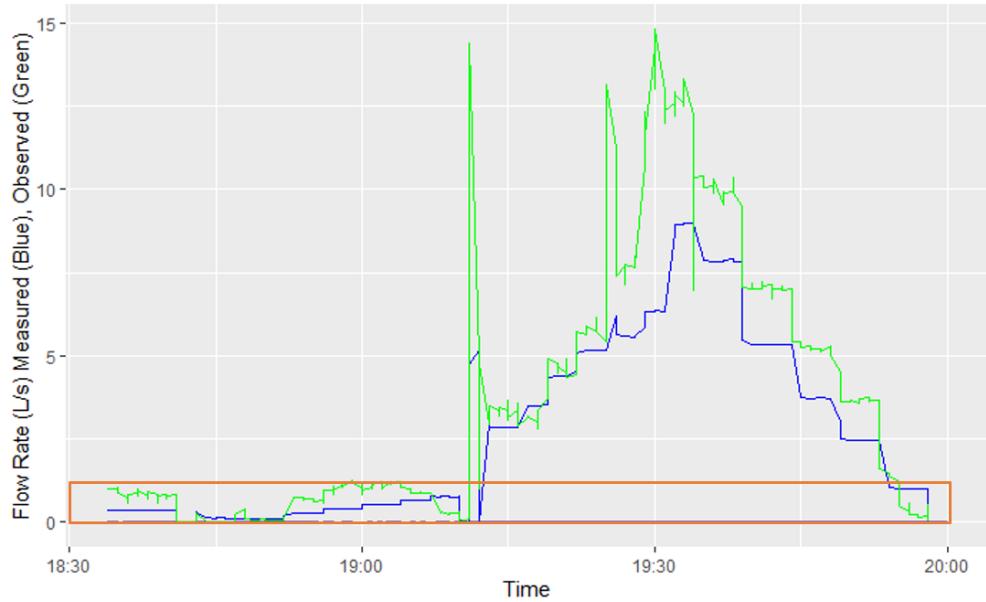


Figure 2.16: Ultrasonic depth plus Blue Siren velocity preliminary flow comparison.

This system will require further tuning to ensure that the flow rate measurements produced are sufficiently accurate. The preliminary results show some evidence of hysteresis, when increasing and decreasing flow rates respond differently, but more controlled lab testing is necessary to identify areas for improving this combination of sensors, and the CRBasic program that combines the data they produce into a flow rate. Further testing at more gradually varied flow rates (smaller jumps in flow rate) will help demonstrate whether this system is ready for site implementation.

Finally, due to the variability of analog signal measurement over the long distances in SMP A, conversion to digital measurement will further increase accuracy and reliability. All lab tests discussed in this chapter have used an analog to digital converter (ADC) between the sensor and data logger as a proof of concept. The device of choice is a Vegetronix SDI-12 Analog Sensor Translator (Figure 2.17), which supports up to four 0-5V sensors attached, and is fully compliant with the SDI-12 protocol. This inexpensive device will greatly simplify the wiring schematics necessary for using two different sensors (Ultrasonic and AV) for measuring flow at the inlets, as well as shorten the length of wire over which an analog signal must travel.

2.5 Conclusions

Measuring flow in natural, uncontrolled conditions is difficult and requires creativity. Understanding the intricacies of sensor configurations, communications, and opportunities for error can help



Figure 2.17: Vegetronix SDI-12 Analog Sensor Translator.

mitigate invalid data. Digital communications significantly increase a sensor's ability to report valid data, ensuring that measurements taken in one location are accurately represented and stored by a data logger several hundred feet away. The increased confidence in data collected will enable future analysis of longer term trends, so investments and efforts made now in upgrading systems can have lasting benefits for GSI studies across all sites. Furthermore, significantly less time will be consumed processing the data in quality control procedures leading to more resources available for actual analysis.

Robust Data Storage

3.1 Background

3.2 VCRWS IDM

3.2.1 Metadata Standardization and Data Formatting

3.2.2 Query Performance

3.3 Next Steps

3.3.1 Streaming Data Ingest

3.3.2 Programmatic Data Retrieval

3.4 Conclusions

Performance Driven Data Analysis

4.1 Background

Green stormwater infrastructure (GSI) systems' performance has historically been difficult to measure due to poor or nonexistent monitoring and inconsistent performance metrics by which to evaluate systems. There are many reasons for this lack of consistency, but much of it stems from the inconsistency of storm events to which GSI systems are subjected, with a wide range of possible precipitation volumes and intensities producing an equally wide variety of responses by these systems. Most systems are designed and constructed using specifications drawn up over the last 25 or more years. The National Pollutant Discharge Elimination System (NPDES) is a common national set of requirements arising from the Clean Water Act (CWA) of 1972 and outlines maximum discharge rates and pollutant loads from certain site types, broken down by activity (industrial, construction, municipal, etc.) (USEPA, 2009). While many states, including Pennsylvania (Department of Environmental Protection, 2006), meet NPDES requirements demanding designs that consider the pre-construction site conditions or require a site to mimic the otherwise pre-anthropogenic environment's hydrology at the site, most do not include a post-construction monitoring plan or extend the NPDES with criteria more specific to and suitable for the local environment.

Additionally, there are few well-defined and agreed upon performance metrics that can be used to quantify a GSI's response to storm events over its lifespan. Infiltration and evapotranspiration capacity of newly constructed systems is occasionally tested for construction acceptance, but these tests are only performed once at discrete randomly selected points within the system and do not convey a wholistic representation of the GSI's capacity. Longer-term, whole-system monitoring and performance analysis using key performance indicators (KPI) that align with and augment existing lab and field tests have the potential to unlock new insights into and understanding of maintenance needs, construction methodology, and design choices by connecting design specifications to real-world performance.

The lack of post-construction monitoring and analysis poses a major roadblock to improving recommendations for the design process that could lead to higher GSI longevity, lower the risk of GSI failure or under-performance, and creating uniform standards for GSI comparisons between geographically distinct sites or projects. Even when these monitoring requirements are in place, diverse site conditions, geographies, and climates necessitate a standardized framework for quantifying performance and comparing between potentially vastly different sites. This chapter outlines proposed key performance indicators (KPI) unique to an infiltration-type rain garden GSI by looking at historical data for SMP A, located in PennDOT's GR2 section of the I-95 Revive project (Figure 2.2). These robust monitoring and analysis techniques will lead to consistent results that can be applied across many sites while ensuring that outside factors do not influence performance measurement results. The following KPI measures are intended as derivatives of existing, more complicated lab

or field tests that are widely accepted, such as infiltration testing, soil texture tests, soil porosity and bulk density tests, to name a few, but are based solely on data collected automatically by a the SMP A sensor network. This usage of existing data to corroborate more difficult testing procedures' results will allow easier comprehension of the KPIs and wider transferability. A standard approach to analysis will open the door to suggested improvements for designs and further exploration of GSI's importance to more sustainable urban environments.

Recession rate, or the change in depth of water ponded over time, is hypothesized to be a potential key performance indicator (KPI) for GSI because it provides an easy to measure proxy for soil health. Soil health, in the context of GSI infiltration, is defined as a lack of compaction, clogging, or other infiltration inhibiting issues (Sokolovskaya et al., 2021). These properties of soil heavily influence hydraulic conductivity and the shape of a soil-water characteristic curve (SWCC) (Figure 4.1). Hydraulic conductivity of soil is the property that defines the ability of soil to pull water from the surface and through the soil column, or the ability to reduce pressure differentials within different sections of soil. The SWCC defines how a given soil responds to varying saturation levels by describing the suction force applied to water in contact with the soil across a range of saturation. Saturation of soil occurs when all the void space is filled with water and the movement of water through the soil column reverts to a gravity driven system that is largely influenced by soil characteristics. Saturated soil generally has the lowest hydraulic conductivity, referred to as K_{sat} , among the range of all volumetric water content (VWC) values possible for that given soil (Eyo et al., 2020). Different soils can have significantly different hydraulic conductivity ranges, and engineered soils with favorable properties (higher saturated hydraulic conductivity) are generally specified for GSI design where infiltration is a desirable treatment method. Infiltration capacity is a function of soil saturation level, and a completely saturated soil will have the lowest infiltration rate, much like it has the lowest hydraulic conductivity. Once fully saturated, infiltration continues at a fixed rate, which is a proxy for saturated hydraulic conductivity, and is therefore also hypothesized to be a KPI for system health as it indicates capacity to store water below the soil's surface. This infiltration continues even after ponding has ended, as the zone of fully saturated soil extends above the water table level where full saturation is essentially permanent. Significant, long-term water table mounding has not been found to occur as a direct result of infiltrating GSI practices (Tu and Traver, 2019; Machusick et al., 2011), meaning that infiltrating water generally recedes until it reaches the normal water table level. Therefore, infiltration should be expected to continue downward at a constant rate resulting in the steady recession of the saturated zone along a 'drying front' similar to the 'wetting front' seen at the beginning of saturation. The following sections discuss the data collected at SMP A used to support these hypotheses, and show that recession rate and soil drying rate can be used as proxies for soil health and extended to overall GSI performance.

