ABSTRACT

MOIN, SHEIDA. Evaluating the Benefits of Near-Continuous Monitoring, Real-Time Control, and SCM Visibility in Performance of Stormwater Control Measures. (Under the direction of Dr. François Birgand and Dr. William F. Hunt, III).

Increase of impervious surfaces due to urbanization affects the natural hydrologic cycle (National Research Council, 2009) by reducing infiltration and evapotranspiration of stormwater and increasing runoff volumes and peak flowrates (Federal Interagency Stream Restoration Working Group 1998). Such changes can lead to flooding, soil erosion, and degradation of water quality in receiving streams (Walsh et al. 2005; Erickson et al. 2013). Traditionally, stormwater was quickly conveyed from urban landscapes to reduce local flooding. Management of stormwater eventually evolved to temporary storage of runoff to reduce downstream flooding (Niemczynowicz 1999). More recently (1980s &1990s), water quality control became an important part of stormwater management (Niemczynowicz 1999). Currently, implementation of stormwater control measures (SCMs) within urban areas is one of the main strategies to mitigate the effects of urban stormwater runoff by restoring some of the characteristics of natural hydrologic cycle (Burns et al., 2012).

The next step in stormwater management is to improve the capacity of current SCMs to accommodate foreseeable future changes, such as (1) growth of urban areas alongside population growth and (2) changing patterns of rainfall frequency and intensity. The world's population will increase from the current 7.8 to 9.7 billion by 2050, and 66% of this population is expected to live in urban areas (UN, 2014). The current percentage of urban dwellers is 54% (UN, 2014). The growth of urban areas and high demand on land would lead to both a shortage of space and a high cost of real estate. Therefore, SCMs placed within the urban environment should optimize hydrologic and water quality mitigation for the occupied footprint. Further, SCM designs should

perhaps accommodate for the changing precipitation patterns (Rosenberg et al., 2010). Currently, SCMs are sized based on the historic precipitation data. However, a change in rainfall frequency, intensity and occurrence of more extreme events (Alexander et al., 2006; Wuebbles et al. 2017) suggests design methodology should likewise change.

In an effort to improve SCM performance, this study investigates three main areas: (1) pollutant dynamics of stormwater runoff at near-continuous resolution, (2) use of real-time control (RTC) technology, (3) and SCM maintenance. Near-continuous water quality sampling of stormwater can bring further understanding of the pollutant dynamics and identify the portion of stormwater that carries the majority of a pollutant. Given such information, the selection on type and sizing of the SCMs can be optimized for targeted pollutant treatment. RTC technology is the controlled release (as opposed to passive release) of water from a basin in response to an incoming volume (i.e., storm event) and can be applied to volume-based SCMs. RTC has proven useful for optimizing the volume capacity in wastewater and sewer systems (Schütze 2004; Pleau et al. 2005). Use of RTC on SCMs can both (1) increase the retention time and thus improve the water quality treatment, and (2) increase volume capacity to capture entire events. Lastly, receiving proper maintenance is critical for long-term performance of SCMs. This study examines the effect of an SCM's visibility on its maintenance and consequent function.

The selected SCM types were a constructed stormwater wetland (CSW) and a wet pond for the first two areas of study; bioretention was added for the third. These are primary SCMs (providing highest pollutant removal) and among the most popular SCMs in North Carolina. Additionally, the large footprint of CSWs and wet ponds can be a limiting factor in their placement and widespread use therefore, these two SCMs can benefit from performance optimization and potentially size reduction.

Evaluating the Benefits of Near-Continuous Monitoring, Real-Time Control, and SCM Visibility in Performance of Stormwater Control Measures

by
Sheida Moin

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Biological and Agricultural Engineering

Raleigh, North Carolina

2021

APPROVED BY:

Dr. François Birgand Committee Co-Chair	Dr. William F. Hunt III Committee Co-Chair
Dr. Natalie Nelson	Ms. Julieta Sherk

DEDICATION

I dedicate this work to my spouse, Ehsan whose endless love and support made this journey possible and whose sacrifices got me to the finish line.

BIOGRAPHY

Sheida Moin's journey to a doctorate degree in stormwater management was a journey of finding passion and following dreams. Sheida was born the third child of Gholamali (Al) Moin and Ilmira (Amie) Chogan Sonboli in Tehran, Iran. Growing up she enjoyed spending time in nature, creating craft art, and learning how machines work. However, within the school programs, she couldn't seem to bring all of these together. Eventually, through the educational system, she picked the math program over art in high school and took part in the national annual university entrance exam for math and engineering.

Ranking among the top students, Sheida picked few majors within two main Universities and after 4 years got her bachelor's degree in Textile Engineering and Fiber Science from Amirkabir University of Technology (Tehran Polytechnic). Following graduation, she gained experience in the industry working on fabric dye, finish, and managing wastewater of those processes. During that time, she worked on improving production efficiency and found a surprising motivator for machine operators to stay on the time track. Gaining access to the garden irrigated by the treated wastewater, during break time has improved the timing of the operators. Fascinated by this finding she pursued ways to engage with art again.

Following her newfound interest and in her father's footsteps to pursue education abroad, Sheida got admitted for graduate studies in the field of Landscape Architecture at Clemson University in South Carolina. During the 3-year program, she searched for a focus area within the field and completed her design thesis on stormwater management. Working on the thesis, she realized that successful design of stormwater features requires collaboration between engineering and landscape architecture in the design phase to both manage stormwater and provide a public

space for human interaction. With this realization and two awards of excellence in research from the department, she pursued a PhD degree in stormwater management.

ACKNOWLEDGMENTS

This work and the resulted dissertation were made possible with support from many individuals and institutions. First and foremost, I would like to thank my advisors Dr. François Birgand and Dr. Bill Hunt for believing in me, valuing my diverse background, encouraging me through the process, and for their incredible support especially through the 'travel ban' and Covid-19 pandemic. I would like to thank the institutions that have funded my education and research, the North Carolina Land and Water Fund (formerly Clean Water Management Trust Fund), and the Department of Biological and Agricultural Engineering at NCSU. Thanks to Dragonfly Pond Works for providing additional funding to support a research fellow and obtain data from additional field sites. I am appreciative of the support provided by OptiRTC on the installation and upkeep of the RTC devices, special thanks to Tatjana Toeldte, Conor Lewellyn, Seth Bryant, Micah Strauss, and Dayton Marchese.

I would like to thank my committee members Dr. Natalie Nelson and Ms. Julieta Sherk and former member Dr. Consuelo Arellano for their guidance and feedback. Among all the faculty and staff, I am especially grateful for the support of Dr. John Classen, Dr. Garey Fox, Heather Austin, Lacy Parrish, Alisha Goldstein (for all the technical and emotional support), Sarah Waickowski, Jeffrey Johnson, Shawn Kennedy (for all the support and laughter during instrument installation and troubleshoot), Samuel Garvey (for help with instrument installation and data collection). Additionally, this extensive data collection was made possible with help of many other students/colleagues, I am thankful to Cyrus Belenky, Bryan Maxwell, Qianyu Hang, Jacquiline Welles, Cheng Chen, Sarah Ferriter, Augustin Thomas, Jack Grimes, Angela Gutierrez, Andrew Yount, Maya Hoon. Lastly a heartfelt thanks to all the friends I have made in the department, who made this journey more enjoyable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: USING OPTIC SENSORS FOR NEAR-CONTINUOUS WATER QUALITY MONITORING OF STORMWATER RUNOFF

Abstract

Spectrophotometer sensors proven to be insightful in different fields of hydrological research, were tested for application on a stormwater wetland for near-continuous water quality monitoring. This paper provides methods, challenges and uncertainties associated with using optics-based spectrometers for stormwater water quality monitoring. Use of spectrophotometer sensors allowed for sampling at 4-minute frequency and sampling 62% of events comprising of 95% of total recorder precipitation volume. Challenges of using spectrophotometers in stormwater included the generally low concentration ranges for nitrate, and the flashiness of the hydrology that rendered calibration of the instruments difficult. Two local calibration methods were tested: simple linear regressions (SLR) and partial least square regressions (PLSR) between concentration values reported by the instruments and those from the laboratory. The uncertainties in predicted concentrations, quantified in this study, were due to the regression type, number as well as stratification or lack thereof, of calibration sample pools. In the end, we were able to show that the annual loads, using PLSR as the calibration method on stratified samples, could be estimated within -11.9% to +9.4% for nitrate, -6.1% to +5.9% for DOC, and -14.3% to +15.6% for TSS. Load estimation on an event basis showed higher uncertainties for some events. Overall, loads were estimated within ±10% for 52% (nitrate) and 62% (DOC) of events, and, within ±20% for 83% (nitrate) and 86% (DOC) of events. SLR models for event load estimates presented lower uncertainty ranges, within ±10% of the median load among 96% (nitrate), 90% (DOC) and 59% (TSS) of events and within ±20% for 86% events (TSS). This article is the first of this kind reporting the use of optic sensors in a stormwater setting and suggests that the

uncertainties on the event loads are a little higher than the ones expected using flow proportional composite samples when used properly. These sensors provide, however, information on the dynamics of nutrient and material export that can be used to better predict the performance of SCMs.

1.1 Introduction

In situ ultraviolet to visible (UV-Vis) spectrophotometers have been introduced to field research in the 1990s (Ojeda and Rojas, 2009). They have been in use for water quality monitoring in wastewater (e.g., Langergraber et al., 2003; Wu et al. 2006; Torres, and Bertrand-Krajewski, 2008), streams in agricultural watershed (e.g., Albert et al., 2013; Liu et al., 2016), streams in urban watersheds (e.g., Duan et al., 2014; Halliday et al., 2016; Cizek et al., 2019), tidal wetlands (Etheridge et al., 2013), reservoir (Birgand et al., 2016), and woodchip bioreactors (Maxwell et al. 2019).

Current methods of water quality monitoring in stormwater often involves flow proportional composite sampling (FPCS) that consists in compositing into one container, samples automatically taken at regular flow volume intervals. In stormwater, flow tends to occur only during and shortly after rainfall events. It is thus customary to analyze stormwater hydrology and water quality at the event scale. As such, FPCS is very well suited to yield water quality information on an event basis, as it yields one representative sample per event, and the concentrations of which are referred to as the event's mean concentration (EMC). EMC can then be compared across systems or used to calculate pollutant loads.

However, FPCS and EMC give no information on pollutant dynamics, e.g., in and out of a Stormwater Control Measure (SCM), leaving little opportunity to understand their internal functioning and/or interpret the observed pollutant removal. Automatic samplers that perform FPCS have the capacity to obtain discrete samples from which pollutant dynamics can be drawn to some extent. But these have in the best cases 24 samples available, leaving the choice of high resolution over short periods of time (2-3 days), or low resolution over longer periods (Harmel et

al., 2003), and in all cases high analytical costs. Depending on the water quality sampling method, low-resolution data collection can lead to high error on event load estimates (Chapter 2).

Spectrophotometers allow monitoring of water pollutant concentrations at the minute resolution scale. This theoretically greatly reduces the measurement uncertainty caused by standard low frequency sampling (e.g., Birgand et al. 2011), and gives the opportunity to provide new insights about pollutant dynamics (Kirchner et al., 2004; Glasgow et al.,2004), and may be the only possibility to capture those in extremely flashy watersheds (e.g., urban).

The spectrophotometer by S::CAN® measures light absorbance in the ultraviolet to visible range (200-730 nm) and is advertised to measure parameters such as total organic carbon (TOC), dissolved organic carbon (DOC), total suspended solids (TSS), and nitrate (NO_3^-). Using an absorbance from the UV to visible range allows measurement of multiple parameters, with higher accuracy (Van Den Broeke, et al. 2006). The spectro::lyser® interprets the absorbance fingerprint, using simplified chemometric techniques embedded on the *in situ* sensor and referred to as the "Global Calibration", to estimate pollutant concentrations. However, local calibration using local samples can largely increase the measurement accuracy (Langergraber et al., 2003; Rieger et al., 2006; Torres and Bertrand-Krajewski, 2008; Etheridge et al., 2013). The mentioned authors suggest, using principal component analysis (PCA) and partial least square (PLS) regression to detect the concentration and absorbance relationship.

This paper reports the use of the spectro::lyser from S::CAN to obtain high-frequency water quality data in an urban watershed. The objectives of the study were: (1) to evaluate the benefits and challenges of, (2) to develop a methodology for, and (3) to evaluate the associate uncertainty in, the use of in situ spectrophotometers in non-sewage stormwater water quality monitoring.

1.2 Methods

1.2.1 Site description

The selected site was a stormwater wetland located on the campus of North Carolina State University. The watershed for this wetland was 9.12 ha, 30% impervious surfaces (*Figure 1.1*).

The location proximity of the selected site facilitated close observation and frequent maintenance of monitoring set up and equipment. The monitoring period reported in this paper covers a period of 20 months in total from December 2017 until August 2019. Two monitoring stations were installed at the wetland at the inlet and at the outlet.

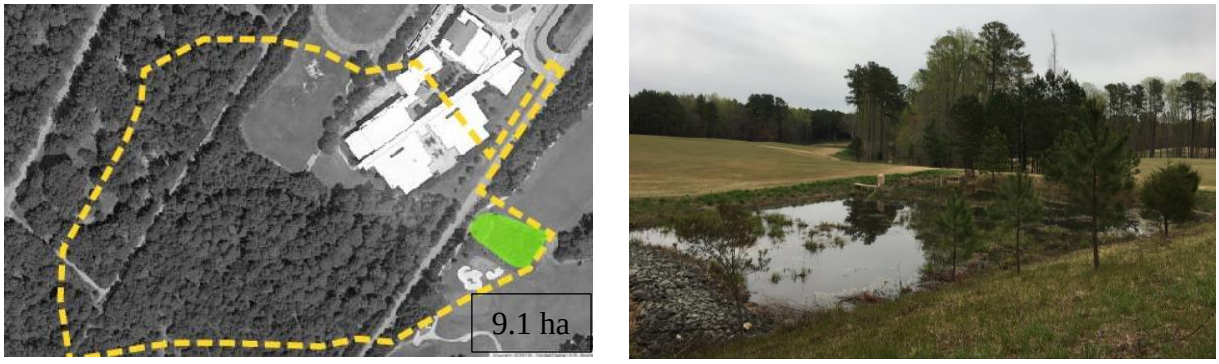


Figure 1.1 Watershed delineation (dotted line) and area for the stormwater wetland in research (highlighted in green)

1.2.2 Monitoring set up

1.2.2.1 Flow Monitoring

At the inlet, a 90-degree wooden weir was installed to measure flow. Flow was calculated from the stage upstream the weir measured using "Teledyne ISCO Bubbler Module" model 730, at 2-minute intervals (*Figure 1.2*) mounted on a 6712 ISCO Sampler. At the outlet, flow was measured using two complementary systems. The first one was an area velocity meter,

"Teledyne ISCO AVM Module" model 750 mounted on a 6712 ISCO Sampler and installed in a culvert downstream of the water level control structure in the wetland (*Figure 1.3*). A minimum of 10 cm of water was kept above the sensor at all times. Area velocity meters (AVM) perform well during high flow, but often fail to obtain accurate measurement at low flows. Hence, to obtain more accurate flow readings, a second system was installed. The outlet orifice was equipped with a 30-degree weir box and a pressure transducer (HOBO U20L-04™) to collect water level readings. The flow measurements obtained from this weir box were used as the outflow of the system for most of the storm events, and, for large events when there was overflow at the water level control structure and bypassing the weir box, AVM data was used for flow calculation.

1.2.2.2 Water Quality Monitoring

Four-minute water quality monitoring was performed using Spectro::lysers® by S::CAN® Messtechnik GmbH with optical path length of 5 mm and measuring light absorbance over the 200-730 nm range. The spectro::lyser probes were fitted with the ruck::sack automated brush activated before each measurement to prevent blockage of optical path length by debris, and minimize fouling on the optics (*Figure 1.2*, *Figure 1.3*).

For local calibration of the spectrometers, discrete water quality samples were obtained using a "Teledyne ISCO portable auto sampler model 6712" with the 24-bottle configuration. For best calibration, it is recommended to cover the entire range of observed concentration (low and high values) (Rieger et al., 2006). To sample along as large a concentration range as possible, the flow proportional sample triggering method available on the ISCO sampler was used. The flow proportional sampling can yield more samples during high flow events,

increasing the chance to obtain concentrations in the upper range values at these times despite their inherent scarcity.

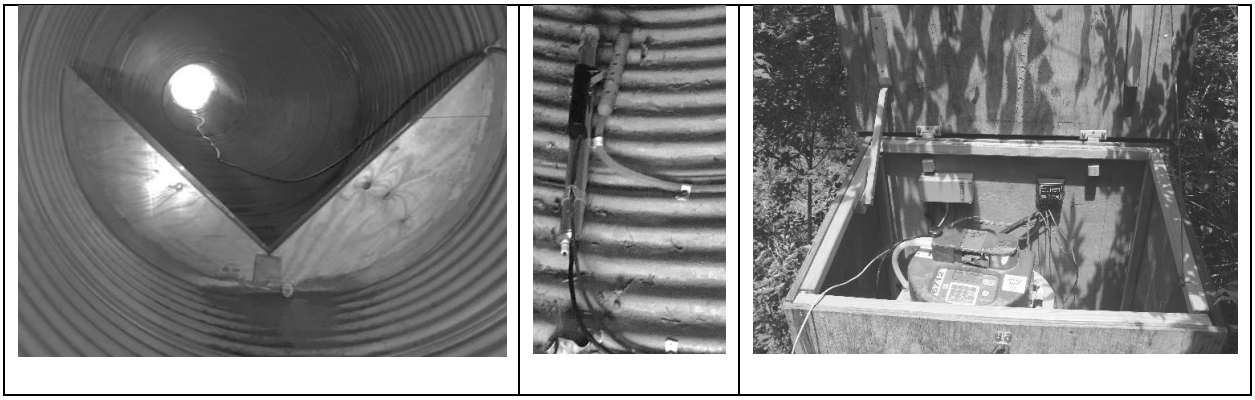


Figure 1.2 Inlet monitoring set up -weir (left), autosampler intake and spectrophotometer (middle), autosampler and spectrophotometer control (right)

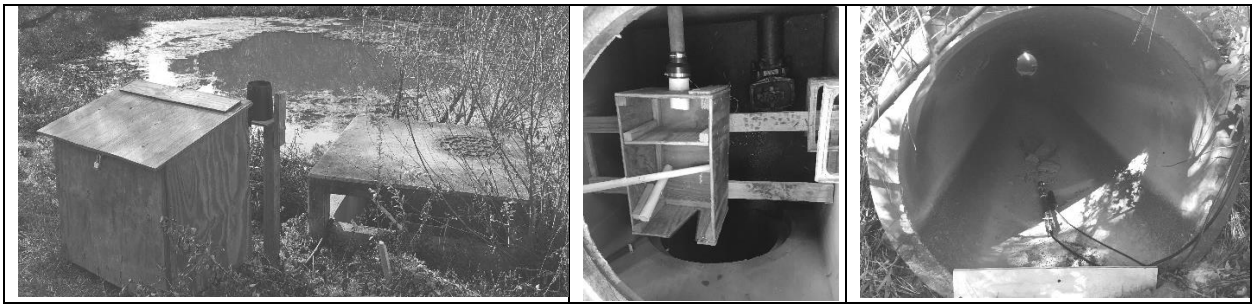


Figure 1.3 Outlet monitoring set up – control box (left), weir box and pressure transducer (middle), spectrophotometer (right)

All instruments were powered by 12 volts batteries charged by solar panels, appropriate for their required power. Spectrophotometer probes used in this study often require solar panels with a higher power (100 W) than ISCO auto samplers (10W) due to their high amperage.

1.2.2.3 Sampling protocol

At the inlet, discrete samples were taken at every 8.5 cubic meter (300 ft^3) using ISCO auto sampler model 6712 with 24-bottle configuration. Selection of best pacing to collect samples in stormwater is often done such that the majority of the hydrograph be sampled (Strecker et al., 2001). However, in this research the objective of sampling was to capture as wide a range of pollutant concentration as possible and specifically samples of high

concentration. Hence, a smaller pacing was selected that did not cover the entire hydrograph but would collect more samples at or near the hydrograph peak.

Discrete absorbance values of runoff were measured every 4-minutes, using the S::CAN spectrophotometer. The 4-minute time interval was selected based on the data storage capacity of the device as well as the possible frequency of maintenance. This time interval allowed for about 4 days of data storage on the device and hence mandatory maintenance after that period for data retrieval. Samples were collected from the site, within 24 hours after the end of a storm event and stored in a fridge at 4°C before filtration and delivery for lab analysis

1.2.2.4 Instrument Maintenance

The tasks presented in the following paragraphs were performed at every site visit to ensure accuracy of collected data and functionality of instruments for future data collection.

Initial maintenance tasks included, data download from all instruments, taking water samples and replacing the auto sampler bottles with clean bottles, checking the battery status and clock on the instruments. Since further process of collected data required comparison of water quality samples with spectrometry readings, the time stamp of the samples was particularly important and there was a need to ensure time alignment between sensors. Additionally, at each site visit, the actual water level was measured at each station manually and in case of discrepancies, it was adjusted on the instrument and recorded on field sheets. Inlet and outlet pipes were cleaned at each site visit to avoid accumulation of debris upstream the weir that could have resulted in false reading of pollutant concentrations.

Maintenance of the spectrophotometer entailed ensuring the cleanliness of path length, automated brush and lenses and followed those described by Etheridge et al. (2013). The automated brush was cleaned with deionized water and a plastic brush. The process of cleaning

the path length and lenses included rinsing with deionized water, scrubbing with brushes and application of a diluted acid (HCl 5%) with cotton swabs left on the optics for about a minute, a second rinse with deionized water and finally drying of the lenses using paper tissues. To evaluate cleanness of lenses, the cleaning procedure was followed by an air absorbance measurement. This task was performed until the air absorbance measurement resulted in a flat reading of less than 10 (m^{-1}), which is close to air measurement taken initially with the instrument with default manufacturer setting. The procedure was repeated until absorbance values in air were within the desired range.

In case of a dry period longer than 2-3 weeks, a maintenance site visit was necessary to ensure cleanness of the spectrophotometer lenses before an event was forecasted. Effort was made for the lenses to be cleaned within a week before an event. This was decided as the automated brush scrubbing of the drying optics lefts traces on the optics that needed to be removed.

1.2.3 Calibration of the *in situ* spectrophotometers

The spectro::lyser probe calculates concentrations of pollutants using its internal standard spectral algorithm referred to as the 'global calibration'. This 'global calibration' uses proprietary chemometrics techniques to calculate concentrations from the absorbance spectra. Our understanding is that the 'global calibration' has been established using thousands of points, and is therefore very robust, but gives an average correspondence between light absorbance and concentrations, which might not exactly fit local conditions. Additionally, the 'global calibration' uses parsimonious algorithms to accommodate limited computing capacities *in situ*. Although the instrument does provide a tool to adjust the global calibration, many articles report the successful use of PLS regression (PLSR) to directly correlate absorbance spectra with concentrations at a

particular station (e.g., Langergraber et al., 2003; Rieger et al., 2006; Torres and Bertrand-Krajewski, 2008; Etheridge et al., 2013).

PLSR was used to calibrate the absorbance spectra with collected local samples. The pollutant load estimates, resulting from this calibration were then compared with those calculated using simple linear regression (SLR) calibration. SLR method was used to calibrate the spectrometer's 'global calibration' results with collected local samples. The first step for calibration of *in situ* spectrophotometer using PLSR method is selection of proper calibration samples. Rieger et al. (2006) identified criteria for local samples to be used in calibration: (1) equal distribution over measured concentration range, (2) include temporal variability, (3) time alignment of spectrophotometer reading and sample, (4) accuracy of laboratory analysis, and (5) higher number of samples to compensate for laboratory errors.

1.2.3.1 Discrete Sample Selection for spectrophotometer calibrations

Among the discrete water quality samples collected from each storm event, only a few were selected for laboratory analysis. The selection criteria were, (1) time alignment with spectrophotometer readings, and (2) variability of concentration range (3) uniform distribution within the concentration range. Indeed, the timings for flow-proportional sampling was independent from the spectrophotometer readings. Only samples collected within a minute of the spectrophotometer readings were kept. To select samples with a variable concentration range, the turbidity values calculated by the instrument were used as indicators of concentration range. To ensure uniform representation over the concentration range, an effort was made to select samples with higher concentrations because of their inherent scarcity.

The selected samples were then prepared for laboratory analysis. A portion of 300 ml of the 900 ml sample, was separated for TSS analysis and the rest of the sample was filtered to

obtain 60 ml sample to analyze for the rest of the pollutants. Analyzed pollutants included Nitrate (NO_3^--N), Ammonium (NH_4^+-N), Total Dissolved Nitrogen (TDN), Total Suspended Solids (TSS), Orthophosphate (PO_4^{3-}) and Dissolved Organic Carbon (DOC). Samples were tested at the Environmental Analysis Laboratory at NC State University. *Table 1.1* summarizes the laboratory analysis methods.

Table 1.1 Laboratory analysis methods

Parameter	Analysis Method
NO_3^--N	Standard Methods 4500-NO3-E or EPA Method 353.2
NH_4^+-N	EPA Method 351.2
DOC	EPA Method 415.1 with Teledyne Tekmar Apollo 9000, 0.45 µm filter
TSS	Standard Methods 2540D or EPA 160.2
PO_4^{3-}-P	Standard Method 4500-P F or EPA Method 365.1

1.2.3.2 Identification and Removal of Outliers for PLS regressions

As previously mentioned, the accuracy of laboratory analysis and the proper alignment of lab and spectrophotometer readings are important for calibration. Therefore, a selection of 'workable samples' was created by detecting and removing 'outliers' or samples for which there was a legitimate concern that they might become 'toxic' for the PLSR. While this is common practice, the extreme flashiness of flow and concentrations of stormflow in these upstream urban and relatively impervious watersheds added some complexity to the process. We observed (details below) that concentrations often significantly varied within seconds, generating a possible mismatch between lab and instrument concentrations when stormwater was physically sampled and scanned *in situ* less than a minute apart.

From the local samples tested in laboratory, a set of workable samples was created for pollutants nitrate, DOC, and TSS. Workable samples were those excluding the outliers and those with concentrations above the laboratory test's detection limits. Samples identified as workable

samples for nitrate, were also used in PLSR calibration for NH_4, TDN, and phosphate. Outlier observations were detected as high residuals of a simple linear regression line between laboratory concentrations and those estimated by the spectrophotometer's "global calibration". Concentrations of nitrate, DOC, and TSS predicted by UV-Vis spectroscopy tend to be robust, as each of these pollutants have a particular absorbance range. Therefore, a linear relationship between laboratory and the spectrophotometer's 'global calibration' concentrations was expected. The absorbance ranges for the mentioned pollutants are 220-230 nm for nitrate (Crumpton et al., 1992; Suzuki and Kuroda, 1987), 250-300 nm for DOC (Rochelle-Newall and Fisher, 2002; Saraceno et al., 2009; Fichot and Benner, 2011), and in the visible range (380-780 nm) for TSS (e.g., Rieger et al., 2006; Torres and BertrandKrajewski, 2008).

For each pollutant with spectrophotometer reported concentration (nitrate, DOC, TSS) laboratory and reported spectrophotometer concentrations were plotted. All samples with residuals to the least-square regression line (Hochedlinger et al., 2006) greater than 0.2 mg/l were deemed to be outliers (e.g., *Figure 1.4*). Other criterion for removing outliers was samples with reported laboratory concentrations below the detection limit of lab test. Laboratory detection limits were 0.1 mg/l for Nitrate and 5 mg/l for TSS. However, the sensor's algorithm measures turbidity rather than TSS. Therefore, the turbidity-TSS relationship was first identified and the same outlier detection method was implemented.

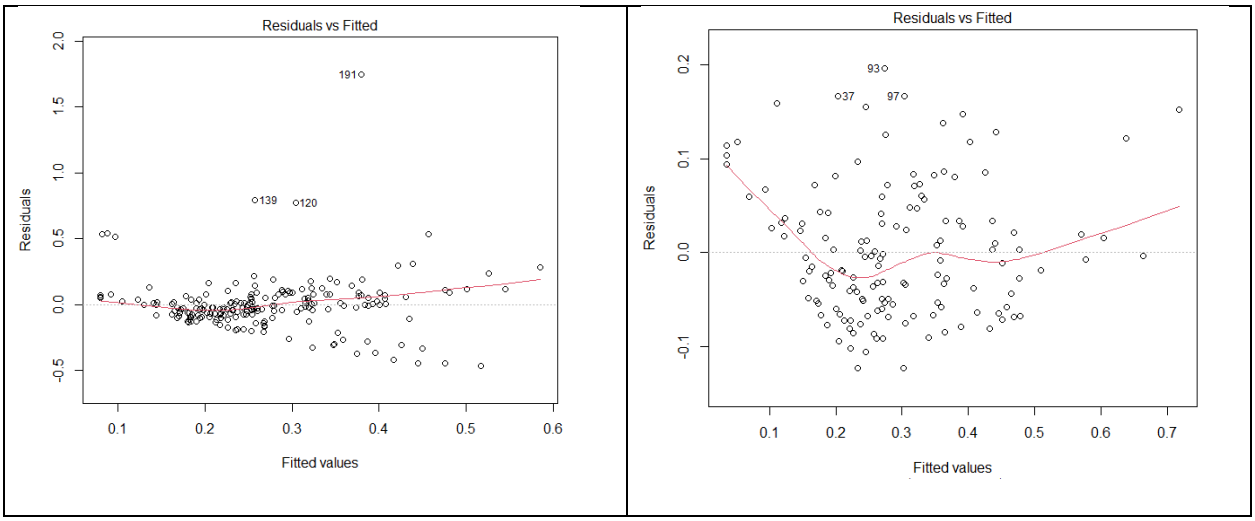


Figure 1.4 Residual values for linear regression model of laboratory and spectrophotometer nitrate values, all observations (left), without outliers (right)

1.2.3.3 PLSR calibration

PLSR was performed using the "pls package" (Mevik et al., 2011) in R language (R Core Team, 2019). PLSR calibration requires as uniform a representation as possible of lab concentrations (Langergraber et al., 2003). Due to the scarcity of high concentration samples, the selected workable samples for each constituent were stratified into two bins of low and high concentrations. Limited number of samples with high concentration values, prevented use of more bins to account for all high values equally. The number of samples in each stratum was selected to be smaller than the total number of samples in the high concentration bin as the high concentration bin included the least number of samples among all bins. This enabled simulations of variable calibration sample pools in each stratum, hence a distribution of regressions and therefore chemographs, from which uncertainties on events and annual loads were calculated. To select the most appropriate number of components for each of the PLSR models, up to 10 components were incorporated for model testing. The selection criteria were a number of components that yielded a model with relatively high R^2 and relatively low RMSE. These values

were plotted against the number of components. The selected value was the number of components at the break point of the RMSE vs component curve, plus one (*Figure 1.5*).

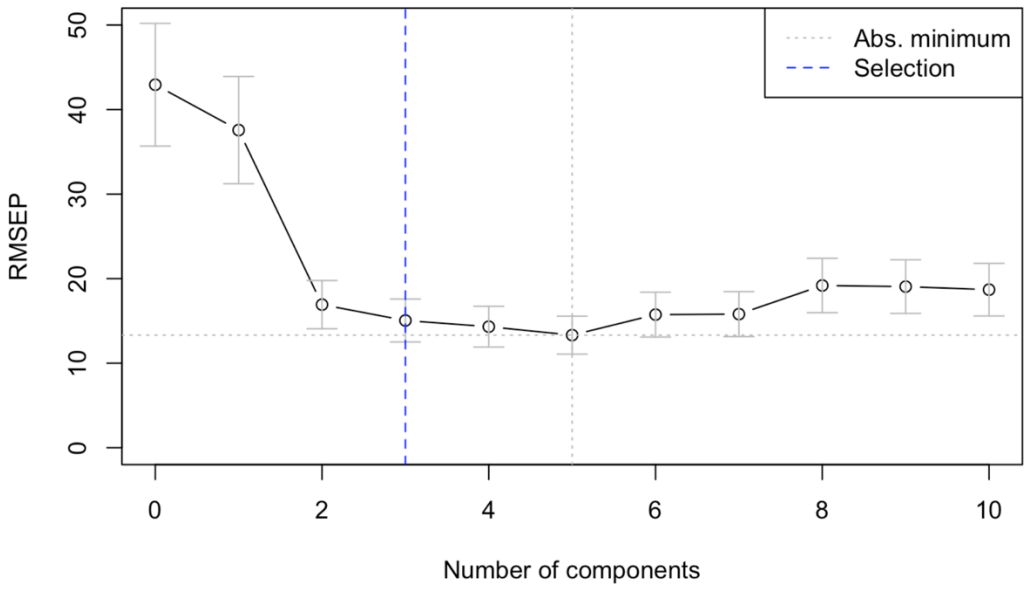


Figure 1.5 Example for selecting the number of components for PLSR model illustrating the breakpoint (ncomp = 2) and the final selection choice (ncomp = 3)

1.2.4 Evaluating uncertainties associated with the continuous water quality sensor

The sources of errors while estimating concentrations from spectrophotometers' absorbance measurements included fouling on the optics, representativeness of the water scanned *in situ*, the calibration method and the distribution of the manual sample concentrations used to perform the calibration. As a result, there were as many estimated chemographs as there were calibrations, i.e., theoretically an infinite number, hence the need to quantify measurement uncertainties for a given confidence interval.

Two indicators were used to calculate uncertainties: the annual and the event loads. Annual loads were chosen because this indicator is often used to evaluate sampling strategies in watersheds (e.g., Littlewood, 1992, 1995; Kronvang and Bruhn, 1996; Webb et al., 1997, 2000; Littlewood et al., 1998; Horowitz, 2003; King and Harmel, 2003; Wang et al., 2003; Moatar and

Meybeck, 2005, 2007; Birgand et al., 2006, 2010, 2011a,b). Event loads were chosen because events are the unit of analysis in stormwater and because uncertainties on event loads, assuming no uncertainty on flow calculations, correspond to uncertainties on the Event Mean Concentrations.

Loads were calculated as the trapezoidal integral of instantaneous loads calculated from the measured two-minute flow and four-minute estimated concentration data (*Equation 1-1*). Two-minute concentration data were obtained by linear interpolation from the 4-min data. The annual and event loads were expressed in kilogram per hectare.

Equation 1-1

$$L_{annual} = \sum_{1}^{i} (Q_i * C_i * t_i)$$

L: annual pollutant load (mg)

Q_i: instant flow (l/s)

C_i: instant pollutant concentration (mg/l)

t_i: time interval (s)

Instantaneous concentrations used for load calculations were estimated using three methods, (1) concentrations calculated internally and given by the spectrophotometer using a global calibration, (2) concentrations calculated using SLR establishing the correlation between the sensor's concentration values with laboratory concentrations, and (3) concentrations calculated using PLSR establishing a regression between absorbance values and laboratory concentrations. SLR and PLSR models were used on (1) all workable samples (presented as reference), (2) from randomly chosen stratified workable samples (much fewer samples), and (3) randomly chosen unstratified workable samples with the same size as the stratified sample pool. Stratification limited the number of samples used in regression models but allowed for random

sampling within each stratum. For comparability reasons, the same number of total samples was also used for random sampling within unstratified workable samples. For each stratified and unstratified calibration pools, 200 simulations of each SLR and PLSR model was performed, each predicting the concentration and load values.

From the 200 simulations, a distribution of chemographs were obtained, as well as a distribution of root mean square error (RMSE) and coefficient of determination (R^2) calculated for each regression.

From these, a distribution of annual load values was obtained, from which the mean, the median, standard deviation and 90% confidence interval were extracted. Using the same simulated chemographs, a distribution of loads was obtained for 29 events, and uncertainties expressed as the mean load and the loads corresponding to the 90% confidence interval. These 29 events were the largest events monitored which cumulative runoff volume comprised 90% of the total runoff volume monitored over 131 events.

1.3 Results

For practical reasons, at the outlet station the S::CAN instrument and the intake for the automatic sampler were placed at two ends of the culvert about 12 m apart, and this made the match between S::CAN readings and the discrete samples almost impossible. This ended up being a major drawback as detailed at the end of this section. Most of the analysis thus concentrates on the data obtained at the inlet station. Additionally, SLR and PLSR methods did not yield reliable results for parameters other than nitrate, DOC and TSS (details herein), hence the focus on these three throughout the result section.

1.3.1 Workable samples and sample stratification

Table 1.2 and *Table 1.3* show the results of lab analysis for all the samples and for workable samples respectively. The following Figure 1.6 visualizes the distribution of these values for nitrate, DOC and TSS. As expected, most samples represent the lower range of concentrations. The extreme nitrate concentrations were deemed as outliers per explained regression outlier detection method and were not included as workable samples. The concentration range of workable samples for nitrate, DOC, and TSS were 0.11-0.87, 1.07-20.84, 5.08-220 respectively. The concentration range of the workable samples defined the range that could reliably be predicted by the regression models.

Table 1.2 Laboratory analysis results for all the tested samples

	NO_3^--N	DOC	TSS	NH_4^+-N	TDN	PO_4^{3-}-P
No. of samples	201	201	195	201	201	201
Mean	0.27	7.69	24.99	0.43	1.65	0.17
SD	0.23	4.22	35.7	1.1	1.57	0.58
Minimum	0	1.07	0	0	0.15	0
Maximum	2.13	20.84	220	7.49	11.49	3.84

Table 1.3 Laboratory analysis results for workable samples

	NO_3^-	DOC	TSS	NH_4^+	TDN	PO_4^{3-}
No. of samples	147	160	114	147	147	147
Mean	0.29	7.84	26.37	0.49	1.76	0.21
SD	0.14	4.07	30.55	1.13	1.53	0.67
Minimum	0.11	1.07	5.08	0	0.26	0
Maximum	0.87	20.84	220	6.12	8.13	3.84

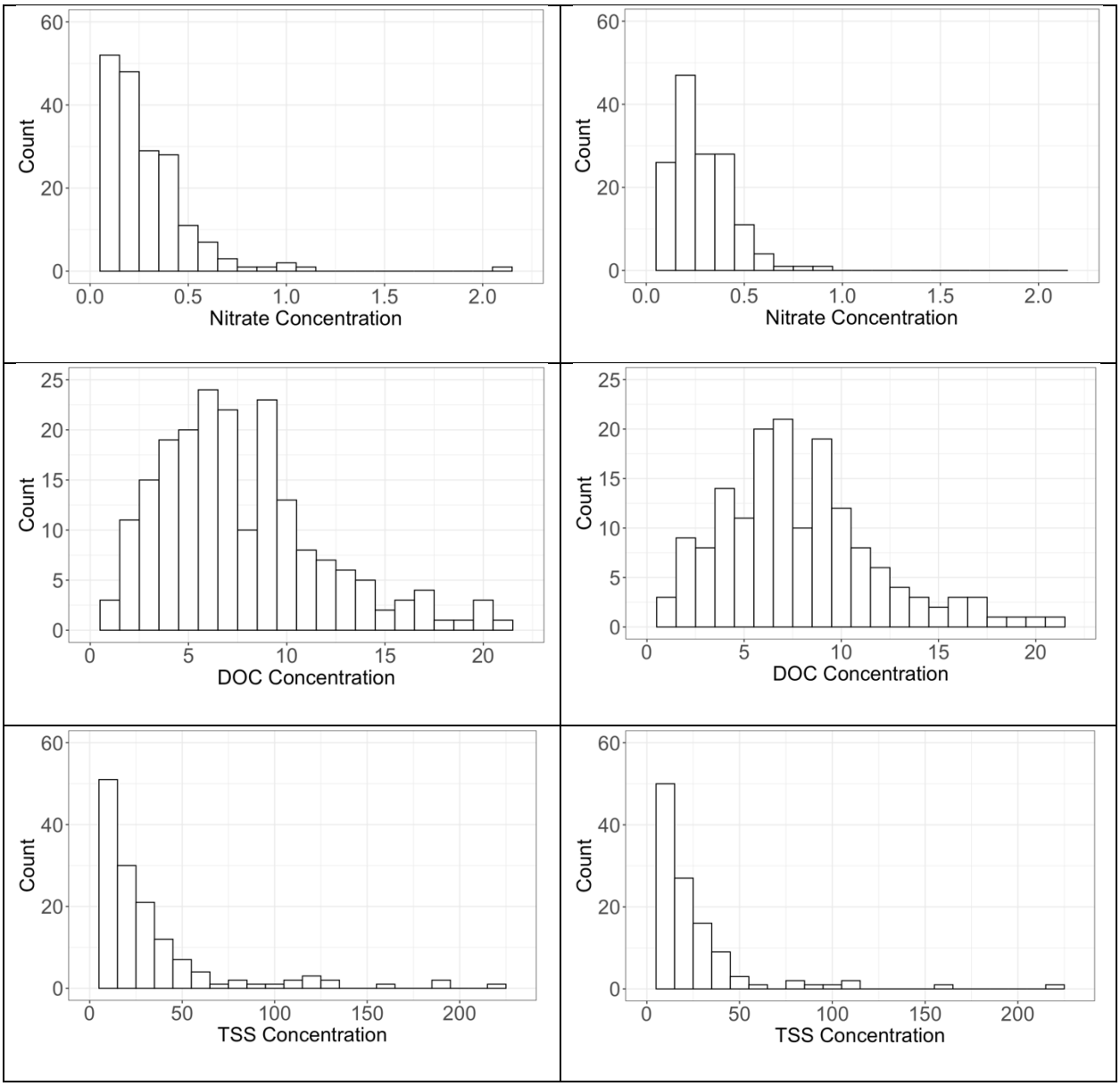


Figure 1.6 Distribution of laboratory concentrations (mg/l) for all samples (left) and workable samples (right)

To achieve a stratified representation of concentration range, workable samples for each pollutant were divided into two bins of upper and lower concentrations. The dividing concentration was selected based on the distribution of concentration for workable samples as 0.4, 10, and 30 mg/l for nitrate, DOC, and TSS respectively. *Table 1.4* shows the resulted bin sizes, as well as the selected strata size. The selected size of strata allowed for random sampling within the upper bin and were similar to those reported by Lin (2017). Based on these values, the

total stratified samples used in PLSR calibration were 48 for nitrate, 52 for DOC, and 36 for TSS. For compatibility purposes the same total number of samples was used for calibration with unstratified sampling.

Table 1.4 Division of concentrations into two bins and selected strata size

	NO_3^-	DOC	TSS
Bin divider (mg/l)	0.4	10	30
Size of upper bin	112	120	87
Size of lower bin	35	40	27
Selected strata size	20	26	18

1.3.2 SLR and PLSR regression models

SLR and PLSR regressions were performed on both stratified and unstratified sample pools. 'Reference regressions' were drawn from all of the unstratified workable samples. 200 stratified/unstratified sampling variation were created using random sampling of the workable samples within a defined limited number of total samples (48 for nitrate, 52 for DOC, and 36 for TSS). The resulting regression parameters are presented in *Table 1.5, Table 1.6, Table 1.7, and Table 1.8.*

Table 1.5 Regression parameters (SLR and PLSR) for Nitrate calibration

Regression	NO_3^--SLR			NO_3^--PLSR		
Sample type	Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
No. of samples	147	48	48	147	48	48
No. of components		-	-	6	6	6
Regression Adjusted R^2 (mean ± SD)	0.76	0.75 ± 0.05	0.78 ± 0.04	0.79	0.83 ± 0.04	0.85 ± 0.03
Residual Standard Error (mg/l) (mean ± SD)	0.07	0.072 ± 0.006	0.075 ± 0.004	0.06	0.057 ± 0.006	0.063 ± 0.005

Some common observations on SLR and PLSR calibration among different pollutants are: (1) both SLR and PLSR calibration resulted in good fits with R^2 greater than 0.75, (2) PLSR calibrations resulted in better fit than SLR models within comparable sample pools with higher R^2 and lower RMSE, (3) calibration with stratified samples resulted in a better fit for both models within comparable sample pools.