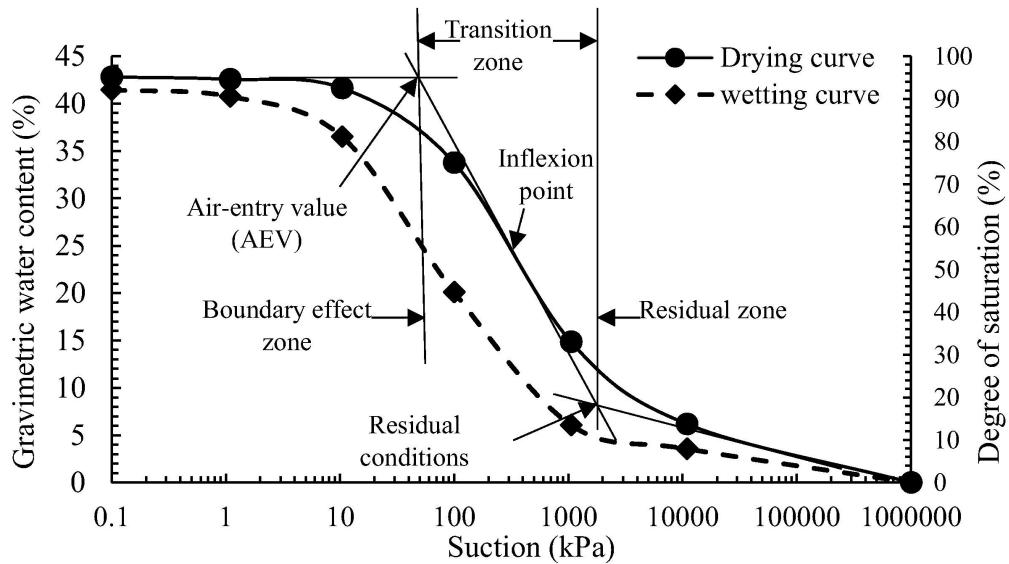


Figure 4.1: Example SWCC.

4.2 Data and Modeling

Data collected is 5 minute averages of 1 minute records (values are measured every minute and stored every 5 minutes as the average of the previous 5 values) over the course of the 3 year study (May 2017 - November 2020). This data resolution captures enough detail without producing an overwhelming number of data records. To better synthesize the data collected, summary statistics for predefined hydrological events are generated based on either a total amount of rainfall followed by a minimum dry time (typically 6 hours), or based on characteristics of other variables of interest - namely ponding and soil moisture values.

4.2.1 Rainfall Events

Using clearly defined rules to identify events is the first step to generating meaningful summary statistics that are applicable across time and space, and useful for comparing GSI with dissimilar properties. In a GSI context, events are storms that result in a measurable quantity of precipitation. Small events typically do not have a pronounced impact on GSI, but the effectiveness of GSI systems is typically greater for smaller or less intense storm events (Liu et al., 2020). However, these smaller events make up the bulk of storm events observed at SMP A and by others (Albright and Schramm, 2018), with 95% of observed total rainfall less than 42mm (Figure 4.2). For this analysis, a 'standard,' or 'normal' storm event begins when rainfall is first recorded and continues until 6 hours after the last observed rainfall, as studies have shown storms separated by 6 or more hours to be meteorologically separate events (Wadzuk et al., 2017). This means the minimum event

duration is 6 hours, and any periods without rainfall of less than 6 hours are continuations of the same storm event. This definition is used as the standard both for generating storm summary data in monthly reports to PennDOT and for all analyses henceforth that do not directly involve soil moisture observations. During the period of interest, there were 414 observed standard events with a mean duration of 12.2 hours, a median of 9.5 hours, and a standard deviation of 7.5 hours. The data are highly right-skewed, with a maximum of 52.2 hours, and appear to follow a roughly chi-square distribution (Figure 4.3).

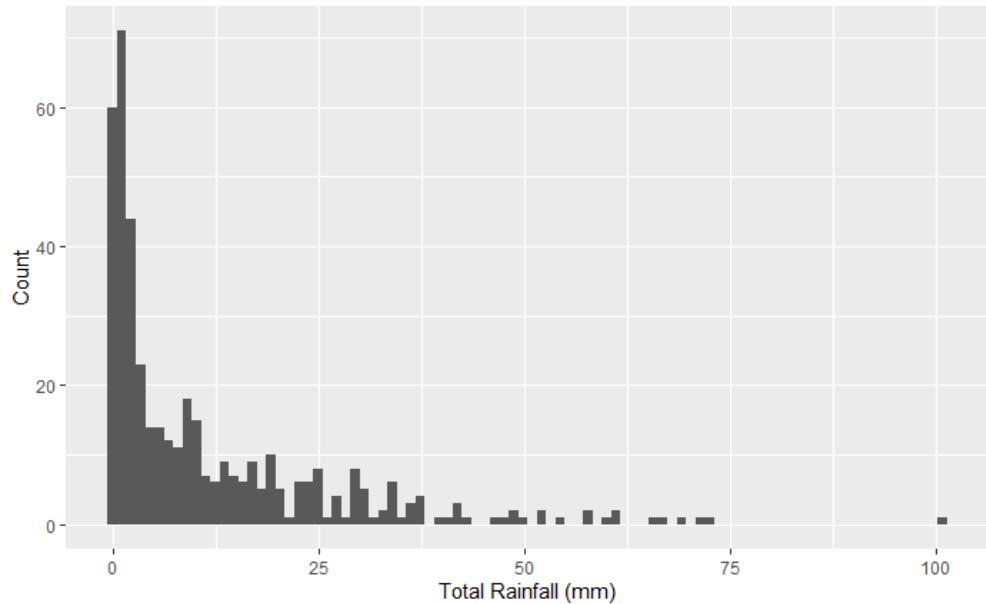


Figure 4.2: Distribution of event total rainfall depths.

Events are distributed throughout a Gregorian 365-day year approximately uniformly, as seen in Figure 4.4, although there are some notable event clusters in both late spring (around ordinal day 150), and during hurricane season (after ordinal day 200). There are notably fewer events during the winter months, but this is likely due to the instrumentation's lack of sensitivity to snowfall. Snowfall events are also less taxing on GSI systems in general, although snowmelt events may have similar characteristics to a rainfall event but have not been studied.

4.2.2 Soil Moisture Events

Events involving soil moisture require a modified approach to event definition due to the prolonged nature of soil moisture's response to storm events. Soil moisture, and by extension ponded water atop saturated soils, take longer than 6 hours to recover after storms of interest. Studies have shown that the recovery period for soils is 72-96 hours (Davis et al., 2009; Wadzuk et al., 2017). To define events for these analyses, both rainfall and soil moisture levels are considered. As before, an event begins with the first observed rainfall. However, the end of the event is defined by a user-defined

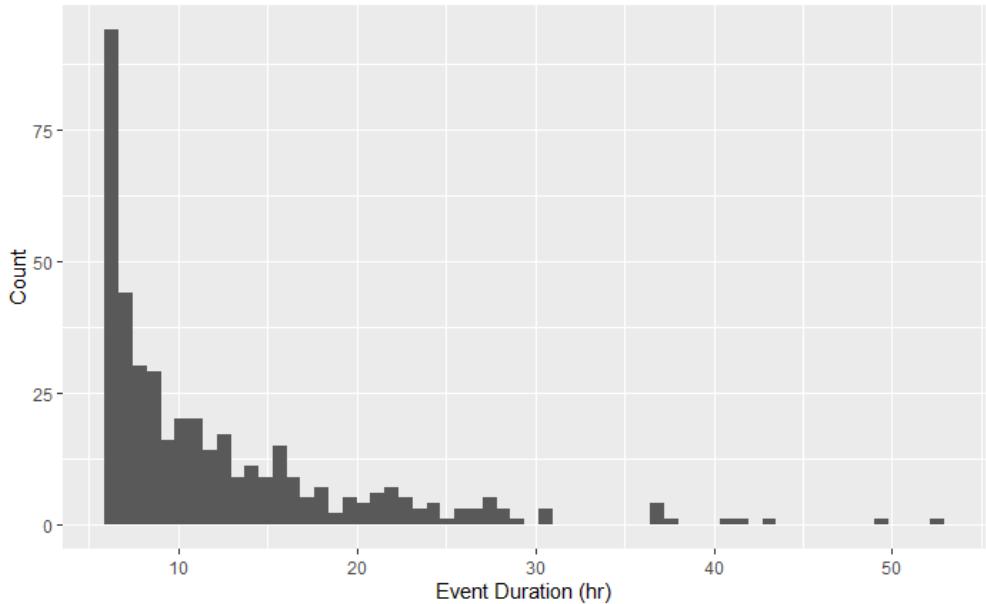


Figure 4.3: Distribution of event lengths.

minimum event time (6 hour default) as well as a threshold ponding level or soil moisture level (either may be used, depending on the type of analysis being performed). A simple moving average of the number of steps comprising the minimum event duration (in other words, the data interval - 5 minutes - divided into the minimum event time of interest) provides a window into the past or the future at a given time step, and can be used to define an event as follows:

$$\begin{aligned}
 \text{Event} = & \{ A = [\text{moving_average}(\text{rainfall}, Q/P) > 0], \\
 & B = [\text{moving_average}(\text{rainfall}, Q/P) == 0 \\
 & \cap \text{moving_average}(K, Q/P) < R] \}
 \end{aligned} \tag{4.1}$$

where A is an indicator for the beginning of an event, B is an indicator for the end of an event, Q is the minimum event duration, P is the observation interval, R is a threshold value, and K is the variable of interest. Timestamps falling between successive A and B values are assigned a unique, random string of 8 characters generated from the first timestamp in the series. The distribution of these specialized events is similar to that of the standard events above, although there are fewer of them due to the increased length of events capturing a greater share of rainfall per event.

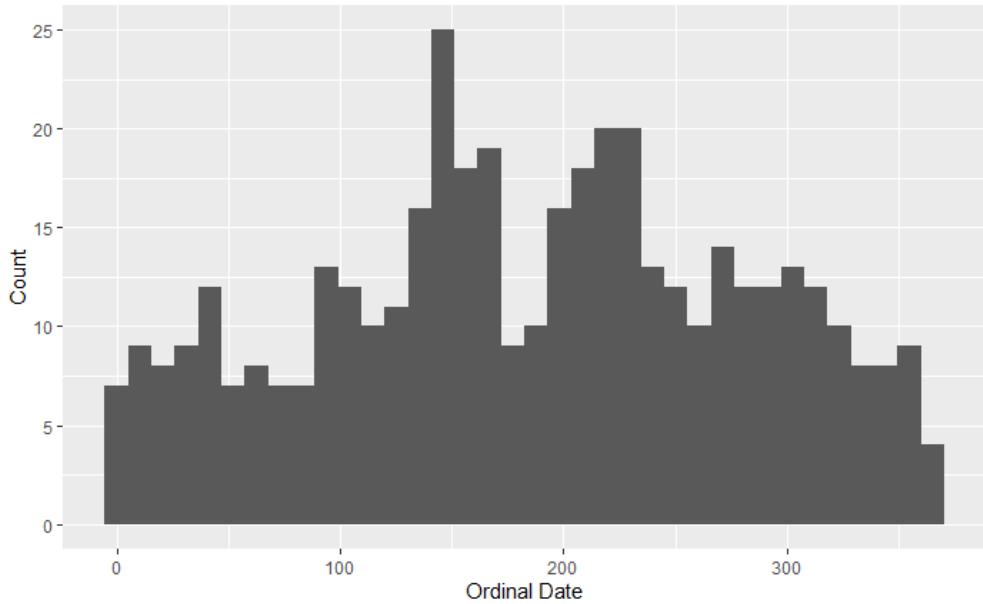


Figure 4.4: Distribution of event ordinal dates.

4.3 Key Performance Indicator Definitions

4.3.1 Recession Rate

Calculating a water level differential across consecutive data records identifies a change in water level that can be used to determine periods of increasing or decreasing volume in the ponded storage. Recession of the pond is indicated by a negative change in level, although no rainfall or inflow can happen simultaneously to ensure that change in level is solely due to infiltration. Additionally, the current water level must be confirmed as being below the overflow point of any outlet structures to ensure, as much as possible, that the recession rate can be attributed solely to infiltration of water into the soil, as opposed to overflow to the CSS. The magnitude of recession is important because it directly correlates to infiltration rate, which is governed by the health of the GSI soils (Gregory et al., 2006; Horton et al., 1994). Larger recession rates indicate a faster drawdown of the water level, while smaller values could be due to prolonged saturation conditions or under-performance due to compaction or clogging of soil pores. Comparing average recession rates calculated over the duration of a storm to other simultaneously recorded atmospheric (temperature, relative humidity, barometric pressure) and GSI state (current pond level) data shows the relationships that have the most significant impact on GSI performance. Changes in the recession rate over the period of a storm or between the average recession of storms over long periods of time indicate changes in the soil health (Jenkins et al., 2010; Zukowski et al., 2016).

The relationships between recession and atmospheric variables or GSI state are complicated by several factors, namely the timing and size of an event, the pre-event geomorphology and hydrologic

state of the GSI, and their collective interaction. The timing and size of a storm event, which can be best described by a combination of time of year, hyetograph, and the length of time since the last storm event, are important because these play the largest role in determining how the pre-storm state of the GSI will react to specific conditions. A GSI system will respond differently at different times of the year to two events with identical hyetographs, while the same GSI with identical starting conditions will respond differently to two different hyetographs. Varying any single variable used to describe an event changes the GSI system's initial response to that event.

The variability of the pre-event state of the GSI necessitates an adaptable baseline to use in evaluating performance. Performance cannot simply be stated as the change from pre-event conditions, because this would make comparisons between dissimilar pre-event states unnecessarily complex. For example, a storm taking place during early spring in the Philadelphia, PA region could have a wide range of pre-storm soil moisture, air and water temperatures, or plant growth conditions, to name a few, all of which will impact how the system responds at the beginning of a storm. Additionally, the fact that suction head is highest when the soil is dry (Eyo et al., 2020, Figure 4.1), means that infiltration loss is greatest at the beginning of an event, no matter the specific initial conditions, because that is when the soil is driest. This inverse relationship means that the most consistent means of comparison will be between the most steady parts of a storm event once the system reaches equilibrium and is treating water at a constant rate. For recession rates, this means looking at the trailing end of a storm, when inflow has ceased, ponding level has reached its peak, and the soil has reached saturation. Allowing the GSI to reach equilibrium means that comparisons will be made strictly between the soil's infiltration performance, without the influence of a specific storm event's timing characteristics. Therefore, healthy performance in the context of recession is the rate at which the system begins reduces ponded volume and approaches recovery from saturation towards fully dry conditions. Full infiltration of ponded water is required to be complete within 72 hours for systems in Philadelphia, but for the majority of storms it is completed within the 6 hour post-rainfall window used to define an event.

4.3.2 Infiltration Drying Rate

Similar to the recession of water ponded on the surface of the garden, the recession of water within the soil column reflects the soil health of the site. Using soil moisture probes at 10, 35, and 60 cm allows the calculation of an average rate of the descent of saturation conditions, or 'drying front', as the soil begins to recover following ponding. This subsurface drying rate should remain consistent between events of different sizes, so long as the soils reach saturation prior to the end of measurable ponding. Measuring the soil moisture values for the duration of a storm event produces a curve that is typically shaped like that in Figure 4.5. The soil moisture curve can typically be expected to respond with a sharp upward jump, 5-20 minutes in length, at the beginning of the storm event's inflow when the wetting front passes each sensor and soil approaches full saturation. This also

coincides with a sharp drop in suction head, as the increased saturation reduces the soil's ability to pull water in quickly. The soil moisture curve has been characterized with a series of seven points in work completed by Matina Shakya, Ph.D. Candidate working on soil moisture measurements for VCRWS/PennDOT. The following is based on the third point in this series. The beginning and end of this positive jump are identified as points A and B, respectively. Point A is typically below 0.40, or 40% volumetric water content (VWC), while point B and C are typically between 0.45 and 0.50, or 45-50% VWC. Soil moisture levels remain at saturation conditions for the duration of ponding, due to continuous infiltration, eventually leading to recession of all ponded water. The reduction in volume from the system due to evaporation, evapotranspiration, or any other means is assumed to be negligible compared to infiltration during a storm and in the hours immediately after. When ponding has ended, infiltration is expected to continue as the saturated zone recedes down into the soil profile (Figure 4.6). Recession of the saturation conditions within the soil behave much like the initial wetting front, but in reverse as water continues to percolate downwards through the soil column. The trailing limb of the saturation curve can be identified by two additional points, C and D in Figure 4.5, identified as the drop off point and inflection point of the soil moisture curve, respectively. In Figure 4.6, the difference in time between steps 2 and 3 is the subject of interest.