For nitrate, SLR models with all workable samples or reference pool resulted in R^2 of 0.76. Use of fewer samples in the model, unstratified sample pool, reduced the coefficient of determination to 0.75 ± 0.05. However, use of stratified sampling increased the coefficient of determination to 0.78 ± 0.04. This increase can be explained by the better representation of the concentration range in stratified sample pool. RMSE values, using any sample pool was around 0.07 mg/l. Despite the low values of RMSE, given the range 0.11-0.87 mg/l and mean 0.29 mg/l nitrate concentrations (*Table 1.3*), this model had relatively large range of error of 24% of the mean concentration.

PLSR model for nitrate using all workable samples resulted in R^2 of 0.79. Use of unstratified sample pool increased the coefficient of determination to 0.83 ± 0.04 and stratified sample pool improved coefficient of determination to 0.85 ± 0.03. All the PLSR models had 6 components. Therefore, under the same conditions the stratified sampling presented the best fit with laboratory tested concentrations. RMSE values with PLSR models were somewhat lower than SLR models at 0.06 mg/l, which corresponded to 20% of the mean concentration. The PLSR models presented a good fit with R^2 in 0.79-0.84 range and RMSE of 0.06 mg/l. However, the R^2 values were smaller than previous reports on PLSR calibration (Lin, 2017; Belenky, 2018). This could be due to the lower range of concentration observed in the studied watershed compared to those studies. Additionally, the range of observed concentrations were below the

spectrophotometer's range of accurate measurements (>1 mg/l) suggested by manufacturer (s::can Messtechnik GmbH, 2019).

Table 1.6 Regression parameters (SLR and PLSR) for DOC calibration

Regression	DOC-SLR			DOC-PLSR		
Sample type	Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
No. of samples	160	52	52	160	52	52
No. of components	-	-	-	6	6	6
Regression Adjusted R^2 (mean ± SD)	0.76	0.75 ± 0.05	0.82 ± 0.03	0.87	0.89 ± 0.03	0.91 ± 0.02
Residual Standard Error (mg/l) (mean ± SD)	1.97	1.98 ± 0.12	1.90 ± 0.14	1.45	1.34 ± 0.13	1.36 ± 0.12

Similar to nitrate, the SLR model on all DOC workable samples, resulted in R^2 of 0.76. Use of unstratified sample pool decreased coefficient of determination to 0.75 ± 0.05 and stratified sample pool, increased this value to 0.82 ± 0.03. RMSE values for this model ranged from 1.97-1.90 mg/l, while the sample concentration range was 1.07-20.84 mg/l with mean of 7.84 mg/l (*Table 1.3*). PLSR models on the same sample pools resulted in coefficient of determination of 0.87 using all workable samples, 0.89 ± 0.03 using unstratified samples, and 0.91 ± 0.02 with stratified samples. Use of stratified samples for calibration, yielded the best model. Additionally, PLSR model resulted in lower RMSE within 20% of the mean observed concentration values. The goodness of fit for DOC concentrations using PLSR model was similar to those of previous studies (Lin, 2017; Belenky, 2018).

Calibration of turbidity values with laboratory measurements of TSS values required use of proper model to explain the relationship between the two. Previous studies have identified this relationship as a first order linear one (Hannouche et al., 2011; Rügner et al. 2013) or log-linear

(Packman et al., 1999). However, visualization of TSS values vs turbidity resulted from global calibration, showed otherwise (*Figure 1.7*). Based on the mechanisms of turbidity measurements with light, it was possible for turbidity values to increase as lower pace as those of TSS concentrations. Turbidity is measured as the light absorbed by particles passing through the path length. During measurement, if different particles align, or a smaller particle is obscured by a larger one, then the instrument cannot detect the second particle. Therefore, a first order linear relationship may not always apply. Observed data in this study suggested a second order linear or log transformed linear relationship between turbidity and TSS values could be used instead. Outliers were taken out because of the uncertainty on the synchronization of the automatic measurements with the discrete sampling.

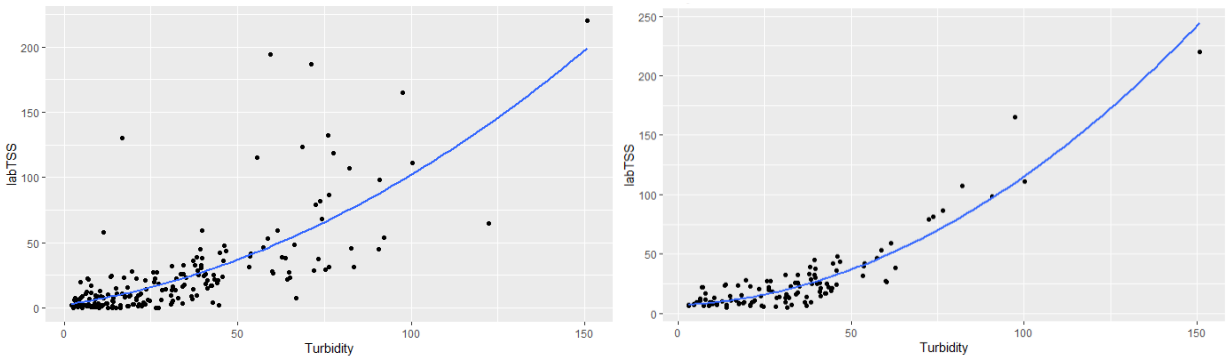


Figure 1.7 TSS laboratory values and spectrophotometer Turbidity values (using "global calibration"). Initial data (left) No-outlier data (right)

Second degree polynomial regressions resulted in coefficient of determination ranging from 0.88 to 0.91. Linear first order regression of log transformed values resulted in coefficients of determination ranging from 0.5 to 0.65. Based on these values the log transformed linear regression did not provide as good a fit to describe the turbidity-TSS correlation.

Table 1.7 Regression parameters (SLR) for TSS calibration

Regression	TSS-SLR (polynomial)			TSS-SLR (log-transformed)		
Sample type	Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
No. of samples	114	36	36	114	36	36
No. of components	-	-	-	-	-	-
Regression Adjusted R^2 (mean ± SD)	0.88	0.86 ± 0.1	0.91 ± 0.3	0.5	0.50 ± 0.1	0.65 ± 0.7
Residual Standard Error (mg/l) (mean ± SD)	10.65	9.37 ± 1.63	11.66 ± 1.84	0.56	0.55 ± 0.05	0.51 ± 0.06

SLR regression of turbidity and TSS also showed best fit with unstratified sample pool ($R^2 = 0.91$). PLSR model offered a better fit with coefficients of determination ranging in 0.86-0.95. Similarly, with the same number of components, use of stratified sample pool offered a better fit with R^2 of 0.95. Both models showed relatively large RMSE, 35-44% of mean concentration for SLR and 27-31% for PLSR (*Table 1.8*), considering the observed TSS concentration range of 5.08-220 mg/l and mean 26.37 mg/l (*Table 1.3*). However, RMSE values for PLSR model were lower.

Table 1.8 Regression parameters (SLR and PLSR) for TSS calibration

Regression	TSS-SLR (polynomial)			TSS-PLSR		
Sample type	Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
No. of samples	114	36	36	114	36	36
No. of components	-	-	-	6	6	6
Regression Adjusted R^2 (mean ± SD)	0.88	0.86 ± 0.1	0.91 ± 0.3	0.86	0.92 ± 0.05	0.95 ± 0.01
Residual Standard Error (mg/l) (mean ± SD)	10.65	9.37 ± 1.63	11.66 ± 1.84	11.45	7.19 ± 1.43	8.27 ± 1.45

Overall, for all parameters the PLSR model with stratified sampling showed higher coefficient of determination, 0.84 for nitrate, 0.91 for DOC and 0.95 for TSS with relatively smaller RMSE.

Spectrophotometer manufacturers suggest use of these sensors for measurement of Nitrate, DOC, TOC and TSS, however some studies were able to detect other pollutants (e.g., Etheridge at al., 2015; Birgand et al., 2016; Lin, 2017). Therefore, same method of PLSR calibration was used with the collected samples on pollutants of Ammonium (NH_4^+), Total Dissolved Nitrogen (TDN) and Orthophosphate (PO_4^{3-}). The Nitrate selection of laboratory sample used for calibration of these constituents. PLSR Calibration results are presented in *Table 1.9*. Based on these results PLSR did not provide a good fit to correlate concentration of Ammonia, TDN and orthophosphate with absorbance spectra. These constituents were not identified to absorb light, however other studies were able to estimate those through spectral data potentially due to co-variability of the color of water with the pollutants (Etheridge et al., 2015).

Table 1.9 PLS regression parameters for NH4+, TDN and PO43− calibration

	NH_4^+ -PLSR	TDN-PLSR	PO_4^{3-}-PLSR
No. of samples	147	147	147
Regression Adjusted R2	0.5722	0.6021	0.6435
Residual Standard Error (RMSE)	0.7384	1.035	0.399

1.3.3 Calibration uncertainties

Uncertainties associated with different calibration methods were evaluated based on annual loads as well as event loads. The regression parameters suggested the PLSR model with stratified sampling as the best fit model that yielded higher R^2 and relatively lower RMSE values. The performance of different calibration methods was evaluated through distribution of load estimates.

1.3.3.1 Annual loads

General observations of annual load estimates for different pollutants were: (1) uncalibrated concentrations obtained directly from the instrument resulted in the highest annual load estimates, (2) SLR calibrations resulted in higher annual load estimates than PLSR calibrations, (3) PLSR calibrations resulted in larger variance of load estimates, (4) calibration with stratified sample pool resulted in a higher load estimates than the unstratified sample of the same model, and (5) calibration with stratified sample pool resulted in smaller variance of load estimates than unstratified sample pool for both SLR and PLSR models.

Nitrate loads

Without calibration, the annual nitrate loads were estimated up to 32% and 22% greater than the ones estimated by PLSR and SLR, respectively, using stratified pools of samples. This highlights the need for calibration. SLR calibration, used a linear regression to adjust the uncalibrated instrument concentrations to the local sample concentrations. PLSR calibrations, used the absorbance values and found the best fit for the workable sample concentrations. The results were thus independent of the concentration values given by the instrument.

Because of the overrepresentation of lower concentrations among the workable sample pool, the unstratified calibration sample pool tended to have more samples in the lower concentration range than in the higher. The results were that for both SLR and PLSR, *the mean load for unstratified pools were lower than those for the stratified pools* (*Table 1.10*). Additionally, because without defining bins, unstratified pools had higher sampling possibilities and variations, the 90% confidence interval for unstratified pools was higher than that for stratified pools (*Table 1.10*). This suggests, and following Rieger et al. (2006) suggestions, that

assembling pools of stratified concentration values is preferable. Discussion herein focuses on results from the stratified pools.

Table 1.10 Annual nitrate load estimates using different calibration methods

Regression	NO_3^--Global	NO_3^--SLR			NO_3^--PLSR		
Sample type		Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
Mean load (kg/ha/yr)	0.835	0.649	0.648	0.680	0.580	0.573	0.626
Relative to reference load (%)			-0.2%	4.8%		-1.2%	7.9%
Median load (kg/ha/yr)			0.647	0.680		0.575	0.628
90% Confidence Interval (kg/ha/yr)			0.605 to 0.687	0.66 to 0.698		0.473 to 0.676	0.553 to 0.687
90% CI compared to median			-6.5% to +6.2%	-2.9% to +2.6%		-17.7% to +17.6%	-11.9% to +9.4%
Standard Deviation (kg/ha/yr)			0.038	0.012		0.063	0.041

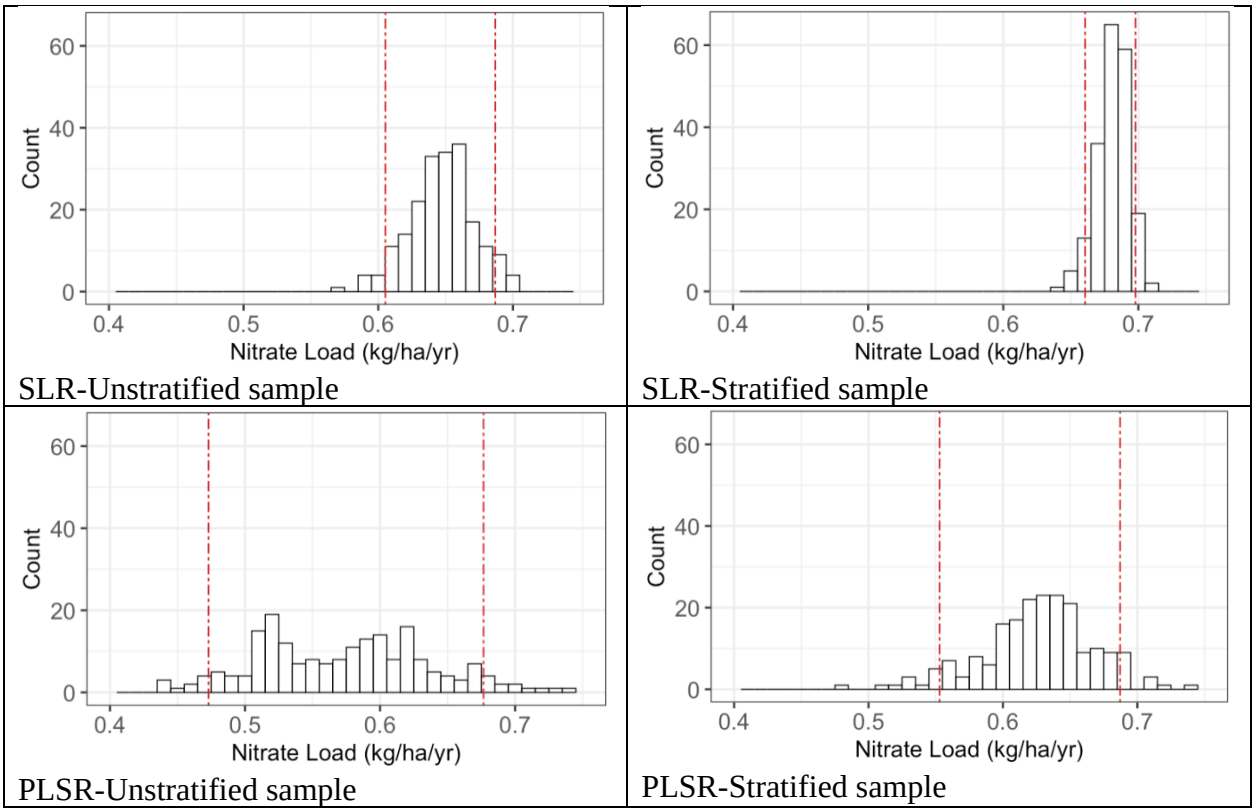


Figure 1.8 Distribution of nitrate annual load estimates and its 90% CI (between the vertical dotted lines) through 200 simulations

Regression indicators of both SLR and PLSR suggest better performance of PLSR versus SLR. The mean load estimated from stratified calibration samples with SLR was 8% greater than PLSR. Moreover, the uncertainty range for SLR (-2.9% to +2.6% of the mean load) was much smaller than that of PLSR (-11.9% to +9.4% of the mean load). This suggests that the regressions developed for PLSR were quite sensitive to the sample pools from which they were derived, a lot more so than those derived from SLR. The apparently better performance of PLSR vs. SLR (using the regression indicators as evidence), may thus have been somewhat misleading. The observed uncertainty range for PLSR was much higher than what has been reported for a coastal plain stream (± 3%; Lin, 2017) where nitrate concentrations were about 10 times higher. However, the uncertainty range for SLR was within that reported.

Regressions robustness depends in part on the number of samples used in the regression and on their stratification (Rieger et al., 2006). The global calibration (although the method for its derivation is proprietary and unknown) embedded in the instrument has been derived from thousands of samples, giving its robustness. As SLR essentially adjusted using linear regression the global calibration to local conditions, it was expected that it be more robust than regressions derived from less than 50 samples like PLSR in our case. Therefore, it is not surprising that the uncertainty range for PLSR was larger than that for SLR.

In the end, it is not clear which of SLR or PLSR should be chosen. One could argue that the 'reference' PLS regression using all 147 samples available would be a lot more robust and should be preferably used because it performs a true local calibration between absorbance and concentrations. However, low concentration values were over-represented (*Figure* 1.6), and therefore calculated load tended to be underestimated, equaling the mean of loads calculated from unstratified pools, compared to loads calculated from stratified pools. Because stratification definitely is a better method for regression, methods that used stratified pools are preferred, but the choice between SLR or PLSR is not obvious, which admittedly is problematic, especially since the mean of the loads differed by about 8%.

DOC loads

The analysis performed on DOC similarly showed that annual loads calculated from the DOC concentrations given by the instrument were largely overestimated (~36%) compared to loads calculated from calibrated concentrations. The analysis also showed that annual loads calculated from unstratified pools or using all workable samples tended to be underestimated by 4 to 5% compared to those calculated from stratified pools (*Table 1.11*), regardless of the regression type. Similar to nitrate, the workable pool of DOC concentrations was skewed

towards lower concentrations, yielding the same bias results compared to stratified pools. Unstratified pools also yielded a lot less accurate estimates of annual loads (-10.6% to +8.8% of the median for PLSR, and -5.5% to +4.8% of the median for SLR) compared to those using stratified pools (-6.1% to +4.9% of the median for PLSR, and -2.6% to +2.6% of the median for SLR).

The mean annual loads using stratified pools differed only by about 2% between SLR and PLSR. The much better agreement between the two methods probably results from the fact that the DOC concentration values used for regressions were much higher (1 to 20 mg C/L vs. 0.1 to 0.6 mg N/L) with a twenty-fold difference between low and high values (vs. 6-fold difference for nitrate). Because the annual load values were much higher, the percentage difference was calculated to be lower.

Table 1.11 Annual DOC load estimates using different calibration methods

Regression	DOC-Global	DOC-SLR			DOC-PLSR		
Sample type		Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
Mean load (kg/ha/yr)	26.52	18.7	18.67	19.45	18.20	17.91	19.28
Relative to reference load (%)			-0.2%	4%		-1.6%	4.6%
Median load (kg/ha/yr)			18.75	19.48		18	19.33
90% Confidence Interval (kg/ha/yr)			17.55 to 19.64	18.83 to 19.98		16.21 to 19.43	18.11 to 20.42
90% CI compared to median			-6.4% to +4.7%	-3.3% to +2.6%		-9.9% to +7.9%	-4.9% to +7.2%
Standard Deviation (kg/ha/yr)			0.636	0.324		1.053	0.66

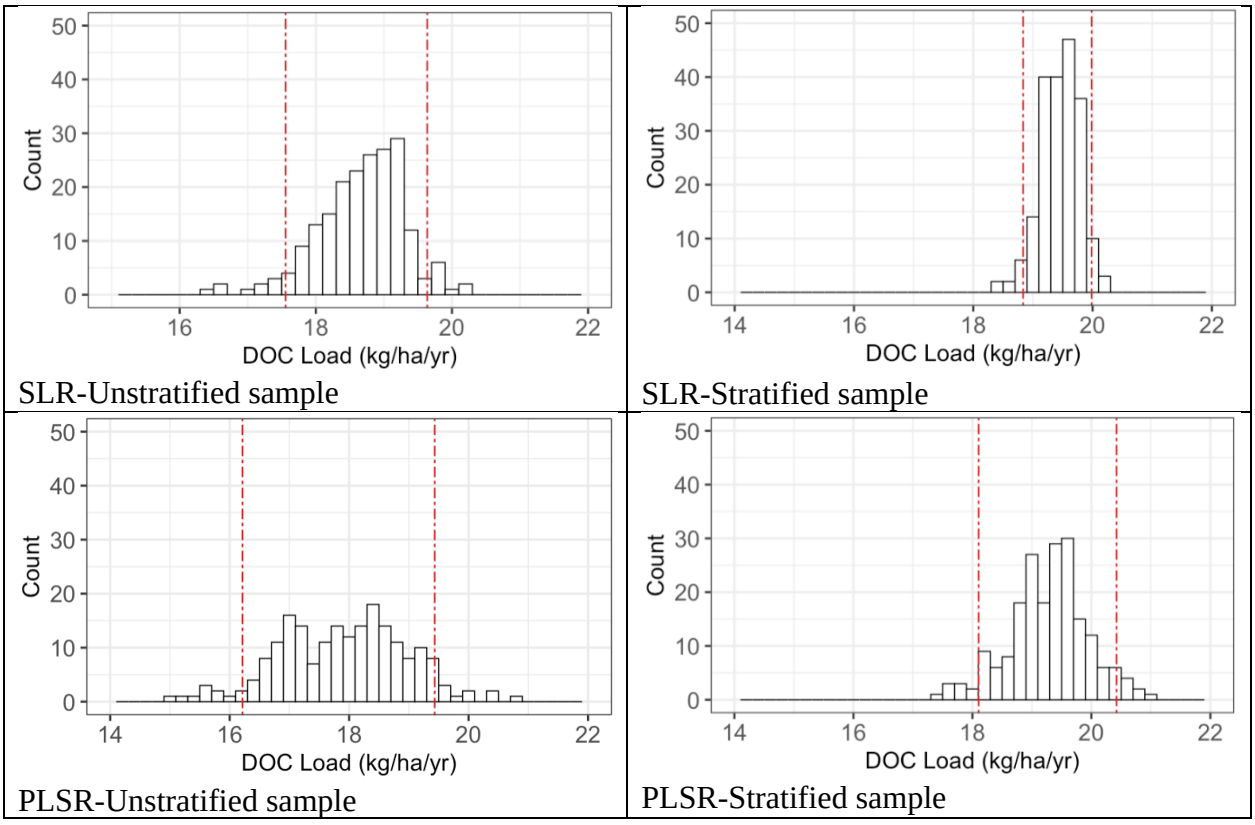


Figure 1.9 Distribution of DOC annual load estimates and its 90% CI (between the vertical dotted lines) through 200 simulations

Table 1.12 Annual TSS load estimates using different calibration methods

Regression	TSS-Global	TSS-SLR (polynomial)			TSS-PLSR		
Sample type		Reference	Unstratified	Stratified	Reference	Unstratified	Stratified
Mean load (kg/ha/yr)	56.65	49.77	51.77	53.81	48.98	46.52	50.82
Relative to reference load (%)			4%	8.1%		-5%	3.8%
Median load (kg/ha/yr)			51.28	53.39		46.36	50.61
90% Confidence Interval (kg/ha/yr)			44.93 to 59.54	50.25 to 58.51		34.42 to 57.36	43.37 to 58.48
90% CI compared to median			-12.4% to +16.1%	-5.9% to +9.6%		-25.8% to +23.7%	-14.3% to +15.6%
Standard Deviation (kg/ha/yr)			4.77	2.55		6.79	4.92

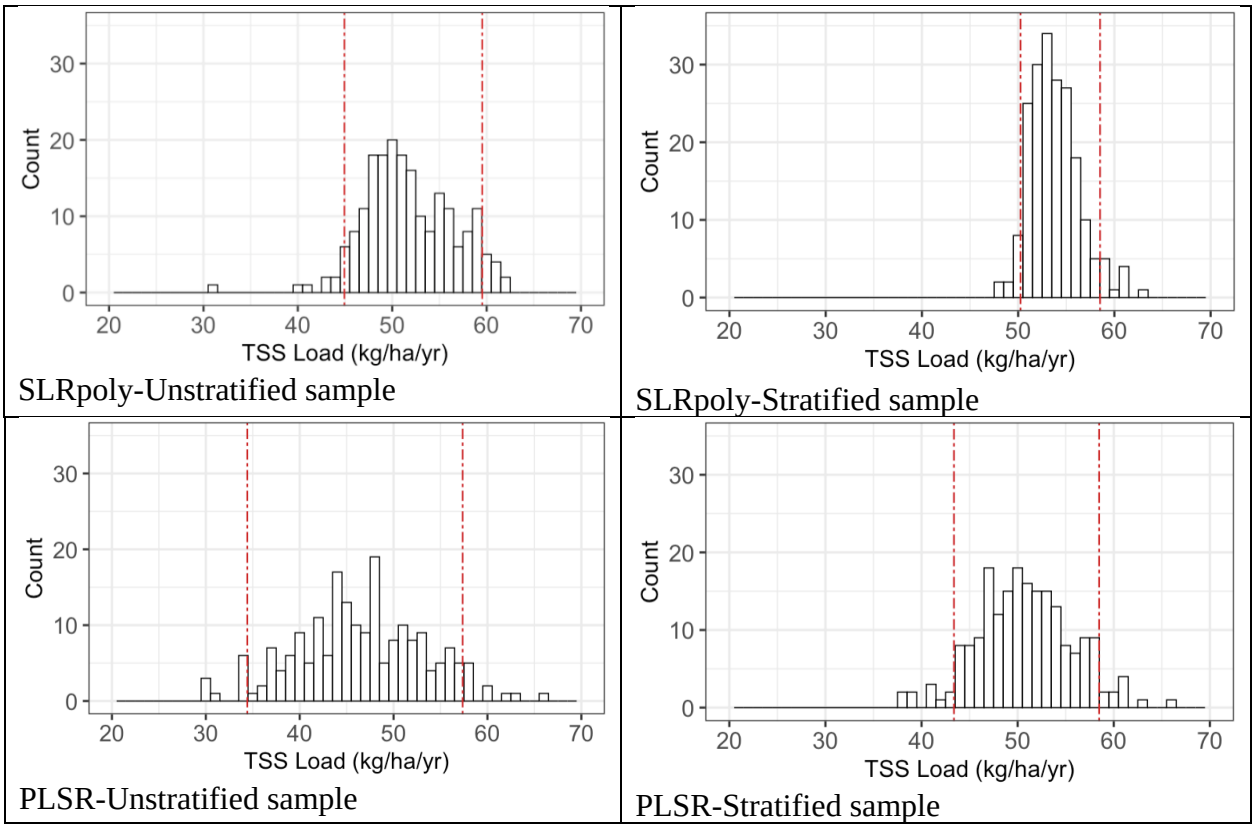


Figure 1.10 Distribution of TSS annual load estimates and its 90% CI (between the vertical dotted lines) through 200 simulations

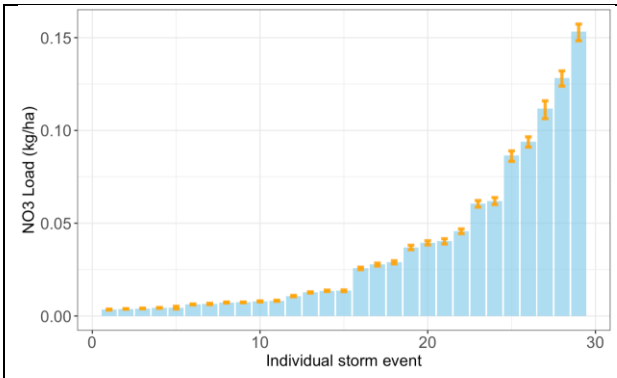
Based on the regression parameters (*Table 1.8*) the SLR model presented better fits than PLSR. SLR model using stratified sampling had a good coefficient of determination at 0.91 ± 0.3 and RMSE of 11.46. PLSR model using stratified sampling presented slightly better regression parameters with R^2 = 0.95 ± 0.1 and RMSE of 8.31 mg/l, but at the cost of load estimation uncertainties greater than ±10%. It is difficult to definitely choose one method over the other.

1.3.3.2 Event loads

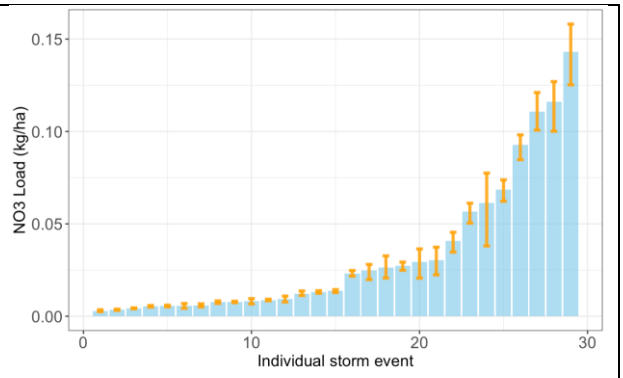
To derive uncertainties at the event scale, a total of 29 events comprising 90% of total cumulative runoff were selected. The runoff depth for these events ranged in 1.15-27.6 mm. Using PLSR and SLR models and random stratified sampling, event loads were estimated for these events using 200 simulations. The results of these simulations are presented in *Figure 1.11*

as mean event load and 90% confidence interval range as well as 90% confidence interval range presented as percentage of the mean load.

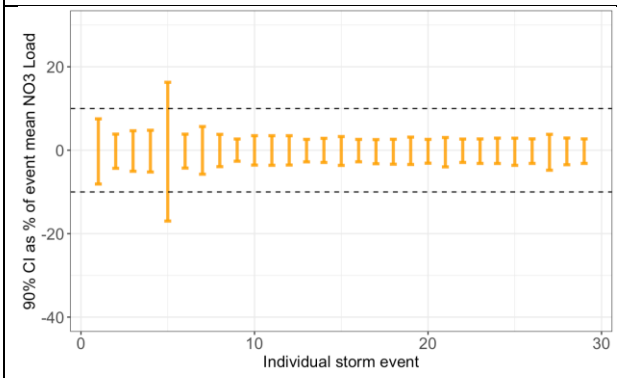
Figure 1.11 Event load estimates and associated 90% confidence interval for nitrate, DOC and TSS



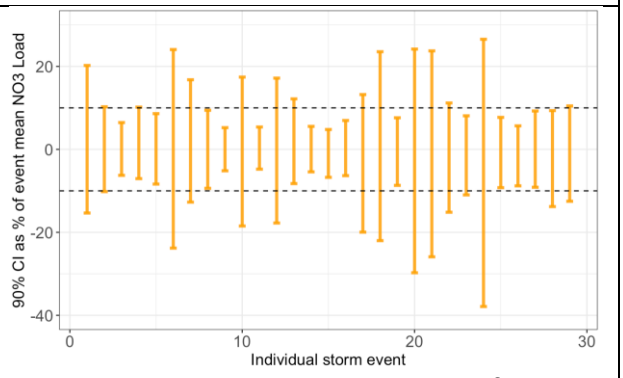
Nitrate-SLR load estimates with 90%CI



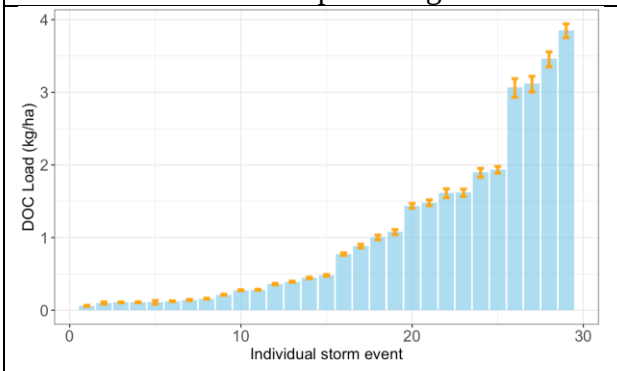
Nitrate-PLSR load estimates with 90%CI



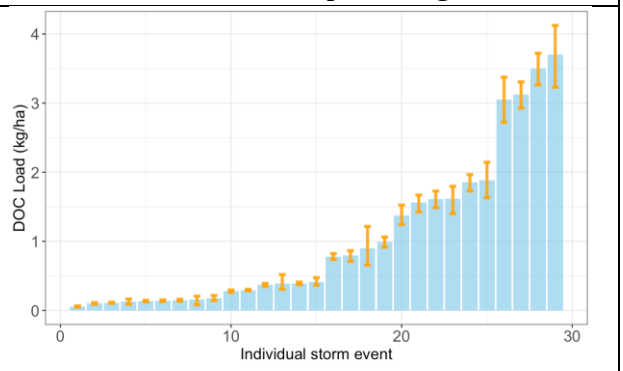
Nitrate-SLR 90%CI as percentage of mean



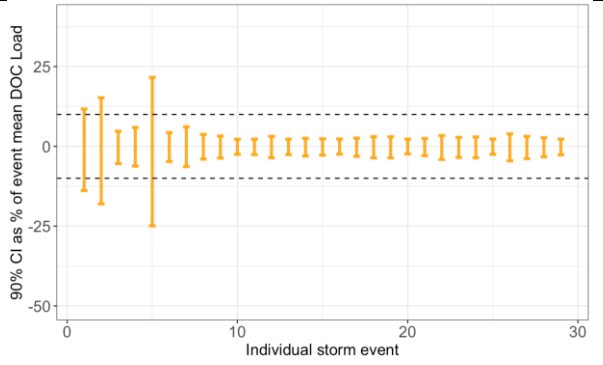
Nitrate-PLSR 90%CI as percentage of mean



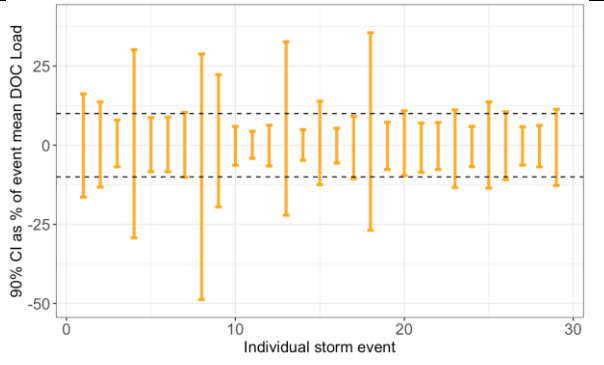
DOC-SLR load estimates with 90%CI



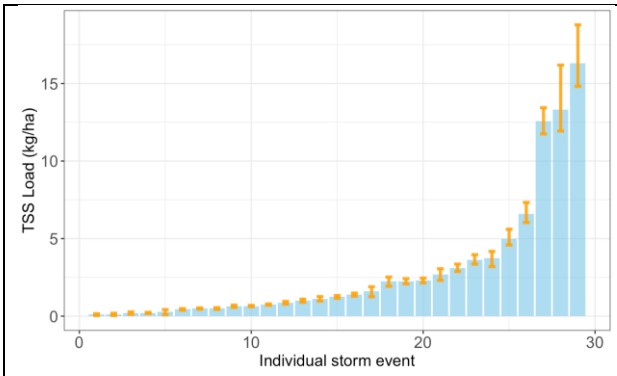
DOC-PLSR load estimates with 90%CI



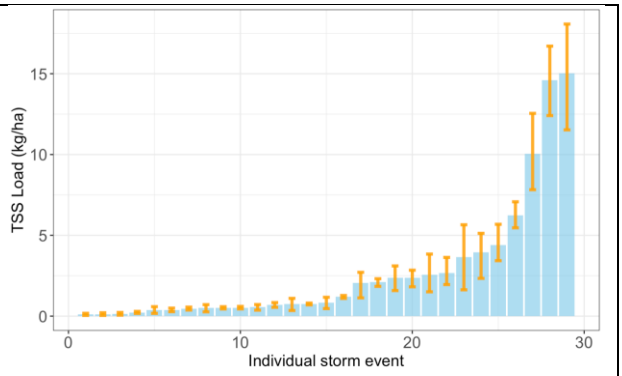
DOC-SLR 90%CI as percentage of mean



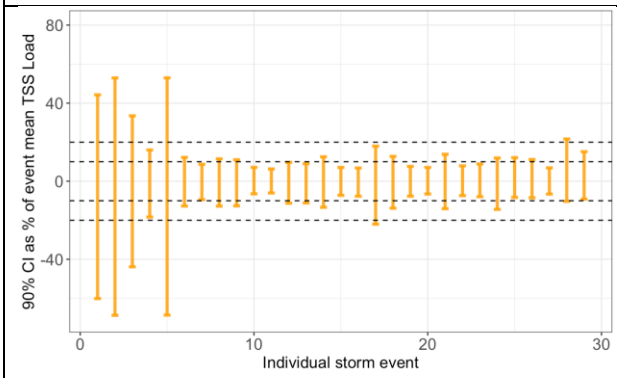
DOC-PLSR 90%CI as percentage of mean



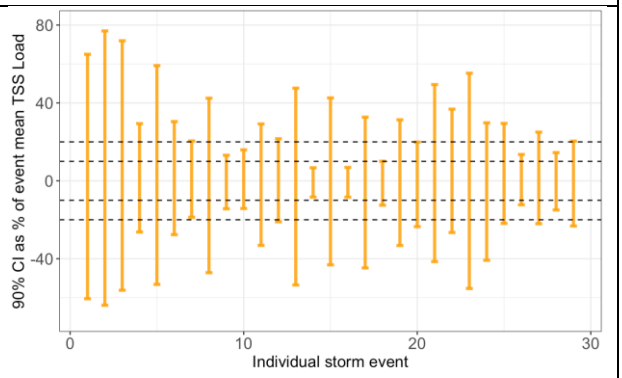
TSS-SLR load estimates with 90%CI



TSS-PLSR load estimates with 90%CI



TSS-SLR 90%CI as percentage of mean



TSS-PLSR 90%CI as percentage of mean

Using PLSR with stratified sampling, event loads were estimated for (1) nitrate within ±10% uncertainty among 52% of events and ±20% uncertainty among 83% of events, (2) DOC within ±10% uncertainty among 62% of events and ±20% uncertainty among 86% of events, (3) TSS within ±10% uncertainty among 10% of events and ±20% uncertainty among 41% of events. Using SLR model with stratified sampling, event loads were estimated with a much better variance, within ±10% uncertainty among 96% of events for nitrate, within ±10% uncertainty among 90% of events for DOC and within ±10% uncertainty among 59% of events and ±20% uncertainty among 86% of events for TSS.

Generally, the uncertainty associated with measurement of smaller values were higher. This was observed in loads estimated by SLR model (*Figure 1.11*). However, the load estimates using PLSR model shows a variation of uncertainty among event loads for all pollutants. This can be explained by the sensitivity of PLSR model to selection of calibration samples. The sample selection for a given calibration represents a certain range of concentration that may not necessarily cover the concentration range observed for an event. In such cases the calibration may result in a higher uncertainty for load estimates. *Figure 1.12* illustrates the nitrate chemograph for an event with concentrations estimated with spectrophotometer's global calibration as well as SLR and PLSR models. The SLR model adjusts the results of global calibration concentrations with local sample concentrations. Therefore, the curves resulting from these two methods are similar. The PLSR model solely uses local sample concentrations to calibrate the spectrophotometer's absorbance values and as a result the calibration is more sensitive to the stratum of local samples used for calibration. In an example calibration showed in *Figure 1.12* compared to global calibration results, the SLR model reduces the concentrations at peak values an aligns with global calibration results at lower values, while the PLSR model

estimates lower concentrations than both at peak values, it estimates higher values than both at the initial lower concentrations.

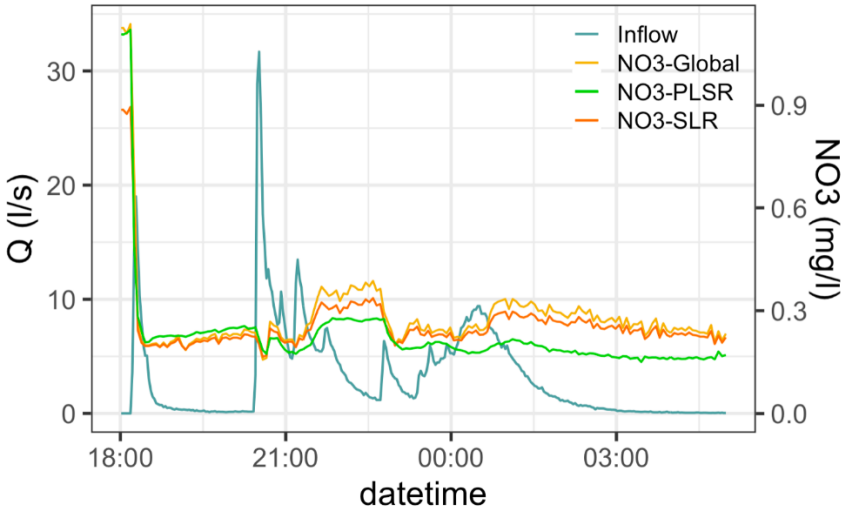


Figure 1.12 Example of hydro- and chemograph for nitrate estimated by global calibration, SLR, and PLSR models (stratified sampling for models)- event occurred on 08/18/2018

The uncertainty associated with event load estimates are higher than uncertainties associated with annual load estimates. Given the explained sensitivity of PLSR model to calibration samples, a single model may overestimate an event and underestimate another. Therefore, these values would compensate for each other on an annual calculation of loads. To ensure the best calibration on an event basis, authors suggest collection of few discrete samples per event and evaluation of calibration results with the laboratory tested concentrations.

1.3.4 Outlet

Table 1.13 shows the results of laboratory analysis of outlet samples. *Figure 1.13* compares the laboratory nitrate concentrations and spectrometer's reported concentrations using global calibration. A similar weak correlation was observed for all pollutants.

Revisiting the monitoring installation, the alignment of local samples and spectrophotometer samples was challenging. The sampling set up for this station faced limitation for proper sampling and maintenance access. The autosampler intake was placed at the ponded end of the culvert to obtain proper flow measurements while the spectrophotometer was placed at the other end, 10 meters apart to allow access for frequent maintenance. Additionally, the culvert lead to an open land, which had created a ponding area. Despite the placement of a weir at the end of the culvert, there was occasional backflow from the ponding area that could interfere with the spectrophotometer measurements as well.

Using the average velocity for each storm event and the distance of the two sensors, the time difference of readings was calculated and used for proper alignment. However, this alignment only slightly improved the correlation of samples.

Without properly aligned samples for calibration of spectrometer, the best fit SLR model from inlet station was used for calibration of outlet data. In the case of TSS the global calibration results were deemed as the best fit. The resulted loads at outlet station are presented in *Table 1.14*.

Table 1.13 Laboratory analysis results-outlet station

	NO_3^-	DOC	TSS	NH_4^+	TDN	PO_4^{3-}
No. of samples	102	102	101	102	102	102
Mean	0.2496	11.076	25.26	0.347	1.753	0.056
standard deviation	0.187	4.203	51.013	0.479	0.969	0.113
Minimum	0.01	4.57	0	0	0.29	0
Maximum	0.71	20.58	391.67	2.66	4.65	0.93

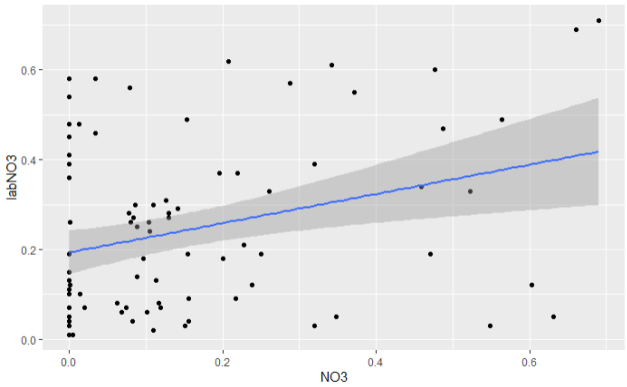


Figure 1.13 NO3− laboratory values and spectrophotometer measurements (using "global calibration")

Table 1.14 Annual load estimates at outlet

Regression	NO_3^-- PAR	NO_3^--SLR Stratified	DOC-PAR	DOC-SLR stratified	TSS-PAR	TSS-SLRpoly stratified
Total load (kg/ha/yr)	0.150	0.144	20.612	14.312	46.531	136.505

Figure 1.14 shows an example of spectrophotometer turbidity measurements, analyzed TSS samples as well as inflow and outflow. Based on this graph the turbidity measurements were affected by precipitation, outflow from the system as well as the turbulence following an outflow.