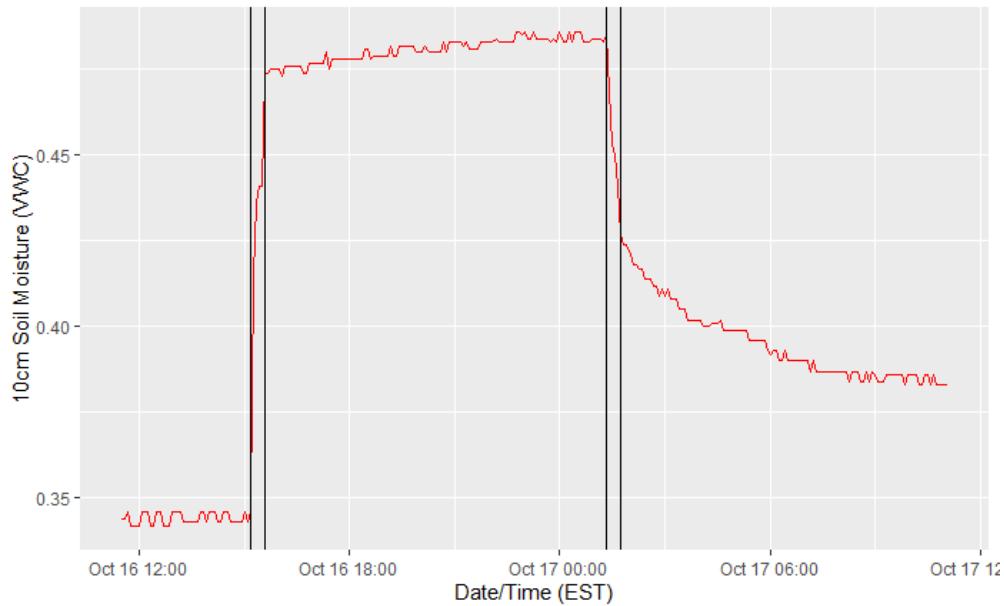


Figure 4.5: Typical soil moisture curve. Vertical black lines denote points A, B, C, and D from left to right, respectively.

The calculation of this metric is slightly more complicated than raw recession rate of the ponded water presented earlier. The inflection points A, B, C, and D are calculated as the 4 primary peaks in the absolute value of the first differential of the soil moisture data:

$$InflectionPoint\{A, B, C, D\} = localmax(|\theta_i - \theta_{i-1}|)_j \quad (4.2)$$

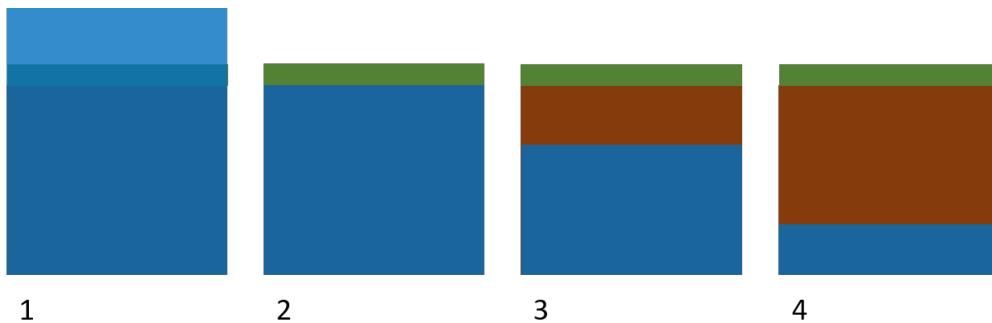


Figure 4.6: Illustration of infiltration drying rate showing 1. ponded water, 2. end of ponding, 3. top-most soil returns to field capacity, 4. saturation zone reaches equilibrium with water table.

where $i = 1,2,\dots,n$ for n observations in a given storm and where $j = 1,2,3,4$ corresponding to A,B,C,D. These points are easily calculated using the ‘findpeaks’ function from the R package ‘pracma’ (`pracma:::findpeaks()`). The function takes a vector of data, the absolute value of the differential soil moisture data applicable to one storm event in this case, and returns the indices of the requested number of local maximums (4 in this case), which can be put in order with a simple sort. Depending on the specific values of soil moisture seen during an event, the four indices may be returned in any order, largely based on the magnitude of difference between any two given points. The basic premise only requires that the four points are properly identified, as they can be easily sorted to identify their correct placement. However, due to the highly variable pre-event state of the GSI, storm events that begin with saturated VWC values (>0.44 VWC) often misidentify some or all of these points, so these events have been disregarded in the context of this analysis.

4.4 Results and Discussion

4.4.1 Recession Rate

Temperature Dependence

The weather station at SMP A records air temperature, and each pressure transducer (PT) records the temperature of water in which it is submerged. The two types of temperature are highly correlated (Pearson correlation = 0.981) For the purpose of this analysis, only water temperature will be considered since water is in direct contact with the soil it is infiltrating and displays some variance of viscosity and other properties at different temperatures. Testing for variations in recession rate across the range of observed water temperatures shows a relationship that is linear (Figure 4.7), as evidenced by the high correlation coefficient of 0.735 and R^2 of 0.54.

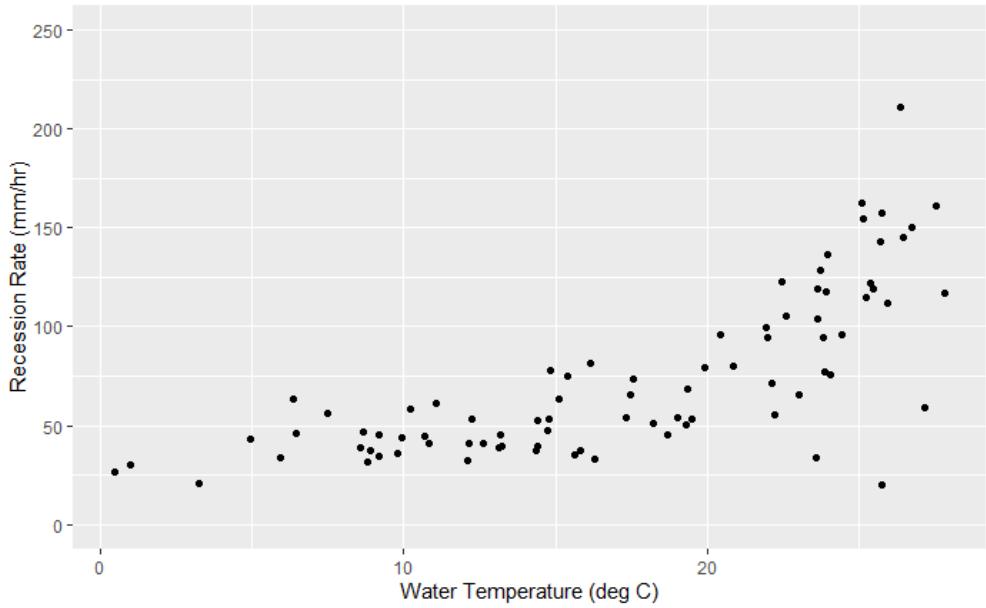


Figure 4.7: Average water temperature vs Recession Rate for SMP A storm event.

The relationship between recession rate and ponded water temperature becomes more apparent when temperature is separated into bins (Figure 4.8). The trend of higher observed recession rates at higher temperatures is clear, as the mean for each bin is higher than that of the previous bin. Five bins were chosen to ensure the number of events per bin was at least 10 for most bins. The lowest bin ((-0.05,5.83]) has only 7 events but follows an identical trend to the upper four bins in magnitude of mean and quartiles (indicated by upper and lower vertical lines). This trend aligns with several physical and hydrologic models, namely that warmer water is less viscous and flows easier, especially through the soil (Emerson and Traver, 2008), such that, since warmer soils contain warmer water, there is less internal resistance to flow by the soil at warmer temperatures. This finding means that GSI performance can be expected to increase during warmer months, both due to increased infiltration capacity and increased water uptake and evapotranspiration by the plant mass (Bartens et al., 2008). Furthermore, it has been shown that soils that have experienced multiple freeze-thaw cycles have higher values of saturated hydraulic conductivity due to the formation of internal ice crystal structures within frozen soils, which can expand pore space and increase the "development of macroscopic cracks and microscopic voids" (Asare et al., 1999). Therefore, soils that experience freezing conditions followed by warm conditions, as is typical in a climate with a distinct harsh winter season, can expect to see a natural "loosening" of the soil via the creation of these additional void spaces. This increase in void spaces has the potential to cause preferential flow paths through the soil, which will lead to increased infiltration rates up to the frost line, where the effect of freeze-thaw cycling ends.

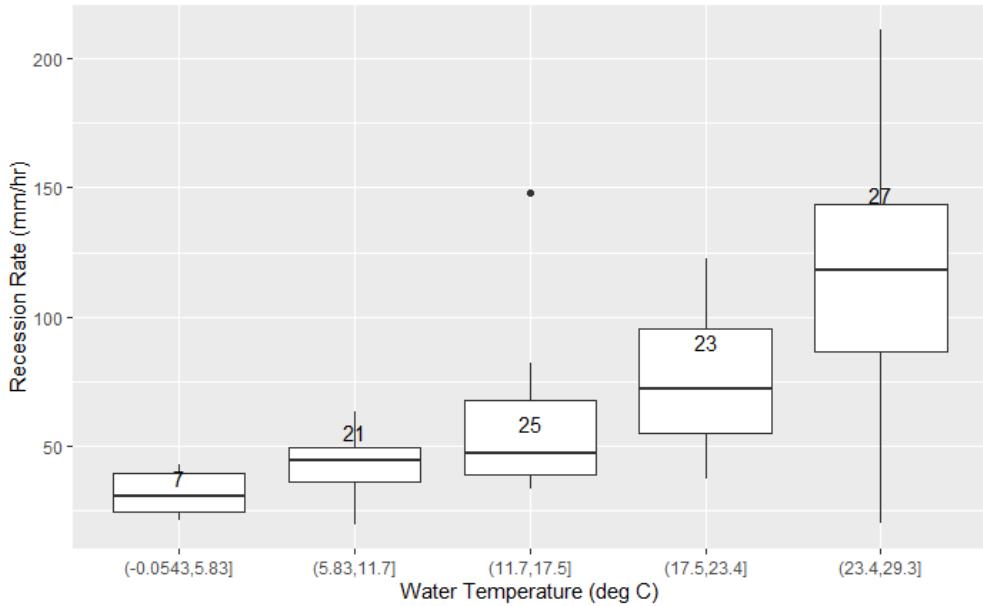


Figure 4.8: Average water temperature vs Recession Rate, separated into bins, n = number storm events per bin.

Insignificant Relationships

Both relative humidity (RH) and barometric pressure, both recorded in a similar fashion to air temperature at SMP A, showed no significant relationship with recession rate, with correlation coefficients of 0.148 and 0.25 and R^2 values of 0.022 and 0.064 respectively. Comparing the average recession rate to RH bins and pressure bins, similar to the approach taken in the previous temperature analysis, shows no correlation between different values for these variables as compared to the average recession rate (Figure 4.9 and 4.10).

From these observations, GSI performance is not dependent on barometric pressure or relative humidity, which largely only impact the soil-air interface. The lack of relationship between the atmosphere and sub-surface infiltration mechanics is expected, since the soil contains enough mass to prevent significant influence from the relatively fast-changing atmosphere.

The level of ponded water collected in the rain garden similarly has no significant effect on the rate of recession, with a correlation of -0.2 and R^2 of 0.043. This runs contrary to expectations that greater head pressure at the soil-ponding interface would lead to higher recession rates. The rate of recession is consistent across nearly the entire ponding depth range (Figure 4.11). The lack of correlation, while surprising, could be due to the fact that sub-surface saturation leads to high enough pore water pressure such that it cancels out the head pressure seen at typical ponding depths of up to a maximum of less than 1 meter. Additionally, the curved profile of the garden bed means that a depth measurement taken in the middle of the garden captures the deepest point, and less significant depths are experienced by the majority of the surface area (Sokolovskaya et al., 2021).

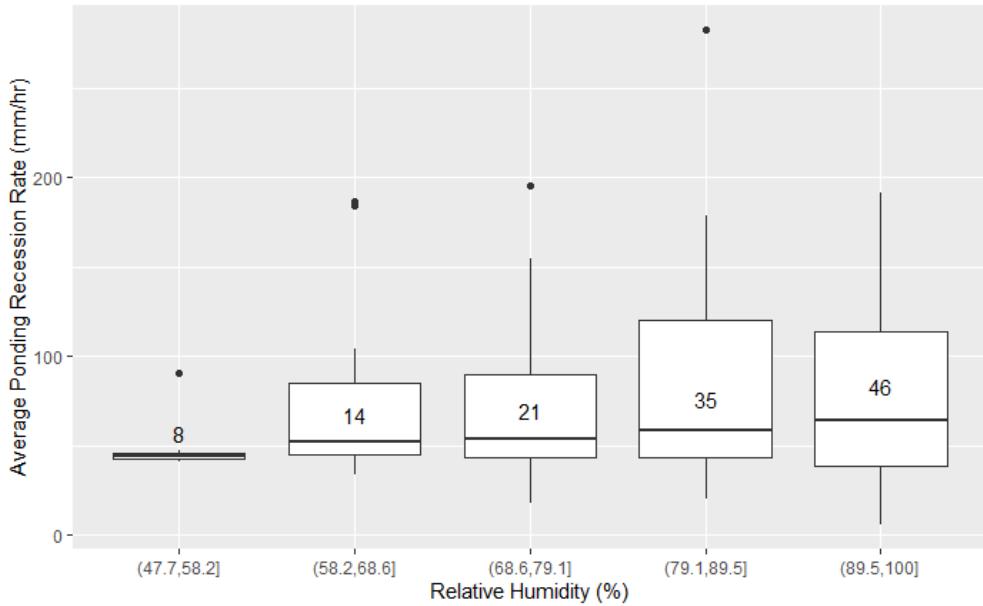


Figure 4.9: Average Relative Humidity vs Recession Rate

While the slope estimate and F-statistic in a linear model for recession rate and pond depth are both significant at the 0.05 level, the low R^2 value means the model is only predicting the mean recession rate.

Regression

Using multiple linear regression, a numeric relationship between the recession rate and pond temperature, max pond depth, ponding duration, total rainfall, and average rainfall intensity can be determined. The expected value of this relationship can determine if a GSI system is functioning nominally. Values that are outside the expected relationship boundaries are cause for concern, and a trend of values outside the norm would indicate that maintenance or further evaluation of the system is required. While the numbers presented here are specific to PennDOT's SMP A, this approach, and indeed the entire scope of data collection, storage, and analysis would be of tremendous value to any system or group of systems with similar properties.

Regressing the following parameters onto ponding recession rate (mm/hr) results in the listed effects (Table 4.1), which results in Equation 4.3, the relationship for water temperature, pond depth, ponding duration, total rainfall, and average rainfall intensity with respect to ponding recession rate. The R-squared value is 0.9023, with an F-statistic (significance test) of 312.1 on 5 and 169 degrees of freedom ($p\text{-value} = < 2.2e - 16$), indicating the partial slopes ($\beta_1 \dots \beta_5$) are not equal to 0. There are, however, normality concerns, as demonstrated by the normal quantile-quantile plot (Figure 4.12), which shows a departure from linearity beyond ± 1.5 quantiles. This is somewhat expected, as the data are not normally distributed, so an additional nonparametric test is required to

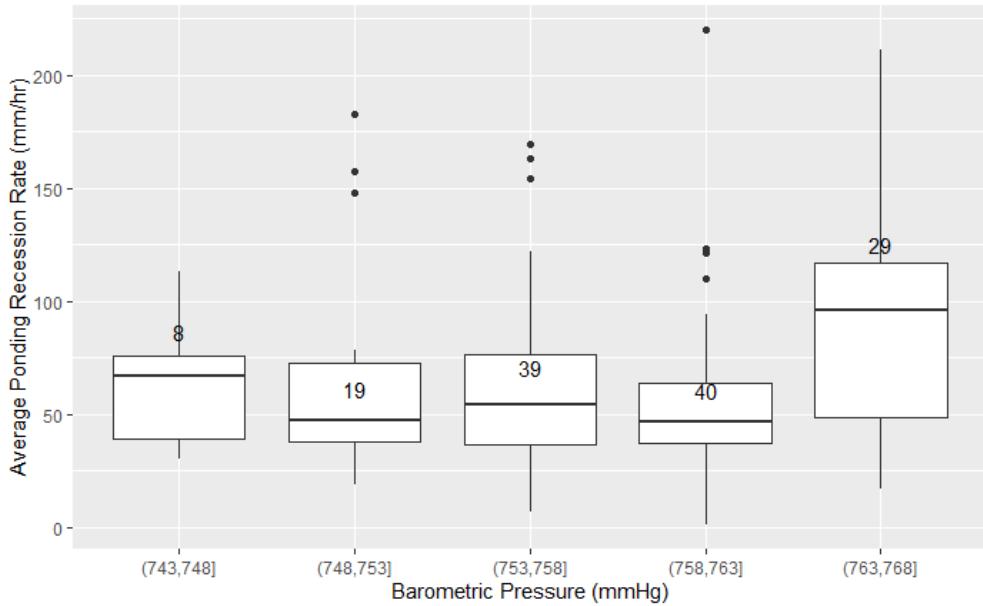


Figure 4.10: Average Barometric Pressure vs Recession Rate

confirm the trend seen. The relationship lacks an intercept because the mean response is expected to be 0 mm/hr, corresponding to no ponding recession when all other variables are also 0.