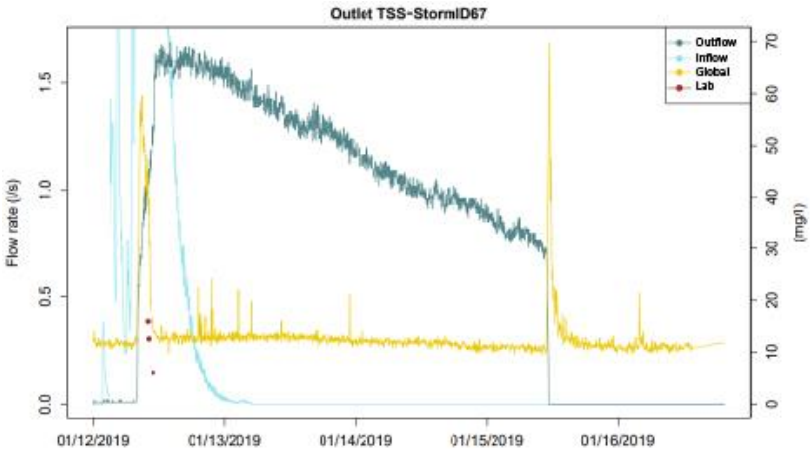


Figure 1.14 TSS laboratory values and spectrophotometer measurements (using "global calibration") showing large discrepancies between sensor and lab data, attributed to the lack of synchronization between the two.

1.4 Discussion

The extreme flashiness of flow and concentrations in stormwater settings, have rendered the description of concentration dynamics difficult using the traditional automated sampling method. *In situ* sensors theoretically present a unique opportunity to access concentration dynamics over long periods and possibly unveil new knowledge in stormwater research and application. *In situ* sensors have been widely used in stormwater associated with Wastewater Treatment Plant and sewage flow (e.g., Torres and Bertrand-Krajewski, 2008; Lepot et al., 2016). To our knowledge, this is the first time the evaluation of *in situ* sensors in non-sewage stormwater flow is reported.

This article suggests that the estimation of concentrations from light absorbance measurements is not particularly simple and that sizable uncertainties may result in the computation of loads. For the 9.2 ha watershed studied, robust calibrations could only be obtained for nitrate, DOC, and TSS, which correspond to the parameters advertised as measurable by the instrument.

In other settings, more parameters have been measured using local calibration based on PLSR (e.g., Etheridge et al., 2014; Frazar et al., 2019; Cizek et al., 2019). In these cases, the ability to measure other parameters with good accuracy was attributed to the stable relationship between the color matrix of the water and the concentrations. In all these cases, there always was baseflow, and the processes generating stormflow were thought to be repeatable enough that the color matrix of the water would covary with concentrations. PLSR was thought to be able to characterize this co-variability. In the case of stormwater in this watershed, there was no baseflow and the processes generating flow could vary radically from a low to high intensity rainfall event, rendering a stability of the relationship between the color-matrix and concentration possibly more random. For parameters not known to absorb light such as ammonium, TDN or phosphate, no robust regressions could be established with PLSR, evidence of a poor stability of the relationship between the color-matrix and concentrations.

For nitrate, DOC, and TSS, which absorb light at known wavelength ranges, acceptable correlations were found in our case. However, the PLSR method widely used as the preferred method to evaluate concentrations from absorbance values appeared to yield relatively large uncertainty ranges for nitrate and TSS, for annual and event loads. For reference, Lin (2017) found that PLSR would yield uncertainty on annual loads within ±3% for nitrate, ±5% for DOC and TSS. For event loads, errors within ±5% may be obtained using FPCS (See Chapter 2).

Values reported in this article appear to be much larger than these references for nitrate and TSS. Overall, calibration with PLSR model and stratified samples resulted in annual load estimations with accuracy of -11.9% to +9.4% of the median load for nitrate, -6.1% to +5.9% for DOC, and -14.3% to +15.6% for TSS. Load estimation using PLSR calibration on an event basis showed higher uncertainties for some events. Overall, loads were estimated within ±10% for

52% (nitrate) and 62% (DOC) of events, and, within ±20% for 83% (nitrate) and 86% (DOC) of events. SLR model for load estimates presented lower uncertainty ranges, within ±10% of the median load among 96% (nitrate), 90% (DOC) and 59% (TSS) of events.

PLSR appears to have been very sensitive to the samples used for calibration. For nitrate, all concentrations were below 0.6 mg/L, leaving a relatively small calibration range, yielding relatively large differences between calibrations. This, added to a probably relatively weak stability of the color matrix/concentration relationship, may explain the relatively large uncertainties using PLSR at the annual and event scales. At the event scale, extremely large uncertainties were observed for TSS from PLSR-derived loads, showing again the high sensitivity of PLSR to the calibration pool. The results for PLSR predicted DOC loads at the event and annual scale showed lower uncertainties closer to the references listed above. There is little reason to believe that the color matrix/concentration relationship would be a lot stronger and more stable for DOC than it was for nitrate and TSS. However, the DOC concentration values and ranges were a lot higher and better stratified than those of nitrate and TSS, which resulted in much lower uncertainties in DOC loads.

However, uncertainties reported with PLSR were calculated using regressions made ‘blindly’. In other words, having one to three lab samples per event and comparing the predicted chemograph to lab samples would lead the user to automatically rule out some of the predicted chemographs, and would probably generate much lower uncertainties. This suggests, however that not one regression model, but a suite of regression models should be systematically run to eliminate bad chemographs.

In our setting, the SLR method seemed to be a lot more robust than the PLSR method, yielding lower uncertainties. This is opposite to what Lin (2017) found. This may suggest that in

a setting where there may be no theoretical reason for a stable color matrix/concentration relationship like in suburban stormwater flow, it might be preferable to rely on a local correction of the embedded 'global calibration', the latter having been derived from hundreds of samples and therefore being more robust.

Despite the limitations listed above, the major advantage of having *in situ* sensors for stormwater water quality monitoring is that the proportion of events captured compared to traditional auto-sampler based monitoring is much greater. In our case 62% of all the rainfall events comprising 95% of recorded volume were monitored. Additionally, the relative errors at the event scale for instance, although significant, still provided concentration dynamics that no other method would generate.

1.5 Conclusions

Use of spectrophotometer sensors in stormwater settings allows for collection of water quality samples at high frequency and eliminates the capacity limitation of traditional automated sampling. Using this method, water quality was monitored for 62% of events, comprising 95% of total recorder precipitation volume. Water quality was monitored at 4-minute frequency which brought further insight into pollutant dynamics and sampling uncertainties (Chapter 2).

To obtain best monitoring results, proper collection of local samples and maintenance of *in situ* spectrophotometers is critical. Obtaining good calibration results and water quality data was critically dependent on (1) monitoring set up, (2) sample time alignment, (3) capturing high concentration samples and (4) meticulous maintenance of spectrophotometer optics.

Considering the flashy behavior of stormwater, the monitoring equipment should be set up to sample at the same time and location. To ensure time alignment of the samples, the instrument clocks should always align, and any delay should be accounted for such as time to

pump based on the length of suction line or cleaning time before measurement. The location of spectrophotometer sensors and intake of autosampler should be placed side by side. For Proper calibration, it is important to capture and use samples with a large and homogenous concentration range. Extrapolating beyond the concentration range of collected samples is statistically incorrect and could potentially yield unreliable results. Within a given monitoring time, high concentration of pollutants occurs in fewer storm events. Therefore, obtaining discrete samples from those events might require extra effort and change in pacing of autosampler. Fouling can occur due to metal precipitation and formation of biofilms (Etheridge et al., 2013), therefore, regular maintenance of the optics is necessary to ensure proper measurements.

PLSR and SLR methods were used for calibrating the instruments. The built-in "global calibration" of the spectrophotometer sensors calibrated with SLR model have proven to provide a good estimate for pollutant parameters ($R^2 > 0.74$). Local calibration of spectra improved the calibration using SLR model ($R^2 > 0.75$) and PLSR ($R^2 > 0.83$). For both SLR and PLSR models, use of stratified sampling improved the regression's coefficient of determination and lowered the RMSE. Using PLSR method, the absorbance spectra were successfully calibrated for measurement of NO_3^-, DOC and TSS as suggested by manufacturer, however, PLSR method could not calibrate for further parameters (NH_4^+, TDN, PO_4^{3-}).

Uncertainty associated with this method was evaluated by comparing cumulative load estimates on annual and event basis. Overall, calibration with PLSR model and stratified samples resulted in annual load estimations with relatively high accuracy of -11.9% to +9.4% for nitrate, -6.1% to +5.9% for DOC, and -14.3% to +15.6% for TSS. Load estimation on an event basis showed higher uncertainties for some events. Overall, loads were estimated within ±10% for 52% (nitrate) and 62% (DOC) of events, and, within ±20% for 83% (nitrate) and 86% (DOC) of

events. SLR model for load estimates presented lower uncertainty ranges, within ±10% of the median load among 96% (nitrate), 90% (DOC) and 59% (TSS) of events.

The higher range of uncertainty on an event basis was due to the apparent higher sensitivity of PLSR models to calibration samples. This was due to high rates of concentration change and lower pollutant concentration range in stormwater versus perennial streams. This sensitivity might present as a higher uncertainty on some event loads, but on the annual basis cumulation of events can compensate for this event basis uncertainty. Inclusion of at least one sample per event for calibration purposes might help in reducing this uncertainty.

Based on the results, the best calibration method could not be clearly selected. The PLSR model offered higher coefficient of determination, but, higher RMSE values and higher variability on annual load estimates and higher uncertainty on event load estimates. When using spectrophotometer sensors for stormwater monitoring, the instrument's global calibration and local SLR calibration might offer more robust results. However, selection of the proper PLSR model can reduce uncertainty associated within the models. Such selection should be made by comparing the calibration results with laboratory tested samples for each event.

1.6 Acknowledgement

This study was supported by the North Carolina Clean Water Management Trust Fund. Special thanks to the following individuals for their notable support:
Shawn Kennedy and Samuel Garvey for their help with instrument installation and troubleshooting. Alisha Goldstein, the project engineer for her support in conducting the study. Cheng Chen for his help with field data collection.

REFERENCES

Aubert AH, Gascuel-Odoux C, Gruau G, Akkal N, Faucheux M, Fauvel Y, Grimaldi C, Hamon Y, Jaffrézic A, Lecoz-Boutnik M, Molénat J, Petitjean P, Ruiz L, Mérot P. 2013. Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. Hydrology and Earth System Sciences 17: 1379–1391. DOI:10.5194/hess-17-1379-2013.

Belenky, C. N. (2018). The Effect of Stream Restoration on Water Quality and Quantity in the Coastal Plain of North Carolina.

Birgand, F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., Bordenave, P., 2010. Uncertainties in Assessing Annual Nitrate Loads and Concentration Indicators: Part 1. Impact of Sampling Frequency and Load Estimation Algorithms. Trans. ASABE 53, 437–446. https://doi.org/10.13031/2013.29584

Birgand, F., Appelboom, T.W., Chescheir, G.M., Skaggs, R.W., 2011a. Estimating Nitrogen, Phosphorus, and Carbon Fluxes in Forested and Mixed-used Watersheds of the Lower Coastal Plain of North Carolina: Uncertainties Associated with Infrequent Sampling. Trans. ASABE 54, 2099–2110. https://doi.org/10.13031/2013.40668

Birgand, F., Faucheux, C., Gruau, G., Moatar, F., Meybeck, M., 2011b. Uncertainties in Assessing Annual Nitrate Loads and Concentration Indicators: Part 2. Deriving Sampling Frequency Charts in Brittany, France. Trans. ASABE 54, 93–104. https://doi.org/10.13031/2013.36263

Birgand, F., Aveni-Deforge, K., Smith, B., Maxwell, B., Horstman, M., Gerling, A. B., & Carey, C. C. (2016). First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency in situ water chemistry measurements at multiple sites. Limnology and Oceanography: Methods, 14(12), 767-783.

Cizek, A. R, J. P. Johnson, F. Birgand, W. F. Hunt, and R. A. McLaughlin (2019). “Insights from using in-situ ultraviolet-visible spectroscopy to assess nitrogen treatment and subsurface dynamics in a regenerative stormwater conveyance (RSC) system”. In: J. Environ. Manage. 252, p. 109656. DOI: 10.1016/j.jenvman.2019.109656.

DiFoggio, R. (2000). Guidelines for applying chemometrics to spectra: Feasibility and error propagation. Appl. Spectroscopy, 54(3), 94A– 113A.

Etheridge, J. R., Birgand, F., Burchell, M.R., Smith, B. T. 2013. Addressing the fouling of in situ ultraviolet–visual spectrometers used to continuously monitor water quality in brackish tidal marsh waters. J. Environ. Qual, 42: 1896–1901.

Etheridge, J. R., Birgand, F., Burchell, M. R. 2015. Quantifying nutrient and suspended solids fluxes in a constructed tidal marsh following rainfall: The value of capturing the rapid changes in flow and concentrations. Ecological Engineering, 78: 41-52.

Frazar, S, A. J. Gold, K. Addy, F. Moatar, F. Birgand, A. W. Schroth, D. Q. Kellogg, and S. M. Pradhanang (2019). "Contrasting behavior of nitrate and phosphate flux from high flow events on small agricultural and urban watersheds". In: Biogeochemistry. DOI: 10.1007/s10533-019-00596-z.

Glasgow, H. B., Burkholder, J. M., Reed, R. E., Lewitus, A. J., Kleinman, J. E. 2004. Realtime remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. J Exp Mar Biol Ecol., 200: 409-448.

Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., ... & Jarvie, H. P. (2015). High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. Hydrological Processes, 29(15), 3388-3407.

Hannouche, A., Chebbo, G., Ruban, G., Tassin, B., Lemaire, B. J., & Joannis, C. (2011). Relationship between turbidity and total suspended solids concentration within a combined sewer system. Water Science and Technology, 64(12), 2445-2452.

Harmel, R. D., King, K. W., & Slade, R. M. (2003). Automated storm water sampling on small watersheds. Applied Engineering in Agriculture, 19(6), 667.

Hochedlinger, M., Kainz, H. & Rauch, W. (2006). Assessment of CSO loads – based on UV/VIS-spectroscopy by means of different regression methods. Water Science and Technology 54 (6–7), 239–246.

Horowitz, A.J., 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. Hydrol. Process. 17, 3387–3409. https://doi.org/10.1002/hyp.1299

King, K.W., Harmel, R.D., 2004. Comparison of Time−based Sampling Strategies to Determine Nitrogen Loading in Plot−scale Runoff. Transactions of the ASAE 47, 1457−1463.

Kirchner JW, Feng XH, Neal C, Robson AJ. 2004. The fine structure of water-quality dynamics: the (high-frequency) wave of the future. Hydrological Processes 18: 1353–1359. DOI:10.1002/hyp.5537

Kronvang, B., Bruhn, A.J., 1996. Choice of Sampling Strategy and Estimation Method for Calculating Nitrogen and Phosphorus Transport in Small Lowland Streams. Hydrol.

Process. 10, 1483–1501. https://doi.org/10.1002/(SICI)1099-1085(199611)10:11<1483::AID-HYP386>3.0.CO;2-Y

Langergraber, G., Fleischmann, N., and Hofstader, F. (2003) A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. Water Sci. Technol., 47: 63–71.

Lepot, M., Torres, A., Hofer, T., Caradot, N., Gruber, G., Aubin, J.-B., Bertrand-Krajewski, J.-L., 2016. Calibration of UV/Vis spectrophotometers: A review and comparison of different methods to estimate TSS and total and dissolved COD concentrations in sewers, WWTPs and rivers. Water Res. 101, 519–534. https://doi.org/10.1016/j.watres.2016.05.070

Lin, C. W. (2017). Characterization of the Solute Dynamics Signature of an Agricultural Coastal Plain Stream, North Carolina, USA. North Carolina State University.

Littlewood, I.G., 1992. Estimating constituent loads in rivers: A review. Report No. 117.

Littlewood, I.G., 1995. Hydrological regimes, sampling strategies, and assessment of errors in mass load estimates for United Kingdom rivers. Environ. Int. 21, 211–220.

Littlewood, I.G., Watts, C.D., Custance, J.M., 1998. Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975–1994). Sci. Total Environ. 210-211, 21–40. https://doi.org/10.1016/S0048-9697(98)00042-4

Liu, W., Birgand, F., Youssef, M., & Chescheir, G. (2016). A novel method to reveal the nitrate transport and fate in agricultural fields. In 2016 10th International Drainage Symposium Conference, 6-9 September 2016, Minneapolis, Minnesota (pp. 1-3). American Society of Agricultural and Biological Engineers.

Lorenz, U., Fleischmann, N., & Dettmar, J. (2002). Adaptation of a new online probe for qualitative measurement to combined sewer systems. In Global Solutions for Urban Drainage (pp. 1-12).

Maxwell, B. M., Birgand, F., Schipper, L. A., Christianson, L. E., Tian, S., Helmers, M. J., ... & Youssef, M. A. (2019). Drying–rewetting cycles affect nitrate removal rates in woodchip bioreactors. Journal of environmental quality, 48(1), 93-101.

Mevik, B., Wehrens, R., and Liland, K. H. (2015). pls: Partial least squares and principal component regression. R package 2.4-3. http://CRAN.R-project.org/package=pls.

Moatar, F., Meybeck, M., 2005. Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. Hydrol. Process. 19, 429–444. https://doi.org/10.1002/hyp.5541

Moatar, F., Meybeck, M., 2007/5. Riverine fluxes of pollutants: Towards predictions of uncertainties by flux duration indicators. C. R. Geosci. 339, 367–382. https://doi.org/10.1016/j.crte.2007.05.001

Ojeda, C. B., & Rojas, F. S. (2009). Process analytical chemistry: applications of ultraviolet/visible spectrometry in environmental analysis: an overview. Applied Spectroscopy Reviews, 44(3), 245-265.

Packman, J., Comings, K., & Booth, D. (1999). Using turbidity to determine total suspended solids in urbanizing streams in the Puget Lowlands.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rügner, H., Schwientek, M., Beckingham, B., Kuch, B., & Grathwohl, P. (2013). Turbidity as a proxy for total suspended solids (TSS) and particle facilitated pollutant transport in catchments. Environmental earth sciences, 69(2), 373-380.

s::can Messtechnik GmbH .(2019). spectro::lyser datasheet. https://www.s-can.at/plugins/system/vm_multiupload_attachment/js/multiupload/server/uploads/spectrolyser_ww_EN.pdf

Strecker, E., Mayo, L., Quigley, M., & Howell, J. (2001). Guidance manual for monitoring highway runoff water quality (No. FHWA-EP-01-022). Federal Highway Administration.

Van Den Broeke, J., Langergraber, G., & Weingartner, A. (2006). On-line and in-situ UV/vis spectroscopy for multi-parameter measurements: a brief review. Spectroscopy Europe, 18(4), 15-18.

Wang, X., Frankenberger, J.R., Kladivko, E., 2003. Estimating Nitrate–N Losses From Subsurface Drains Using Variable Water Sampling Frequencies. Transations of the ASAE 46, 1033–1040.

Webb, B.W., Phillips, J.M., Walling, D.E., Littlewood, I.G., Watts, C.D., Leeks, G.J.L., 1997. Load estimation methodologies for British rivers and their relevance to the LOIS RACS(R) programme. Sci. Total Environ. 194-195, 379–389. https://doi.org/10.1016/S0048-9697(96)05377-6

Webb, B.W., Phillips, J.M., Walling, D.E., 2000. A new approach to deriving “best-estimate” chemical fluxes for rivers draining the LOIS study area. Sci. Total Environ. 251-252, 45–54.

Wu, J., Pons, M., and Potier, O. (2006) Wastewater fingerprinting by UV-visible and synchronous fluorescence spectroscopy. Water Sci. Technol., 53: 449– 456.

Zamora, D., & Torres, A. (2014). Method for outlier detection: a tool to assess the consistency between laboratory data and ultraviolet–visible absorbance spectra in wastewater samples. Water science and technology, 69(11), 2305-2314.

CHAPTER 2: UNCERTAINTY ANALYSIS FOR STORMWATER FLOW-INTERVAL COMPOSITE SAMPLING

Abstract

Flow-interval composite sampling is a popular sampling method in stormwater monitoring. Several studies have previously evaluated this sampling method and revealed valuable information. However, none of them used *measured* and *independent* reference pollutant loads/concentrations from continuous hydro- and chemographs for their evaluation. In this study, minute scale discharge and water quality monitored using an *in situ* spectrophotometer, served as reference data. The latter were numerically subsampled to simulate and evaluate the sources and magnitude of uncertainty associated with flow-interval composite sampling. The reference data showed that during events, the majority of TSS was transported within the first 40% of event flow volume, while the majority of nitrate and DOC loads occurred in the last 60-100% of event flow volume. The identified sources of uncertainty included (1) the sampling start threshold, (2) the number of samples per event, and (3) the percentage of event sampled. Our results suggest that very high percentages, i.e., at least 85% (for nitrate and DOC), to 95% (for TSS) of an event should have been sampled to limit systematic bias, respectively under- and overestimating event loads. This is a lot more than the previously accepted 60%. Our results also show that the first sample due to be pumped and mixed within the composite bottle can induce very large (>50%) error in the final Event Mean Concentration (EMC) or event load. We recommend to program automatic samplers not to take the first sample. Finally, ±5% error on event load or EMC could be obtained using proper sampling interval, if at a minimum the composite bottle would be 25% full for nitrate and DOC, and about 30% full for TSS. This study

provides revised guidelines to design flow-interval composite sampling for stormwater and/or to assess the level of uncertainty researchers might encounter and report.

2.1 Introduction

Runoff from urban land use has intensified characteristics such as increased volume, peak flow and pollutant loads (Leopold 1973, Line et al. 2002, Walsh, et al. 2005). Stormwater control measures (SCMs) are designed to mitigate these extreme hydrologic and water quality characteristics. Assessment of their mitigation impact is an important tool for (1) regulatory evaluation of discharged pollutants from a site, (2) identifying and managing potential sources of pollutants, (3) monitoring the SCM performances, and (4) for improvement of SCM design. Proper assessment of SCM impact, requires accurate measurement of hydrology and water quality passing through. However, obtaining accurate water quality measurements are particularly difficult due to the fluctuating and flashy nature of stormwater flow and pollutant concentrations, as well as limitations of current monitoring methods in obtaining high resolution data. To obtain reliable pollutant removal percentages, the uncertainty of water quality measurements must account for no more than a fraction of the removal calculated.

Several studies have evaluated different sampling types (discrete/composite), techniques (manual/automatic), and intervals (flow/time) to measure their accuracy, precision, and configurations to achieve optimal pollutant concentration and load estimates (King & Harmel, 1998; Stone et al., 2000; Leecaster et al., 2001; Harmel & King, 2003; Maestre & Pitt, 2006; Lee et al., 2007; Ma et al., 2009; Ackerman et al., 2010; Harmel et al., 2010; McCarthy et al., 2018). Among different sampling methods, flow-interval composite sampling is most popular, and has been assessed to yield most accurate results amongst most other methods tested (Leecaster et al. 2001; Lee et al. 2007). While discrete sampling reveals information on pollutant dynamics with a

good accuracy, the cost associated with it prevents its popularity (Harmel & King, 2005). Flow-interval composite sampling offers a balance between accuracy and sampling costs (Ackerman et al., 2010), and allows for longer sampling duration (US EPA, 1992). However, this sampling method does not provide information on pollutant dynamics. *Figure 2.1* represents flow-interval sampling where samples are collected at constant volume interval, calculated by flow measurement, these samples are all collected in one sampling jar and composited.

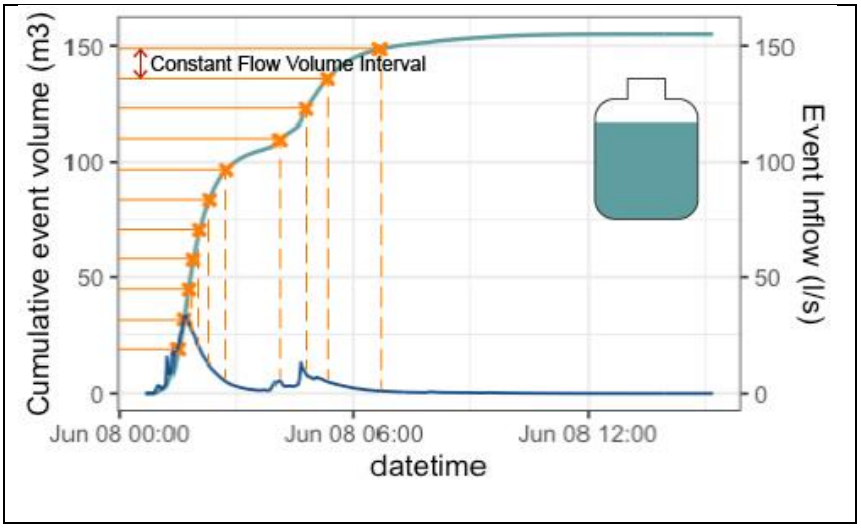
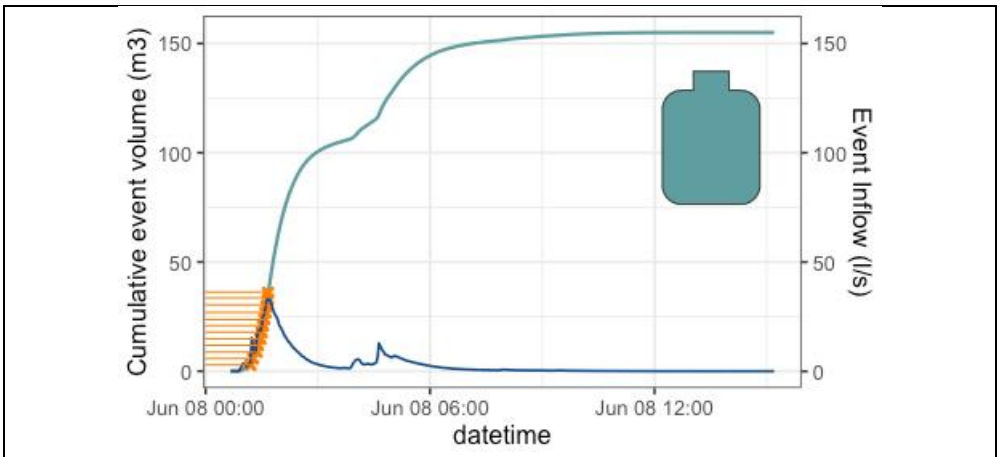
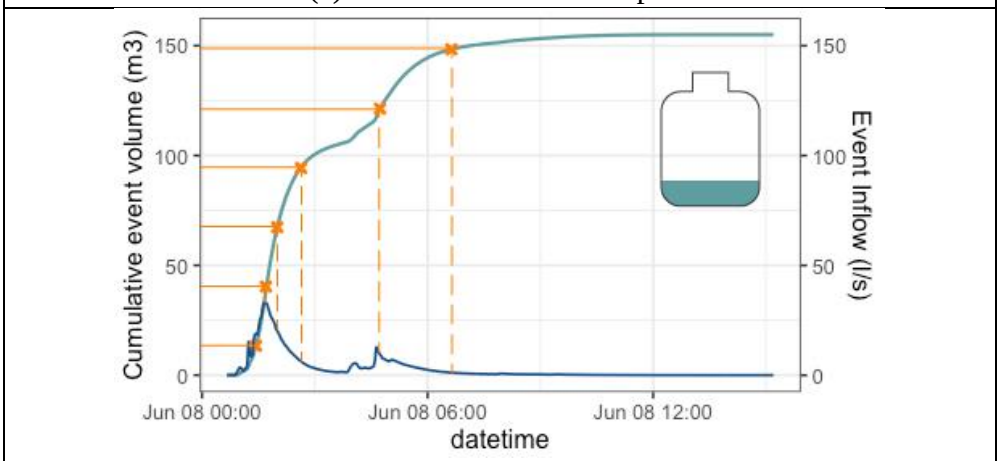


Figure 2.1 Graphical representation of flow-interval sampling collected at constant flow volume intervals and composited in one sampling jar. Each cross represents sampling instances along the cumulative flow volume curve (green curve) and also the timing of the sampling along the hydrograph (blue curve)

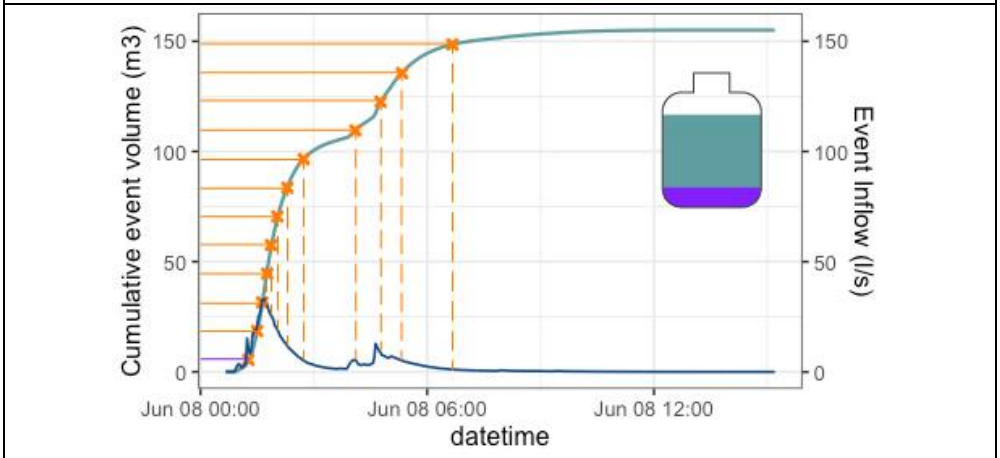
Despite the popularity of flow-interval composite sampling compared to other sampling methods, this method has been shown to sometimes overestimate loads (e.g., Stone et al., 2000; Ackerman et al., 2010). Nonetheless, sampling guidelines (i.e. US EPA, 1992; Burton & Pitt, 2001; Strecker et al., 2001; US EPA, 2002) and best practices have been established using this flow-interval composite sampling (i.e. King & Harmel, 1998; Leecaster et al., 2001; Harmel & King, 2003; Maestre & Pitt, 2006; Lee et al., 2007; Ackerman et al., 2010). These have identified the main sources of uncertainty, which include (1) the duration of event sampled, (2) the number of samples, and (3) the sampling start threshold.



(a) Duration of event sampled



(b) Number of samples



(c) Sampling start threshold with/without first sample in composite bottle

Figure 2.2 Illustration of the common sources of uncertainties in flow-interval sampling strategies to evaluate EMC and loads in stormwater. a) the bottle is full much before the end of the event because the flow volume interval is too small; b) too large a flow volume interval may lead to too few samples in the composite bottle; c) event is well sampled and composite bottle contains the first sample

2.1.1.1 Duration of event sampled

Guidelines suggest sampling of the entire duration of storm (US EPA, 1992; Maestre & Pitt, 2006; *Figure 2.1*; *Figure 2.2*). However, in practice, sampling the entire storm duration might not always be attainable. It is commonly recommended to, at a minimum, emphasize sampling before and during peak flow or at least over 60% of the duration of the event flow (Strecker et al., 2001), when sampling the entire duration of event is not possible. This approach assumes that the majority of pollutants is transported during the first part of a storm event. However, this assumption may not apply to all types and sizes of watersheds, nor for all pollutants (Deletic, 1998). Therefore, there is a need to further evaluate the relationship between duration of event sampled and the accuracy of load estimates as it can affect the error in concentration estimates (Ackerman et al., 2010).

2.1.1.2 Number of samples

Too few samples may yield inaccurate load estimates (Maestre & Pitt, 2006) and increasing the number of samples has been shown to reduce the sampling error (Leecaster et al. 2001; Haan, 2002; *Figure 2.2*). Stormwater sampling guidelines offer graphs to select number of samples, which are based on statistical formulas that calculates this number based on error range, variance of measurement, degree of confidence, and statistical power (Burton & Pitt, 2001 (page231); US EPA, 2002 (page 70, Appendix B)). Based on these guidelines, to derive statistically reasonable conclusions within 20% error, coefficient of variation of 0.5, 95% confidence, and 80% power, approximately 30 samples should be collected (Burton and Pitt, 2001). Empirical studies show collection of 12-16 composite samples can result in event mean concentration (EMC) estimates within 10-20% error (Strecker, et al., 2001).

2.1.1.3 Sampling start threshold

The first part of a storm often washes the accumulated pollutants from urban surfaces, and may contain higher pollutant concentrations than the rest of a storm event (McCarthy et al., 2018; Hvitved-Jacobsen et al., 2010; Pitt et al., 2004; *Figure 2.2*). Although sampling should start early within an event (Maestre & Pitt, 2006), collection of samples very early during inflow may lead to overestimation of EMC (Maestre & Pitt, 2006; Lee et al. 2007) as samples may be more reflective of the first flush rather than the entire event (Deletic, 1998; Khan et al., 2006; Lee et al. 2007). First flush is defined as the higher pollutant load in the initial part of runoff (Pitt et al., 2004). Khan et al. (2006) have shown that sampling from the first flush may result in a bias towards overestimation of EMC. Lee et al. (2007) have shown that sampling the middle of an event would include a more representative sample and sampling early in the event would be biased towards the concentration peak often leading to overestimation.

There is no reason to *a priori* doubt these guidelines. However, none of them have been established from measured continuous hydro- and chemographs. Studies that have evaluated uncertainties from flow-interval sampling have compared the EMC of different sampling methods with a best estimate of the reference EMC. The reference EMC were often calculated by collecting further samples, which is subject to the mentioned sampling limitations (Leecaster, et al. 2001; Harmel & King, 2005; Harmel et al., 2010; McCarthy et al., 2018), or were calculated from models simulating chemographs (King & Harmel 2004; Ma et al., 2009; Ackerman et al. 2010). While these studies have revealed valuable information, lack of access to *measured* and *independent* reference loads from continuous hydro- and chemographs is a shortcoming of the previously reported work.

Stormwater is notoriously flashy with flow and concentrations sometimes varying by orders of magnitude within minutes. *In situ* spectrophotometers offer the possibility to collect concentration data at a frequency compatible with minute-scale changes at an affordable cost of laboratory testing (Torres & Bertrand-Krajewski, 2008; Etheridge et al., 2013; Liu et al., 2016; Halliday et al., 2016; Birgand et al., 2016; Maxwell et al. 2019; Chapter 1). Eliminating the capacity shortage and the loss of information on water quality dynamics associated with automated sampling, there theoretically are, with *in situ* spectrophotometers, much fewer limits on capturing the entirety of an event and on sampling multiple events. Using such sensors thus eliminates the uncertainty associated with low frequency sampling (Birgand et al., 2010).

However, these sensors carry uncertainties of their own and only measure a limited number of parameters, yet. In particular, because their calibration depends on the sample calibration pool (details in chapter 1), the actual chemograph lies among a suite of computed chemographs. Therefore, the actual reference chemograph is not known, but the results presented in chapter 1 suggest that the concentration values and dynamics computed are likely very close to actual ones. As such, a chemograph computed with the optimum conditions defined in chapter 1 may serve as an acceptable surrogate for a reference chemograph.

In this study the sources and magnitude of uncertainty associated with flow-interval composite sampling were evaluated using near-continuous flow and water quality, called reference data herein, collected using an *in situ* spectrophotometer. The objectives were to:

1- Explore pollutant dynamics in stormwater through high resolution chemographs and identify the most polluted part of inflow.

2- Evaluate the error range associated with different sources of uncertainty, i.e., duration of event sampled, number of samples, and sampling start threshold.

3- Offer guidelines to minimize the uncertainty associated with the identified three main sources.

2.2 Methods

2.2.1 Site description

Inlet of a stormwater wetland located on the campus of North Carolina State University was selected for this study.

This wetland was within a watershed of 9.12 ha with 30% impervious surfaces and a ratio wetland surface area/drainage area of 1.39% (*Figure 2.3*). The watershed consisted of wooded area (~86%), roads (~2%), and roof surfaces (~12%).

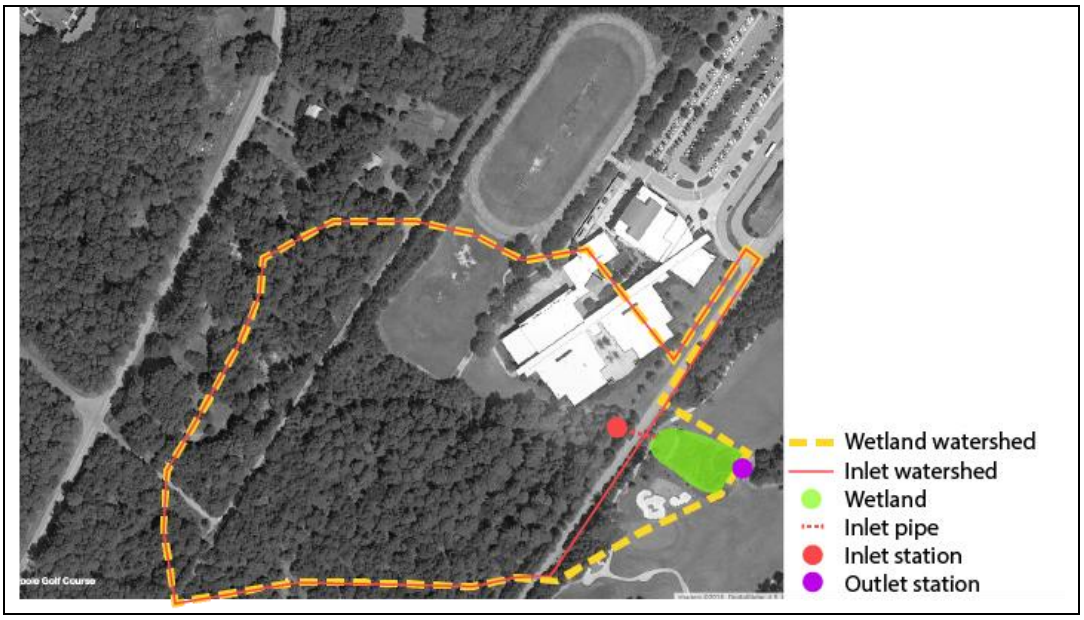


Figure 2.3 Watershed delineation (dashed yellow line), portion of the watershed draining to inlet station (red line), and area for the stormwater wetland in research (highlighted in green)

2.2.2 Reference data acquisition

Reference data were acquired over 20 months between December 2017 and August 2019.

2.2.2.1 Flow reference data

Flow rates and cumulative volume were calculated using 2-minute stage data measured using a Teledyne ISCO® Bubbler Module installed upstream a 90-degree weir, itself installed inside a 120-cm diameter corrugated pipe upstream the wetland. The flow measurement module is traditionally used to continuously compute flow rates and cumulative flow volumes, used for flow-proportional composite sampling. After a given flow-interval has been reached, this system sends a signal to the automatic sampler to draw a water sample, then added in one composite jug where all samples are mixed (*Figure 2.1*). In this study, the same system was used to draw flow-interval samples, not to obtain flow-proportional composite samples, but to obtain discrete samples (i.e., one sample per bottle for a maximum of 24 samples/bottles per event) later used for calibration of the spectrophotometer (details below and in Chapter 1).

2.2.2.2 Water Quality Monitoring and Sampling

Water quality was measured using an *in situ* UV-Vis spectrophotometer, Spectro::lyser™ produced by S::CAN® Messtechnik GmbH with an optical path length of 5mm. The spectrophotometer sensor was placed behind the weir such that it was submerged in the presence of flow and collected water quality data at 4-minute intervals. Monitored parameters included nitrate (NO_3^-), total suspended solids (TSS), and dissolved organic carbon (DOC). This probe measures light absorbance over the 200-730 nm range. Before each measurement, an automatic cleaning of the optics was performed using the Ruck::sack™ scrapper system from the same manufacturer. Additionally, the spectrophotometer sensor was manually and frequently cleaned using dilute acid (HCl 2%) followed by deionized water and to ensure cleanness of lenses and minimize fouling (Ethridge, et al. 2014).

The S::CAN spectrophotometer collects absorbance and converts the data to concentration values using a built-in "global calibration" algorithm. However, the "global calibration" algorithm that comes with the instrument cannot accommodate all types of waters and the manufacturer recommends local calibration. The literature suggests that best results are obtained after chemometric calibrations between absorbance data and local concentration values (Langergraber et al., 2003; Rieger et al., 2006; Torres and Bertrand-Krajewski, 2008; Etheridge et al., 2013, Chapter 1). For local calibration of spectrophotometers, flow-interval discrete water quality samples were obtained using Teledyne ISCO® portable auto sampler model 6712 with 24 bottle configurations as described above. It is best for calibration, to use a pool of samples covering and equally representing the entire range of observed concentrations (low and high values; Rieger et al., 2006). High concentrations tend to occur during high flows in stormwater (Soeur, et al. 1995; Brown, et al. 1995). Therefore, flow-interval sampling method was thought to be best to obtain samples of variable concentrations. Detailed description of data collection and calibration of the spectrophotometer is provided in Chapter 1. Results showed that the method provided a range of chemographs slightly differing from one another. In this study, it was assumed that the concentration dynamics were very similar among all predicted hydrographs, and that any of them could be used as 'reference' dataset. In the end, the concentrations predicted by PLSR model using stratified samples served as reference chemographs for nitrate, DOC, and TSS.

2.2.3 Simulations of Flow-interval Composite Sampling

The 4-minute absorbance-derived concentrations were first linearly interpolated to obtain and align with the 2-min flow data. The 2-min concentration and flow data then served as the reference and numerically sampled to simulate flow-interval composite sampling on a per storm

event basis. All numerical sampling simulations were performed using the R language (R core team, 2020).

2.2.3.1 Targeted Storm events

An individual storm event was defined as a rainfall event that occurred following at least 6 hours of dry weather. A total of 131 individual storm events were recorded generating from 1.2 to 27.6 mm of runoff. Using the spectrophotometer sensor, water quality data was available for 81 of these storm events. Among these, storm events were ranked based on their cumulative runoff volume, and in the end 29 events that accounted for a cumulative 90% of the total annual runoff volume were selected for sampling simulations and uncertainty analysis (*Figure 2.4*).

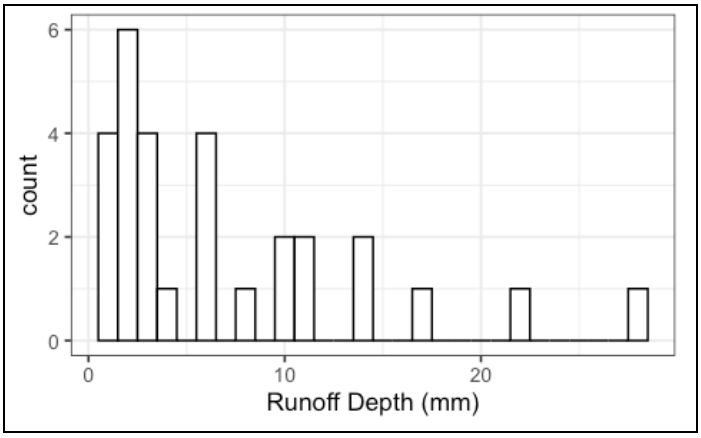


Figure 2.4 Distribution of runoff depth for the 29 selected events

2.2.3.2 Sampling Volume Capacity

In practice obtaining flow-interval composite samples requires a bottle to collect and composite samples. Therefore, for simulation purposes a 20-liter bottle was assumed as the maximum volume capacity, collecting samples of 200 mm aliquots. This assumption would allow for collecting a maximum of 100 samples per event, such that the results reported herein be directly translated into percentage values.

2.2.3.3 Start Threshold to define the First Sample

While the sampling interval defines the flow-volume interval between consecutive samples, the start threshold in flow-interval sampling corresponds to the threshold conditions set on autosamplers by the user to trigger the very first sample after the operator leaves the field prior to an event. It can be set to as a flow rate that is above zero, a given flow-volume, or another specified baseline. Selection of a proper start threshold has shown to directly affect measurement uncertainty and it is suggested to be set at a low value to ensure sampling of small events (Harmel et al., 2003). *Figure 2.5* illustrates two variations of start thresholds for a given sampling flow-interval of 25 m^3. While the flow-interval is the same, changing the start threshold would affect timing of the samples collected within the hydrograph (depicted as vertical lines), and therefore would likely affect the concentrations of the composite sample. By changing the start threshold, the first sample was always within the first passing of flow interval (i.e. the first 27 m^3 of the storm in Figure 2.5).

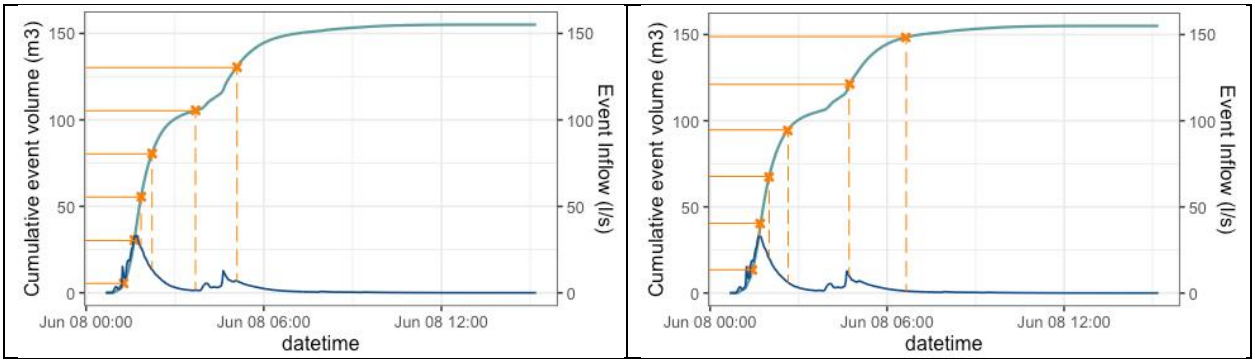


Figure 2.5 Both figures illustrate sampling at 27 m^3 flow interval, while the start threshold on the left figure is at 5 m^3 and the start threshold on the right figure is at 15 m^3

To quantify its impact on load and EMC uncertainties, a range of start threshold values were simulated. For this, for each event, the *start period* was defined as the period between the

beginning of flow and the time the flow-interval volume was reached. Within this *start period*, and based on the available water quality data for each storm event, up to 200 starting times were randomly selected to correspond to the timing of the *first sample*. For events with less than 200 available water quality data points within the *start period*, all available data points were used in the simulations.

Because the first sample has been shown to sometimes play a disproportional role in the final flow-weighted concentrations (Maestre & Pitt, 2006; Khan et al., 2006; Lee et al. 2007; Bach et al., 2010), paired simulations were conducted by in- or excluding the first sample in the computed composite sample concentration for each simulation.

2.2.3.4 Sampling Interval

Each of the storm events was numerically sampled for a range of flow-intervals. This range of intervals was reflective of the range of runoff volumes for the selected storm events and the capacity limitation of 100 samples, so that the entire event was sampled. Cumulative volume of the selected storms ranged from 105-2516 m^3 of runoff (1.15-27.42 mm), based on this range flow-intervals of 1-30 m^3 (0.01-0.33 mm) at 1m^3 (0.01 mm) increment, were selected for simulations.

For each sampling interval, up to 200 simulations were performed corresponding to the selection of start thresholds. This resulted in a total number of 64,060 simulations of flow-proportional composite sampling, each resulting in one composite sample with an EMC value for nitrate, dissolved organic carbon and total suspended solids.

2.2.4 Statistical analysis

2.2.4.1 Description of estimators

Studies in stormwater often measure and evaluate water quality in terms of concentrations and mass loads (Strecker et al., 2001). Flow-interval composite sampling gives, at the event scale, an estimate of the flow-weighted concentration, usually referred to as the Event Mean Concentration (EMC). The pollutant or nutrient load during the event can then be calculated as the product of the EMC by the cumulative flow volume (*Equation 2-1*) as shown in *Equation 2-2* (Strecker et al., 2001; Urbonas, 1994; Lee, et al., 2002).

Equation 2-1 $$V = \sum_{i=1}^{i=n-1} \left[\left(\frac{Q_i + Q_{i+1}}{2} \right) * (t_{i+1} - t_i) \right]$$

Where:

V: event volume (L)

Q_i: Instantaneous flow (L/s) at time stamp t_i

n: total number of datapoints

$t_{i+1} - t_i$: time interval (s)

Equation 2-2 $$L = EMC \times V$$

Where:

L: event pollutant load (mg)

EMC: event mean concentration (mg/L)

V: event volume (L)

For each simulation, the deviation from, or the 'error' on, the reference load was calculated as the percentage difference between the estimated and the reference loads (*Equation*

2-3). For a given sampling interval, up to 200 simulations were run generating a distribution of errors. The mean, the 5th and 95th percentiles (defining the 90% confidence interval) of the distributions were calculated. The choice of 200 was determined as the distribution indicators did not change significantly for more simulations (data not shown). Although root mean square error (RMSE) is used in similar studies (Leecaster et al., 2002; Ma et al., 2009; Harmel et al., 2003), authors believe that this estimator can mask valuable information and sometimes cause false interpretation of results.

Equation 2-3 $$Relative\ error = \frac{Load\ estimate - Reference\ load}{Reference\ load} * 100$$

2.2.4.2 Reference data

Reference load was calculated based on the 2-minute hydrology and water quality dataset, using *Equation 2-4*.

Equation 2-4 $$L_{Ref} = \sum_{i=1}^{i=n-1} \left(\frac{Q_i C_i + Q_{i+1} C_{i+1}}{2}\right) * (t_{i+1} - t_i)$$

Where:

L: Event reference pollutant load (mg)

Q_i: Instantaneous flow (L/s) at time stamp t_i

C_i: Instantaneous concentration (mg/L) at time stamp t_i

n: total number of datapoints

$t_{i+1} - t_i$: time interval (s)

2.2.4.3 Pollutant dynamic analysis

Pollutant dynamics were analyzed for the 29 events accounting for 90% of the total cumulative flow over the monitoring periods, using dimensionless cumulative curve of pollutant load vs volume (Bertrand-Krajewski et al., 1998; Lee and Bang 2000; McCarthy 2009).

In dimensionless cumulative curve load vs. volume method, curves plotted above the one-to-one line indicate occurrence of higher portion of load earlier and in smaller portion of inflow volume. Using *Equation 2-1* and *Equation 2-4*, cumulative volume and cumulative load were calculated for each storm at each data point (t_j). The cumulative values at each data point were then normalized by diving by the total cumulative values (*Equation 2-5*; *Equation 2-6*).

Equation 2-5 $$V_{tj} = \frac{\sum_{i=0}^{i=j} V_i}{V_{event}}$$

Equation 2-6 $$L_{tj} = \frac{\sum_{i=0}^{i=j} L_i}{L_{event}}$$

Additionally, to identify the most polluted part of inflow, the first flush evaluation method suggested by Bach et al. (2010) was used with modification of evaluating loads instead of concentrations and normalized cumulative runoff depth instead of non-normalized. In this method normalized cumulative runoff was divided in 5 parts with increments of 20 percentage points (e.g. 0-20%, 20-40%, etc.) and the distribution of and average pollutant loads in each part was evaluated through box plots.

2.2.4.4 Uncertainty analysis

The error indicators were calculated for each of the 29 events accounting for 90% of the total cumulative flow over the monitoring periods. The error indicators were then compared and analyzed as a function of (1) sampling intervals, (2) number of samples and whether the first sample was included, and (3) percent of event sampled. The statistical significance of effect of including the first sample, was tested using paired t-test.

2.3 Results and discussion

2.3.1 Stormwater pollutant dynamics: typical chemographs

The typical relationship between flow and concentrations observed is illustrated in *Figure 2.6*. Before each event, the area upstream the weir was carefully emptied such that measured concentrations would correspond to new water and not residual water from the previous event. At the onset of the majority of events, the instrument measured nitrate and DOC concentrations that were relatively high compared to the rest of the event. This water likely corresponded to the very first liters of water arriving at the station. These concentrations were quickly diluted with sizable flow rates of several liters per second, creating an apparent concentration 'trough' at the onset of the rising limb of the hydrograph. Nitrate and DOC concentrations gradually increased from then until they peaked after the flow peak, and receded to concentrations close to (e.g., nitrate) or about 75% of those of the peak (e.g., DOC) until flow receded completely. The apparent delay in arrival of dissolved constituents such as nitrate and DOC seems to suggest that for this watershed, there was no first flush effect with these constituents.

Conversely, TSS concentrations exhibited a 'concentration effect' where the TSS concentrations were very small at the onset of the hydrograph, and rapidly increased along with flow. TSS concentration peak would then occur before the flow peak (as illustrated in *Figure 2.6*), at the flow peak, but also sometimes right after the peak. This dynamic was more typical of what was expected for particulates or particulate bound pollutants, as with flow increase the shear stress and transport capacity of water increases as well. The receding of the chemograph afterwards might be associated with lower erosion/transport energy and/or the lowering of erodible/transportable material.

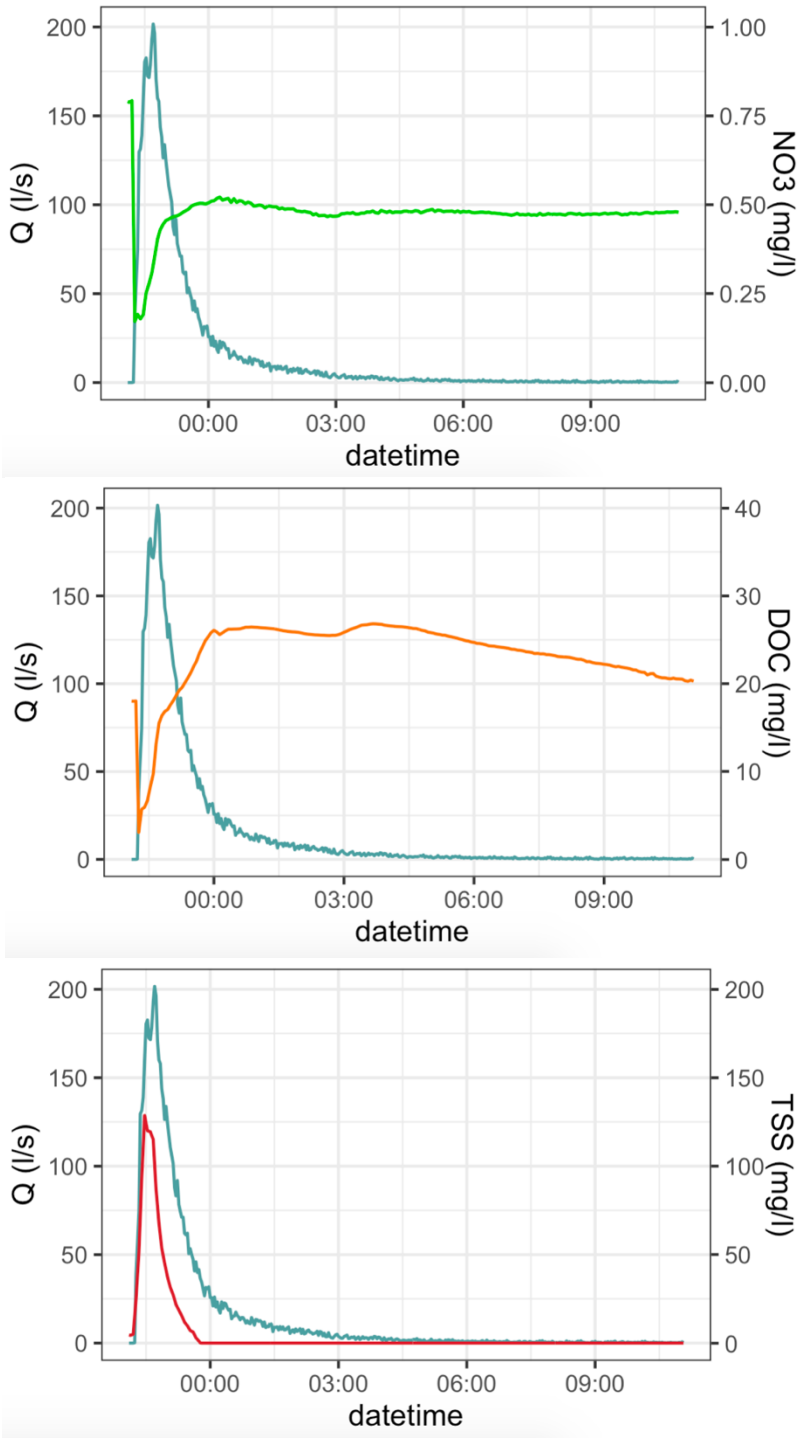


Figure 2.6 Typical chemographs and associated chemograph for event with 8.4 mm runoff that occurred on 08/21/2018 showing and apparent 'dilution effect' of the nitrate and DOC concentrations with respect to flow rates, and a 'concentration effect' of TSS concentrations

2.3.2 Stormwater pollutant dynamics: cumulative load analysis

The chemograph analysis provides a qualitative approach to stormwater pollutant dynamics. The cumulative load analysis completes it on the quantitative side.

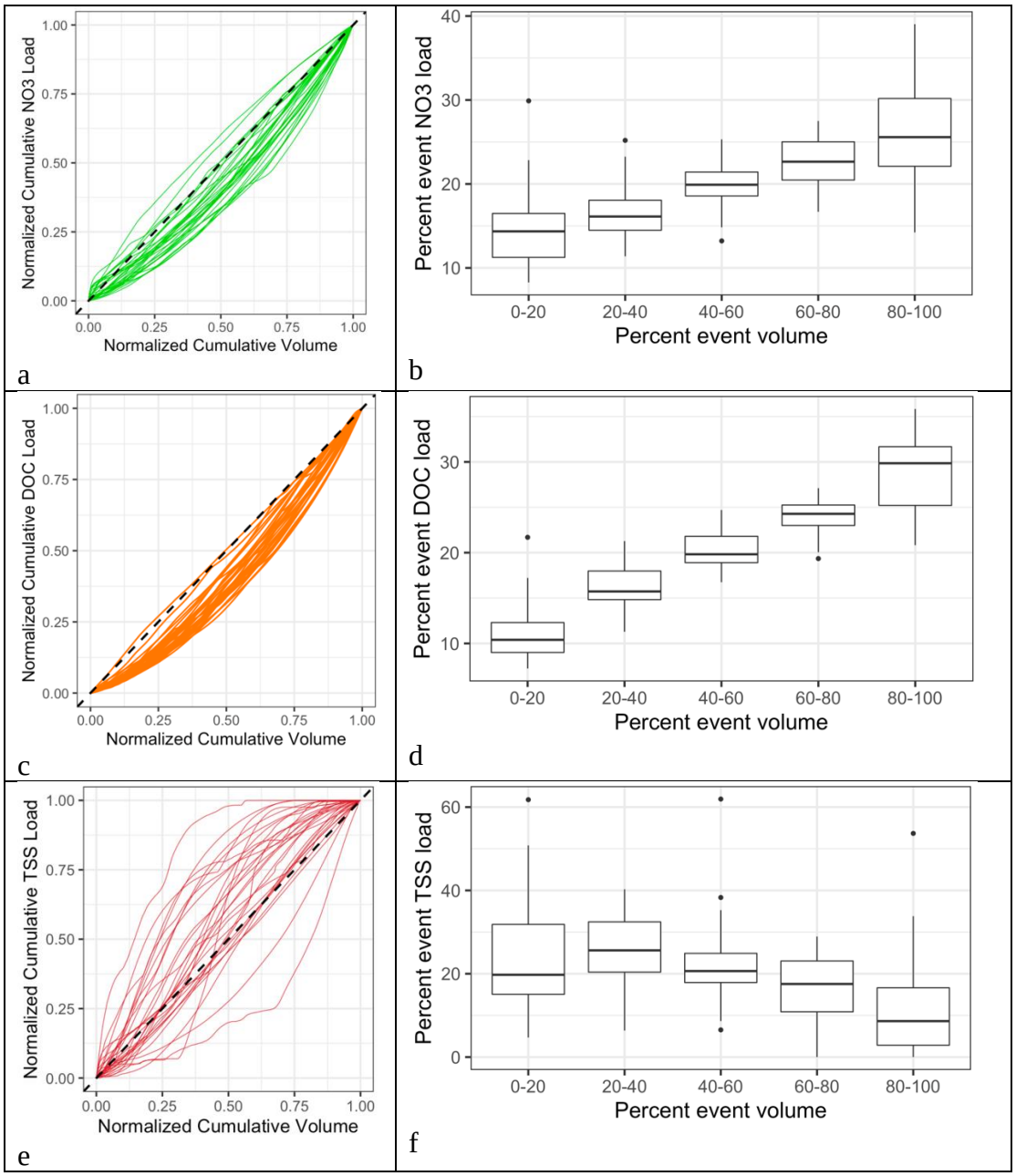


Figure 2.7 Load-volume curves (left) and boxplots of normalized mass per portion of event volume (right) to describe the dynamics of NO_3, DOC and TSS

Figure 2.7 shows the plots of normalized event cumulative load vs normalized event cumulative volume (left) and the percentage of total load carried in each of the 5 parts of the

inflow volume (right). In these normalized plots, the slope of the dotted line is 1. In non-normalized plots, the EMCs would correspond to the slope of the dotted line. Instantaneous concentrations at any particular event would correspond to the slope of the tangent to each solid line. For a majority of events, the highest proportion of the nitrate and DOC loads occurred after 50% of the flow volume had passed corresponding to concentration peaks occurring after flow peaks (*Figure 2.6*; *Figure 2.7*a; *Figure 2.7*c). A few events show alignment with the one-to-one line suggesting instantaneous concentrations were very close to the EMC through most of the event (*Figure 2.7*a; *Figure 2.7*c). Two events did have higher nitrate load transported early in the event (*Figure 2.7*a) corresponding to nitrate peak occurring prior to peak flow. The boxplots for both of these pollutants show a consistent increase in the proportion of the load towards the latter portion of event volume (*Figure 2.7*b; *Figure 2.7*d). The last 20 percentile of volume (i.e. 80-100% of cumulative event volume) had highest variability of load and on average carried about 25% of total nitrate load and about 30% of total DOC load (*Figure 2.7*b; *Figure 2.7*d). Based on these observations collecting samples only at the early part of the event or 60% of the event would have resulted in load underestimation for these pollutants.

The relationship of TSS load transport with event volumes showed more variability: for the majority of the events, the concentration peak occurred before the peak flow (e.g., *Figure 2.6*), for fewer, the TSS peaks and flow peaks were generally synchronous (alignment on the one-to-one line in *Figure 2.7*e), and in even fewer instances the TSS peak occurred after the flow peak (*Figure 2.7*e). Such variability in the relationship of TSS loads and flow volumes was also reported by Metadier and Bertrand-Krajewski (2012). The boxplots show highest variability of TSS load in the first 20 percentile of the inflow volume and highest average loads were observed in the second 20 percentile (i.e. 20-40% *Figure 2.7*f). Based on these observations at early parts

of the events carry the most TSS loads and sampling only the early part of event could have resulted in an overestimation of the TSS loads, as also suggested by Lee et al. (2007).

2.3.3 Evaluation of uncertainties associated with flow-proportional composite sampling

Comparing the relative error at different sampling flow intervals, suggests that a too high or a too low flow-interval can result in significant relative error. *Figure 2.8* illustrates the relative errors for nitrate load with sampling flow intervals varying from 1 to 30 m^3 for one particular event. Up to 200 simulations were run for each sampling interval, and each point represents the calculated error for one simulation. The errors of load estimates can be divided into 3 sections of sampling intervals as illustrated in *Figure 2.8*, (1) too low a sampling flow interval may result in filling the composite bottle before the end of the event (section 1 in *Figure 2.8* ; *Figure 2.2*a) and in the case of nitrate may result in an overall underestimation of loads, (2) 'desirable' sampling intervals when 100% of the event was sampled when there were 'enough' samples in the bottle generating acceptable errors (±5% , section 2 in *Figure 2.8*) , and (3) high sampling interval volume may result in too few samples in the composite bottle and an increasing range of relative error with the increase of sampling interval (section 3 in *Figure 2.8*; *Figure 2.2*b), in the case of nitrate this may result in overestimation of the loads.

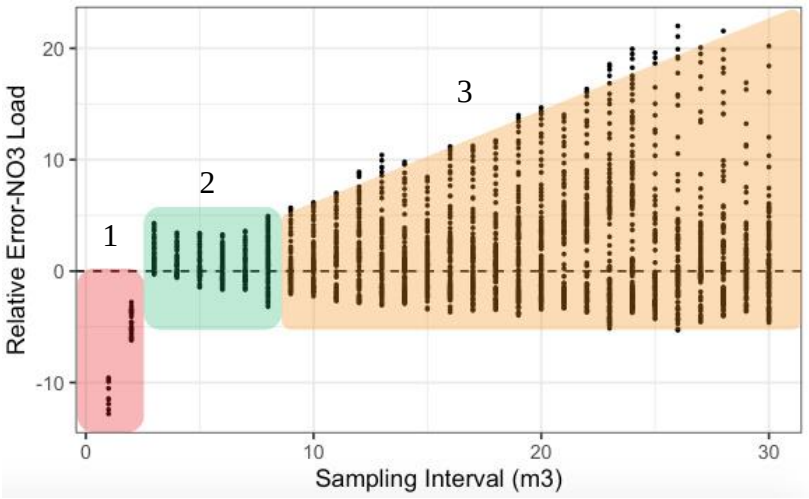


Figure 2.8 An example of relative nitrate load error per sampling interval. Each dot represents the relative error that was computed for one simulation. Up to 200 simulations were computed for a given sampling interval. The red area represents the consequences of having a bottle full before the end of an event; the green area represents the sampling intervals for which the uncertainties are minimum; the bisque area represents the consequences of too few samples

Each event had its own 'uncertainty signature'. For small events, the bottle would not get full but significant uncertainties might come from too few samples when large sampling intervals were simulated. For large events, largest uncertainties would be associated with small sampling intervals and the bottle getting full before the end of the event. To harmonize the results, the uncertainties were plotted as a function of the number of samples per bottle, and as a function of the percentage of the event that was sampled, and drivers of uncertainties were extracted.

2.3.3.1 Uncertainty associated with sampling start threshold and the first sample

Plotting uncertainties as a function of the number of samples per composite bottle, or as a function of sampling intervals, showed very large uncertainties (sometimes over 100%). Further analysis revealed that these large uncertainty values were associated with simulations when the first sample was part of the composite bottle (*Figure 2.9*). It is important to remember that for a given sampling volume interval, the first sample would occur during the *start period.* Results illustrated in *Figure 2.9* suggest that the sample taken during this start period could, for some

events have a dramatic negative effect on uncertainties. Including the first sample in composite sampling, increased the relative error of Nitrate load estimation mainly towards overestimation especially at higher sampling intervals (*Figure 2.9*a) where fewer number of samples were collected (*Figure 2.9*b).

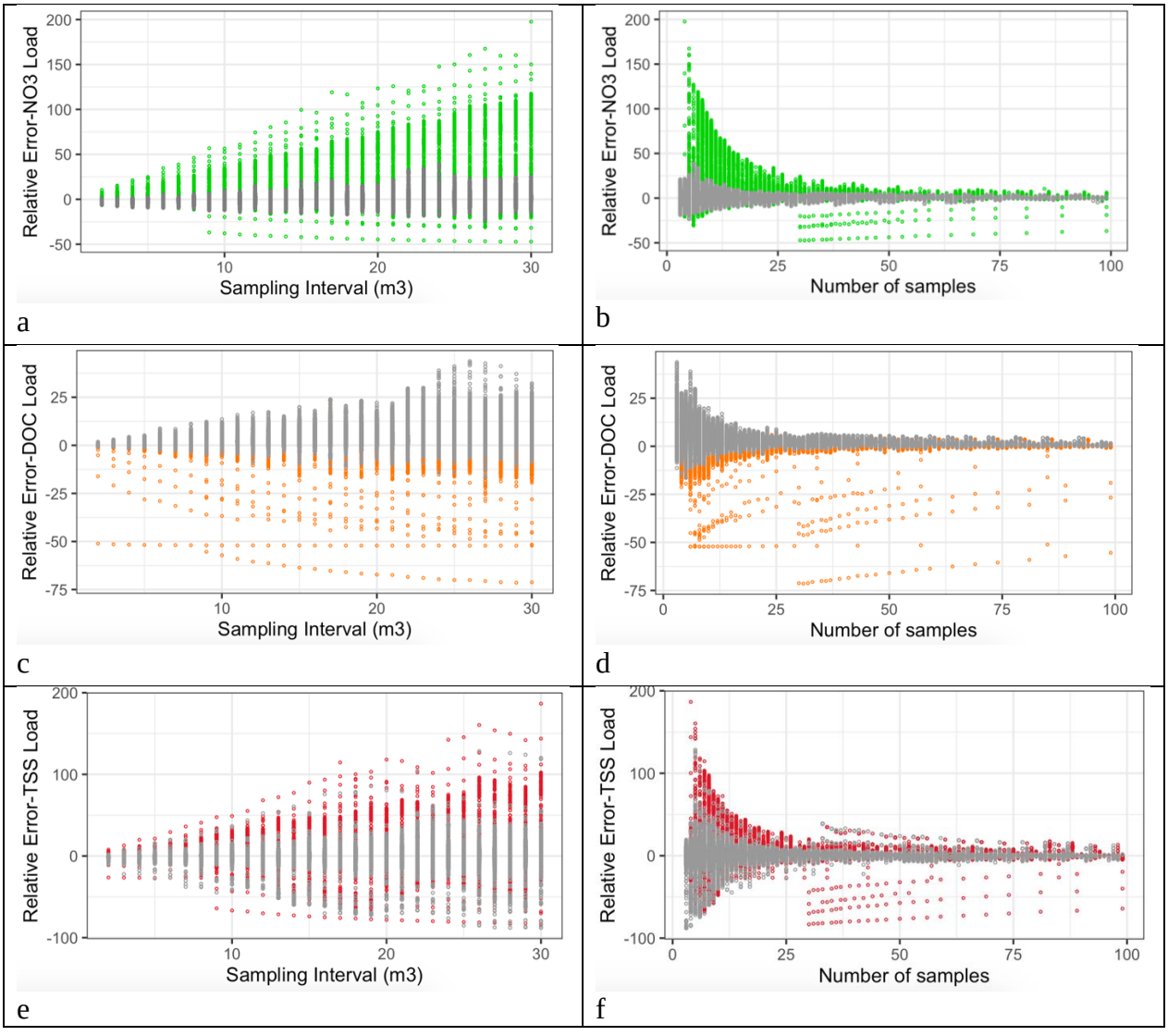


Figure 2.9 Comparing relative error of event load estimations including the first sample (presented in color), and excluding the first sample (presented in gray), as a function of sampling interval (left) and the number of samples per bottle (right) for all studied pollutants

In a few observations, inclusion of the first sample resulted in underestimation of event's nitrate load (negative relative error), these values were observed at different sampling intervals

(*Figure 2.9*a) and number of samples (*Figure 2.9*b). Inclusion of the first sample could result in extreme negative outlier errors for DOC at collection of any sampling interval (*Figure 2.9*c) and number of samples (*Figure 2.9*d).

Similar to nitrate and DOC load estimates, presence of the first sample resulted in extreme outlier negative relative errors especially at high sampling interval (*Figure 2.9*e) where larger number of samples were collected (*Figure 2.9*f) for TSS. Overall exclusion of the first sample in composite sampling could increase the accuracy of TSS load estimates.

While these results were puzzling, it became important to observe why, for some events, including the first sample had such dramatic effect on the uncertainty, e.g., extreme high range of uncertainty values at high sampling interval and outlier relative errors throughout the event (*Figure 2.9*). A closer look at the hydrographs of such events, showed several peaks or pulses of inflow (i.e. *Figure 2.11, Figure 2.13*) as opposed to a textbook hydrograph. Other studies suggest sampling too early within the event can cause overestimation of concentrations and loads due to the first flush effect (Khan et al 2006; Lee et al 2007). *Figure 2.10* and *Figure 2.12* illustrate for two events the relative error of load estimates with regular sampling scheme and excluding the first sample in the composite sampling, for the events illustrated in *Figure 2.11* and *Figure 2.13* respectively. The results suggest that excluding the very first sample from the composite bottle could dramatically reduce the outlier error values of load estimation and also reduce the range of uncertainty at higher intervals where fewer number of samples would be collected.

As shown previously, the studied watershed tended to have the highest TSS concentrations prior to the peak flow (*Figure 2.7*e; *Figure 2.7*f). The particular event illustrated in *Figure 2.11*a had an initial flow after which the flow approached zero before another peak occurred, this initial flow carried higher TSS loads (*Figure 2.11*b). Sampling from this portion

of the event could have resulted in overestimation of loads, especially at higher sampling intervals where fewer samples would be collected, and less dilution could occur in the composite bottle (*Figure 2.10*a). The increasing high range of uncertainty values at high sampling intervals is reflective of the decrease in number of samples by increase of sampling interval. The sampling was simulated again excluding the first sample, where the possible start thresholds after the passing of the first sampling interval were tested. The load estimation error for this simulation showed a drastic reduction of uncertainty at high sampling intervals.

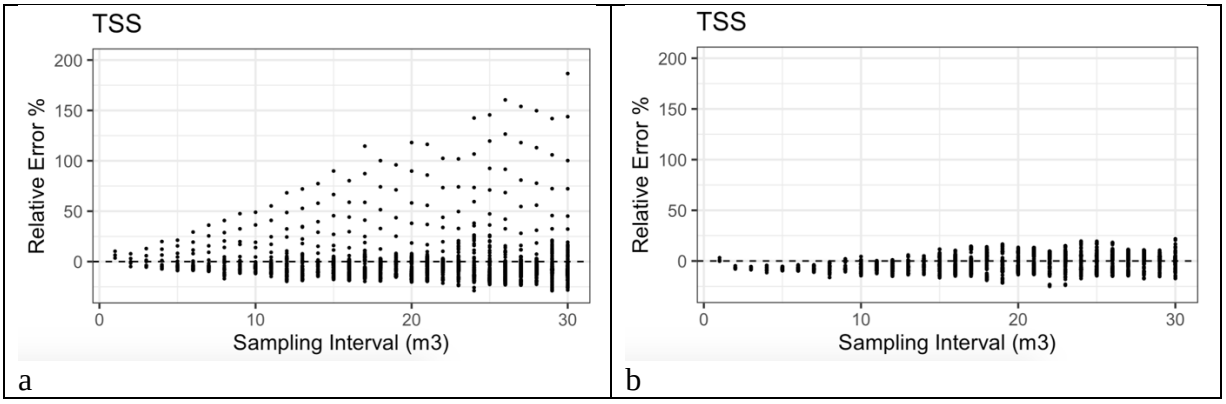


Figure 2.10 Relative error of load estimates at different sampling intervals for 3 pollutants, including all samples (a) and excluding the first sample (b)

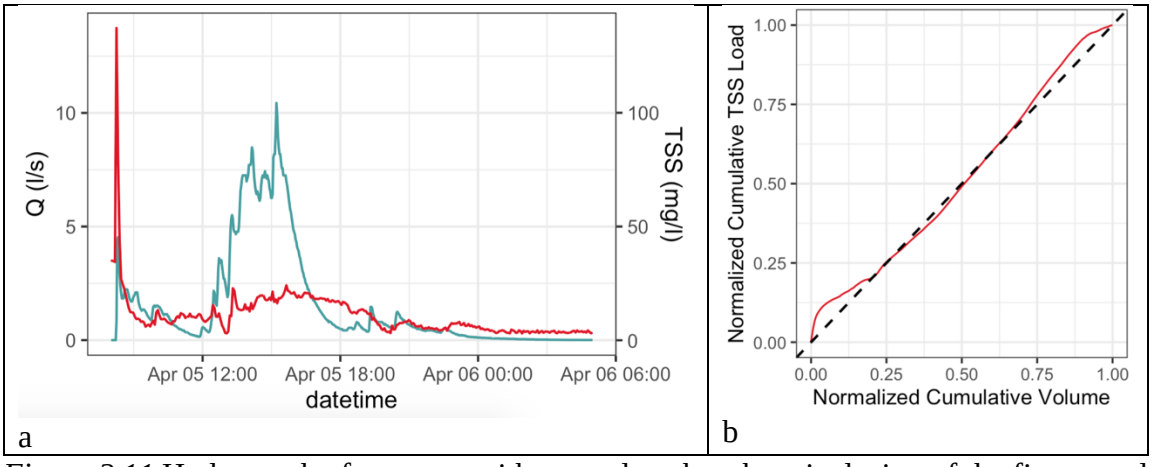


Figure 2.11 Hydrograph of an event with several peaks where inclusion of the first sample can have high impact on the relative error of load estimates

Another example event generating this time, large load underestimation is illustrated in *Figure 2.13*a. This event also had some small flow variations before the major peak occurred, however, these were rather diluted and did not carry much of the TSS load (*Figure 2.13*b). Sampling within the *start period* potentially led to underestimation of TSS loads at higher sampling intervals where fewer number of samples would be collected, and also at lower sampling intervals where a smaller portion of event would be sampled (*Figure 2.12*a). Similarly, removal of the first sample from simulation has eliminated these outlier error in load estimation (*Figure 2.12*b).

The discussed examples illustrate uncertainties in estimating TSS loads, but a similar pattern was also observed for both nitrate and DOC. Overall, these results suggest that preventing the first sample to reach the composite bottle is preferable as it lowers the risk of very high uncertainties in the final estimation of EMC or loads. For this reason, all analyses herein consider the case where simulations were performed excluding the first sample.

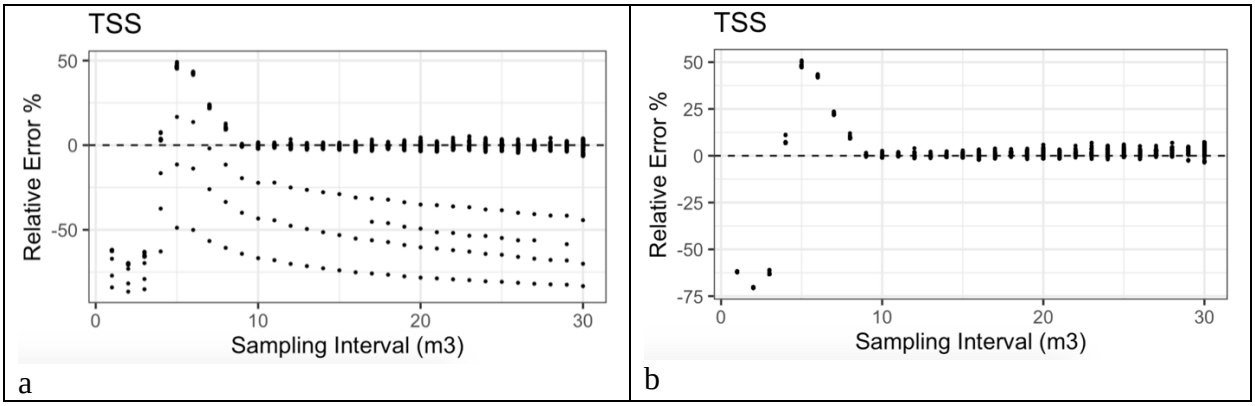


Figure 2.12 Relative error of load estimates at different sampling intervals for 3 pollutants, including all samples (a) and excluding the first sample (b)

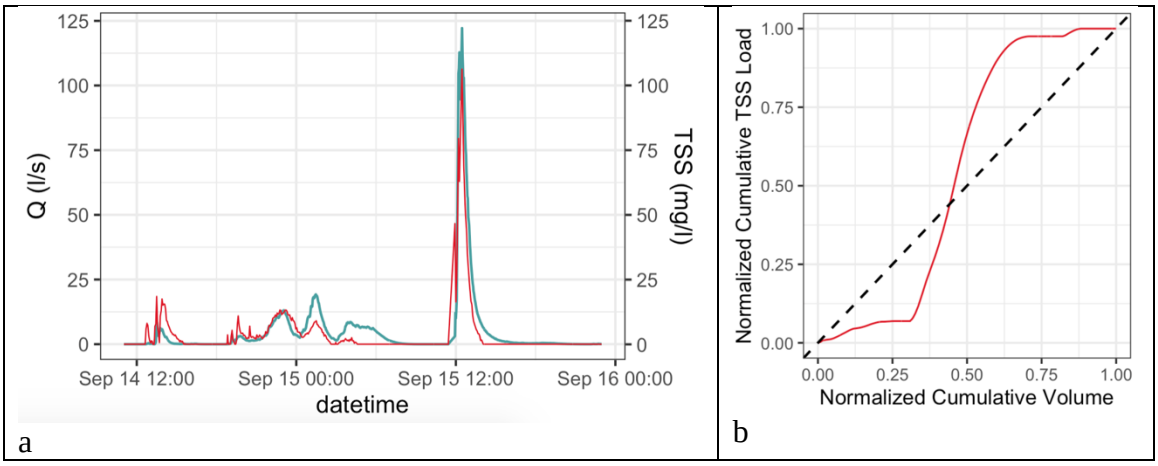


Figure 2.13 Hydrograph of an event with several peaks where inclusion of the first sample can have high impact on the relative error of load estimates

2.3.3.2 Uncertainty associated with the number of samples per composite bottle

As a reminder, for each simulated combination of event × sampling interval (ranging from 1 to 30 m^3, or 0.01-0.33 mm), a maximum of 100 samples per composite bottle was allowed. The load estimates as a function of the number of samples per composite bottle is illustrated herein. Each estimate for each simulation for all 29 events is represented as a dot on *Figure 2.9*. For each value of the number of samples, there is a distribution of dots, from which the uncertainty can be characterized. We chose to use the mean as an indicator of the bias (solid colored line), and the 90% confidence interval (grey area) to assess the overall range of uncertainty. These results exclude the first sample as suggested in the previous section.

In general, the relative error decreased with an increasing number of samples in a composite bottle. Very few samples per bottle generated largest uncertainties, accompanied with high variability in the uncertainty range (*Figure 2.14*). This resulted from the variable number of distribution points for less than 10 samples per bottle.

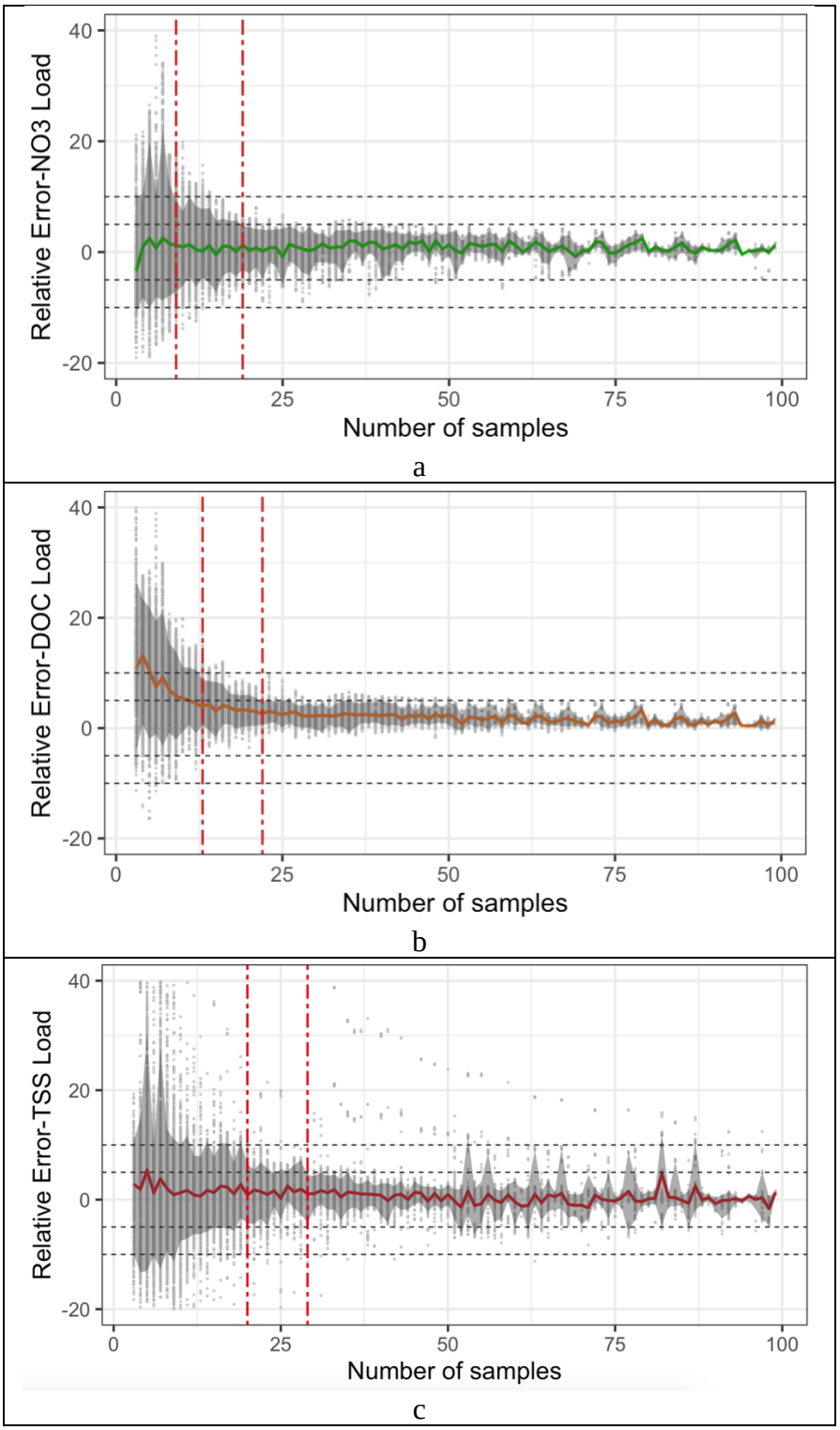


Figure 2.14 Relative error of event's pollutant load estimation per number of samples, excluding the first sample, with gray area representing the 90-percentile range and the mean value shown in colored line. 5 and 10% error range are marked as horizontal lines and the selected minimum number of samples are presented as vertical lines

Based on the 90% confidence interval criterion, it is possible to derive the minimum number of samples for which there would be 90% chance for event loads to be within ±5% and ±10% of the actual (i.e., reference) load. For our watershed and using the measured hydrographs and chemographs, at least 9, 13, 20 samples (or 9, 13, 20% of the bottle capacity) should have been collected to obtain nitrate, DOC and TSS loads, respectively, within ±10% of the actual load. A minimum of 19 and 22 samples (or 19 and 22% of the bottle capacity) for nitrate and DOC, respectively, would have been necessary to obtain loads within ±5% of the actual load.