Table 4.1: Ponding recession rate regression results.

	Coefficient	Estimate	Std. Error	t value	Pr(> t)
β_1	B1 pond mean temp (°C)	-0.238352	0.085110	-2.801	0.005697
β_2	B1 pond max depth (mm)	-0.117213	0.005684	-20.622	<2e-16
β_3	B1 ponding duration (hr)	0.981035	0.224939	4.361	2.24e-05
β_4	Total Rainfall (mm)	0.182069	0.063980	2.846	0.004980
β_5	Average Intensity (mm/hr)	-0.377106	0.110401	-3.416	0.000797

$$y = -\beta_1 * 0.2384 - \beta_2 * 0.11721 + \beta_3 * 0.98104 + \beta_4 * 0.1821 - \beta_5 * 0.3771 \quad (4.3)$$

The regression model in Equation 4.3 was chosen for its high R-squared value, which reflects the percent of variation in the response (recession rate) explained by the total variation in the model parameters, as well as its overall significance (high F-statistic with low p-value). Parameters kept in the model (listed in Table 4.1) have individual p-values < 0.05 . Estimates of recession rate predict the mean response for each unit change of the model parameters multiplied by their respective coefficients. Other models considered included some combination of the final parameters, water temperature, event duration, and ordinal day. While the equation predicts the mean response given a set of inputs and will not always be precise, a confidence interval of 95% can indicate values that are outside the expected range. Consecutive events with abnormally high or low recession values will indicate a need for inspection or intervention.

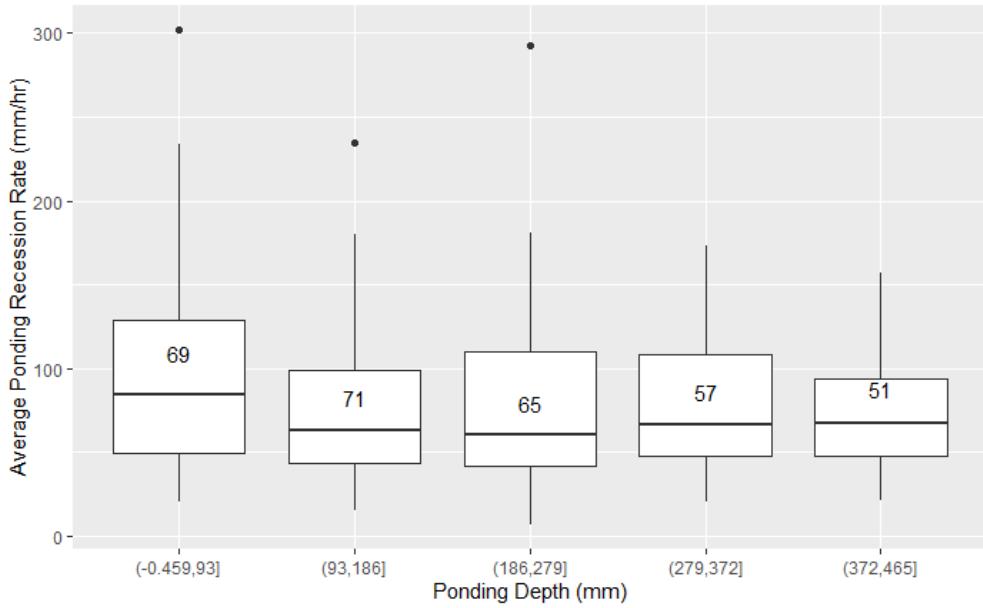


Figure 4.11: Average Ponding Depth vs Recession Rate

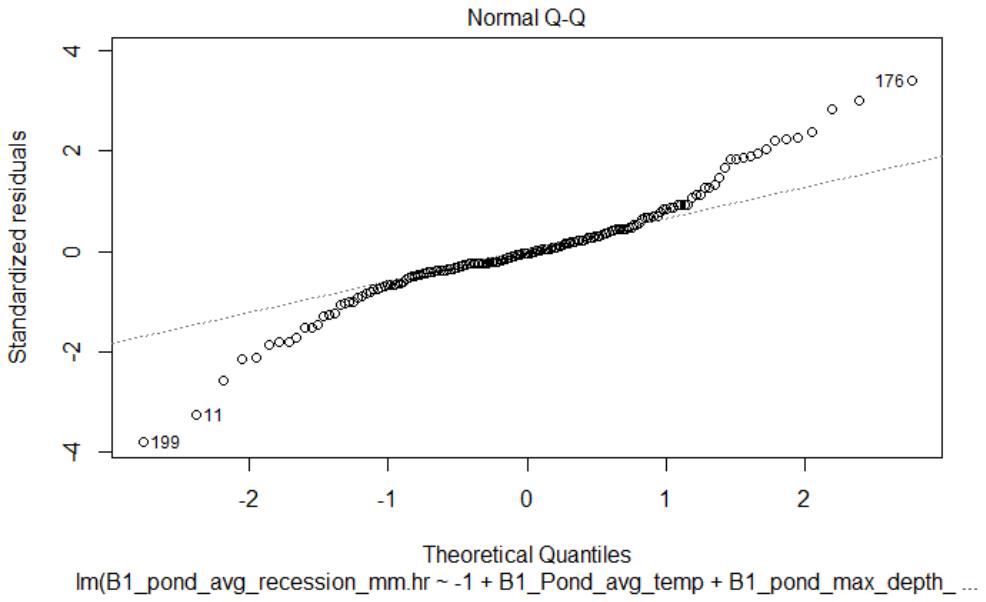


Figure 4.12: Normal Q-Q plot for ponding recession regression.

The model is not without its shortcomings, however, and does not account for departures from normality of the training data. The Kendall seasonal trend statistic, τ , tests for seasonal trends in monotonic, nonparametric data. The τ statistic is a ratio of the probabilities of the observed order of the data to that of a different ordering. That is, assuming two variables (X and Y, for example) are independent, the expected value of τ is 0, as the likelihood of the observed ordering is equal to that of any other ordering for two independent series Abdi, 2007. Calculating τ for comparison between B1 recession rate and the variables found in the regression equation 4.3 results in the values

found in Table 4.2. All the relationships are significant enough to reject the claim that the pairwise comparisons are independent, suggesting the trends observed, and by extension their inclusion in a normal multiple regression, are significant, despite the lack of normality.

Table 4.2: Kendall seasonal trend results.

Variable	Kendall τ	p-value
B1 pond mean temp ($^{\circ}\text{C}$)	0.1226526	0.001711
B1 pond max depth (mm)	0.9150601	<2.2e-16
B1 ponding duration (hr)	0.8294355	<2.2e-16
Total Rainfall (mm)	0.5483617	<2.2e-16
Average Intensity (mm/hr)	0.4102984	<2.2e-16

4.4.2 Infiltration Drying Rate

There are 76 valid events during the period of interest, essentially a subset of the standard events, quantified with different ending criteria. For each event, the rate of drying for the 3 zones between the surface and the three soil moisture sensors at progressively deeper placements is calculated as distance/time. Figures 4.13, 4.14, and 4.15 demonstrate that this rate can be expected to fall within 0-5 cm per hour. The 95% percentile falls at 9.5 hours between 10 and 35cm (approx 2.63cm/hr), and 7.4 hours between 35 and 60cm (approx. 3.37cm/hr), meaning 95% of events can be expected to exhibit drying rates higher than this. The distribution does not display correlation with water temperature (Pearson correlation 0.047, p-value 0.34), and has remained consistent over time, as displayed in Figure 4.16.