The 90% confidence interval for TSS in *Figure 2.14* is a bit noisier than for nitrate and DOC, suggesting that ±5% accuracy and within 90% confidence interval might not be attainable. But this is contingent upon the number of samples from which the 90% confidence interval was calculated. The underlying tendency, however, suggested that ±5% accuracy could be reliably attained for 29 samples (of 29% of the bottle capacity). All these values are based on the premise that the first sample was excluded.

2.3.3.3 Uncertainty associated with percentage of event sampled

The last main source of uncertainty illustrated in *Figure 2.8* occurred when the composite bottle gets filled up before the end of an event, leaving a percentage of the event flow volume unsampled. A potentially large source of uncertainty was due to the percentage of event sampled, when it is less than 100%.

Figure 2.15 shows the relative error of load estimation for all events vs. the percentage of event sampled. Due to the few data points at low values of percent event sampled, the presented data was aggregated for every 5% of event sampled (*Figure 2.15*-left) and illustrated in detail at higher percentage of event sampled with aggregation at every 1% (*Figure 2.15*-right).

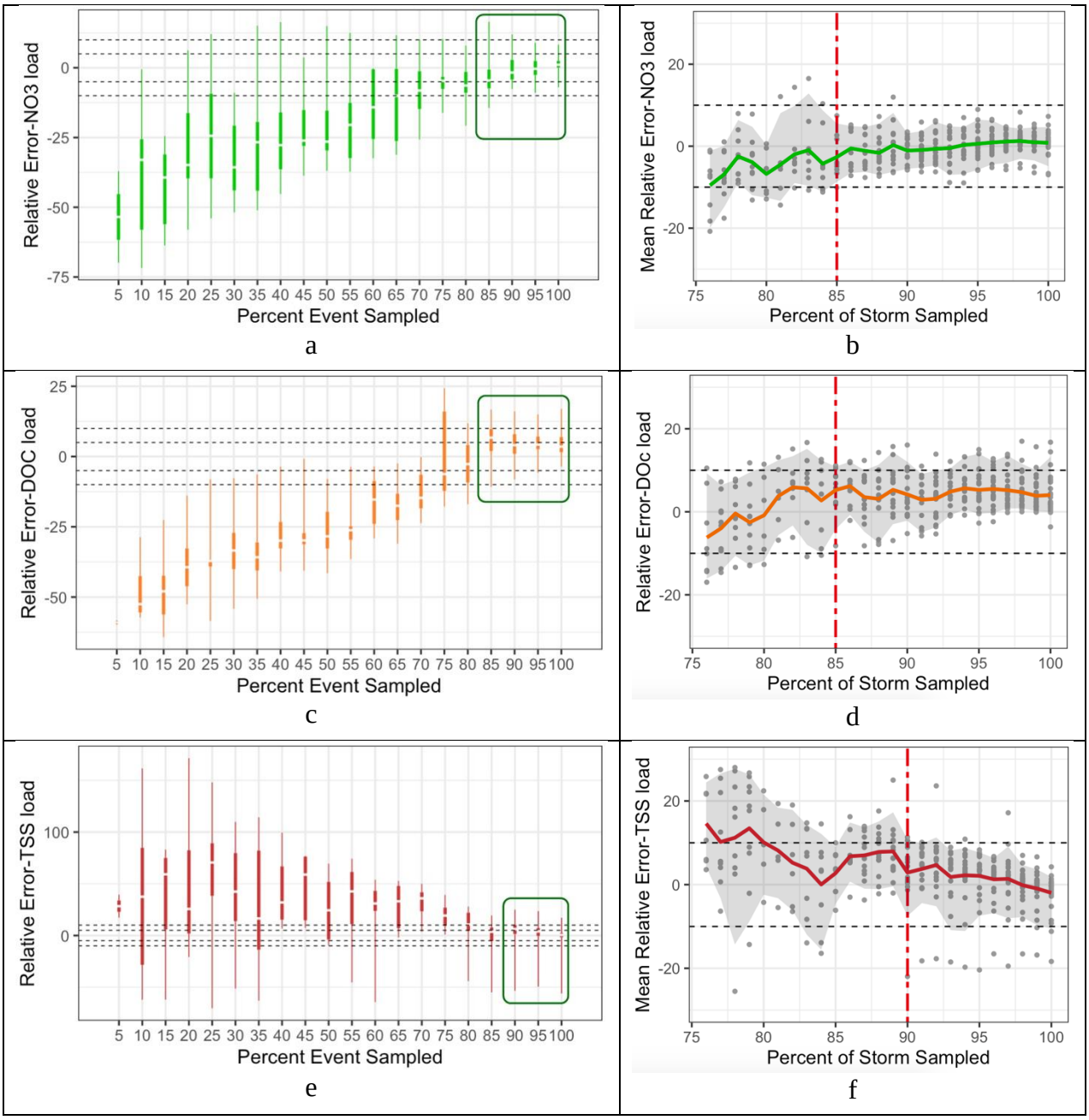


Figure 2.15 Relative errors of event load estimation as a function of percent of event sampled aggregated at every 5 % (left column). The results highlighted in the green rectangles are illustrated at a 1% for higher percentages of event sampled (right column). The 90% confidence intervals and the average errors are represented with a grey area and a solid line. Suggested minimum percent event sampled to achieve ± 10% are shown with a vertical dotted line (right column)

In general sampling a low percentage of event resulted in systematic underestimation of nitrate and DOC loads and systematic overestimation of TSS loads (Figure 2.15-left). Sampling

less than 100% of the event resulted in a systematic bias towards overestimation and underestimation of the normalized and therefore actual EMCs for constituents with, respectively, a ‘concentration effect’ (e.g., TSS) or a ‘dilution effect’ (e.g., nitrate and DOC) (*Figure 2.16*).

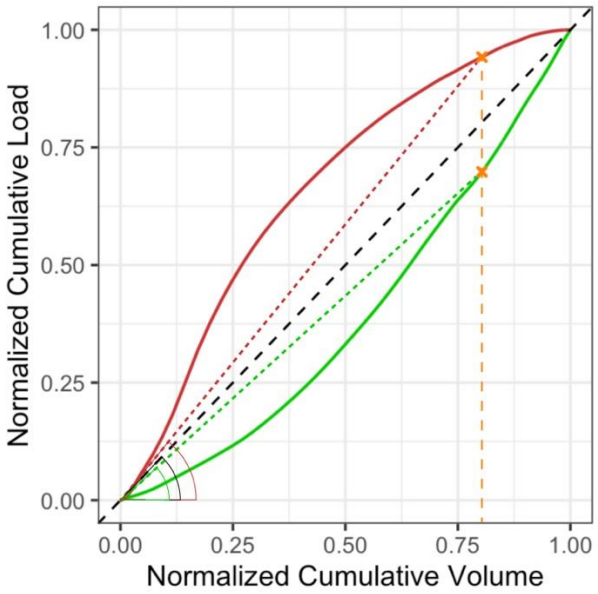


Figure 2.16 Illustration of the over- and underestimation of the normalized EMC for TSS, and, nitrate when only 80% of an event would be sampled. The cumulative curves for TSS, nitrate represent the general behavior observed in Figure 1 6 Actual and apparent EMC

Figure 2.15-right shows the relative errors (grey dots) and the 90% confidence interval (grey area) calculated from the data available across all the event × sampling intervals combinations. These plots correspond to the zoomed in version for percent event sampled greater than 75% at one percentage point resolution. Results show that for nitrate, DOC, and TSS, there was 90% confidence that the errors would be within ±10% when at least 85%, 85%, and 90%, respectively, of the event would be sampled, and very significant errors would occur with lower percentages. For DOC, even with more than 85% of event sampled, the results show a systematic overestimation of +3% on average (*Figure 2.15*d).

For TSS, sampling 85% of event would have resulted in load estimates within a ±20% accuracy (*Figure 2.15*f). Our results show that with the commonly accepted minimum of 60% of the storm sampled (Strecker et al., 2001), this would, in the case of the studied watershed, lead to potentially very large errors (*Table 2.1*). The suggested 60% threshold is based on the TSS load that are thought to often be transported earlier within the event as sediment removal is often the main goal in stormwater water quality treatment. However, our results show that for TSS and probably for all pollutants associated with suspended particles such as total phosphorus, a much higher threshold should be sought. Given the uncertainty variability at different percent event sampled and high uncertainty associated with TSS estimation, it seems appropriate for reports on stormwater water quality to include percent event sampled as an important variable along the reported removal rates.

Table 2.1 Range, mean, and 90% confidence interval (CI) of relative error at each percent event sampled (i.e., percentage of total inflow volume sampled)

Pollutant	Error range	Percent Error for Percent Event Sampled					
		50%	60%	70%	80%	90%	100%
NO_3^-	total range	-37 _ +15	-33 _ 0	-26 _ +10	-21 _ +8	-8 _ +12	-7 _ +8
	90% CI	-36 _ +6	-32 _ 0	-22 _ +5	-12 _ 0	-7 _ +6	-5 _ +5
	mean	-21	-14	-8	-6	-1	1
DOC	total range	-42 _ -8	-30 _ -4	-24 _ 0	-17 _ +12	-9 _ +16	-4 _ +17
	90% CI	-38 _ -11	-28 _ -4	-22 _ 0	-11 _ +11	-2 +12	0 _ +13
	mean	-26	-16	-13	-3	5	5
TSS	total range	-10 _ +70	-65 _ +54	+4 _ +50	-45 _ +28	-54 _ +25	-56 _ +18
	90% CI	-9 _ +65	-55 +33	+8 _ +40	-4 _ +22	-8 +11	-10 +6
	mean	27	24	32	11	3	-2

2.3.3.4 Practical recommendations to choose best sampling volume interval

In light of all our results, it is possible to propose recommendations to calculate optimum sampling volume intervals. Ideally, one would be able to change these intervals to best fit forecasted events. Practically, this would very rarely happen. Instead sampling flow intervals are set once and for all, or sometimes may be changed seasonally.

The three new pieces of information that this work provides is that uncertainties are acceptable (±10%) provided that 1) the first sample not be sampled (most automatic samplers have this feature preprogrammed), 2) more than 90% of an event volume be sampled, and 3) that one can afford to have as low as 30% of the composite bottle filled. Obviously, these results are somewhat contingent upon the peculiarities of the studied watershed, and it is possible that different values would have been recommended from continuous data from another watershed. Nonetheless, we believe most of the general results would hold.

The appropriate sampling interval for a given event, based on the results can be calculated using *Equation 2-7*. The site's runoff volume for a given event can be calculated based on the size and imperviousness of watershed and the size of the event.

Equation 2-7 $$Sampling\ interval = \frac{Runoff\ volume * percent\ event\ to\ be\ sampled}{Number\ of\ samples}$$

To select a sampling interval that can accommodate a range of event sizes for a given site, the following steps are suggested:

- Defining a range of event sizes to be monitored at the site. This range should exclude the extreme events.
- Calculation of the min and max runoff volume for the site.

- Calculation of the appropriate sampling interval to capture 100% of the minimum runoff volume with the minimum number of samples using *Equation 2-7* (this is the proper sampling interval).
- Calculation of the number of samples collected from the maximum runoff volume captured at desired percentage (i.e. 90% as suggested for TSS with ±10% error range) using *Equation 2-7*. This would result in the highest number of samples that should be captured in the composite bottle.
- Calculation of the volume of each sample to be composited by dividing the maximum capacity of composite bottle by the maximum number of samples to be collected (resulted from previous step).

In many instances, extreme rainfall events can be somewhat predicted, like in the case of hurricanes or tropical storms in the southeast US. In these cases, the best approach would be to change the sampling volume intervals to very high values to capture near 100% of the storm. Since there is no practical way to find the sampling intervals that would fit all events all the time, this work can also be used to assess whether the EMC found for all monitored events have a high chance to be stained with errors. The first test is certainly the percentage of the flow sampled. Less than 90% has a very high chance to yield systematically biased EMCs. On the other side of the spectrum, when less than 25% of the bottle would be filled, it would be possible to verify the timing of the sampling with respect to flow. Having a sample very early in the event should lead the operator to question the validity of the EMC obtained.

2.4 Conclusions

Hydrology and water quality of an urban watershed of 91.2 square kilometers with 30% impervious surfaces was studied. Water quality monitoring was performed using minute scale

flow and water quality data, using an *in situ* spectrophotometer for the latter. Use of the spectrometer eliminated the limited sampling capacity issue of autosampler and allowed for high-frequency water quality sampling. Using the spectrophotometer, 68% of events were sampled during monitoring period of 20 months. The resulted high resolution chemographs were numerically sampled to simulate flow-interval composite sampling, measure the uncertainties associated with this sampling method, and derive additional guidelines for stormwater sampling. Flow-interval composite sampling method is considered the most robust sampling method in stormwater (Lee, et al, 2007, Ma et al. 2009). Sampling uncertainty was measured in terms of relative error of event load estimates. The identified sources of uncertainty were (1) the sampling start threshold, (2) the number of samples composited per event, and (3) the percentage of event flow volume sampled.

Our results provide/update three new pieces of information on the generation of uncertainties associated with Flow proportional composite sampling.

1) The first scheduled sample should not be taken (most automatic samplers have this feature preprogrammed). Based on the results, sampling early within the event and specially sampling the small peaks that reside to zero before the main peak of event, can result in high relative error for load estimation of all pollutants. Our simulations show that eliminating the first sample that would be triggered by the automatic sampler dramatically reduces the risk of large uncertainties on EMC and event loads.
2) To achieve an accurate load estimation, a large portion of a storm event volume should be sampled along with a representative number of samples. Exhaustion of autosampler capacity early within an event would result in systematic underestimation of constituents exhibiting dilution effects (e.g., dissolved constituents such as Nitrate and DOC) and

overestimation of constituents exhibiting concentration effects (e.g., particulates and particulate bounds constituents such as TSS). To obtain accurate measurement of Nitrate and DOC within ±5% range and with 90% confidence interval, at least 85% of an event should be sampled. Measurement of TSS load at the same accuracy may not be attainable, however sampling 95% of an event would result in ±10% accuracy with 90% confidence interval.

3) A relatively small percentage of the composite bottle can be filled to obtain satisfactory uncertainties: to obtain uncertainty within ±5%, less than 25% of the bottle capacity is enough for dissolved constituents like nitrate and DOC, and 30% for particulate constituents like TSS. To obtain uncertainties within ±10%, less than 15% and 20% of the bottle capacity is enough for, respectively, dissolved constituents like nitrate and DOC, and for particulate constituents like TSS.

To achieve these levels of accuracy in sampling at a given watershed, the decision on flow-interval should be made based on sampling capacity and predicted amount of runoff. Amount of runoff can be estimated through watershed characteristics and historical rainfall data. To decide on a sampling flow-interval 3 factors should be considered: (1) the smallest event size to be sampled should result in a composite bottle filled to about at least 20% of its volume, (2) the largest event size to be sampled should not exceed the autosampler's capacity, and (3) the weather forecast should be monitored and in case of extreme events, the flow-interval should be adjusted accordingly. Operators could also use our results to assess the chances for significant errors. In particular, we suggest it is best to report the percentage of event sampled as well as the number of samples per event, to account for the potential uncertainties associated with the results in future stormwater studies.

2.5 Acknowledgement

This study was supported by the North Carolina Clean Water Management Trust Fund. Special thanks to the following individuals for their notable support:

Shawn Kennedy and Samuel Garvey for their help with instrument installation and troubleshooting. Alisha Goldstein, the project engineer for her support in conducting the study. Cheng Chen for his help with field data collection.

REFERENCES

Ackerman, D., Stein, E. D., & Ritter, K. J. (2011). Evaluating performance of stormwater sampling approaches using a dynamic watershed model. Environmental monitoring and assessment, 180(1-4), 283-302.

Bach, P. M., McCarthy, D. T., & Deletic, A. (2010). Redefining the stormwater first flush phenomenon. Water Research, 44(8), 2487-2498.

Bertrand-Krajewski, J. L., Chebbo, G., & Saget, A. (1998). Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon. Water research, 32(8), 2341-2356.

Birgand, F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., & Bordenave, P. (2010). Uncertainties in assessing annual nitrate loads and concentration indicators: Part 1. Impact of sampling frequency and load estimation algorithms. Transactions of the ASABE, 53(2), 437-446.

Birgand, F., Aveni-Deforge, K., Smith, B., Maxwell, B., Horstman, M., Gerling, A. B., & Carey, C. C. (2016). First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency in situ water chemistry measurements at multiple sites. Limnology and Oceanography: Methods, 14(12), 767-783.

Brown T., Burd W., Lewis J., and Chang G., 1995. Methods and Procedures in Stormwater Data Collection. In Stormwater NPDES Related Monitoring Needs. ASCE. p. 198

Burton Jr, G. A., & Pitt, R. (2001). Stormwater effects handbook: A toolbox for watershed managers, scientists, and engineers. CRC Press.

Deletic, A. (1998). The first flush load of urban surface runoff. Water research, 32(8), 2462-2470.

Erickson, A. J., Weiss, P. T., & Gulliver, J. S. (2013). Optimizing stormwater treatment practices. A Handbook of Assessment and Maintenance, 1(1), 1-337.

Etheridge, J. R., Birgand, F., Osborne, J. A., Osburn, C. L., Burchell, M. R., & Irving, J. (2014). Using in situ ultraviolet-visual spectroscopy to measure nitrogen, carbon, phosphorus, and suspended solids concentrations at a high frequency in a brackish tidal marsh. Limnology and Oceanography: Methods, 12(1), 10-22.

Etheridge, J. R., Birgand, F., Burchell, M.R., Smith, B. T. 2013. Addressing the fouling of in situ ultraviolet–visual spectrometers used to continuously monitor water quality in brackish tidal marsh waters. J. Environ. Qual, 42: 1896–1901.

King, K. W., & Harmel, R. D. (1998). Considerations in selecting a water quality sampling strategy. In 2001 ASAE Annual Meeting (p. 1). American Society of Agricultural and Biological Engineers.

Haan, C. T. (2002). Statistical methods in hydrology (second version). Iowa State: Blackwell Publishing.

Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., ... & Jarvie, H. P. (2015). High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. Hydrological Processes, 29(15), 3388-3407.

Harmel, R. D., King, K. W., & Slade, R. M. (2003). Automated storm water sampling on small watersheds. Applied Engineering in Agriculture, 19(6), 667.

Harmel, R. D., & King, K. W. (2005). Uncertainty in measured sediment and nutrient flux in runoff from small agricultural watersheds. Transactions of the ASAE, 48(5), 1713-1721.

Harmel, R. D., Slade Jr, R. M., & Haney, R. L. (2010). Impact of sampling techniques on measured stormwater quality data for small streams. Journal of environmental quality, 39(5), 1734-1742.

Hvitved-Jacobsen, T., Vollertsen, J., Nielsen, A.H., 2010. Urban and Highway Stormwater Pollution. Taylor & Francis Inc, 347 pp.

Khan, S., Lau, S. L., Kayhanian, M., & Stenstrom, M. K. (2006). Oil and grease measurement in highway runoff—sampling time and event mean concentrations. Journal of Environmental Engineering, 132(3), 415-422.

Langergraber, G., Fleischmann, N., and Hofstader, F. (2003) A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. Water Sci. Technol., 47: 63–71.

Lee, J.H., Bang, K.W., 2000. Characterization of urban stormwater runoff. Water Research 34, 1772–1780.

Lee, J. H., Bang, K. W., Ketchum Jr, L. H., Choe, J. S., & Yu, M. J. (2002). First flush analysis of urban storm runoff. Science of the Total Environment, 293(1-3), 163-175.

Lee, H., Swamikannu, X., Radulescu, D., Kim, S. J., & Stenstrom, M. K. (2007). Design of stormwater monitoring programs. Water research, 41(18), 4186-4196.

Leecaster, M. K., Schiff, K., & Tiefenthaler, L. L. (2002). Assessment of efficient sampling designs for urban stormwater monitoring. Water research, 36(6), 1556-1564.

Leopold, L. B. (1973). River channel change with time: an example. Geological Society of America Bulletin, 84(6), 1845-1860.

Line, D. E., White, N. M., Osmond, D. L., Jennings, G. D., & Mojonnier, C. B. (2002). Pollutant export from various land uses in the Upper Neuse River Basin. Water Environment Research, 74(1), 100-108.Litton, R. B. (1968). Forest landscape description and inventories-a basis for land planning and design. Res.Paper PSW-RP-09. Albany, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture; 88 P, 49

Liu, W., Birgand, F., Youssef, M., & Chescheir, G. (2016). A novel method to reveal the nitrate transport and fate in agricultural fields. In 2016 10th International Drainage Symposium Conference, 6-9 September 2016, Minneapolis, Minnesota (pp. 1-3). American Society of Agricultural and Biological Engineers.

Ma, J. S., Kang, J. H., Kayhanian, M., & Stenstrom, M. K. (2009). Sampling issues in urban runoff monitoring programs: Composite versus grab. Journal of Environmental Engineering, 135(3), 118-127.

Maestre, A., & Pitt, R. E. (2006). Identification of significant factors affecting stormwater quality using the national stormwater quality database. Journal of Water Management Modeling.

Métadier, M., & Bertrand-Krajewski, J. L. (2012). The use of long-term on-line turbidity measurements for the calculation of urban stormwater pollutant concentrations, loads, pollutographs and intra-event fluxes. Water research, 46(20), 6836-6856.

McCarthy, D.T., 2009. The first flush of E. coli in urban stormwater runoff. Water Science and Technology 60 (11), 2749–2757.

McCarthy, D. T., Zhang, K., Westerlund, C., Viklander, M., Bertrand-Krajewski, J. L., Fletcher, T. D., & Deletic, A. (2018). Assessment of sampling strategies for estimation of site mean concentrations of stormwater pollutants. Water research, 129, 297-304.

Pitt, R., Maestre, A., & Morquecho, R. (2004, February). The national stormwater quality database (NSQD, version 1.1). In 1st Annual Stormwater Management Research Symposium Proceedings (pp. 13-51).

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rieger, L., Langergraber, G., & Siegrist, H. (2006). Uncertainties of spectral in situ measurements in wastewater using different calibration approaches. Water science and technology, 53(12), 187-197.

Sansalone, J. J., & Buchberger, S. G. (1997). Partitioning and first flush of metals in urban roadway storm water. Journal of Environmental engineering, 123(2), 134-143.

Soeur C., Hubka J., Chang G., 1995. Methods for Assessing Urban Storm Water Pollution, in Stormwater NPDES Related Monitoring Needs. ASCE. p. 558.

Stone, K. C., Hunt, P. G., Novak, J. M., Johnson, M. H., & Watts, D. W. (2000). Flow-proportional, time-composited, and grab sample estimation of nitrogen export from an eastern Coastal Plain watershed. Transactions of the ASAE, 43(2), 281.

Strecker, E., Mayo, L., Quigley, M., & Howell, J. (2001). Guidance manual for monitoring highway runoff water quality (No. FHWA-EP-01-022). Federal Highway Administration.

Torres, A., & Bertrand-Krajewski, J. L. (2008). Partial Least Squares local calibration of a UV–visible spectrometer used for in situ measurements of COD and TSS concentrations in urban drainage systems. Water Science and Technology, 57(4), 581-588.

Urbonas, B. R. (1995). Recommended parameters to report with BMP monitoring data. Journal of Water Resources Planning and Management, 121(1), 23-34.

US EPA. (1992). NPDES storm water sampling guidance manual. US EPA office of water. https://www3.epa.gov/npdes/pubs/owm0093.pdf

US EPA. (2002). Urban stormwater BMP performance monitoring. US EPA office of water. https://www3.epa.gov/npdes/pubs/montcomplete.pdf

Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., & Morgan, R. P. (2005). The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society, 24(3), 706-723.

CHAPTER 3: QUANTIFYING THE BENEFITS OF REAL-TIME CONTROL OF A WET POND AND A WETLAND IN NORTH CAROLINA

Abstract

Real-time control (RTC) of stormwater control measures (SCMs) can increase the storage capacity of existing SCMs, resulting in a potential increase of retention time leading to hydrologic and water quality benefits. Additionally, some studies suggest RTC as a solution to equip existing SCMs for changing precipitation patterns and unprecedented extreme events. This research, in collaboration with OptiRTC, examines the performance of RTC on a constructed stormwater wetland (CSW) and a wet pond in North Carolina. These two SCMs were retrofitted with RTC in predictive mode, with performance being evaluated by comparing hydrology and water quality before and after the retrofit. RTC implementation increased the number of no outflow events for smaller storms (accounting for less than 16% of total inflow volume) while increasing the number of higher volume events. Cumulative volume reduction of both SCMs remained the same pre- and post-retrofit. RTC underestimated the incoming event's volume for 86% of total inflow volume at the CSW and 65% at the wet pond, resulting in intra-event overflow. Mean peak outflow rates remained fairly the same. Retention time post-retrofit (1) decreased at the CSW, due to occurrence of overflow, and (2) increased by a factor of 4 at the wet pond. Both SCMs had higher than average volume and pollutant reduction pre-retrofit. Water quality results were mixed, reflecting RTC's impacts on retention time. (1) The CSW was unaffected by RTC implementation, (2) while RTC improved pollutant retention of the wet pond by 29%, 17%, and 7% for TSS, DOC, and PO_4^{3-}-P respectively.

3.1 Introduction

Urbanization and expansion of impervious surfaces has led to increased runoff volume and flooding (Leopold 1973; Hollis 1975), peak flow (Leopold 1973; Leopold 1994; Jennings and Jarnagin 2002), and pollutant loads (Trimble 1997; Line et al. 2002) in urban streams and receiving waters (Walsh et al. 2005; Walsh 2000). Stormwater wetlands and wet ponds are two popular stormwater control measures (SCMs) that mitigate volume and peak flow and offer water quality improvements through the retention of runoff and slow release of it (MDE 2009; PADEP 2006). This slow discharge of water is normally passively-released through an orifice in the outlet structure and is driven by intra-event water level pressure head (NCDEQ 2018; *Figure 3.1*). The outlet is designed to provide a release rate that ensures available volume in the SCM to capture the majority of annual runoff (NCDENR 2009; MDE 2009), while providing retention time for water quality treatment (NCDEQ 2018). Research shows that increase in retention time can improve water quality (Papa et al. 1999; Smolek et al. 2015) by (1) encouraging sedimentation and concomitant removal of phosphorus and heavy metals with fine particles (Pettersson 2002; Vaze & Chiew 2004), (2) increasing UV exposure (Vergeynst et al. 2012) and (3) extending plant contact that can encourage biological mechanisms and nutrient uptake (Greenway 2004). An important factor in retention time is SCM storage capacity (Walker, 1998). However, with passive outflow, increasing the retention time requires elongating the time water is stored in the basin, which reduces the SCM's capacity to capture subsequent rain events (Guo, 2002). Use of active outflow potentially increases the retention time without sacrificing the hydraulic capacity of the SCM (Marsalek 2005; Gaborit et al. 2016).

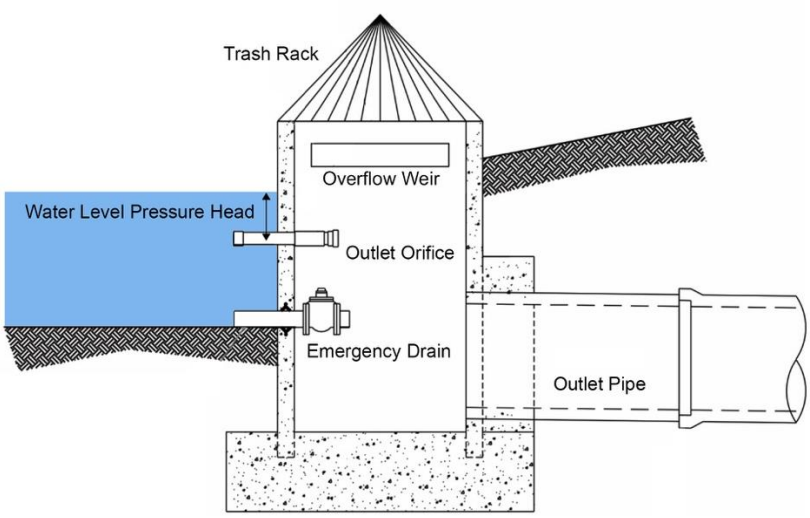


Figure 3.1 Outlet structure of wetland/wet pond with passive water release through an orifice, modified from NCDEQ 2018

Active outflow is the controlled (manual or automatic) release of water from a basin with the goal of increasing its hydraulic retention time and water quality treatment capacity. It entails a controlled closure of the outlet with a valve or gate to retain the captured runoff in the SCM for extended periods and then release the water when pre-defined critical levels are reached (e.g. overflow level; Gaborit et al. 2013). The release of outflow can be either reactive, in response to a received storm event (existing condition) or predictive, in response to a predicted runoff volume (future condition) (Gaborit et al. 2016). In the predictive approach, the water level in the SCM is monitored to calculate the available storage capacity for the predicted event in real-time and, if required, advanced release of some water to provide more capacity (Gaborit et al. 2016; Kerkez, et al. 2016). Active outflow has been studied and implemented under different names such as active control, dynamic control, adaptively controlled, real-time control (RTC), and continuous monitoring and adaptive control (CMAC).

RTC has been widely used in management of wastewater and combined sewer systems (Schütze 2004; Pleau et al. 2005; Stinson & Vitasovic 2006). The use of RTC in stormwater management has been gaining attention due to potential hydrologic and water quality benefits

resulting from the increased storage capacity. Some studies focused on the use of RTC as a method to adapt current SCMs to the changing precipitation patterns and unprecedented extreme events (Quigley et al., 2008; Kerkez, et al. 2016; Lefkowitz, et al. 2016; Shishegar et al. 2018). SCM size is based on the historic precipitation data. However, studies suggest rainfall frequency and intensity are changing (Alexander et al., 2006; Kunkel et al., 2013; Kendon et al., 2014; Mann et al., 2017; Wuebbles et al. 2017). Perhaps SCM design should account for this change (Rosenberg et al., 2010)? Active control outflow has the potential to increase an SCM's capacity without changing the SCM's size.

Table 3.1 Summary of studies on the active outflow, including SCM type and location, method of comparison, and duration of the study

SCM type	**Author/s**	**Location**	**Method**	**Study Duration**
Wet retention pond	Muschalla et al., 2014	Quebec City, Canada	Passive-site monitoring Active-simulation	2 months (Jul-Aug 2008)
	Strauss and Lefkowitz, 2017	Montgomery County, MD	Passive-conventional data Active-site monitoring	11 months (Jan-Nov 2016) 5 events
	Gilpin and Barrett., 2014	Travis county, TX	Passive-site monitoring * Active-site monitoring	8 months (Mar-Oct 2013) 10 events (2 paired)
Dry detention pond	Gaborit et al., 2013	Quebec City, Canada	Passive- site monitoring Active-simulation	6 consecutive summers (2010)
	Middleton and Barrett, 2008	Austin, TX	Passive-conventional data Active-site monitoring	10 months (Jul-May 2006) 13 events
	Jacopin et al., 2001	France	Passive-site monitoring * Active-site monitoring	1 year 8 events
	Marchese, 2018	Beckley, WV	Passive-simulation Active-site monitoring	6 months
	Carpenter et al., 2013	Quebec City, Canada	Passive-site monitoring * Active-site monitoring	6 consecutive summers (2010) 15 events
	Schmitt et al., 2020	Blacksburg, VA	Passive-simulation Active-simulation	15 years of precipitation data
	Bilodeau et al., 2018	Quebec City, Canada	Passive-simulation Active-simulation	6 months 4 events
Rainwater harvesting	Gee and Hunt, 2016	Craven county, NC	Passive-site monitoring * Active-site monitoring	Passive-15months, 90 events Active-14 months, 51 events
	Roman et al., 2017	New York City, NY	Passive- conventional data Active-simulation	10 years of precipitation data 652 events
GI treatment train	Lewellyn and Wadzuk, 2017	Villanova, PA	Passive-site and simulation Active-site and simulation	2 years 201 events

*Passive and active control studies on two different sites

Active outflow has been evaluated for different SCM types such as dry ponds (Jacopin et al., 2001; Middleton and Barrett, 2008; Carpenter et al., 2013; Gaborit et al., 2013; Bilodeau et

al., 2018; Marchese, 2018; Schmitt et al. 2020), wet ponds (Gilpin and Barrett., 2014; Muschalla et al., 2014; Strauss and Lefkowitz, 2017), rainwater harvesting (Gee and Hunt, 2016; Roman et al., 2017), and a treatment train of different SCMs (Lewellyn and Wadzuk, 2017). *Table 3.1* provides a summary of previous studies of active outflow on different SCM types and locations, including their evaluation method and duration of study. These studies have evaluated reactive and predictive approaches as well as manual or digital control, through both simulation or pilot studies, and all have suggested improved volume reduction, peak attenuation (*Table 3.2*), and pollutant removal efficiency (*Table 3.3*) with the use of active outflow. Wet pond-focused studies have suggested flow peak attenuation of 50% (Muschalla et al., 2014) and volume reductions of 67% (Strauss and Lefkowitz, 2017). However, neither study reported expected performance, had the ponds been passively controlled. Wet pond studies further reported TSS removal efficiency would improve by 8-12% (Muschalla et al., 2014), 13-16% (Strauss and Lefkowitz, 2017), and 27% (Gilpin and Barrett., 2014; *Table 3.3*) with implementation of RTC. Nitrate-nitrite could improve as well ([8-32% (Strauss and Lefkowitz, 2017) or 94% (Gilpin and Barrett., 2014)], *Table 3.3*). In these studies water quality improved even with different retention times associated with different RTC configurations. Tested RTC configurations included 24-hour retention time (Gilpin and Barrett., 2014), 48-hour retention time (Strauss and Lefkowitz, 2017) and maximum retention time unless exceeding 80% of the aquatic life span of mosquitoes (Muschalla et al., 2014). The limited number of observed events (<5) is a shortcoming of these studies (*Table 3.1*).

Table 3.2 Summary of hydrologic benefits of active outflow reported in the literature

Author/s	**Configuration**	**Hydrologic benefit**
Muschalla et al., 2014	Maximum retention time unless exceeding 80% of the aquatic life span of mosquitoes	50% peak attenuation
Strauss and Lefkowitz, 2017	Increased retention time from 24 hr to 48 hr	67% volume reduction
Jacopin et al., 2001	On/off system set to zero outflow during rain	Increased max storage capacity site1: 17% to 34% site2: 43% to 50%
Schmitt et al., 2020	Reactive strategy to maximize detention time, avoid overflow, provide smooth drawdown	Peak attenuation for small to moderate events, (return period < 2 years) Maintained or increased peak flow for larger storm rates.
Bilodeau et al., 2018	Increase detention time an average of 36 hrs	46% peak attenuation 22% less frequent use of the downstream collector
Gee and Hunt, 2016	Predictive strategy to minimize overflow during an event	91% volume reduction 93% peak attenuation
Roman et al., 2017	Predictive strategy to minimize overflow during an event	Increased max storage capacity 76.6%
Lewellyn and Wadzuk, 2017	If no rainfall, pump water from the cistern to the treatment train	33% overflow volume reduction

The expansion of RTC into stormwater requires further evaluation of RTC performance and implementation in different locations and on different types of SCMs. In collaboration with OptiRTC ® (Opti Boston, MA, USA), this study has evaluated the company's proprietary RTC system on a stormwater wetland and a wet pond in Raleigh, North Carolina. The research objectives are: (1) quantifying the hydrologic and water quality benefits of using RTC in a stormwater wetland and wet pond in NC, and (2) providing suggestions for the future design of these SCMs vis-à-vis RTC to optimize performance. This study is the first to evaluate the RTC on a stormwater wetland.

Table 3.3 Summary of water quality benefits of active outflow reported in the literature

Author/s	**Configuration**	**Removal-passive**	**Removal-active**
Muschalla et al., 2014	Maximum retention time unless exceeding 80% of the aquatic life span of mosquitos	68.5-86.0 % TSS	additional 8-12% TSS removal
Strauss and Lefkowitz, 2017	Increased retention time from 24 hr. to 48 hr.	20-36% NO3-N 40-72% TSS	28-68% NO3-N 53-88% TSS
Gilpin and Barrett., 2014	24 hr. retention	39% E.coli 0% NO3-N 71% TSS	88% E.coli 94% NO3-N 98% TSS
Gaborit et al., 2013	Max detention time, avoid overflow, slow drawdown, limit retention to 3-4 days to avoid mosquitos.	46% TSS	90% TSS
Middleton and Barrett, 2008	12 hr. retention time, 12 hr. drawdown time	75% TSS conventional removals were lower than active outflow results	91% TSS 58% NO3-N, NO2-N 35% TKN 55% T-Copper 69% T-Lead 62% T-Zinc 34% COD 52% TP
Jacopin et al., 2001	On/off system set to zero outflow during rain	6% TSS 1 14% TSS 2	47% TSS 1 57% TSS 2
Marchese, 2018	48 hr. retention time, 12 hr. drawdown time No outflow during an event, unless to avoid overflow.	0% TN* 0% TP* 0% TSS*	31% TN 48% TP 62% TSS
Carpenter et al., 2013	Max detention time, avoid negative effects of overflow and particle resuspension	39% TSS 10% NH3-N 20% Zinc	90% TSS 84% NH3-N 42% Zinc

*These low removal values are potentially due to the reported low residence time (1hr) for the simulated passive control.

3.2 Methods

3.2.1 Site description

A wet pond and constructed stormwater wetland (CSW) located on the NC State campus were selected for instrumentation and monitoring (*Table 3.4*). The CSW treated runoff from a 9.12-ha, 30% impervious watershed that drained a wooded area, rooftop, landscape, and roadway

(*Table 3.2*). The CSW surface area at permanent pool elevation was 0.13ha, with a forebay comprising ~13% of this area. The outlet structure was a 1.7-m square riser with a 50-mm diameter drawdown orifice located 530mm below the overflow weir. The emergency drain was a 150-mm diameter orifice with a valve placed at the outlet structure side (*Figure 3.9*d).

Table 3.4 Characteristics of the studied wet pond and CSW, location, watershed characteristics, and design features

	CSW	Wet pond
Location (Latitude and longitude)	35°45'25.5"N, 78°41'15.4"W	35°46'33.8"N, 78°40'51.5"W
Year constructed	2001	2004
Watershed area (ha)	9.12	1.7
Watershed land use	Wooded area, rooftop, landscape, road	Rooftop, parking lot, landscape
Watershed imperviousness (%)	30	83
Permanent pool surface area (ha)	0.13	0.05
Forebay area (m^2)	157	93
Storage volume at permanent pool elevation (m^3)	408	437
Storage volume at temporary pool elevation (m^3)	1158	760
Temporary pool depth (m)	0.53	0.52
Outlet structure	1.7 m square riser	1.6 m square riser decreasing to 1 m at the top
Drawdown orifice diameter (mm)	50	25
Emergency drain diameter (mm)	150	150

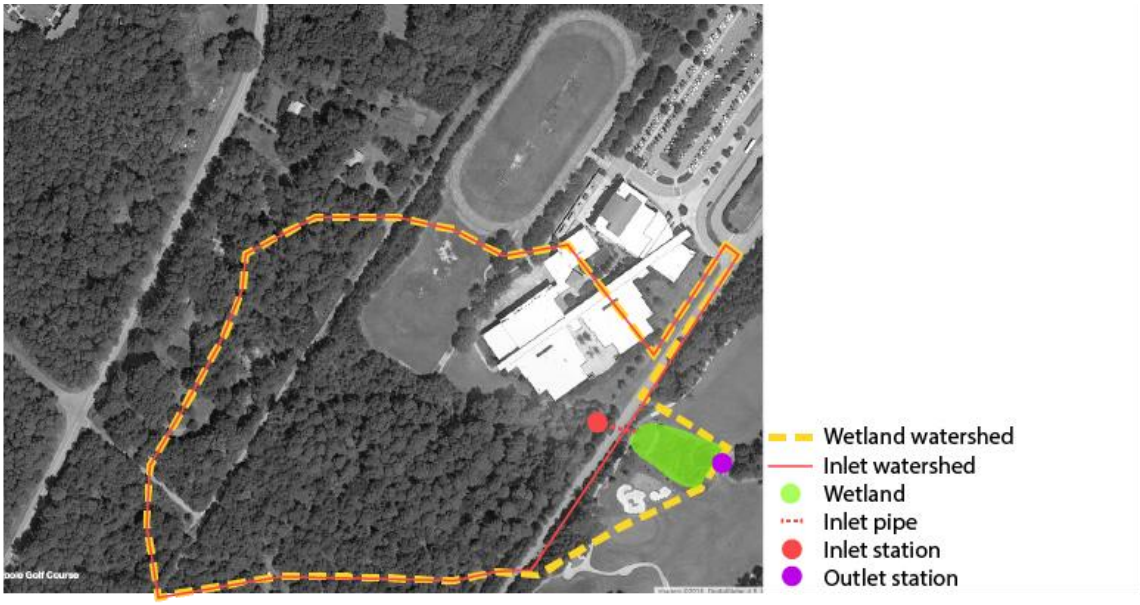


Figure 3.2 Watershed delineation (dotted line) and surface area of the CSW (highlighted in green)

The wet pond treated runoff from a 1.7-ha, 83% impervious watershed that drained a rooftop, parking lot, and landscape (Figure 3.3). Most (89%) runoff flows from a rooftop and parking lot, while the latter is pretreated by a swale. A fraction (11%) of runoff flows from a parking lot and surrounding landscape that passed through a CSW before entering the pond (Figure 3.3). The pond surface area at permanent pool elevation was 0.05ha, with a forebay comprising ~20% of this area. The outlet structure was a 1.6-m square riser gradually decreasing to 1-m at the top with a 25-mm diameter drawdown orifice located 520mm below the overflow weir. The emergency drain was a 150-mm diameter orifice with a valve placed at the pond side (*Table 3.16*b).