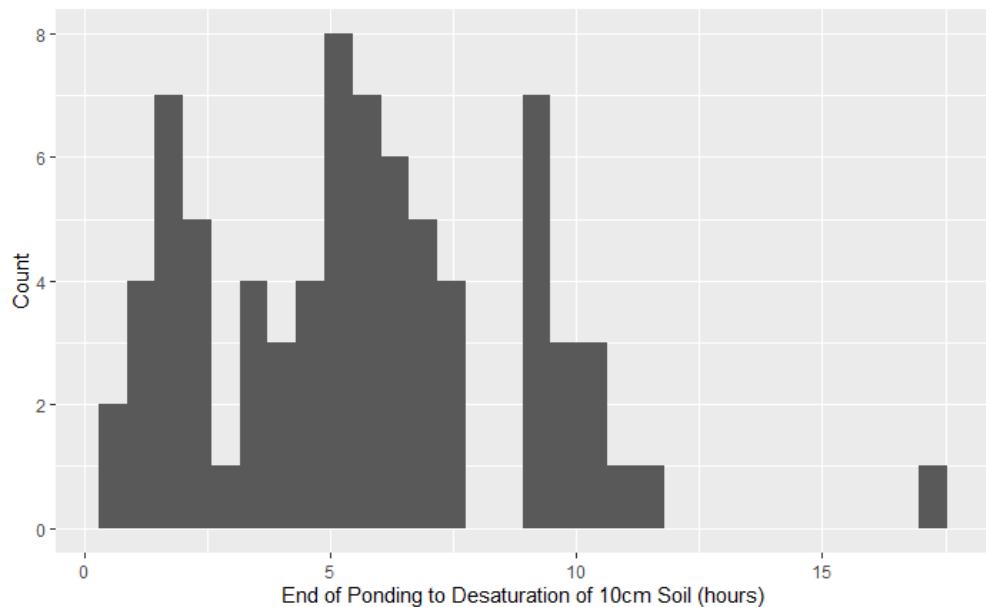


Figure 4.13: Time to Desaturation between end of ponding and 10cm.

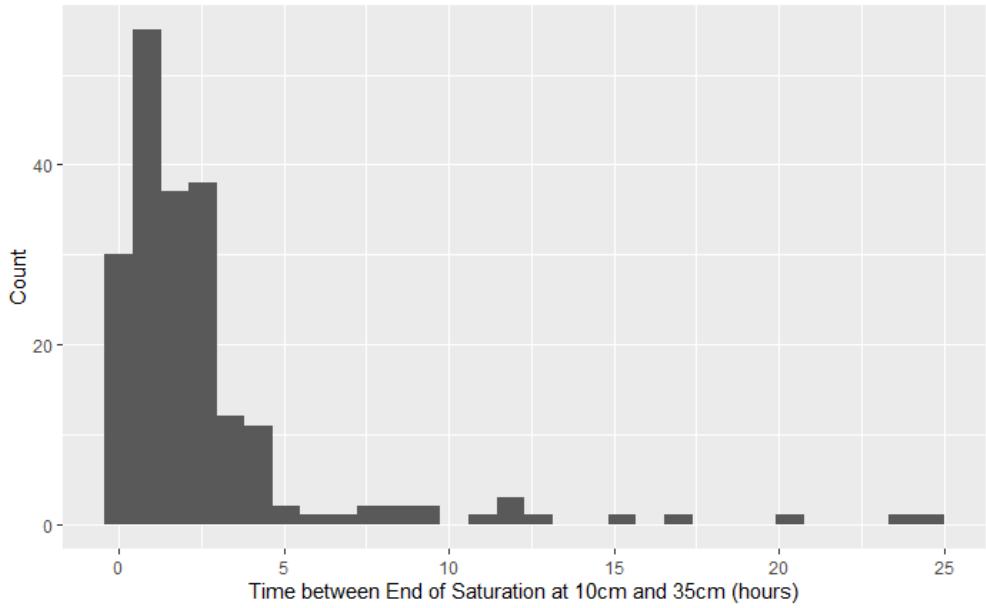


Figure 4.14: Time to Desaturation between 10 and 35cm.

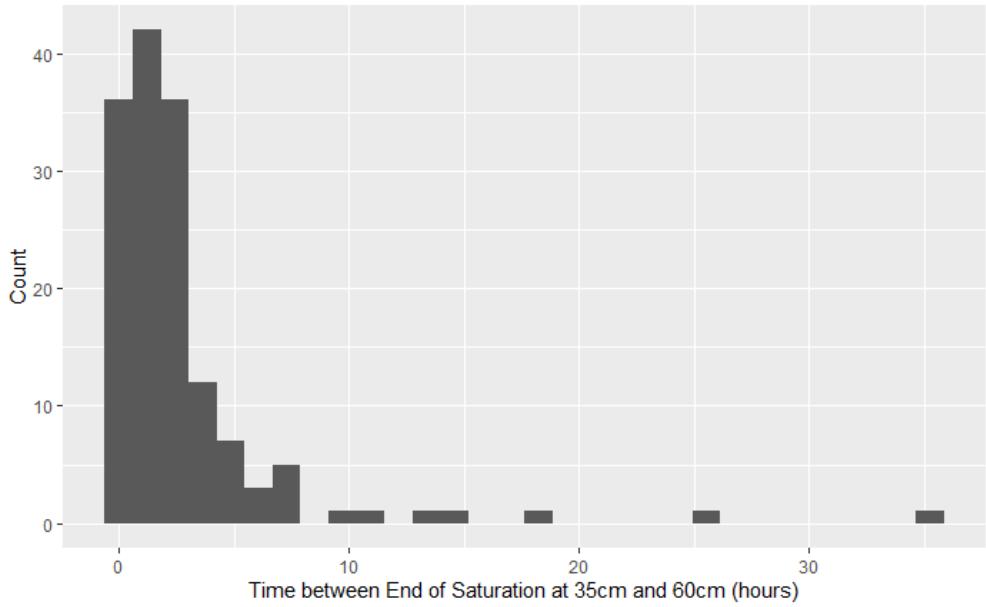


Figure 4.15: Time to Desaturation between 35 and 60cm.

Infiltration drying rate compares favorably to Saturo constant-head infiltration testing performed at the site across three growing seasons. These tests showed a mean K_{sat} value of 1.3-6.2 cm/hour, which compares favorably with the mean infiltration drying rate of 1.8-9.5 cm/hour. Both metrics have large variances, and it would be helpful to do more direct comparison testing to investigate if the variability is consistent between them.

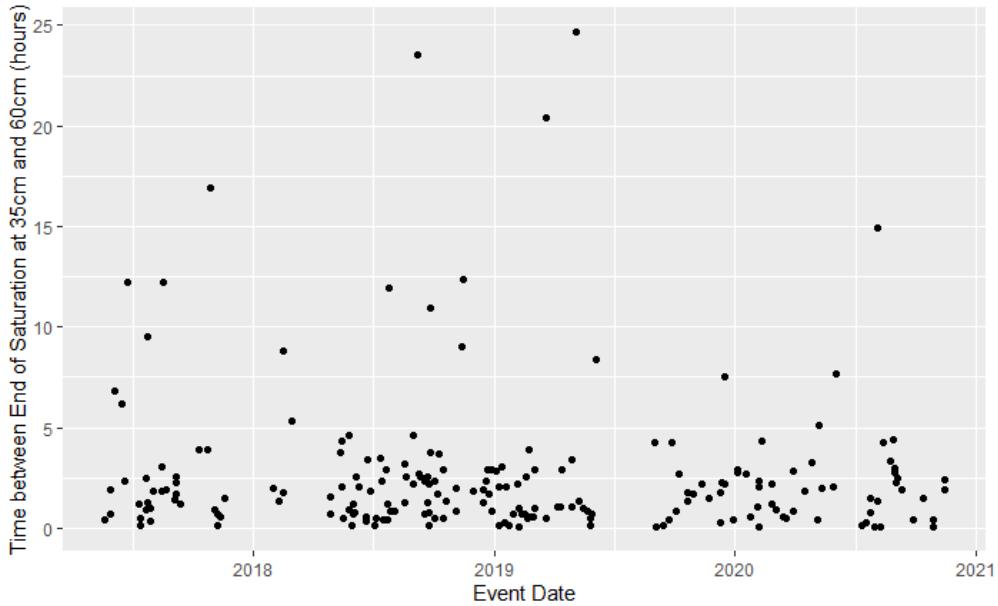


Figure 4.16: Event Date vs Drying Time between 10 and 35cm.