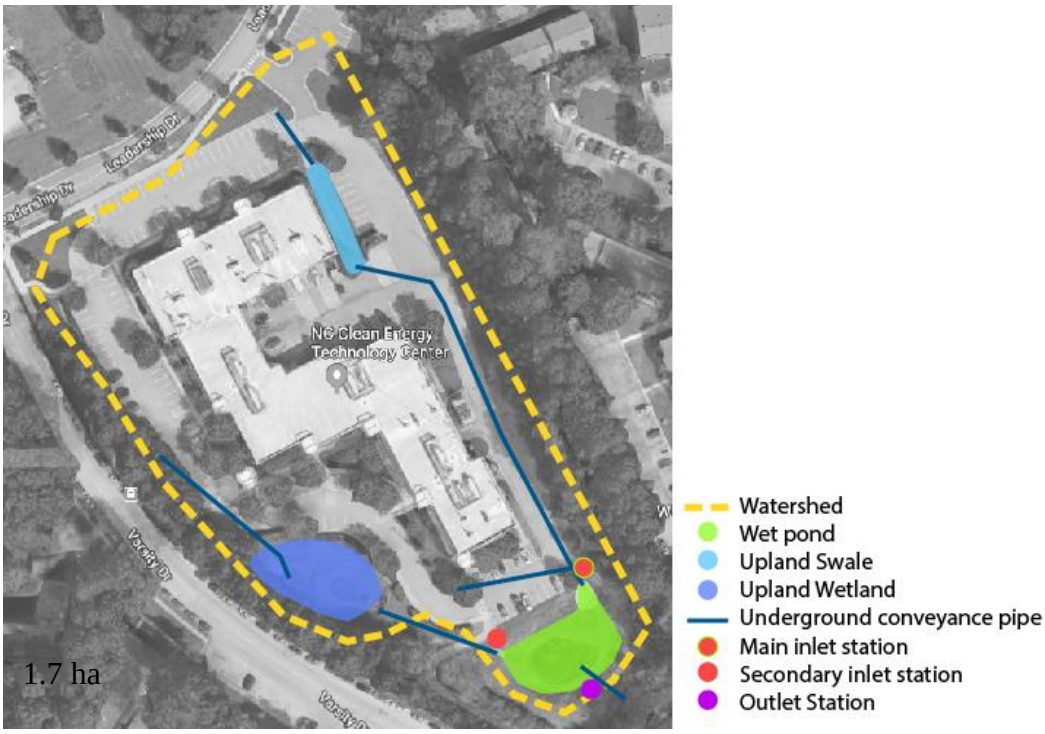


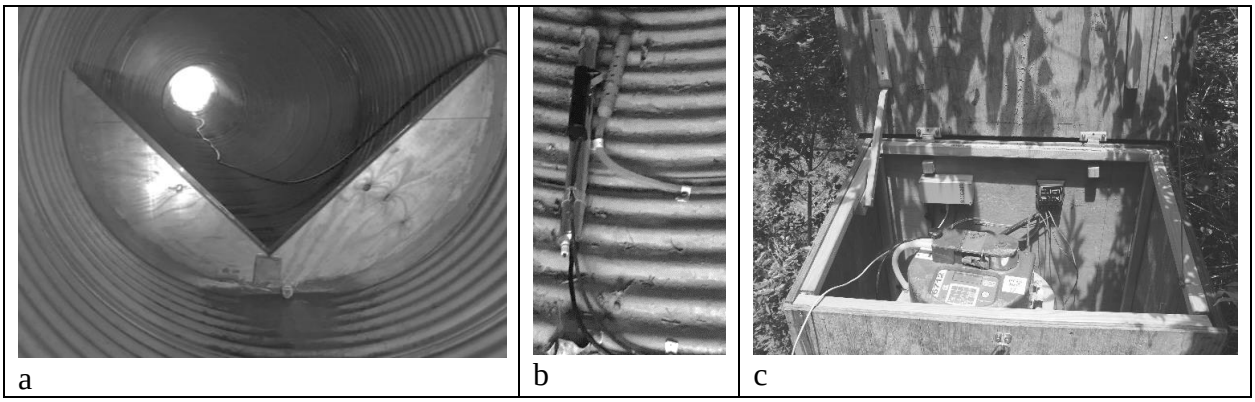
Figure 3.3 Watershed delineation (dotted line) and surface area of the wet pond (highlighted in green), including the upland pre-treatment SCMs

3.2.2 Flow and water quality monitoring

3.2.2.1 CSW

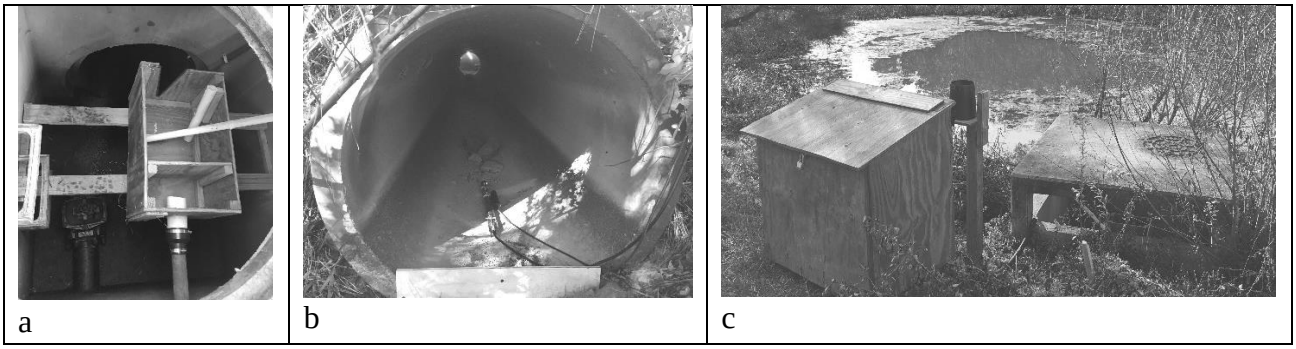
Flow and water quality of the wetland were monitored for 12 months (December 2017 – December 2018), and the effect of RTC retrofit was monitored for 7 months (January – August 2019). Inflow passed through an underground 1.52-m corrugated reinforced concrete pipe (RCP; *Figure 3.2*). A 90° V-notch wooden weir was placed at the entry of this pipe, and flow rate was calculated using the upstream flow depth using a "Teledyne ISCO® Bubbler Module" model 730, at 2-minute intervals, mounted on a 6712 ISCO autosampler (*Figure 3.4*). Data were collected on 2-minute intervals. At the outlet the flow rate was calculated using two complementary methods. Flow depth and velocity were first measured using an area velocity meter (AVM) "Teledyne ISCO AVM Module" model 750, mounted on a 6712 ISCO autosampler, placed at the entry of a 0.9-m diameter outlet RCP. The AVM was placed in the

0.91m RCP outlet pipe. Secondly, low flow depths were again measured using a pressure transducer (HOBO U20L-04™) placed inside a 22.5° v-notch wooden weir box installed at the outlet orifice (*Figure 3.5*). AVMs ideally measure higher flows, but often fail to accurately measure low flows (also observed by Middleton and Barrett, 2008). Thus, the need for pressure transducer-measured low flows. Pressure transducer flow measurement was only employed pre-retrofit, because once the SCM was retrofitted, the outlet orifice was closed. The measured flow rate was used by the autosampler to calculate cumulative flow volume and trigger water quality sampling.



a b c

Figure 3.4 Inlet monitoring configuration -weir (left), autosampler intake and spectrometer (middle), autosampler and spectrophotometer control (right)



a b c

Figure 3.5 Outlet monitoring configuration – weir box and pressure transducer (left), spectrophotometer (middle), autosampler box (right)

Water quality was measured at 4-minute intervals using *in-situ* UV-Vis spectrophotometer, Spectro::lyser UV produced by S::CAN® Messtechnik GmbH with an optical

path length of 5mm and measuring light absorbance over the 200-730 nm range. These sensors allowed for water quality monitoring of any size event. Therefore, water quality was obtained for all events with a minimum of 1 mm rainfall. To prevent blockage of optical path length by debris, and minimize fouling on the optics, the spectro::lyser probes were fitted with the "ruck::sack" automated brush (from the same manufacturer) activated before each measurement (*Figure 3.4*b). Additionally, the spectrophotometer was cleaned during each site visit using deionized water and diluted acid (HCL 2%) (Ethridge, et al. 2013). At the inlet, the spectrophotometer was placed behind the weir beside the ISCO intake such that it was submerged in presence of flow (*Figure 3.4*b). At the outlet the spectrophotometer was placed at the end of the outlet pipe for maintenance access purposes, behind a broad crested weir (*Figure 3.5*b).

For local calibration of the spectrometers, discrete water quality samples were obtained using a Teledyne ISCO portable autosampler model 6712 with the 24-bottle configuration. For best calibration, it is recommended to cover the entire range of observed concentrations (low and high values) (Rieger et al., 2006). To sample along as large a concentration range as possible, a focus was made to obtain samples at high peak flows. A detailed description of spectrophotometer operation, selection of local samples for calibration, and calibration process were presented in Chapter 1.

The water quality samples were collected from the site, within 24 hours of the previous event's conclusion. Samples were transferred to the research refrigerator that was 4.3km from the project site. From the 1-liter discrete sample bottle, a 300-mL plastic bottle was filled for total suspended solids (TSS) analysis, and the rest of the sample was filtered twice through a 12µm and a 0.22µm Fisher Scientific® filter into a 60-mL glass bottle to analyze the remaining

pollutants. Analyzed pollutants included Nitrate (NO_3^--N), Ammonium (NH_4^+-N), Total Dissolved Nitrogen (TDN), Total Suspended Solids (TSS), Orthophosphate (PO_4^{3-}), and Dissolved Organic Carbon (DOC). Samples were tested at the Environmental Analysis Laboratory at NC State University, using the analysis methods provided in *Table 3.5*. Among the monitored pollutants, nitrate, TSS, and DOC absorb light and therefore are measurable by spectrophotometers (s::can Messtechnik GmbH 2019). Other studies were able to estimate NH_4^+-N and PO_4^{3-}-P through spectral data, potentially due to co-variability of the color of water with the pollutants (Etheridge et al., 2015). Although, in this research, such co-variability was not observed (Chapter 1), and those pollutants were not included in the data analysis. All instruments were powered by 12-volts batteries charged by solar panels (10W for the ISCO and 100W for the spectrophotometer).

Table 3.5 Laboratory analysis methods

Parameter	Analysis Method
NO_3^--N	Standard Methods 4500-NO3-E or EPA Method 353.2
NH_4^+-N	EPA Method 351.2
DOC	EPA Method 415.1 with Teledyne Tekmar Apollo 9000, 0.45 µm filter
TSS	Standard Methods 2540D or EPA 160.2
PO_4^{3-}-P	Standard Method 4500-P F or EPA Method 365.1

3.2.2.2 Wet pond

Due to delays in selecting the wet pond, the pre-retrofit monitoring period was shorter than that of the CSW. Flow and water quality of the wet pond were monitored for approximately three months (October 2018-January 2019). During pre-retrofit monitoring, the RTC system was installed. The effect of this retrofit was monitored for 8 months (January - August 2019). Monitoring stations were installed at the two inlets and the outlet. A 90° v-notch wooden weir was placed at the entry of the 0.6-m RCP discharging the main inflow (*Figure 3.6* left). At the

outlet, a compound weir (15-cm tall, 45° v-notch lower portion and 61-cm wide rectangular upper portion) was placed inside the outlet structure and before the outlet pipe, covered from the top with a baffle to direct the overflow behind this weir (*Figure 3.7*). At the main inlet and the outlet, flow rate was calculated by measuring the water level just upstream of the weir by Teledyne ISCO® Bubbler Module model 730, mounted on a 6712 ISCO autosampler. Secondary inflow passed through a 0.46-m RCP (*Figure 3.6* b). A 60° v-notch wooden weir box was placed at the end of this pipe, and flow rate was calculated using the depth of flow upstream of this weir as measured by a pressure transducer (HOBO U20L-04™; *Figure 3.6* b). All water level data was collected on 2-minute intervals.

Water quality was collected as flow proportional composite samples using the ISCO autosampler with a 20-liter bottle. The autosampler was configured to collect representative samples from rainfall depths ranging from 7mm through 28mm. A representative composite sample included samples from at least 90% of event duration with a minimum of 10 samples per event resulting in load estimation with a ±20% error range (Chapter 2). Within 24 hours of an event's cessation, the composite sample was transferred to the research refrigerator, 1.3km from the project site, treated, and analyzed using the same procedures explained in the CSW section.

a b

Figure 3.6 Monitoring configuration at the main pond inlet [weir placed at the entry of pipe(a)] and the secondary inlet [Weir placed at the end of pipe draining the upland wetland (b)]

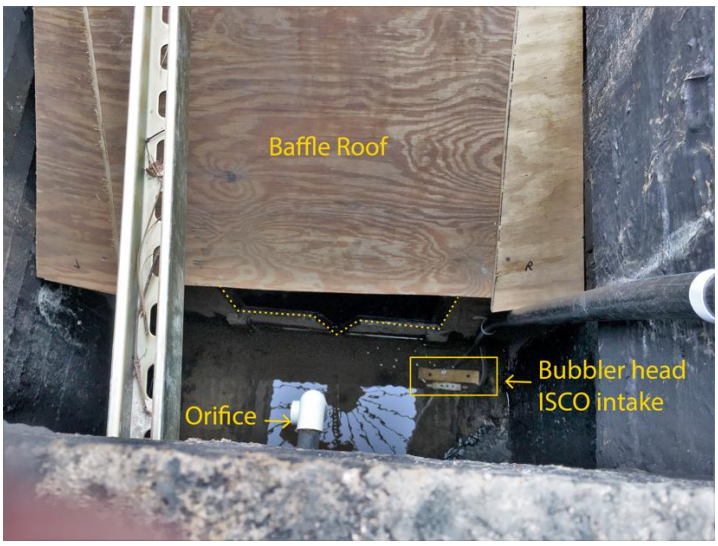


Figure 3.7 Monitoring set up at the outlet: a composite weir covered by a baffle to direct overflow behind the weir

3.2.3 Real-time control set up and configuration

The real-time control system used in this study was the proprietary package, OptiNimbus offered by OptiRTC, Inc. This system includes a water level sensor (Levelgage, Keller America Inc), a rain gauge, an electronically actuated butterfly valve (P series, ProMation Engineering, Inc), solar power kit (24V), and a control panel (OptiThunder) connected to the Opti's cloud-based software (OptiNimbus) through cellular data (*Figure 3.9*, *Figure 3.10*). Using the design

plans and a land survey of each SCM, a stage-storage relationship was developed (*Table 3.4*). The water level sensor submerged in the pond collected the stage data, which was calibrated using manual site measurements (*Figure 3.9* e,f). Using this information, the RTC system calculated the storage capacity of the SCM in real-time. The RTC system was studied in predictive mode. Therefore, the precipitation forecast obtained from the National Weather Service, coupled with watershed characteristics (*Table 3.4*) was used by the RTC to calculate the incoming runoff volume. The RTC was configured to respond to a 24-hour precipitation forecast that had (1) a minimum occurrence probability of 70% and (2) a depth that exceeded 1.27 mm (*Table 3.6*). Based on the anticipated runoff volume and the storage capacity of the SCM, the RTC would release water through the actuated valve, to provide storage capacity and fully capture the incoming event's runoff volume. To regulate the water level in the basin, the RTC identified the weather status (wet weather, dry weather, pre-event) and controlled the valve based on the selected critical water level associated with its weather status (*Table 3.7*). All the collected data were accessible online through OptiRTC's web dashboard.

Table 3.6 RTC configurations for weather forecast

Control parameter	**Description**	**Value**	**Unit**
Forecast Duration	how far into the future to look at the weather forecast	24	hr
Forecast Buffer (after storm event)	how long after the forecast shows no precipitation, to wait before believing it is dry weather	6	hr
Forecast Probability	The threshold over which to believe forecast precipitation	70	%
Qualified precipitation	Minimum forecasted precipitation that causes the RTC to react	1.27	mm

Table 3.7 Description of weather status and critical water levels used for RTC configuration

Weather status and critical water levels	**Description**
Wet weather	Precipitation is detected by the rain gauge
Wet weather target water level	Avoid an increase of water level above this value, during wet weather
Dry weather	6 hours after no precipitation is detected by the rain gauge
Dry weather target water level	Maintain a wet pond at this level, during dry weather
Pre-event	An event with a minimum probability of 0.7 and minimum depth of 1.27 mm is forecasted
Pre-event water level	Limit of drawdown to provide storage capacity for incoming event

In this study both SCMs were retrofitted with an actuated valve installed at the emergency drain (*Figure 3.8*). The valve was placed inside the outlet structure of the CSW (*Figure 3.9*), and outside of the outlet structure of the wet pond (*Figure 3.10*). A perforated upturn elbow with a trash rack was placed at the pond side of the drain of both SCMs to prevent clogging (*Figure 3.10* right). The invert of this elbow was 0.45 m (18 in) higher than the bottom of the SCM. This elevation was considered as a failsafe, to prevent complete drainage of the SCM in case of system malfunction. The selected RTC configurations that increased retention time and selected critical water levels for each SCM are explained in the following paragraphs.

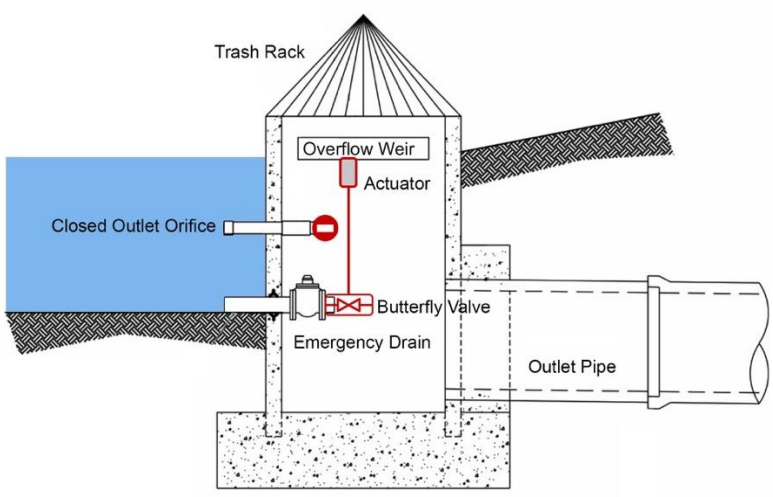


Figure 3.8 Outlet structure after the RTC retrofit with closed orifice and an actuated butterfly valve installed on the emergency drain, modified from NCDEQ 2018

Figure 3.9 Real-time control configuration at wetland, solar panel (a), Opti control panel, power box, and rain gauge (b), valve actuator (c), butterfly valve (d), water level sensor (e,f)

a

b

Figure 3.10 Real-time control configuration at wet pond, Opti control panel and power box connected to a solar panel (left), actuated valve and upturn elbow (right)

3.2.3.1 CSW

Vegetation has a critical role in wetlands' pollutant removal. Therefore, the retention times and critical water levels were selected to maintain the integrity of the plants. Wetland plants can tolerate water fluctuations and submergence for short periods, require saturated soil to grow and, after maturing, can tolerate further dry and wet conditions (personal conversation with Mellow Marsh Farm, Inc. on Feb 19, 2019). Three critical water levels (wet weather target, dry weather target, and pre-event) and their descriptions are provided in (*Table 3.7*). During the growth season (Spring), the captured runoff volume was retained for 3 days after which was discharged until the normal pool level was reached (orifice level). This provided a saturated condition with low submergence to encourage growth. After the plants fully matured (during summer and fall), the post-event retention time was maintained at 3 days but, the dry weather water level elevated (7.6-cm (3in) above the normal pool level). This modestly elevated dry weather water level compensated for water lost through evapotranspiration, which helped avoid plant stress during warm dry seasons. After plant senescence in fall/winter, the dry weather target

water level was set equal to the wet weather target. Thus, providing maximum retention time by maintaining the valve closed until a forecasted event for which there was not enough storage in the CSW. A detailed algorithm of the CSW RTC is provided in *Table 3.8* and illustrated in *Figure 3.11*. The selected wet weather target levels were 5 cm (2in) below the overflow weir (*Table 3.8*). This selection was made to limit overflow and allow measurement of all outflow based on the RTC valve. Finally, drawdown time was set to 48 hours, and the minimum water level to draw down was set to the bottom of the shallow pool.

The seasonal dry/wet weather target water levels were selected based on (1) the condition of the SCM as a matured wetland with vegetation at an equilibrium, and (2) the climate of the site. Selection of proper RTC configuration for a given wetland requires consultation with plant specialists to identify tolerable water levels, duration of inundation, and the plant growing season (Casanova & Brock 2000).

Table 3.8 RTC configurations of retention time, drawdown, and critical water levels for the CSW during different seasons

Control Parameter	**Configuration**		
Duration	Spring	Summer/Fall	Winter
	Apr 1st- Jun 1st	Jun 1st- Nov 1st	Nov 1st- Apr 1st
Retention time	3 days	3 days	Maximum-until next event
Wet weather target water level	5 cm (2 in) below the overflow	5 cm (2 in) below the overflow	5 cm (2 in) below the overflow
Dry weather target water level	current orifice level	7.6 cm (3 in) above orifice level	5 cm (2 in) below the overflow (or maintain the valve closed while in dry weather)
Pre-event water level to provide storage capacity	bottom of the shallow pool	bottom of the shallow pool	bottom of the shallow pool
Post-retention Drawdown time	48 hrs	48 hrs	48 hrs

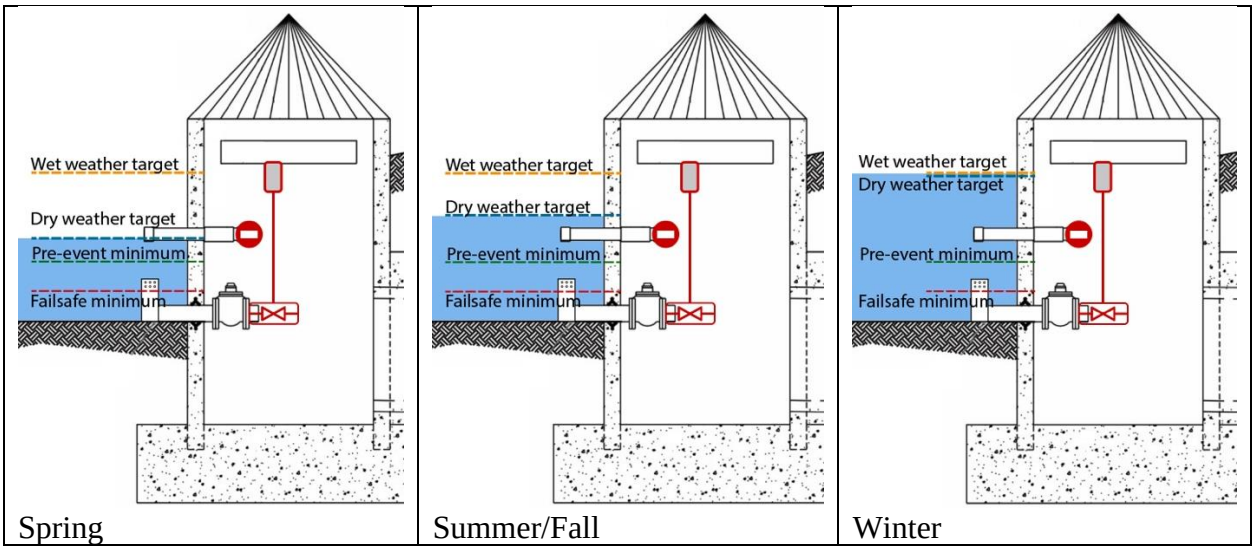


Figure 3.11 CSW critical water levels for different seasons, modified from NCDEQ 2018

3.2.3.2 Wet pond

Wet ponds have limited vegetation, therefore increasing the retention time at any time was possible without harming SCM function. Hence, the RTC was set to provide the maximum possible retention time. The RTC configurations (1) allowed the water level during wetter seasons to elevate within 5 cm (2 in) of the overflow, and (2) maintained water in the SCM during dry weather (6 hr. post-precipitation) unless an event was forecasted for which the runoff volume exceeded the available capacity of the pond. The drawdown time to release water was set to 48 hours, and the pre-event drawdown level to provide storage capacity was set to the valve invert [45cm (18 in) from the bottom].

Table 3.9 Wet pond RTC configurations for retention time, drawdown, and critical water levels

Control Parameter	Configuration
Retention time	Maximum-until next event
Wet weather target water level	5 cm (2 in) below the overflow
Dry weather target water level	5 cm (2 in) below the overflow (or maintain the valve closed while in dry weather)
Pre-event water level to provide storage capacity	Valve invert [or failsafe level of 45cm (18 in) above the bottom]
Post-retention Drawdown time	48 hrs

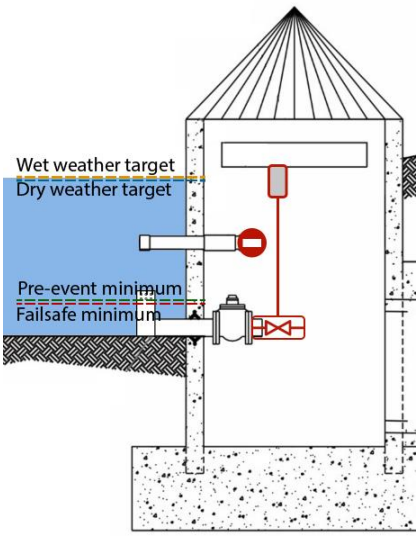


Figure 3.12 Wet pond critical water levels, modified from NCDEQ 2018

3.2.4 Data analysis

An individual storm event was defined as an inflow event that occurred following at least 6 hours of dry weather. Inflow depth was calculated by dividing the cumulative volume of an individual event (m^3) by the watershed area (m^2). Events that occurred separately based on inflow but had overlapping (or continuous) outflow were assumed as one. After the RTC retrofit, the outflow occurring before and during an event was associated with that event. Because the RTC was set on predictive mode, the volume of water released from the SCM was in response to a forecasted event. This definition was used for the analysis of volume, peak flow and pollutant

reduction. Retention times were calculated as the time difference (hr) between the inflow and the outflow centroids (Hancock et al. 2010).

Weir equations were used to calculate flow rate based on the water level above the weir (Grant and Dawson 2001):

Equation 3-1 $Q = 1838 * H^{2.5}$, 90° v-notch weir

Equation 3-2 $Q = 796.7 * H^{2.5}$, 60° v-notch weir

Equation 3-3 $Q = 571.5 * H^{2.5}$, 45° v-notch weir

Equation 3-4 $Q = 274.4 * H^{2.5}$, 22.5° v-notch weir

Equation 3-5 $Q = 1838(L - 0.2L)H^{1.5}$, *rectangular weir*

Where Q is flow rate (L/s), H is the head of water approaching (or 1 m upstream of) the weir (m) and L is the rectangular weir length (m). Instantaneous flow rates were calculated based on the 2-minute interval level data, and the cumulative event volume was calculated using *Equation 3-6*.

Equation 3-6 $$V_{event} = \sum_{i=1}^{i=n-1} \left[\left(\frac{Q_i + Q_{i+1}}{2} \right) * (t_{i+1} - t_i) \right]$$

Where:

V: event volume (L)

Q_i: Instantaneous flow (L/s) at time stamp t_i

n: total number of data points

$t_{i+1} - t_i$: time interval (s)

At the wetland pre-retrofit outflow was mainly calculated based on the level data collected by the pressure transducer installed in a weir box at the orifice. Post-retrofit, the outlet orifice was closed, and outflow was calculated based on the water level and velocity data collected by the AVM. Two approaches were used to calculate flow rates: Manning's equation and the area-velocity method. Both methods estimated cumulative outflow volume greater than

the SCM's cumulative inflow, which was assumed as an overestimation because the same was not observed pre-retrofit. The area-velocity method resulted in less overestimation than using Manning's equation and thus was used for final flow calculation (*Equation 3-7*).

Equation 3-7 $$Q = VA * 1000$$

Where Q is flow rate (L/s), V is the velocity of water (m/s) and A is the area of water passing through the pipe (m^2) calculated using *Equation 3-8*.

Equation 3-8

If h < r $$A_{water} = A_{segment} - A_{triangle} = \frac{\theta}{360}\pi r^2 - \frac{r^2}{2} sin\,\theta$$

If h > r $$A_{water} = A_{circle} - \left(A_{segment} - A_{triangle}\right) = \pi r^2 - (\frac{\theta}{360}\pi r^2 - \frac{r^2}{2} sin\,\theta)$$

Where r is the radius of the pipe (m) and θ is the angle of the circle segment (degrees) calculated using *Equation 3-9*. When to use these equations is illustrated by *Figure 3.13*.

Equation 3-9

If h < r $$\theta = 2 * (cos^{-1}\left(\frac{r-h}{r}\right)) * {}^{180}/_{\pi}$$

If h > r $$\theta = 2 * (cos^{-1}\left(\frac{h-r}{r}\right)) * {}^{180}/_{\pi}$$

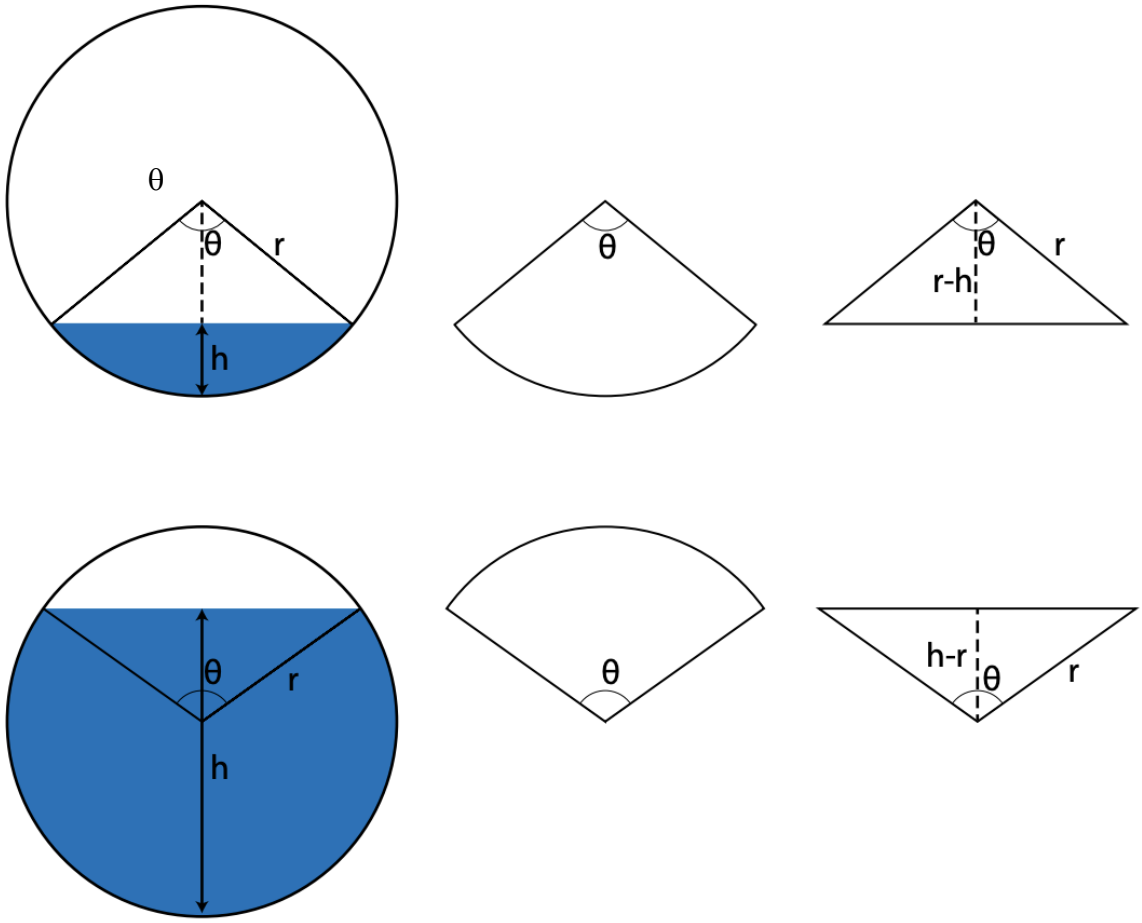


Figure 3.13 Cross-section of water in the outlet pipe and the geometry used for area calculation

Because the outflow volume based on the AVM (at the CSW) was higher than that of inflow, these values were compared with the RTC's estimated discharge for correction purposes. The RTC data were obtained from the online dashboard that included water level, water volume in the basin, valve status (and percentage of opening), and estimated discharge. These reported values were used for data analysis with unit adjustments.

Water quality was evaluated based on the collected pollutant concentration values and load values calculated using concentrations and flow volumes. At the wet pond, the monitoring methods yielded event mean concentration (EMC) values. Event pollutant loads were calculated using *Equation 3-10* and *Equation 3-6*. At the wetland the monitoring methods resulted in

instantaneous concentration values. Therefore, event loads were first estimated by instantaneous concentrations using *Equation 3-11* and thus EMC was back calculated using *Equation 3-10*. For calculation purposes, the 4-minute instantaneous concentrations were extrapolated to a 2-minute frequency using simple linear regression. This transformation brought both flow and water quality data to the same frequency.

Equation 3-10 $$L_{event} = V_{event}.EMC_{event}.10^6$$

Equation 3-11 $$L_{event} = \sum_1^{i-1}\left(\frac{Q_iC_i+Q_{i+1}C_{i+1}}{2}\right).(t_{i+1}-t_i).10^6$$

Where:

L: event pollutant load (Kg)

V: event cumulative volume (L)

EMC: event mean concentration (mg/L)

Q_i: instantaneous flow (L/s) at time t_i

t_i: timestamp (s)

C_i: instantaneous concentration (mg/L) at time t_i

i: the row of data

As previously noted, two different methods were used for water quality monitoring at the CSW (*in-situ* spectrophotometer) and the wet pond (flow-proportional composite sampling tested in a lab). The collected water quality data at the wetland required local calibration; this calibration was performed using simple linear regression (SLR) and partial least square regression (PLSR; Chapter 1). Both methods have estimated loads within a ±20% error range, with PLSR resulting in lower uncertainty on cumulative annual load estimations and SLR

resulting in lower uncertainty on event load estimations (Chapter 1). Therefore, the loads presented for the wetland were calculated using both calibration methods.

3.2.5 Statistical analysis

Hydrology and water quality at each study site were statistically analyzed to compare paired influent and effluent volumes, peak flow rates, pollutant concentrations, and pollutant loads on an event basis. This analysis was performed separately on pre-retrofit and post-retrofit data for each site. Paired datasets were first tested for normality using Shapiro-Wilk and Kolmogorov-Smirnov tests. Tests that showed normal distribution with both methods were analyzed using the student's t-test. Tests that were not normally distributed were tested for symmetry using the m-out-of-n bootstrap symmetry test (Miao et al, 2006). Data with symmetrical distribution were analyzed using Wilcoxon signed-rank test; asymmetrical datasets were analyzed using the paired samples sign test. Statistical analysis was performed at $\alpha = 0.05$ significance level. R version 4.0.2 was used to perform statistical and data analyses (R Core Team, 2020).

Datasets with non-normal distribution are often tested for log-normality and analyzed using student's t-test. In this study due to the use of RTC, the outflow dataset contained several zeros (zero outflow). Log transformation of datasets containing zero values would require treatment of those values as $\log(0) = \infty$. The common practice in dealing with zero values for log-transformation is adding a small constant to all values or only to the zero values (Bellego & Pape, 2019). Due to the behavior of log transformation that enlarges low values and compresses high values these methods can misrepresent the relationship between observations and introduce a large bias (Bellego & Pape, 2019). Therefore, the log-transformation was not implemented in the data analysis herein.

The hydrologic and water quality effect of the retrofit was evaluated by comparing removal efficiencies (*Equation 3-12*) of event volume, peak flow rate, cumulative volume, event concentrations, event loads, and cumulative loads (Martin & Smooth 1986; Strecker et al., 2001). Additionally, summary statistics were tabulated and explored using boxplots for the aforementioned parameters as well as retention times. Cumulative volumes account for the variability among events, and they were compared using exceedance probability plots. Pollutant loads account for both concentration and volume and are considered a more robust metric to evaluate SCM water quality performance than pollutant concentration (Barrett, 2005; Lenhart & Hunt, 2011). Additional parameters used to assess retrofit water quality benefits included annual load reduction per watershed area (kg/ha/yr), annual SCM exported load per watershed area (kg/ha/yr), and daily load reduction per SCM area (g/m^2/day). Such parameters normalize the results based on the data collection period SCM size and watershed area. Because hydrology and pollutant load data are typically skewed right (as illustrated in *Figure 3.14*), medians (in addition to mean) are presented as summary statistics.

Equation 3-12

$$Removal\ Efficiency\% = (\sum Influent - \sum Effluent) / \sum Influent * 100$$

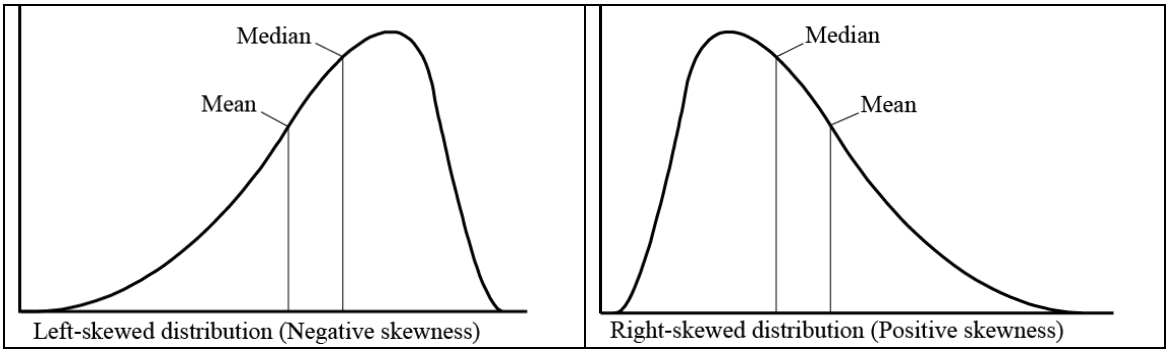


Figure 3.14 For both types of skewed distribution, median is a more representative summary statistics than mean

3.3 Results

Some storm events were removed from the data analysis due to issues in data collection. These issues included (1) power disconnection and shutdown of the instruments that occurred during prolonged overcast conditions where the batteries couldn't be recharged by the solar panels, (2) inaccurate data due to sensor failure, and (3) problems with outflow measurement due to outlet clogging by vegetation, frozen water in the orifice, and small animals (namely a baby turtle) trapped in orifice. In the CSW 6 of 106, and the wet pond 2 of 51 monitored events were removed due to the aforementioned circumstances.

3.3.1 Hydrology

3.3.1.1 CSW

3.3.1.1.1 Monitored events

Due to outlet hydrology monitoring challenges, post-retrofit outflows were estimated based on RTC discharge (Supplemental material 3-A). Sixty-four events were monitored for hydrology during pre-retrofit with inflow volumes, ranging from 0.2-3384-m^3 (0.002-37.1-mm); 37 post-retrofit events were monitored, ranging from 1.3-1969-m^3 [0.014-21.6-mm (*Table 3.10*)]. During pre-retrofit, some extreme events were observed, including hurricane Florence (3384m^3), although small volume events were more frequent (*Figure 3.15*; *Figure 3.16*). This resulted in a smaller interquartile range of inflow volumes, but a larger minimum-maximum range compared to those of post-retrofit events (*Table 3.10*; *Figure 3.16*).

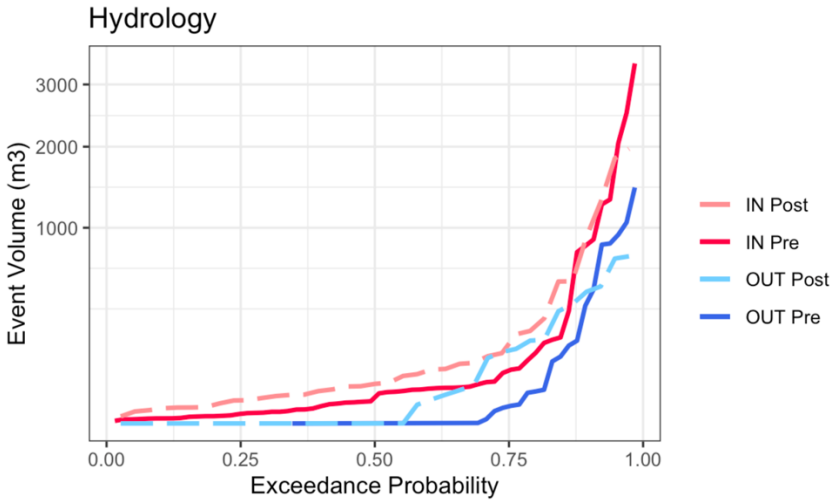


Figure 3.15 Exceedance probability of event volumes pre-retrofit and post-retrofit for CSW

Table 3.10 Summary of event inflow and outflow volumes pre- and post-retrofit at CSW

	Station	Min (m^3)	1st Quartile (m^3)	Median (m^3)	Mean (m^3)	3rd Quartile (m^3)	Max (m^3)
Pre-retrofit	In	0.2	2.6	18.2	233.5	68.8	3383.9
	Out	0.0	0.0	0.0	103.4	6.8	1453.9
Post-retrofit	In	1.3	14.2	40.9	239.2	129.2	1968.5
	Out	0.0	0.0	0.0	106.7	133.3	729.6

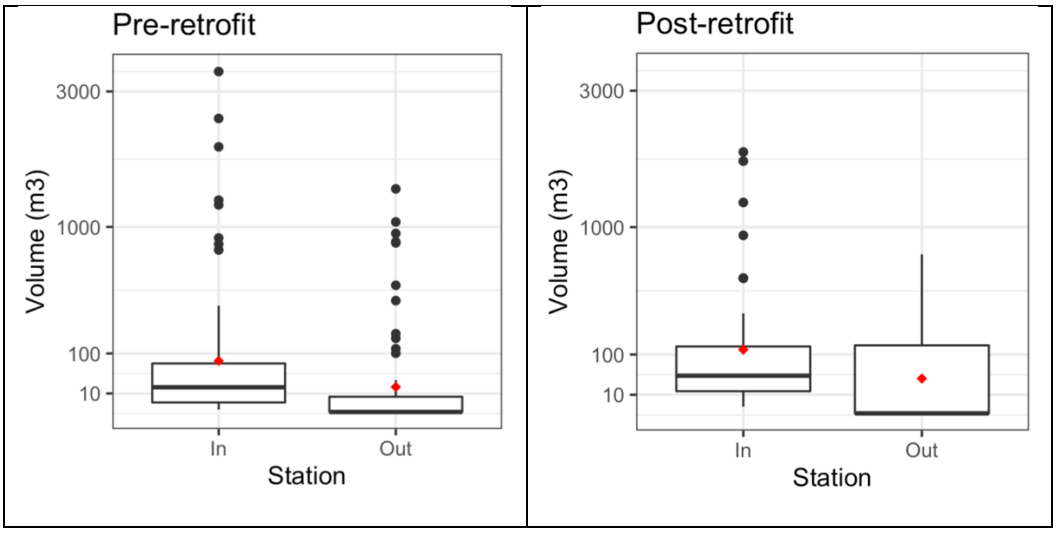


Figure 3.16 Event volumes pre-retrofit and post-retrofit at CSW

3.3.1.1.2 Volume reduction

Implementing RTC led to an increase in events with no outflow. Pre-retrofit, events with 100% volume reduction, ranged from ~0.2 to 183 m^3 of inflow, accounting for 5.4% of total inflow (72% of events). After retrofitting with RTC, storms with up to 228m^3 of inflow were fully captured; this accounted for 12.4% of total inflow volume (51% of events). Decreasing the number of outflow events can reduce the frequency of flow disturbance on receiving streams which, in turn lessens the impact on the stream ecosystem (Walsh et al. 2009). Although fewer events produced outflow post-retrofit, the number of higher volume outflows increased, resulting in a larger interquartile range of post-retrofit outflow volumes (*Figure 3.16*).