4.5 Conclusions

The methods described here attempt to evaluate methods for comparing storm events across time and space. Using recession and infiltration KPIs to evaluate a system of GSI performance across an entire region will enable insights into long term design successes, and early detection of system errors. The data have shown that performance is most highly impacted by temperature, which oscillates with an annual seasonal period. GSI can generally be quantified in terms of recession and drying rates for ponding level and soil moisture, respectively. Higher rates for each quantify indicate faster transfer of water into the subterranean water table, and faster reduction in soil moisture that will lead to better preparation for the next storm.

In general, average infiltration rates between 40 and 120 mm/hr can be expected, depending on time of year and water temperature. Average drying rates centered around 9.5 cm/hr can be expected, although this value would be expected to vary from GSI to GSI based on soil design criteria. These statistics are not dependent on GSI parameters such as surface area, loading ratio, or number of inlets or outlets, and are therefore independent of site configuration. The statistics can be calculated for any GSI system, or part of a system, to determine if infiltration rates and recovery rates fall within the expected range. Thanks to the few parameters needed, namely ponding depth and soil moisture level, few sensors are necessary, keeping costs for this kind of monitoring low, making it accessible to be deployed at a larger number of sites, or at multiple locations within a single site. While this work intentionally does not address spatial variability within a single site for infiltration or drying rates, the data collected are assumed to represent the average conditions throughout SMP

A, and the same assumption would be valid for the statistics calculated at other sites in a similar manner.

The recession and infiltration rates are affected by conditions both above and below the soil surface and provide an acceptable proxy for soil health. The recession rate is shown to be affected by temperature and could be reduced by clogging of the soil surface or compaction of lower soil layers, leading to reduced rates. Similarly, the infiltration drying rate is shown to be consistent across a wide variety of GSI conditions and could be adversely affected by changing soil properties that make the GSI function less efficiently. The established relationships and expected values for each are useful as indicators for long term performance, and departures from expectations are cause for further investigation.

Conclusion

Bibliography

- Abdi, Hervé (2007). "The Kendall Rank Correlation Coefficient". In: *Encyclopedia of Measurement and Statistics* (cit. on p. 31).
- Akhavan, Bloorchian Azadeh and Zhou Jianpeng (2015). *Alternative Treatment of Flow Monitoring Data to Evaluate the Impact of Green Infrastructure on Stormwater Volume Reduction in Combined Sewers* (cit. on p. 2).
- Albright, Cara Melissa and Harrison Schramm (2018). "Improvements and applications in climate data analysis for determining reference rainfall years". In: *Journal of Applied Meteorology and Climatology* 57.2, pp. 413–420 (cit. on pp. 10, 20).
- Asare, S. N., R. P. Rudra, W. T. Dickinson, and G. J. Wall (1999). "Effect of freeze-thaw cycle on the parameters of the green and ampt infiltration equation". In: *Journal of Agricultural and Engineering Research* 73.3, pp. 265–274 (cit. on p. 27).
- Bartens, Julia, Susan D. Day, J. Roger Harris, Joseph E. Dove, and Theresa M. Wynn (2008). "Can Urban Tree Roots Improve Infiltration through Compacted Subsoils for Stormwater Management?" In: *Journal of Environmental Quality* 37.6, pp. 2048–2057 (cit. on p. 27).
- Burcin, Becerik-Gerber, Siddiqui Mohsin K., Brilakis Ioannis, et al. (2014). "Civil Engineering Grand Challenges: Opportunities for Data Sensing, Information Analysis, and Knowledge Discovery". In: *Journal of Computing in Civil Engineering* 28.4, p. 4014013 (cit. on p. 2).
- Callahan, Bernadette M (2019). "Green Stormwater Infrastructure for CSO Control". In: *World Environmental and Water Resources Congress 2019*, pp. 184–189 (cit. on p. 2).
- Calt, Elizabeth (2018). "Comparing the Hydrologic Performance of a Linear Cascading Bioswale to Traditional Bioinfiltration in a Highly Urbanized Setting". Master's Thesis. Villanova University (cit. on p. 10).
- Davis, Allen P., William F. Hunt, Robert G. Traver, and Michael Clar (2009). "Bioretention Technology: Overview of Current Practice and Future Needs". In: *Journal of Environmental Engineering* 135.3, pp. 109–117 (cit. on p. 21).
- Department of Environmental Protection (2006). *Pennsylvania Stormwater BMP Manual*. Tech. rep. (cit. on p. 18).
- Eargle, John M (2002). "Resistance Change with Temperature for Copper". In: *Electroacoustical Reference Data*. Boston, MA: Springer US, pp. 106–107 (cit. on p. 8).
- Emerson, Clay H. and Robert G. Traver (2008). "Multiyear and Seasonal Variation of Infiltration from Storm-Water Best Management Practices". In: *Journal of Irrigation and Drainage Engineering* 134.5, pp. 598–605 (cit. on p. 27).
- Eyo, E. U., S. Ng'ambi, and S. J. Abbey (2020). *An overview of soil–water characteristic curves of stabilised soils and their influential factors* (cit. on pp. 19, 24).
- Gregory, J H, Michael Dukes, Pierce Jones, and Grady Miller (2006). "Effect of Urban Soil Compaction on Infiltration Rate". In: *Journal of Soil and Water Conservation* 61, pp. 117–124 (cit. on p. 23).
- Heffernan, Taylor, Stephen White, Tyler Krechmer, et al. (2016). "Green Stormwater Infrastructure Monitoring of Philadelphia's Green City, Clean Waters Program". In: *World Environmental and Water Resources Congress 2016*, pp. 115–124 (cit. on p. 2).

- Horton, R., M. D. Ankeny, and R. R. Allmaras (1994). "Effects of Compaction on Soil Hydraulic Properties". In: *Developments in Agricultural Engineering*. Vol. 11. C. Elsevier, pp. 141–165 (cit. on p. 23).
- Jenkins, Jennifer K. Gilbert, Bridget M. Wadzuk, and Andrea L. Welker (2010). "Fines Accumulation and Distribution in a Storm-Water Rain Garden Nine Years Postconstruction". In: *Journal of Irrigation and Drainage Engineering* 136.12, pp. 862–869 (cit. on p. 23).
- Liu, Wen, Qi Feng, Ravinesh C. Deo, Lei Yao, and Wei Wei (2020). "Experimental Study on the Rainfall-Runoff Responses of Typical Urban Surfaces and Two Green Infrastructures Using Scale-Based Models". In: *Environmental Management* 66.4, pp. 683–693 (cit. on p. 20).
- Machusick, Matthew, Andrea Welker, and Robert Traver (2011). "Groundwater Mounding at a Storm-Water Infiltration BMP". In: *Journal of Irrigation and Drainage Engineering* 137.3, pp. 154–160 (cit. on p. 19).
- Mays, L W (2010). *Water Resources Engineering*. Wiley (cit. on pp. 3, 12).
- Philadelphia Water Department (2018). *5 Down, 20 to Go: Celebrating 5 Years of Cleaner Water and Greener Neighborhoods* (cit. on p. 2).
- (2021). *City of Philadelphia: Watershed Protection* (cit. on p. 3).
- Sokolovskaya, Natalya, Ali Ebrahimian, and Bridget Wadzuk (2021). "Modeling Infiltration in Green Stormwater Infrastructure: Effect of Geometric Shape". In: *Journal of Sustainable Water in the Built Environment* 7.2, p. 04020020 (cit. on pp. 19, 28).
- Stevens Water (2021). *HydraProbe Data Sheet* (cit. on p. 7).
- Tu, Min-Cheng and Robert Traver (2019). "Water Table Fluctuation from Green Infrastructure Sidewalk Planters in Philadelphia". In: *Journal of Irrigation and Drainage Engineering* 145.2, p. 05018008 (cit. on p. 19).
- USBR (2001). *Water Measurement Manual - Chapter 7 - Weirs*. Tech. rep. (cit. on p. 7).
- USEPA (2009). *Urban stormwater management in the United States*, pp. 1–598 (cit. on p. 18).
- Wadzuk, Bridget M., Conor Lewellyn, Ryan Lee, and Robert G. Traver (2017). "Green Infrastructure Recovery: Analysis of the Influence of Back-to-Back Rainfall Events". In: *Journal of Sustainable Water in the Built Environment* 3.1, p. 04017001 (cit. on pp. 20, 21).
- Zukowski, Zachary, Clay H. Emerson, Andrea L. Welker, and Brendon Achey (2016). "Evaluation of Field Hydraulic Conductivity Data: Comparing Spot Infiltrometer Test Data to Continuous Recession Data". In: American Society of Civil Engineers (ASCE), pp. 517–526 (cit. on p. 23).