The CSW significantly reduced volume pre-retrofit (*Table 3.11*). The difference between inflow and outflow post-retrofit was also significant (*Table 3.11*). However, 6 (20% of) events added outflow (represented by negative removal efficiencies). Post-retrofit the outflow was released prior to an event with its discharged volume calculated based on the weather forecast; outsized overflow indicates an overestimated forecast event size. Such observation could also occur due to sensor failure that resulted in a miscalculation of SCM's available storage capacity. Inflow for these 6 events was low (6.9- 207.6-m^3), comprising only 5% of total volume. *Figure 3.17* illustrates such an event.

For 12 events, the RTC's prepared storage capacity was less than the actual inflow volume, leading to outflow during the event (i.e., overflow). These 12 events ranged from 30.3-1968.5m^3 inflow volume, comprising 85.6 % of total inflow volume. *Figure 3.18* illustrates two such events. Several factors can cause such overflow: (1) events that last longer than the 24-hour precipitation forecast window and therefore are not predicted in time for the RTC to react, (2) uncertainty of forecasted precipitation depth (Zaremba et al., 2019; *Figure 3.18*a), (3) sensor

failure causing miscalculation of SCM's available capacity, (4) latency of cellular data delivery, causing the valve closure after the start of an event (*Figure 3.18*b), and (5) RTC configuration that maintained water level at full storage capacity at all times, which therefore resulted in intra-event overflow due to (1) inaccurate volume estimation of a forecasted event, or (2) not enough time pre-event to release water and provide the required capacity [24-hour forecast (winter configuration used in this study)]. If the RTC goal is to prevent intra-event overflow, providing some free storage capacity at all times can compensate for some errors in timing and accuracy of forecasted event volume.

Table 3.11 Event volume reduction pre-retrofit and post-retrofit in CSW

	Min event volume reduction efficiency %	Median event volume reduction efficiency %	Mean event volume reduction efficiency %	Max event volume reduction efficiency %	Inflow to outflow volume p-value	Type of significant difference test
Pre-retrofit	0	100	88	100	2.22e-16	Samples sign test
Post-retrofit	-867	100	20	100	4.126e-05	Samples sign test

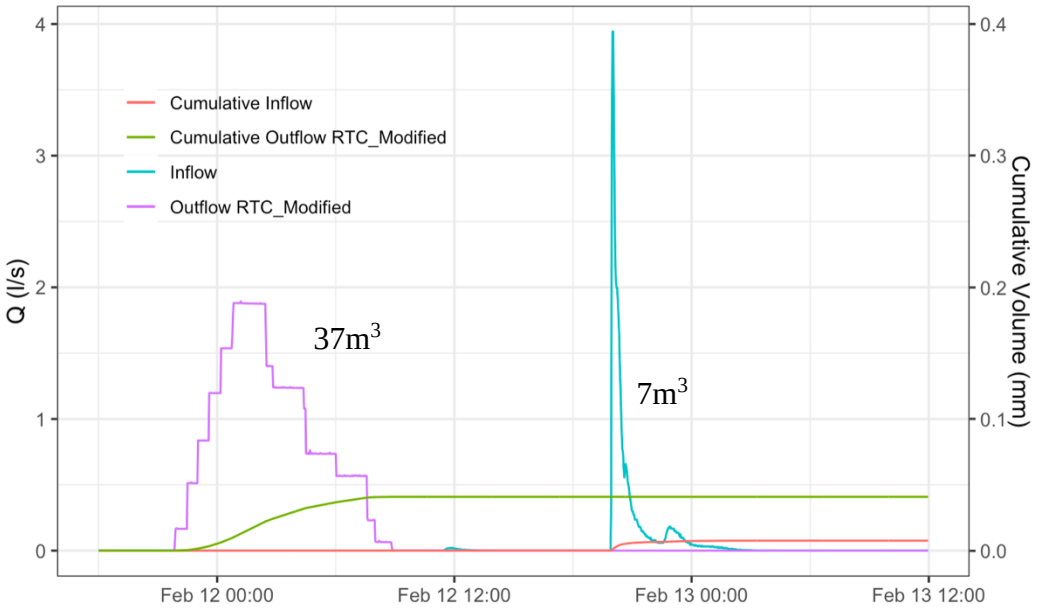


Figure 3.17 Example of an event with negative volume removal efficiency, i.e. higher outflow volumes than inflow due to advance release of water at CSW

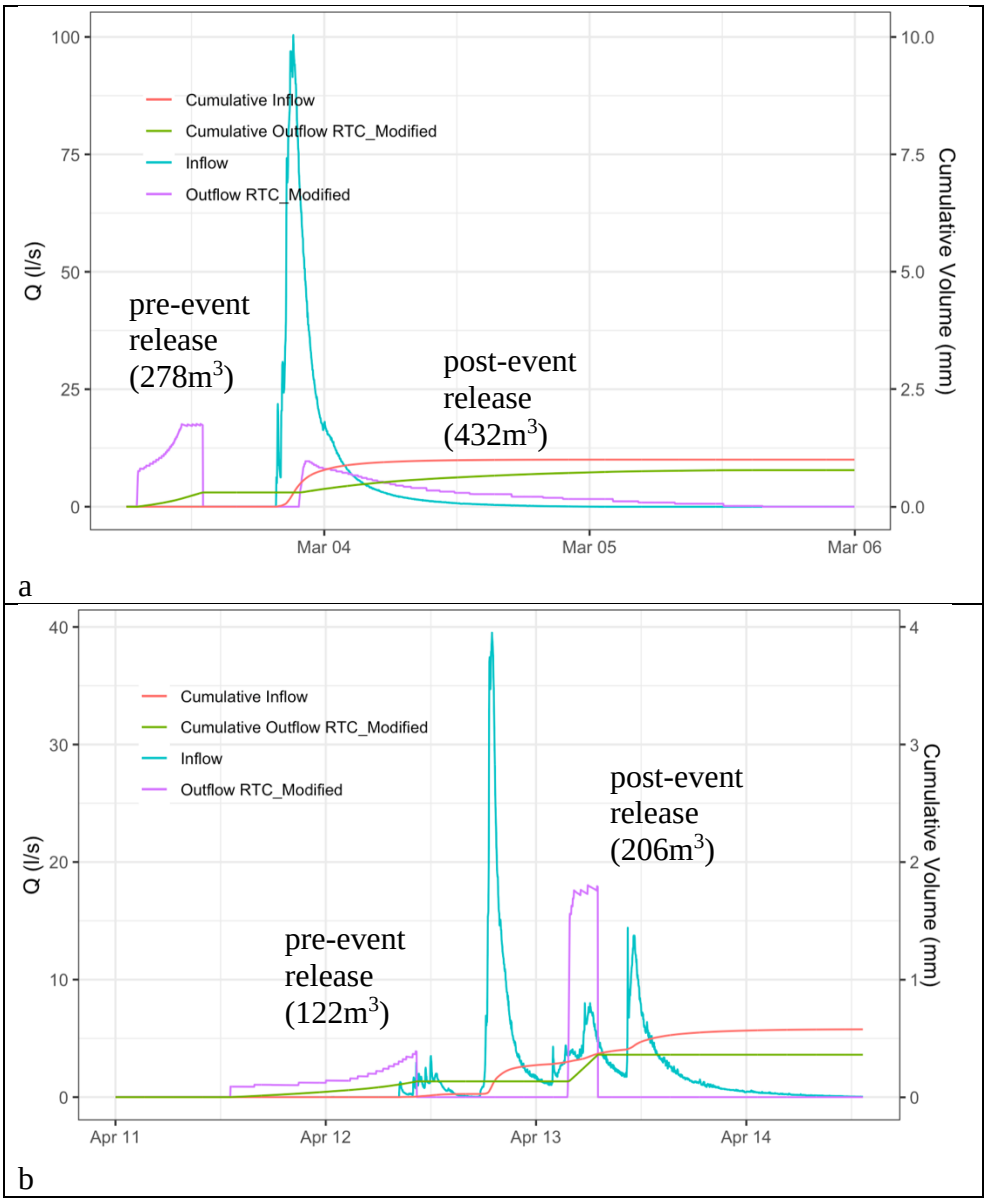


Figure 3.18 Example of two post-retrofit events at CSW with overflow during an event due to (a) miscalculation of event size and available storage capacity potentially due to the uncertainty of the forecasted event, (b) occurrence of an event that lasted longer than 24-hour forecast window

Figure 3.19 illustrates the cumulative inflow and outflow of the CSW pre- and post-retrofit. The volume reduction of CSW was generally in the same range as previous CSW studies in the region both on event (Line et al. 2008) and cumulative (Merriman & Hunt 2014) basis. The cumulative volume reduction of CSW remained the same after the implementation of RTC (*Table 3.12*). Given the water balance of the CSW (Equation 3-13), this suggests that the

extended retention time has not appreciably affected water loss through evapotranspiration. Because the event volume reduction can vary greatly depending on the status (e.g. capacity) of the SCM and the size of the event, the cumulative volume reduction emphasizes the importance of continuous long-term evaluation of RTC.

Equation 3-13 $V_B = V_{in} - V_{out} = Q_{in} + G_{in} - Q_{out} - G_{out} - ET$

Where

V_B: Water volume in the basin (m^3)

Q_{in}: inflow (m^3)

G_{in}: groundwater inflow (m^3)

Q_{out}: outflow (m^3)

G_{out}: infiltration (m^3)

ET: evapotranspiration (m^3)

Table 3.12 Cumulative volume reduction pre-retrofit and post-retrofit for wetland

	Cumulative volume In (m3)	Cumulative volume Out (m3)	Cumulative volume reduction efficiency %
Pre-retrofit	14943	6616	55.7
Post-retrofit	8851	3949	55.4

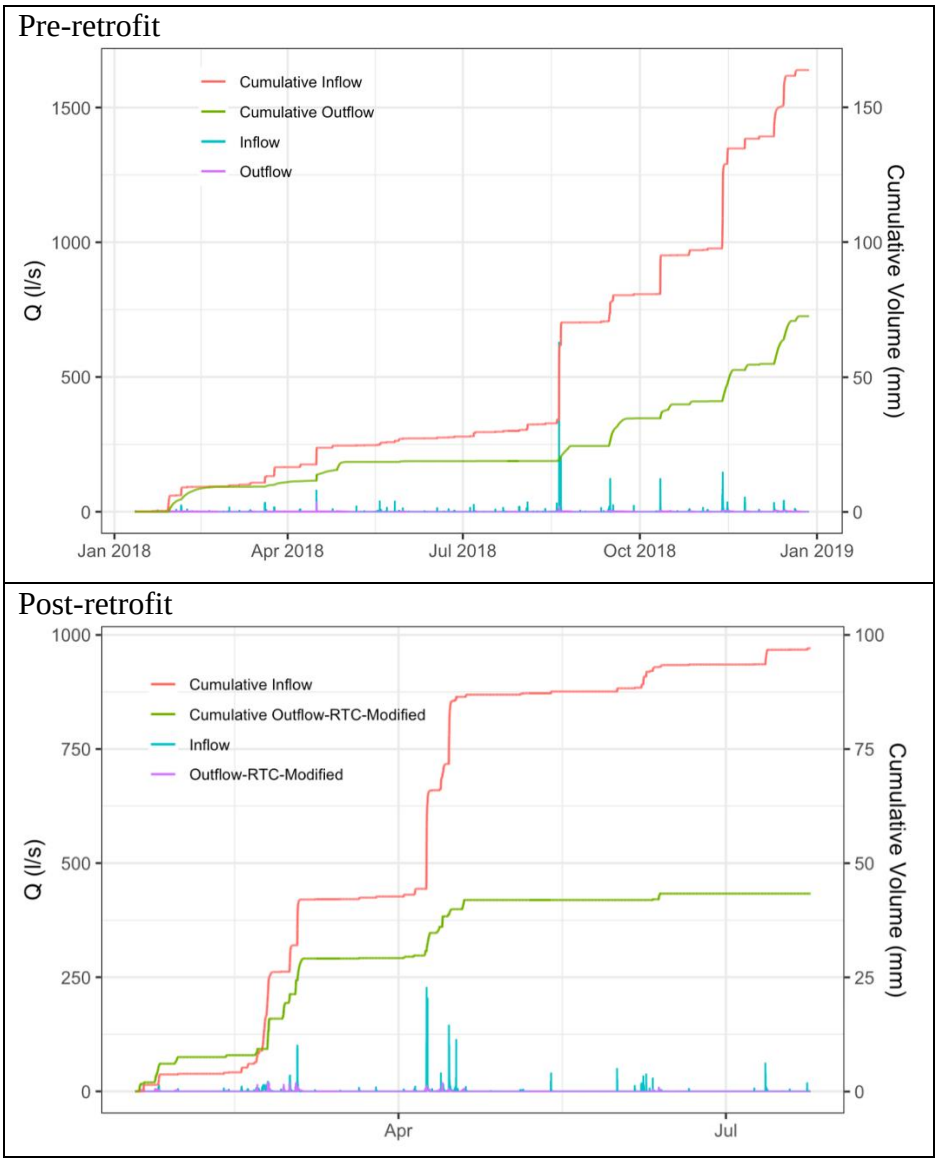


Figure 3.19 Cumulative volume reduction, pre- and post-retrofit at CSW

3.3.1.1.3 Peak flow mitigation

Mean inflow peaks did not change pre- and post-retrofit. Low peak inflows were most frequent and were interspersed with occasional extremely high inflows (*Table 3.13*). Pre-retrofit, peak outflows had a small interquartile range with some extreme observations (*Figure 3.20*). Post-retrofit peak outflows were more variable, with a slightly higher mean (*Figure 3.20*). The increase of mean peak flow was due to the release of higher volume outflow events in a short

duration prior to an event (24-hour weather prediction window). Despite the higher mean outflow peak post-retrofit, the maximum peak flow was substantially reduced.

Table 3.13 Summary statistics of peak flow at inlet and outlet of CSW during pre- and post-retrofit

	Station	Min (l/s)	1st Quartile (l/s)	Median (l/s)	Mean (l/s)	3rd Quartile (l/s)	Max (l/s)
Pre-retrofit	In	0.11	1.7	7	29.7	19	627.7
	Out	0	0	0	1.3	0.2	36.4
Post-retrofit	In	1.7	5.1	10.4	29.7	34.8	227.1
	Out	0	0	0	3.4	5.3	18

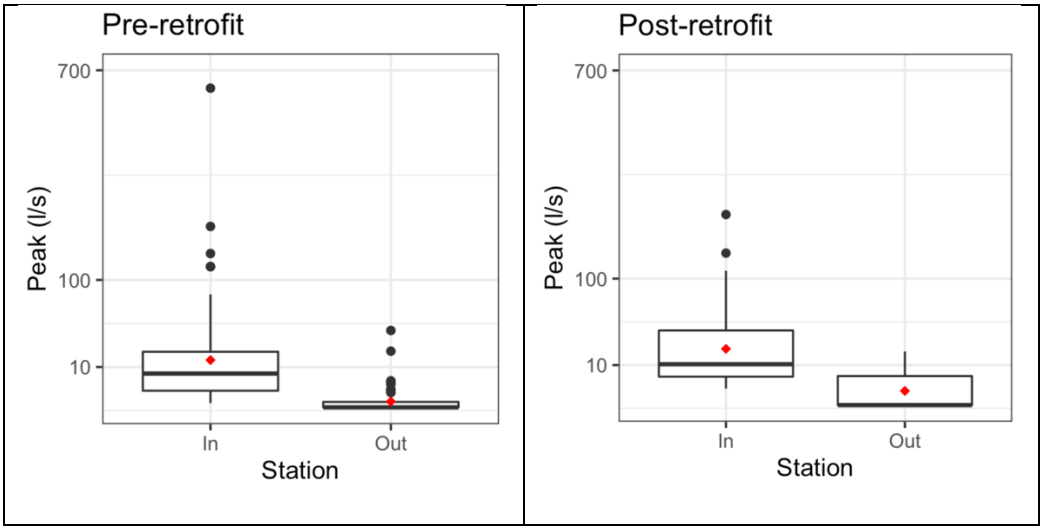


Figure 3.20 Peak flows at inlet and outlet of CSW pre- and post-retrofit

During both periods the SCM significantly mitigated peak flow (*Table 3.14*). Pre- and post-retrofit peak flow mitigation was generally very high, no matter the period. However, post-retrofit one event had negative peak flow mitigation (i.e., outflow peak higher than inflow peak) during one small storm, because inflow volume had been overestimated by the RTC.

Table 3.14 Summary statistics of peak flow reduction efficiency pre- and post-retrofit

	Min event peak mitigation efficiency %	Median event peak mitigation efficiency %	Mean event peak mitigation efficiency %	Max event peak mitigation efficiency %	Inflow to outflow peak p-value	Type of significant difference test
Pre-retrofit	54	100	98	100	2.20E-16	Samples sign test
Post-retrofit	-154	100	81	100	5.53E-10	Samples sign test

3.3.1.1.4 Retention time

As noted in the methods section, retention time was calculated as the time difference between the centroids of inflow and outflow. However, such association of in/outflows post-retrofit was not always as straightforward. Although RTC is set to release water prior to an event, for some events such release extended beyond commencement of inflow (*Figure 3.22*), plus outflow could occur as overflow during an event. Therefore, post-retrofit retention times are presented in three categories: events with outflow (1) preceding inflow, (2) preceding subsequent event inflow, and (3) no outflow (*Figure 3.22*). For events with two outflow types, retention time calculations were based on the outflow period accounting for the majority of the volume.

Post-retrofit the retention time decreased relative to pre-retrofit except for a few outlier observations and had less interquartile variability than that of pre-retrofit (*Table 3.15*; *Figure 3.21*). This decrease in retention time relates to the occurrence of overflow during 12 of the 18 outflow-producing events. The drivers of overflow post-retrofit are discussed in volume reduction section.

Two extreme (perhaps odd) retention times were reported during pre-retrofit (*Table 3.15*). These two observations were associated with either: (1) two events with less than 6 hours of dry weather in between, or (2) continuous outflow. The former example resulted in a negative

time of retention, while the latter lead to a longer than usual retention time (illustrated in Supplemental material 3-B).

Table 3.15 Retention time pre-retrofit and post-retrofit at CSW

	Number of events	Min event retention time (hr)	1st Quartile event retention time (hr)	Median event retention time (hr)	Mean event retention time (hr)	3rd Quartile event retention time (hr)	Max event retention time (hr)
Pre-retrofit	18	-15.4	7.8	15.0	25.1	30.3	151.4
Pre-retrofit-No outflow	46	-	-	-	-	-	-
Post-retrofit-outflow preceding subsequent event	12	-2.1	6.7	15.0	23.0	25.3	79.0
Post-retrofit-outflow preceding inflow	6	-14.8	-0.1	1.9	5.6	7.4	36.1
Post-retrofit-No outflow	19	-	-	-	-	-	-

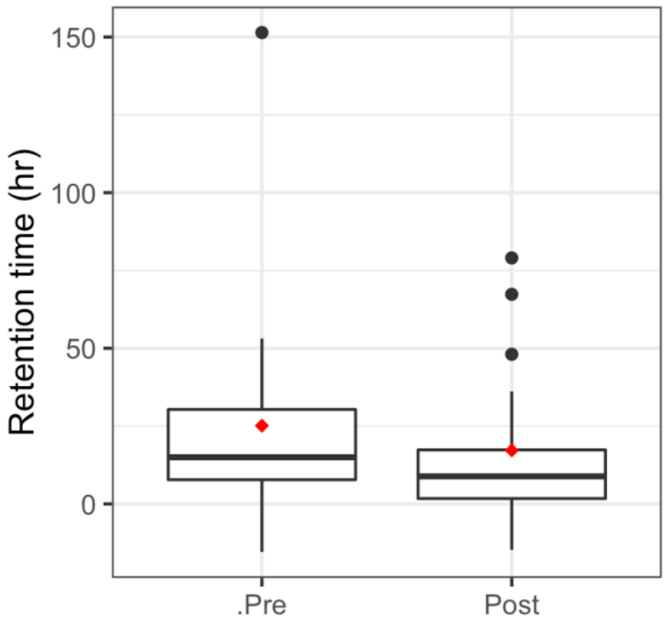


Figure 3.21 Retention time pre- and post-retrofit at CSW

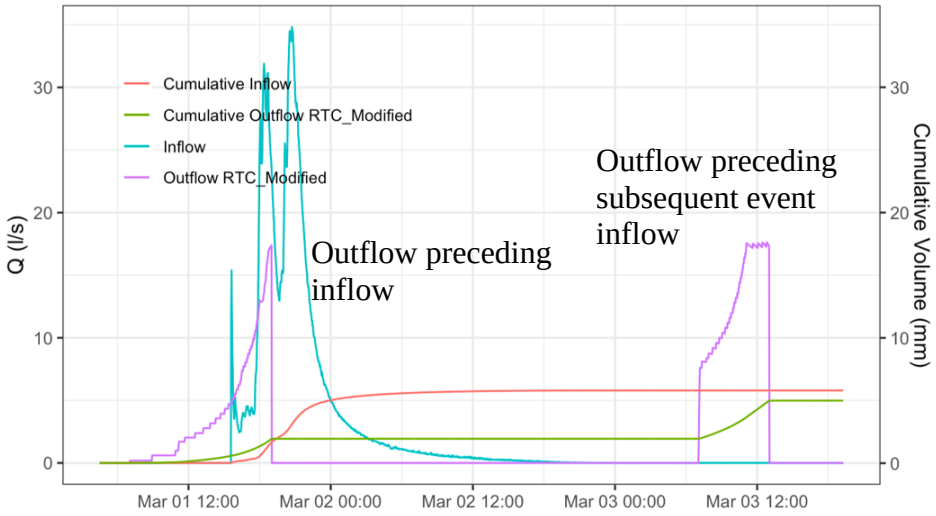


Figure 3.22 Example of a post-retrofit event with outflow during the event

3.3.1.2 Wet pond

3.3.1.2.1 Monitored events

Ten events were monitored for hydrology pre-retrofit and 39 post-retrofit. Delays in selection of a study site resulted in fewer monitored events pre-retrofit, but a sufficient number of events were monitored to draw reasonable volume and load removal estimations in the author's view. May and Sivakumar (2009) suggested the collection of 5-7 events were needed to attain a relatively accurate estimate of site mean concentrations (SMC), and Maniquiz-Redillas et al. (2013) suggested the collection of 6-8 events to estimate TSS SMC within a 20% relative standard error range.

Inflow volumes of monitored events ranged from 131-1408m^3 (7.7-82.6mm) pre-retrofit and 0.4-763m^3 (0.02-44.8mm) post-retrofit. In general, pre-retrofit events had higher variability (larger interquartile range) and included higher volume events than those of post-retrofit (*Figure 3.23*; *Figure 3.24*). With a median of 80m^3, more than 50% of post-retrofit events had volumes smaller than the minimum pre-retrofit inflow event (131m^3).

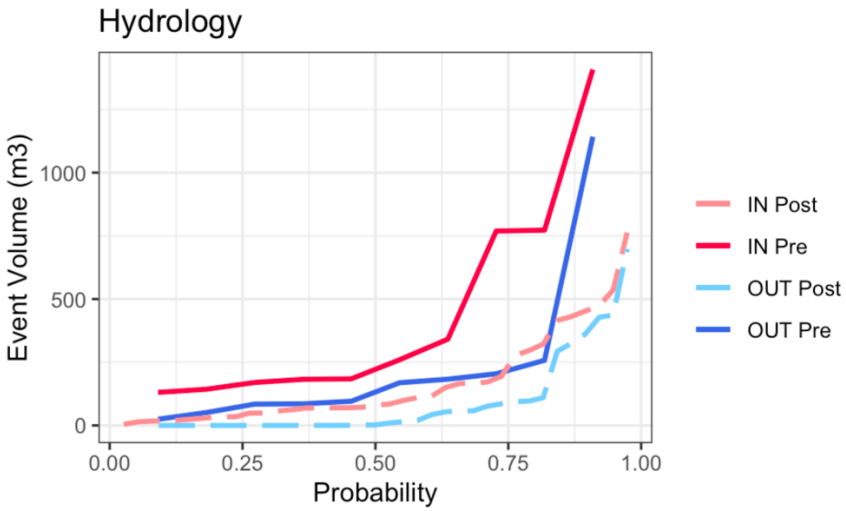


Figure 3.23 Exceedance probability of event volumes pre-retrofit and post-retrofit for wet pond

Table 3.16 Summary of event inflow and outflow volumes pre- and post-retrofit at the wet pond

	Station	Min (m^3)	1st Quartile (m^3)	Median (m^3)	Mean (m^3)	3rd Quartile (m^3)	Max (m^3)
Pre-retrofit	In	131	173	222	436	662	1408
	Out	25	85	132	230	199	1142
Post-retrofit	In	0.4	34	80	156	183	763
	Out	0	0	1.5	84	76	808

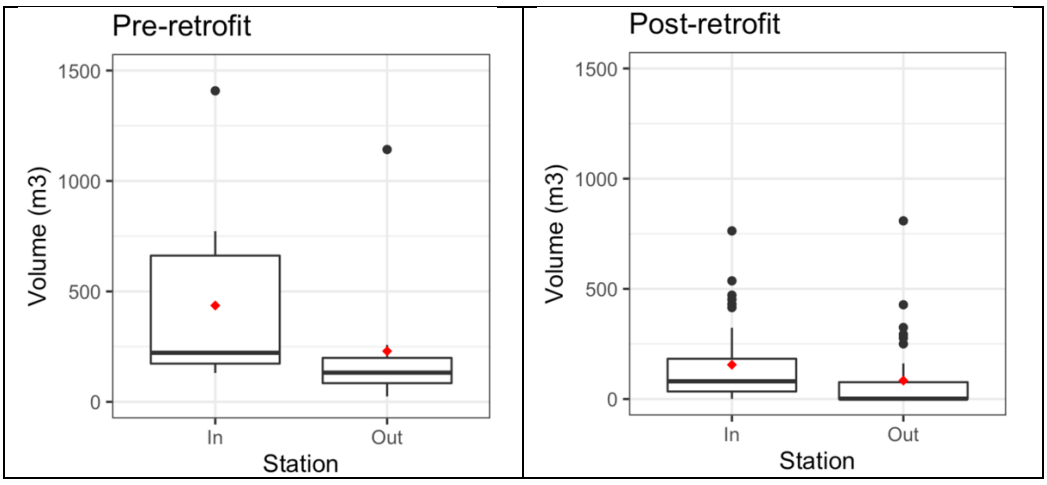


Figure 3.24 Event volume pre-retrofit and post-retrofit at the wet pond

3.3.1.2.2 Volume reduction

Implementing RTC led to an increase in events with no outflow. Pre-retrofit, all monitored events produced outflow. Post-retrofit events up to 114.6-m^3 of inflow had 100% volume reduction, accounting for 16% of total inflow. The mean outflow volume post-retrofit was lower than that of pre-retrofit, potentially due to higher number of no-outflow events and the relatively smaller inflow volumes post-retrofit (*Table 3.16*). Despite the lower post-retrofit inflow volumes and lower average outflow volumes, a number of higher volume outlier outflow events occurred (*Figure 3.24*).

The wet pond significantly reduced volume pre-retrofit (*Table 3.17*). The difference between inflow and outflow post-retrofit was also significant (*Table 3.17*). However, 4 (10% of) events added outflow (represented by negative removal efficiencies). Excess outflow resulted from overestimated forecast event size (*Figure 3.25*).

For 11 of 39 events, the RTC's prepared storage capacity was less than the actual inflow volume, leading to outflow during the event. These 11 events ranged from 7.8-763-m^3 inflow volume, comprising 65% of the total inflow volume. *Figure 3.26* illustrates two such events, where overflow occurred due to (a) occurrence of an event that lasted longer than the 24-hour precipitation forecast window and (b) uncertainty of forecasted precipitation depth.

Table 3.17 Event volume reduction pre-retrofit and post-retrofit in the wet pond

	Min event volume reduction efficiency %	Median event volume reduction efficiency %	Mean event volume reduction efficiency %	Max event volume reduction efficiency %	Inflow to outflow volume p-value	Type of significant difference test
Pre-retrofit	0	53	49	87	0.01024	Student's t-test
Post-retrofit	-439	96	49	100	1.186E-07	Student's t-test

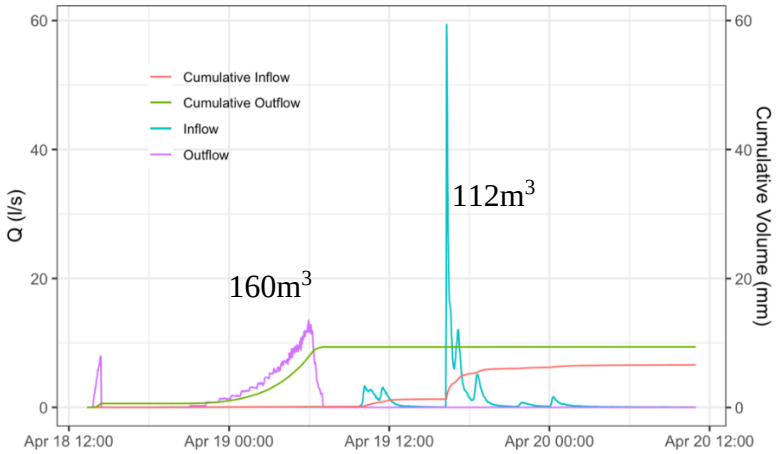


Figure 3.25 Example of an event with negative volume removal efficiency, i.e. higher outflow volumes than inflow due to advance release of water at the wet pond

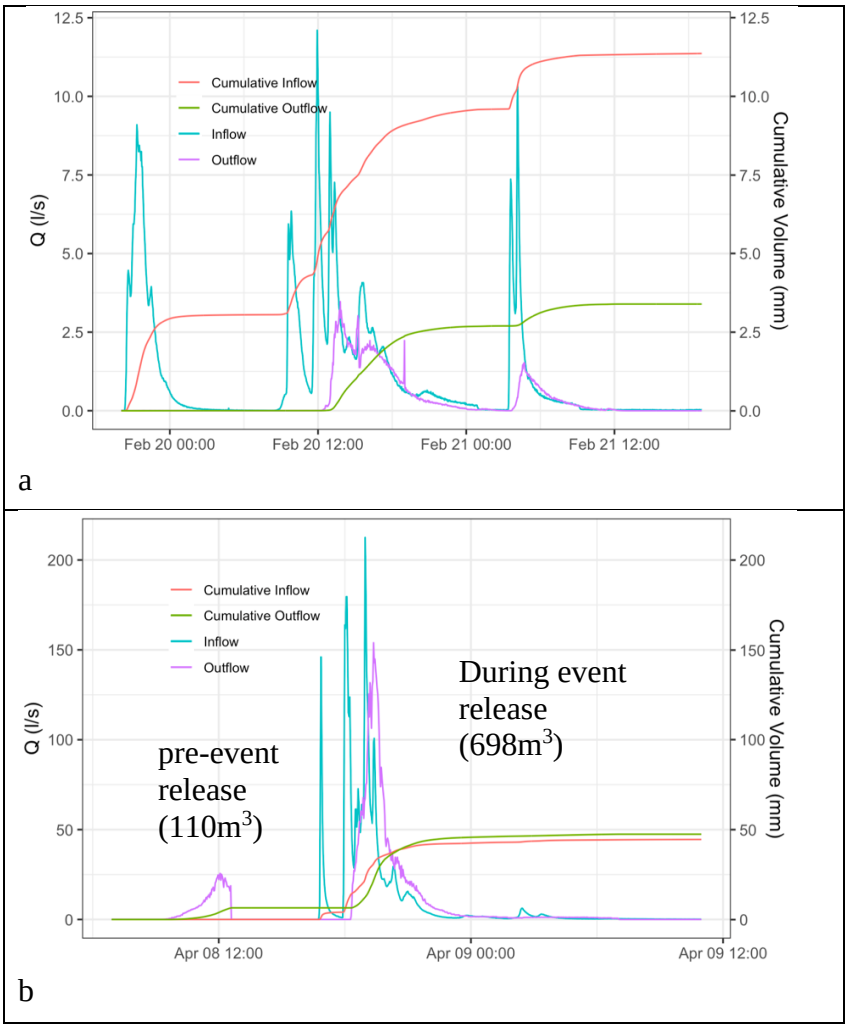


Figure 3.26 Example of two post-retrofit events at the wet pond with overflow during an event due to (a) occurrence of an event that lasted longer than 24-hour precipitation forecast window, and (b) miscalculation of event size and available storage capacity potentially due to the uncertainty of the forecasted event

Figure 3.27 details the cumulative inflow and outflow of the wet pond pre- and post-retrofit. The volume reduction of this wet pond was significantly higher than those reported in the literature (Hancock et al. 2010; Schwartz et al. 2017), higher than a reported leaking wet pond in the same region (Line et al. 2012), and lower than an infiltrating wet pond in NC (Baird et al. 2020). Examination of the water level data post-retrofit showed a mean daily intra-event water loss of 3.9 cm (range of 1.9-8.22 cm/day) to evapotranspiration and infiltration. The cumulative volume reduction of the wet pond has remained fairly unchanged pre- and post-retrofit (*Table 3.18*). Given the water balance of the wet pond (*Equation 3-13*), this suggests that the potential increase in water head and increased surface area due to the RTC retrofit has not affected water loss through evapotranspiration or infiltration. The similar cumulative volume removal during both periods, can explain why certain post-retrofit events had increased outflow volume while more events had no outflow.

Table 3.18 Cumulative volume reduction pre-retrofit and post-retrofit for the wet pond

	Cumulative volume In (m3)	Cumulative volume Out (m3)	Cumulative volume reduction efficiency %
Pre-retrofit	4361	2297	47.3
Post-retrofit	6066	3264	46.2

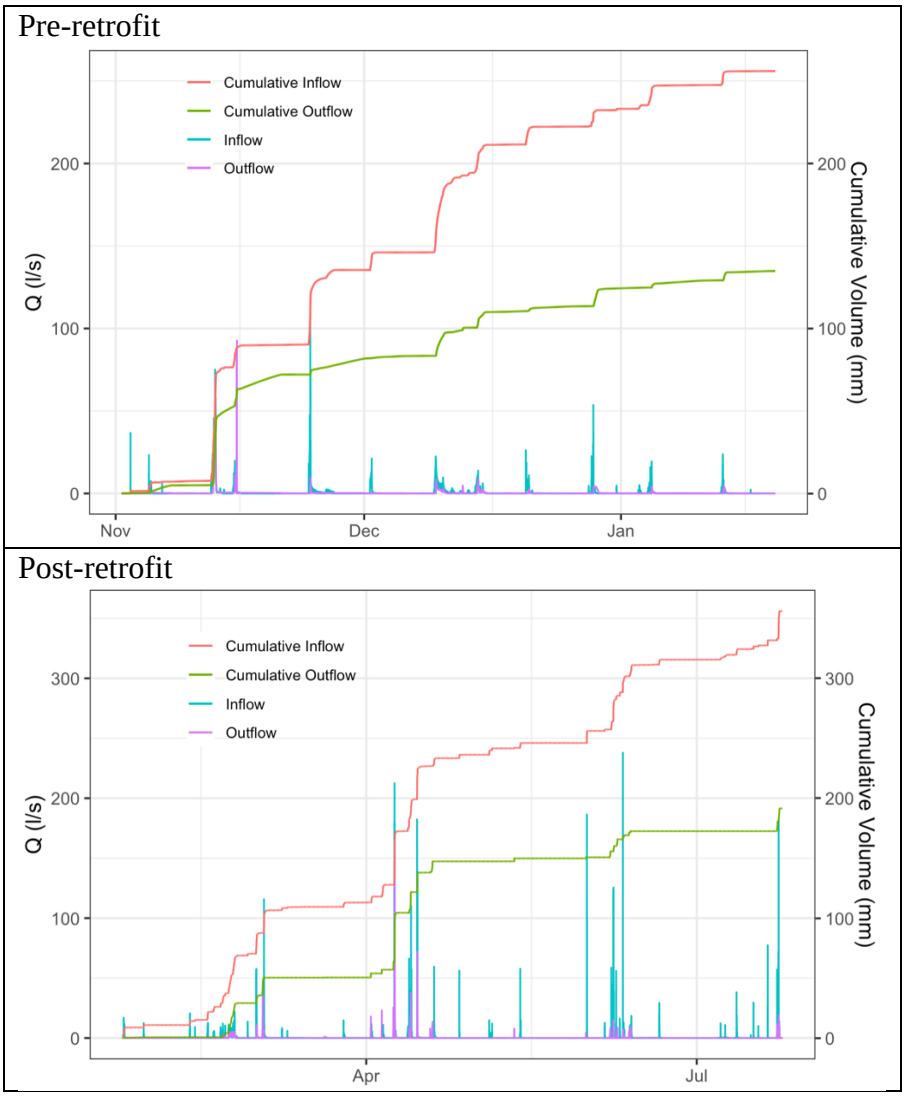


Figure 3.27 Cumulative volume reduction during pre- and post-retrofit at the wet pond

3.3.1.2.3 Peak flow mitigation

Post-retrofit peak inflows and outflows had higher variability than those of pre-retrofit, with some extreme events observed (*Figure 3.28*). Despite the higher mean peak inflow during post-retrofit, the mean peak outflow was slightly lower than that of pre-retrofit (*Table 3.19*). Implementing RTC led to a decrease in the mean and median and increase in variability of outflow peaks (*Figure 3.28*). The increased variability of outflow peaks post-retrofit was due to (1) more no-outflow events and (2) RTC configuration that allowed release of large volumes of

water in a short duration that led to high peak outflows (i.e., 100% open valve with 6-in diameter).

Table 3.19 Summary statistics of peak flow at inlet and outlet of the wet pond during pre- and post-retrofit

	Station	Min (l/s)	1st Quartile (l/s)	Median (l/s)	Mean (l/s)	3rd Quartile (l/s)	Max (l/s)
Pre-retrofit	In	13.9	21.5	25	39.9	49.4	106.9
	Out	0.1	3.1	4.5	13.7	8.8	92.7
Post-retrofit	In	0.5	11.8	18.5	52.9	58.6	237.9
	Out	0	0	0.3	11.6	10.3	154

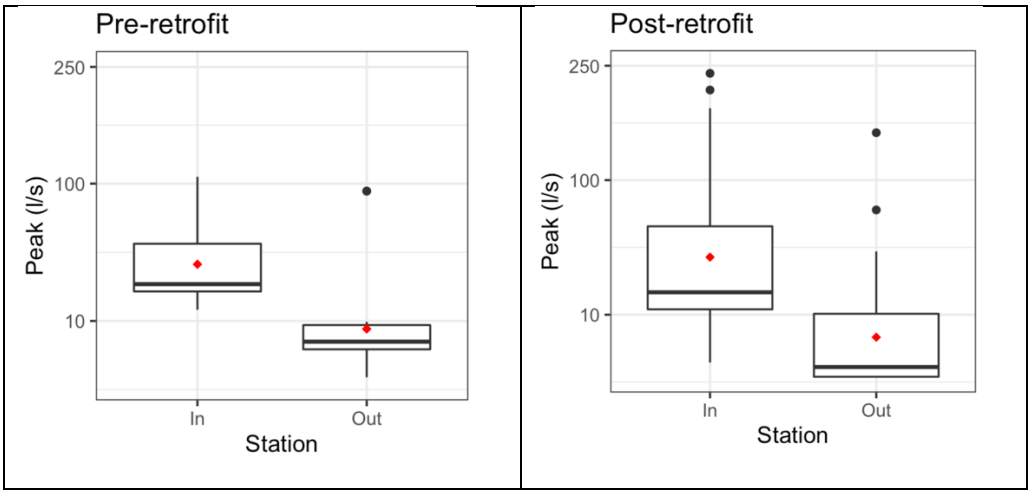


Figure 3.28 Peak flows at inlet and outlet of the wet pond pre- and post-retrofit

During both periods the SCM significantly mitigated peak flow (*Table 3.20*). Although, post-retrofit, 3 events had peak outflows higher than peak inflow resulting in negative peak mitigation efficiency. These events were among the smallest events (2.5-20.3-m^3 of inflow volume). Post-retrofit the peak flow mitigation efficiency remained in the same range, with decreased mean (by 13%) and increased median (by 14%) compared to those pre-retrofit (*Table 3.20*). The decreased mean can be an artifact of the described negative efficiency values and the increased median can be explained more no-outflow events post-retrofit.

Table 3.20 Summary statistics of peak flow reduction efficiency pre- and post-retrofit

	Min event peak mitigation efficiency %	Median event peak mitigation efficiency %	Mean event peak mitigation efficiency %	Max event peak mitigation efficiency %	Inflow to outflow peak p-value	Type of significant difference test
Pre-retrofit	23	85	70	99	0.02364	Student's t-test
Post-retrofit	-576	99	57	100	3.353E-07	Samples sign test

3.3.1.2.4 Retention time

Implementation of RTC has increased both mean retention time by a factor of ~4 (from 15.3 to 62.2hr) and the median time by a factor of ~2 [from 8.5 to 15.4 hr (*Table 3.21*)]. Post-retrofit the retention time had larger interquartile variability with few extreme high observations (*Figure 3.29*). The improved retention times were still within the suggested range to avoid mosquito breeding (Knight et al., 2003). Despite the improved mean retention time, 4 events have shown negative retention time (i.e. outflow centroid earlier than inflow centroid). All of these events had overflow during the event, in other words, their size was not predicted properly, and one had negative volume reduction (i.e. overestimation of event size). *Figure 3.30* illustrates two examples for such events that had overflow during the event resulted from an underestimation of event size (a) and high outflow prior to the event that lasted during an event due to overestimation of event size (b).

Table 3.21 Retention time pre- and post-retrofit at the wet pond

	Number of events	Min event retention time (hr)	1st Quartile event retention time (hr)	Median event retention time (hr)	Mean event retention time (hr)	3rd Quartile event retention time (hr)	Max event retention time (hr)
Pre-retrofit	10	1.9	3.7	8.5	15.3	20.9	56.9
Pre-retrofit-No outflow	0	-	-	-	-	-	-
Post-retrofit-outflow preceding subsequent event	18	-16.9	0.9	15.4	62.2	61.0	434.0
Post-retrofit-outflow preceding inflow	2	-0.9	-0.5*	-0.1*	-0.1*	0.3*	0.7
Post-retrofit-No outflow	19	-	-	-	-	-	-

*Summary statistics calculated for 2 observations.

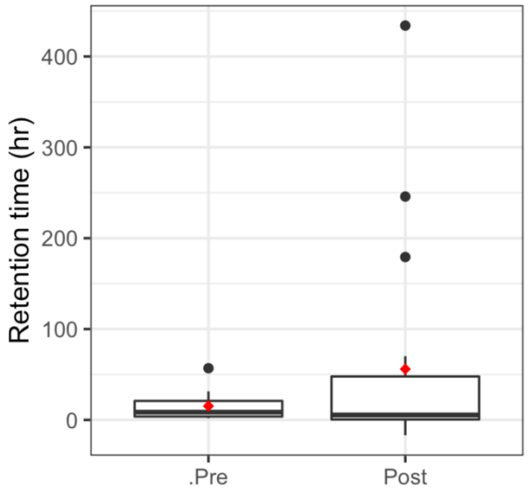


Figure 3.29 Retention time pre- and post-retrofit at the wet pond

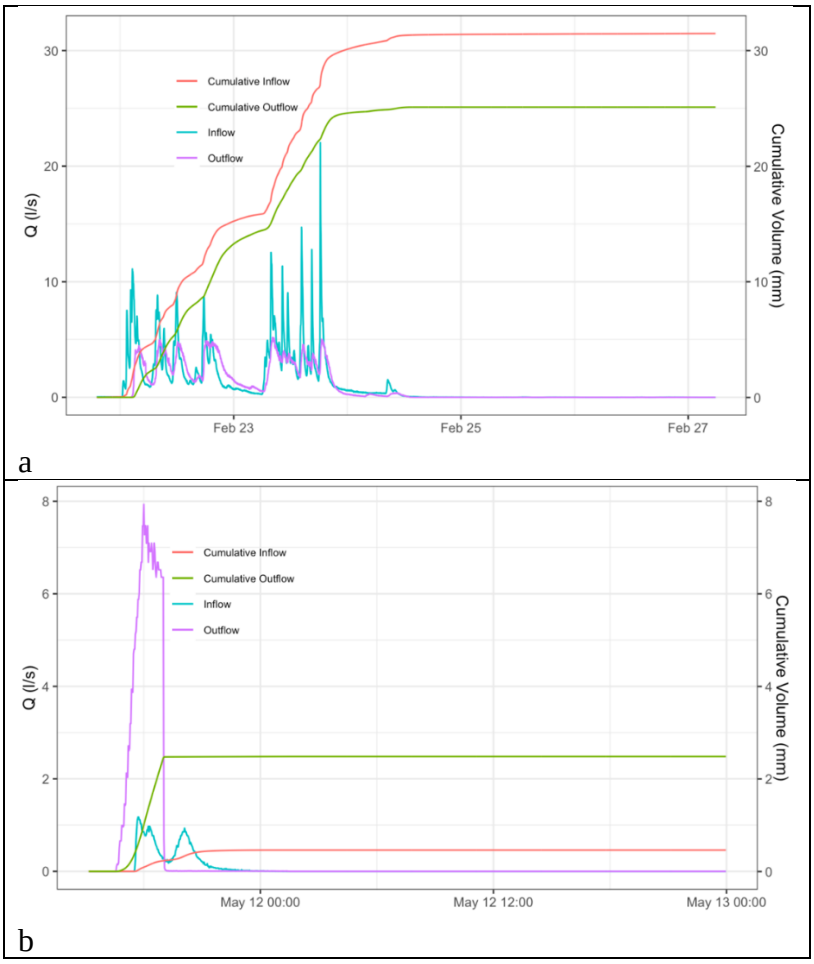


Figure 3.30 Post-retrofit events at the wet pond with negative retention time (a) overflow during the event due to underestimation of the event size, (b) high outflow prior to the event that lasted during an event due to overestimation of event size

3.3.2 Water quality

3.3.2.1 CSW

Among the pre-retrofit events monitored for hydrology, 32 were sampled for water quality including paired in/outflow samples for 14 events. Post-retrofit monitoring included 29 events (water quality samples) and 12 (paired samples). Due to challenges introduced by location of the spectrophotometer and autosampler intake at the outlet station, (1) the concentration values were calibrated only using the SLR model, and (2) captured concentrations may represent both outflow from the CSW and the backflow from the end of the outlet pipe (Chapter 1).

The SLR model has shown lower load estimation error range on event basis and PLSR offers lower load estimations error on cumulative basis (Chapter 1). Based on both calibration methods, CSW significantly removed NO3− -N and TSS concentrations during both monitoring periods and DOC concentrations during pre-retrofit (Table 3.22). Because during both periods the majority of events did not produce outflow (i.e., no paired samples), the median outflow concentration was zero and, thus, had 100% median removal efficiency (*Figure 3.31; Table 3.23*). The mean removal efficiencies were generally comparable to those pre-retrofit for NO3− -N and TSS (*Table 3.23*). The decrease of mean DOC removal efficiency post-retrofit was due to high variability among events and observation of some (8) events with high DOC export (~-300% removal efficiency), while fewer pre-retrofit events (6) had DOC export.

The cumulative load removal efficiency for NO_3^--N and TSS were essentially the same (±3%) during both periods but decreased by ~10% for DOC post-retrofit (*Table 3.24*). Given that retention time remained almost the same post-retrofit, it was expected that removal rates would also remain the same. The load reductions slightly increased post-retrofit, which is likely due to the generally lower cumulative inflow pollutant loading during that period (*Table 3.24*; *Figure 3.32*).

The CSW provided high pollutant removal both pre- and post-retrofit. The cumulative load removal for NO_3^- and TSS was higher than those reported by Merriman & Hunt (2014), and concentration reduction for these pollutants were in the upper range suggested by Line et al. (2008). Comparison of event basis TSS removal with NC required performance standards (NCDEQ 2017), also suggest that (1) tests of efficiency would have been considered invalid by NCDEQ (2017) for about half of monitored paired events, and (2) the CSW performance with regards to TSS decreased post-retrofit (Supplemental material 3-C).

Table 3.22 CSW's EMC for all pollutants pre- and post-retrofit

	Pollutant	Median event Conc. In (mg/l)	Mean event Conc. In (mg/l)	Mean event[1] Conc. Out (mg/l)	Inflow to outflow Conc. p-value	Type of significant difference test
Pre-retrofit	NO_3^- - SLR	0.39	0.53	0.05	**4.66E-10**[2]	Samples sign test
	NO_3^-- PLSR	0.37	0.53		**4.66E-10**	Samples sign test
	TSS - SLR	15.24	20.31	7.42	**1.54E-08**	Wilcoxon signed rank
	TSS - PLSR	29.56	31.37		**2.98E-07**	Wilcoxon signed rank
	DOC - SLR	5.77	7.62	5.55	**2.10E-03**	Samples sign test
	DOC - PLSR	4.97	6.93		**9.91E-03**	Wilcoxon signed rank
Post-retrofit	NO_3^-- SLR	0.37	0.37	0.04	**3.73E-09**	Wilcoxon signed rank
	NO_3^-- PLSR	0.39	0.42		**3.73E-09**	Wilcoxon signed rank
	TSS - SLR	16.69	24.19	8.41	**2.61E-07**	Wilcoxon signed rank
	TSS - PLSR	32.96	34.51		**6.62E-09**	Student's t-test
	DOC - SLR	6.34	7.71	6.32	2.30E-01	Wilcoxon signed rank
	DOC - PLSR	5.68	7.64		2.38E-01	Wilcoxon signed rank

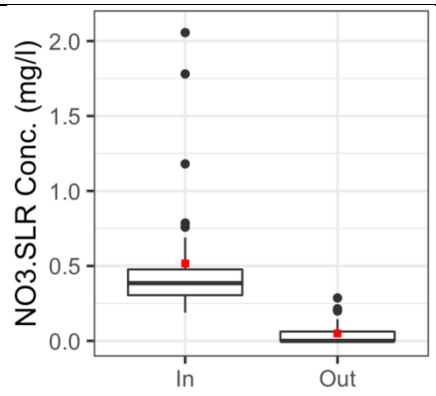
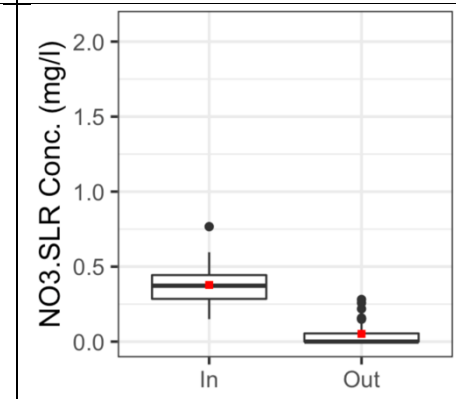
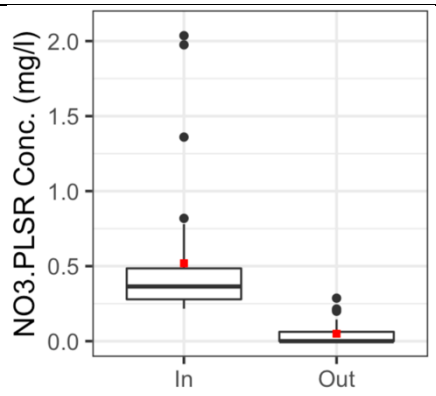
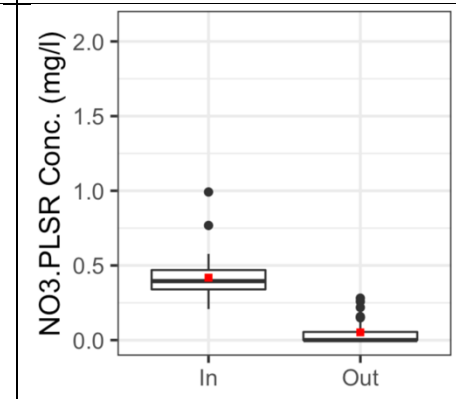
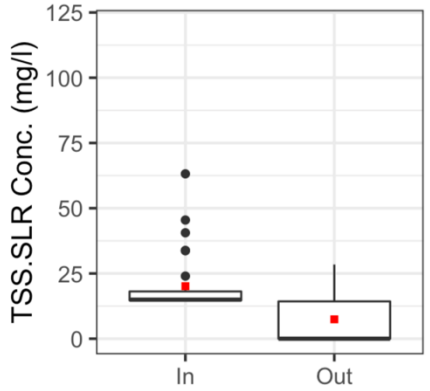
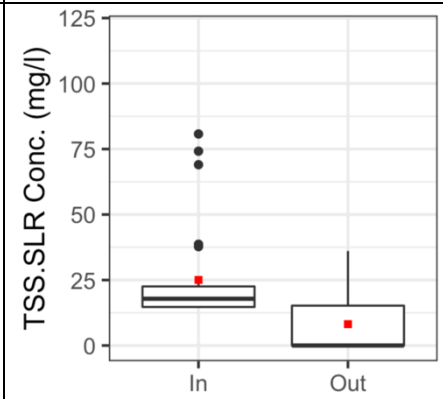
[1] Median event Conc. out = 0.

[2] Bold values are significant.

Table 3.23 CSW's summary statistics of event pollutant removal efficiencies (RE) pre- and post-retrofit

	Pollutant	Min event RE%	1st quartile RE%	Median event RE%	Mean event RE%	3rd quartile event RE%	Max event RE%
Pre-retrofit	NO_3^-- SLR	49	79	100	88	100	100
	NO_3^-- PLSR	39	75	100	86	100	100
	TSS - SLR	-46	37	100	69	100	100
	TSS - PLSR	-106	24	100	58	100	100
	DOC - SLR	-474	14	100	39	100	100
	DOC - PLSR	-462	14	100	37	100	100
Post-retrofit	NO_3^-- SLR	55	83	100	90	100	100
	NO_3^-- PLSR	46	86	100	89	100	100
	TSS - SLR	-62	18	100	64	100	100
	TSS - PLSR	-47	41	100	71	100	100
	DOC - SLR	-315	-30	100	1	100	100
	DOC - PLSR	-359	-26	100	1	100	100

Figure 3.31 CSW's EMCs for all pollutants using both calibration methods pre- and post-retrofit

	Pre-retrofit Conc.	Post-retrofit Conc.
NO_3-SLR		
NO_3-PLSR		
TSS-SLR		

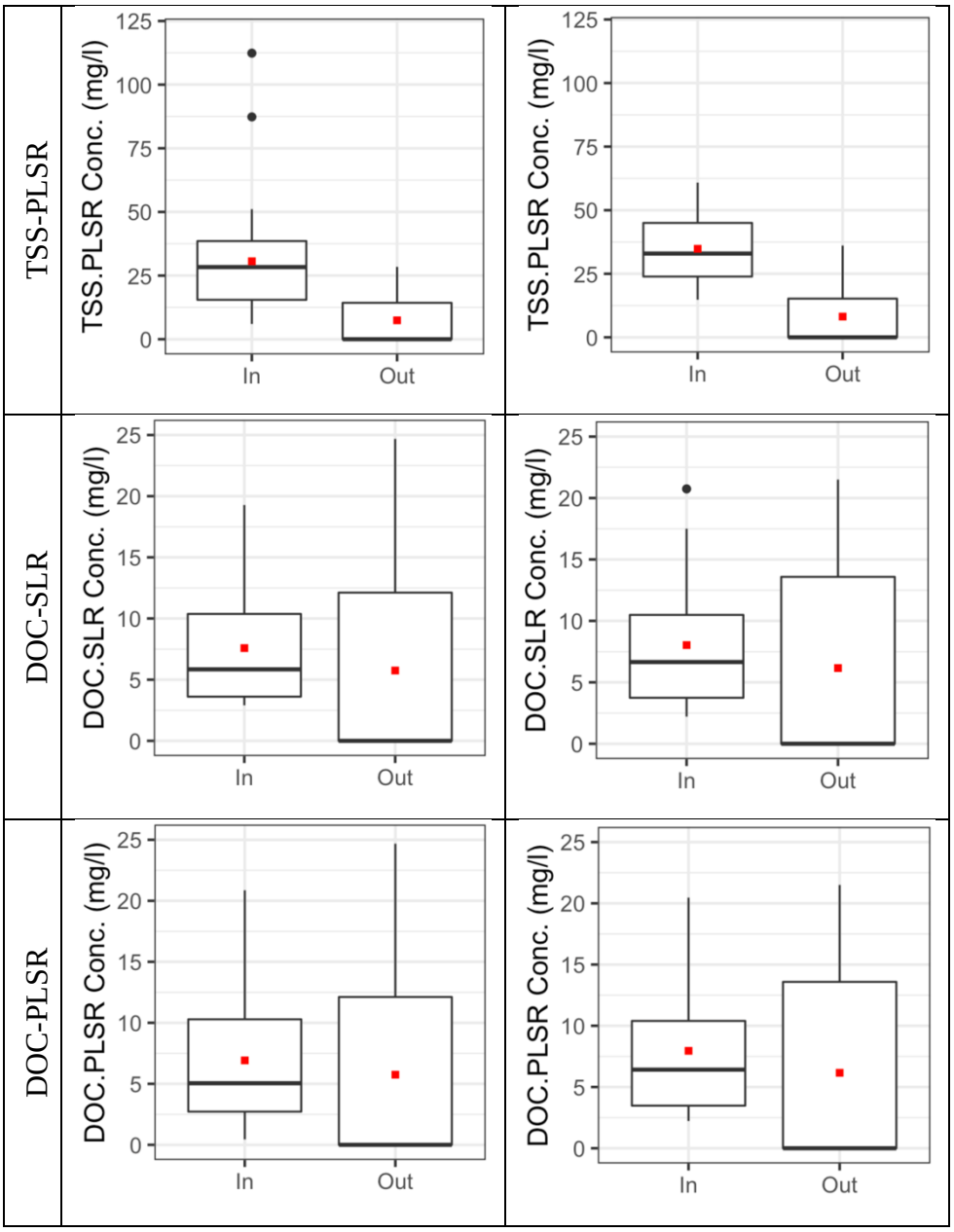

TSS-PLSR
DOC-SLR
DOC-PLSR
TSS.PLSR Conc. (mg/l)
DOC.SLR Conc. (mg/l)
DOC.PLSR Conc. (mg/l)
In
Out

Table 3.24 CSW's cumulative loads and removal efficiencies (RE) pre-retrofit and post-retrofit

	Pollutant	Cumulative load In kg/ha/yr	Cumulative load Out kg/ha/yr	Cumulative load RE%	Load reduction kg/ha/yr	Load reduction per SCM area g/m2/day
Pre-retrofit	NO_3^-- SLR	0.62	0.08	88	0.5	0.007
	NO_3^-- PLSR	0.52		86	0.4	0.005
	TSS - SLR	38.15	9.32	76	28.8	0.348
	TSS - PLSR	37.22		75	27.9	0.337
	DOC - SLR	16.55	6.85	59	9.7	0.117
	DOC - PLSR	16.42		58	9.6	0.116
Post-retrofit	NO_3^-- SLR	0.36	0.05	85	0.6	0.007
	NO_3^-- PLSR	0.33		84	0.5	0.007
	TSS - SLR	34.26	7.81	77	52.3	0.632
	TSS - PLSR	29.84		74	43.6	0.526
	DOC - SLR	10.43	5.29	49	10.2	0.123
	DOC - PLSR	10.62		50	10.5	0.127

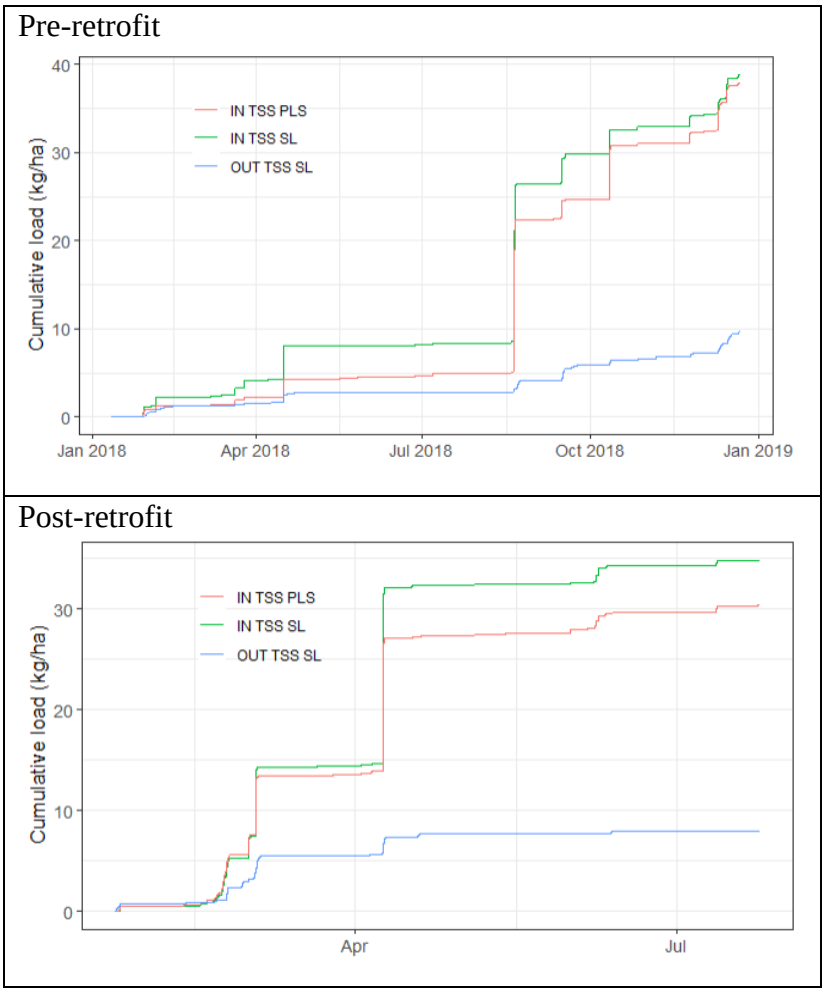


Figure 3.32 CSW's cumulative TSS load reduction pre- and post-retrofit

3.3.2.2 Wet pond

Pre-retrofit 7 events were monitored for water quality all with paired outflow samples, and post-retrofit 24 events were monitored for water quality with 15 paired samples. Generally, post-retrofit inflow events had higher median concentrations for all pollutants except for DOC, and outflow events had lower median concentrations except those for TDN (Table 3.25; *Figure 3.33*). The outflow EMC were more variable post-retrofit (*Figure 3.33*), which can be due to concomitant variability of retention times during that period (Figure 3.29).

Comparison of in/outflow EMCs for all pollutants showed a significant removal for all pollutants except DOC post-retrofit (*Table 3.25*). Pre-retrofit, the wet pond significantly removed

only TDN (*Table 3.25*). Median event concentration removal efficiencies improved post-retrofit by 7% for TDN, 25% for NH_4^+-N, 33% for NO_3^--N, 100% for PO_4^{3-}-P, 18% for TSS, and 8 % for DOC. Mean event concentration removal efficiencies increased post-retrofit only for PO_4^{3-}-P, TSS, and DOC, but decreased for all forms of nitrogen (*Table 3.25*). The improvement in median concentration removal efficiencies is a result of the number of no-outflow events post-retrofit (i.e., events with 100% removal efficiency). Additionally, the decreased mean concentration removal efficiencies for some pollutants can be ascribed to negative volume reduction for select events post-retrofit. Pre-retrofit pollutant removals were insignificant (*Table 3.25*), likely due to (1) lower inflow concentrations [removal efficiency is a function of inflow concentrations, and lower inflow concentrations are often related with lower removal efficiency (Barrett, 2005)], (2) inflow concentrations being closer to the irreducible concentrations of the SCM (Schueler and Holland 2000; Strecker et al. 2001), and (3) a limited number of paired samples.

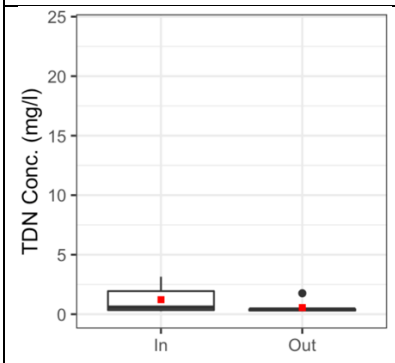
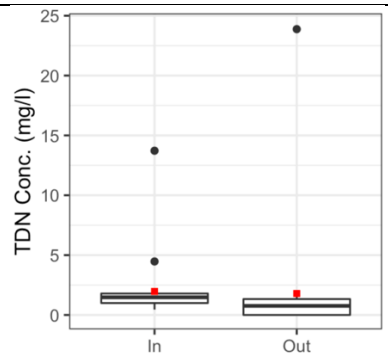
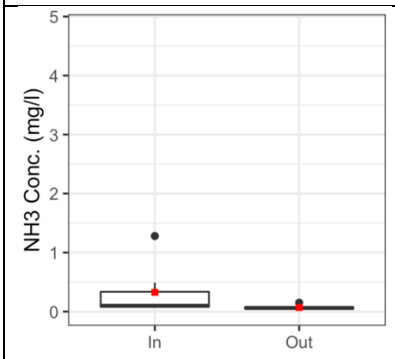
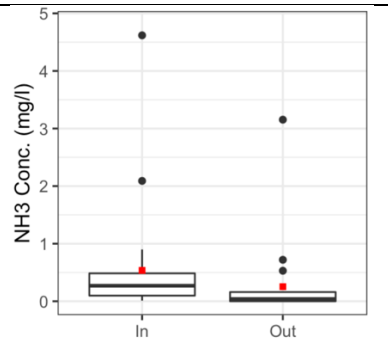
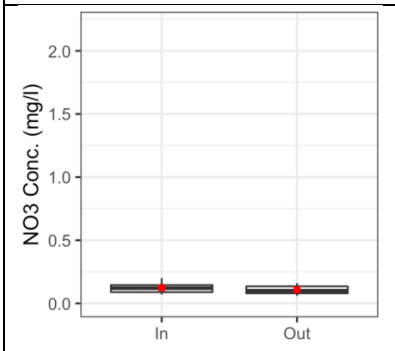
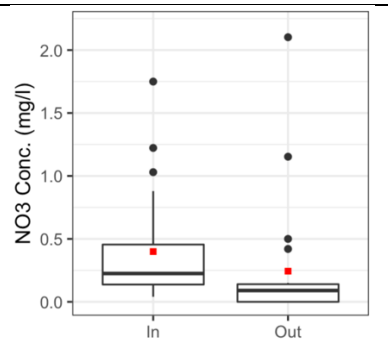
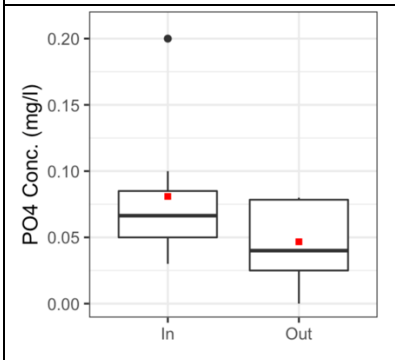
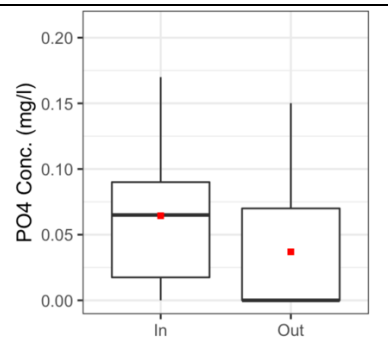
Cumulative pollutant load removal efficiencies decreased post-retrofit, while the annual load reduction per watershed area was similar pre- to post-retrofit (*Table 3.26*; *Figure 3.34*). Although, implementing RTC has increased the retention time by a factor of ~4 on average, the presented results do not support a clear improvement in pollutant removal. Comparison of event basis TSS removal with NC required performance standards suggest tests of removal efficiency would have been considered invalid by NCDEQ (2017) due to low influent concentrations (Supplemental material 3-C).

Table 3.25 Comparison of EMC and RE values of all pollutants pre and post retrofit at wet pond

	Pollutant	Median event Conc. In (mg/l)	Mean event Conc. In (mg/l)	Median event Conc. Out (mg/l)	Mean event Conc. Out (mg/l)	Median event Conc. RE%	Mean event Conc. RE%	Inflow to outflow Conc. p-value	Type of significant difference test
Pre-retrofit	TDN	0.57	1.22	0.41	0.56	28	33	**3.12E-02**[1]	Samples sign test
	NH_4^+-N	0.10	0.33	0.06	0.07	50	38	1.25E-01	Samples sign test
	NO_3^--N	0.12	0.12	0.10	0.11	12	10	8.99E-02	Student's t-test
	PO_4^{3-}-P	0.06	0.08	0.04	0.05	0	21	2.44E-01	Student's t-test
	TSS	12.50	16.52	6.72	6.92	66	30	6.06E-02	Student's t-test
	DOC	3.95	4.95	4.10	4.55	2	0	5.04E-01	Student's t-test
Post-retrofit	TDN	1.47	1.96	0.76	1.79	35	30	**9.22E-03**	Wilcoxon signed rank
	NH_4^+-N	0.27	0.54	0.04	0.25	75	19	**1.18E-02**	Samples sign test
	NO_3^--N	0.23	0.40	0.09	0.24	45	5	**3.09E-03**	Wilcoxon signed rank
	PO_4^{3-}-P	0.07	0.06	0.00	0.04	100	28	**4.78E-02**	Wilcoxon signed rank
	TSS	15.17	26.94	2.59	7.61	84	61	**4.01E-05**	Samples sign test
	DOC	3.62	5.09	3.11	4.96	11	17	5.08E-02	Wilcoxon signed rank

1Bold values are significant.

Figure 3.33 Wet pond's in/outflow EMCs during pre- and post-retrofit

Pre-retrofit concentration	Post-retrofit Concentration
	
	
	
	

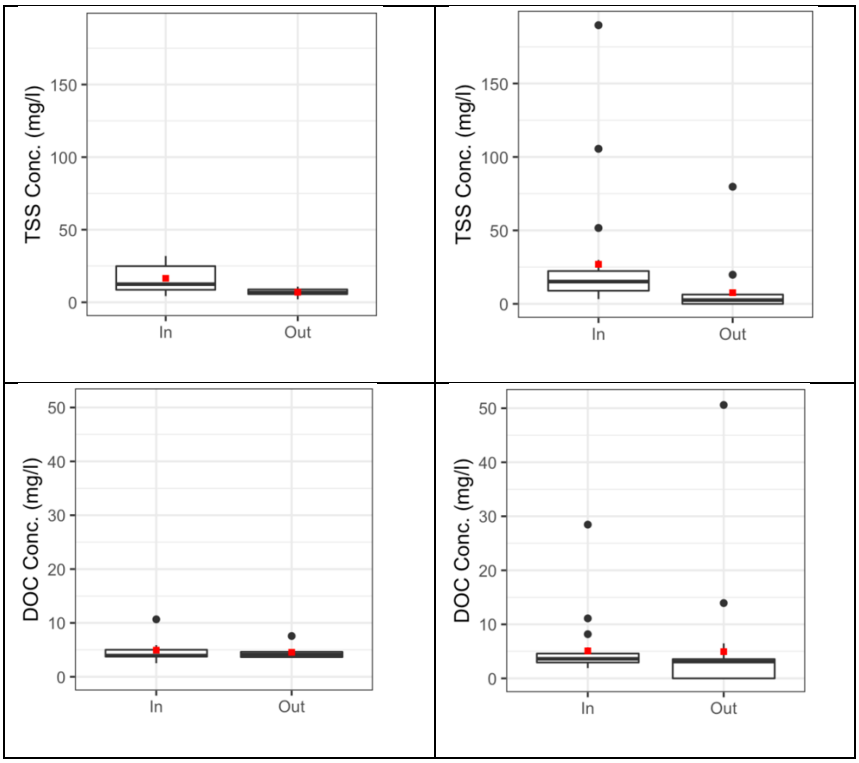

TSS Conc. (mg/l)
150
100
50
0
In
Out
TSS Conc. (mg/l)
150
100
50
0
In
Out
DOC Conc. (mg/l)
50
40
30
20
10
0
In
Out
DOC Conc. (mg/l)
50
40
30
20
10
0
In
Out

Table 3.26 Wet pond cumulative loads and removal efficiencies (RE) pre-retrofit and post-retrofit

	Pollutant	Cumulative load In kg/ha/yr	Cumulative load Out kg/ha/yr	Cumulative load RE%	Load reduction kg/ha/yr	Load reduction per SCM area g/m2/day
Pre-retrofit	TDN	9.38	2.46	74	6.9	0.0337
	NH_4^+-N	1.66	0.37	78	1.3	0.0063
	NO_3^--N	1.07	0.51	53	0.6	0.0028
	PO_4^{3-}-P	0.76	0.34	56	0.4	0.0021
	TSS	160.12	36.75	77	123.4	0.6004
	DOC	43.04	25.32	41	17.7	0.0862
Post-retrofit	TDN	13.20	6.53	51	6.7	0.0325
	NH_4^+-N	3.35	1.09	68	2.3	0.0110
	NO_3^--N	1.76	1.03	41	0.7	0.0035
	PO_4^{3-}-P	0.30	0.20	33	0.1	0.0005
	TSS	154.05	34.92	77	119.1	0.5798
	DOC	29.50	17.66	40	11.8	0.0576

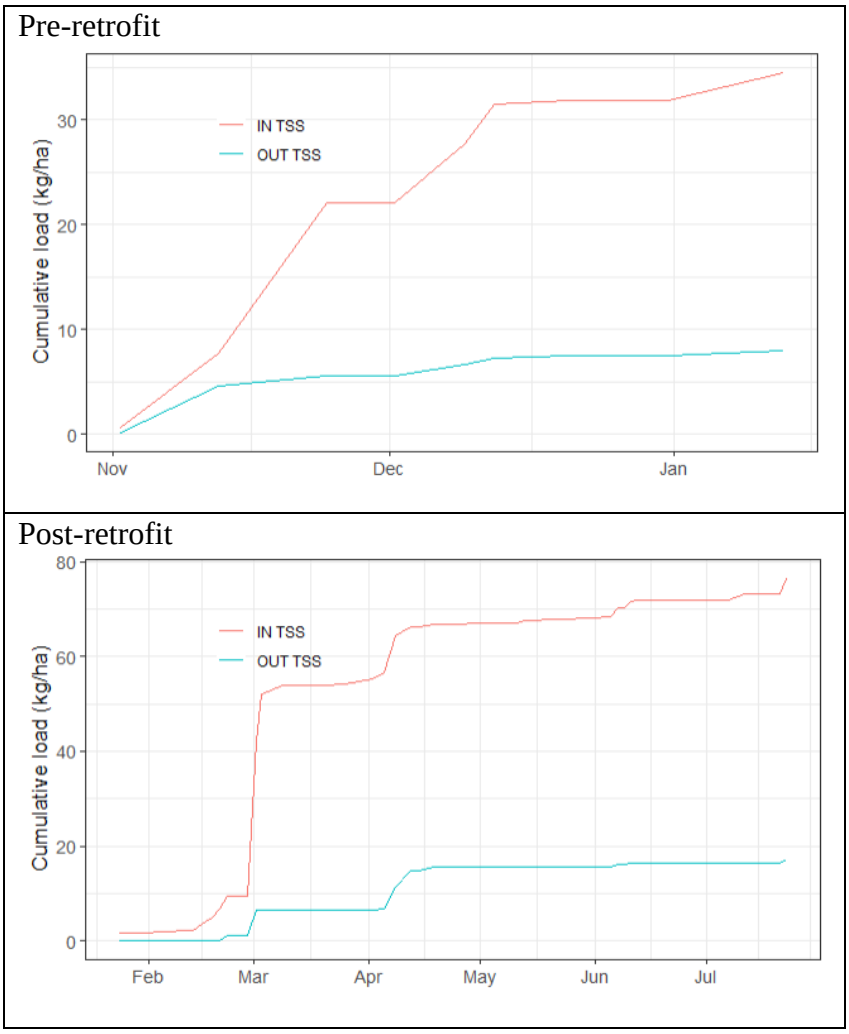


Figure 3.34 Wet pond cumulative TSS load reduction pre- and post-retrofit

3.3.3 RTC installation and operation challenges

The installation of the RTC's actuated valve might require the emptying of an SCM depending on the outlet structure and selected RTC configuration. In this study the RTC valve was installed on the emergency drain valve to provide full control of the entire SCM storage capacity. However, due to the outlet structure configuration at the wet pond (the drain valve placed at the pond side), the RTC valve had to be installed upstream of the outlet. This required the complete drainage of the wet pond for installation purposes. Given the size of the wet pond and precipitation pattern at that time, the pond was filled shortly after RTC installation.

However, completely emptying an SCM can impose further challenges in SCM functioning and monitoring.

Installation of the RTC's valve on the emergency drain can potentially cost more compared to installation on the orifice. The emergency drain of the CSW was buried under accumulated sediment/muck and required clean up and excavation. This was due to neglect of CSW maintenance.

Installation of an upturn elbow to provide a failsafe water level is critical when using RTC in CSWs. The RTC continuously monitored the SCM's status through the water level sensor and controlled the actuated valve based on its data. In this study, the water level sensor failed twice at the CSW, which led to (1) a partial and then (2) a complete drain of the CSW. The first was caused by a short circuit due to a cut in the wire covering at the end of the protective conduit, which might have occurred during installation. The cause for the other incident was unclear and the sensor provider suggested a lightning strike might have short circuited the sensor. In the case of the unexpected "complete" emptying of the CSW, the installed upturn elbow performed as a fail-safe and retained some water to support the aquatic ecosystem.

The RTC overestimated effluent volumes for some events which led to a large cumulative overestimation. RTC estimates effluent volume based on the percentage of valve open and the water level in the pond. Wet pond RTC-estimated discharges had cumulative effluent volumes higher than those of influent (effluent ≈ 1.13*influent; *Figure 3.35*). However, the cumulative effluent volume calculated using the bubbler sensor (at the wet pond) was much lower (effluent ≈ 0.54*influent; *Figure 3.35*). The RTC's overestimation of outflow can be caused by different factors such as (1) latency in cellular data transfer and thus less accurate time stamps of valve status and (2) less accurate readings of the valve opening due to accumulation of

debris. The effluent overestimation does not interfere with the RTC performance, because the RTC valve is controlled based on the storage capacity in the SCM and not the effluent volume. However, accurate measurement of outflow is important in evaluation RTC's hydrologic performance. Therefore, the authors suggest the use of a reliable separate sensor (such as a bubbler sensor) to collect local outflow data. Do not solely rely on RTC-estimated outflow.

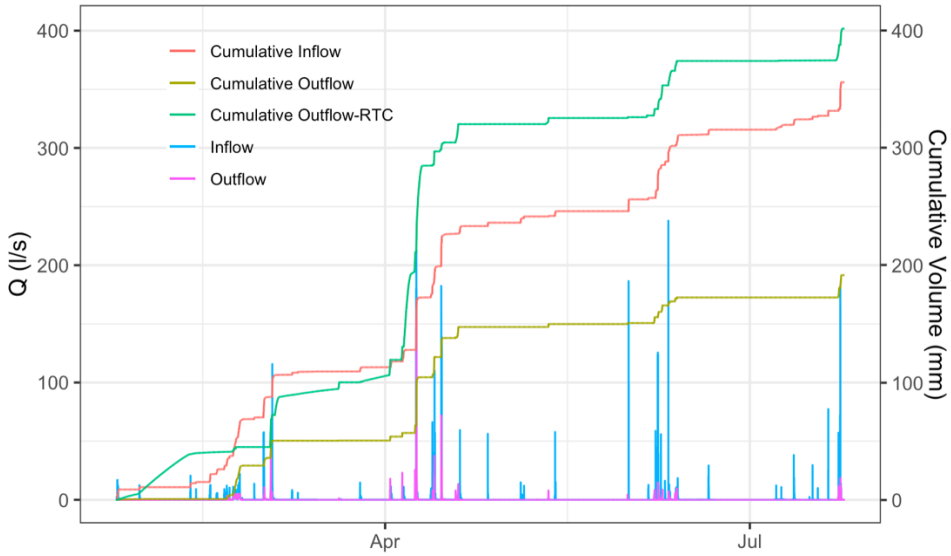


Figure 3.35 Post-retrofit wet pond cumulative influent and effluent, calculated from the data collected by the Bubbler sensor and RTC estimate

3.4 Conclusions

RTC implementation did not have any effect on the cumulative water balance of the studied SCMs. RTC prevented outflow from smaller events comprising 12.4% of total inflow volume at the CSW and 16% at the wet pond. Contradicting this benefit, some larger events had more outflow post-retrofit than pre-retrofit. Mean peak outflow rates remained essentially unchanged pre- to post-retrofit.

Post-retrofit the retention time (1) slightly decreased at the CSW, and (2) increased on average by a factor of ~4 at the wet pond. The observed decrease at the CSW was due to a high frequency of overflow. RTC underestimated event sizes for 86% of total inflow volume at the

CSW and 65% at the wet pond, which resulted in overflow during an event. Assuming the RTC runoff estimation algorithm is unlikely the cause, such underestimation can be due to (1) events that last longer than the 24-hour precipitation forecast window and therefore are not predicted in time for the RTC to react, (2) uncertainty of forecasted precipitation depth (Zaremba et al., 2019), (3) sensor failure causing miscalculation of SCM's available capacity, and (4) latency of cellular data delivery, causing the valve closure after the start of an event. Such underestimation and overflow resulted in shorter retention time for the events where the majority of outflow occurred as overflow.

Given the slight change in retention time at CSW, the retrofit had little to no effect on water quality treatment. However, the improved retention time at the wet pond led to improvement of SCM's average concentration removal efficiency for TSS, DOC and PO_4^{3-}P by 29%, 17%, and 7%, respectively. The cumulative annual load removal per watershed area remained the same post-retrofit for TSS and DOC.

Both SCMs had high volume, peak flow, and pollutant removals pre-retrofit as well as low influent concentrations. Both impacted the retrofit evaluation. The CSW reduced 55% of volume, mitigated 98% of peak flow, sequestered 58% TSS on a mean event basis, and lowered annual load per watershed area by 75%, similar values for the wet pond were 47%, 70%, 66%, and 77%, respectively. Given the SCMs performed so well pre-retrofit, the implementation of the retrofit had minor effect on their function. Additionally, the TSS effluent concentration were low (< 20 mg/l) for a number of events, which, per the state of North Carolina's assessment protocol (NCDEQ,2017) rendered pollutant removal tests ineffective.

The RTC improved retention time at the wet pond, and coupled with good performance of this SCM, suggests the potential for decreasing the SCM size. Post-retrofit the retention time

at the wet pond has improved by a mean factor of ~4 (a median factor of ~2). Therefore, it seems possible to substantially reduce the size of such wet ponds (with infiltration characteristics) while using RTC and maintain the same level of performance.

3.5 Acknowledgement

This study was supported by the North Carolina Clean Water Management Trust Fund. Special thanks to the following individuals for their notable support:
Shawn Kennedy and Samuel Garvey for their help with instrument installation and troubleshooting. Alisha Goldstein, the project engineer for her support in conducting the study. Mellow Marsh Farm, Inc for providing consultation on plant health and appropriate seasonal water level. OptiRTC for their technical support during the project. Cheng Chen for his help with field data collection.

REFERENCES

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., Vazquez-Aguirre, J.L., 2006. Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res. Atmos. https://doi.org/10.1029/2005JD006290.

Baird, J., Hunt, III, W., & Winston, R. (2014). Evaluating the Hydrologic and Water Quality Performance of Infiltrating Wet Retention Ponds. In World Environmental and Water Resources Congress 2014 (pp. 145-154).

Baird, J. B., Winston, R. J., & Hunt, W. F. (2020). Evaluating the hydrologic and water quality performance of novel infiltrating wet retention ponds. Blue-Green Systems, 2(1), 282-299.

Barrett, M. E. (2005). Performance comparison of structural stormwater best management practices. Water Environment Research, 77(1), 78–86. doi:10.2175/106143005X41654

Bean, E. Z. (2005). A field study to evaluate permeable pavement surface infiltration rates, runoff quantity, runoff quality, and exfiltrate quality.

Bellego, C., & Pape, L. D. (2019). Dealing with Logs and Zeros in Regression Models. Série des Documents de Travail, (2019-13).

Bilodeau, K., Pelletier, G., & Duchesne, S. (2018). Real-time control of stormwater detention basins as an adaptation measure in mid-size cities. Urban Water Journal, 15(9), 858-867.

Carpenter, J. F., Vallet, B., Pelletier, G., Lessard, P., & Vanrolleghem, P. A. (2014). Pollutant removal efficiency of a retrofitted stormwater detention pond. Water Quality Research Journal of Canada, 49(2), 124-134.

Casanova, M. T., & Brock, M. A. (2000). How do depth, duration and frequency of flooding influence the establishment of wetland plant communities?. Plant Ecology, 147(2), 237-250.

Etheridge, J. R., Birgand, F., Burchell, M.R., Smith, B. T. 2013. Addressing the fouling of in situ ultraviolet–visual spectrometers used to continuously monitor water quality in brackish tidal marsh waters. J. Environ. Qual, 42: 1896–1901.

Gaborit, E., Muschalla, D., Vallet, B., Vanrolleghem, P. A., & Anctil, F. (2013). Improving the performance of stormwater detention basins by real-time control using rainfall forecasts. Urban water journal, 10(4), 230-246.

Gaborit, E., Anctil, F., Pelletier, G., & Vanrolleghem, P. A. (2016). Exploring forecast-based management strategies for stormwater detention ponds. Urban Water Journal, 13(8), 841-851.

Gee, K. D., & Hunt, W. F. (2016). Enhancing stormwater management benefits of rainwater harvesting via innovative technologies. Journal of Environmental Engineering, 142(8), 04016039.

Gilpin, A., & Barrett, M. (2014). Interim report on the retrofit of an existing flood control facility to improve pollutant removal in an urban watershed. In World Environmental and Water Resources Congress 2014 (pp. 65-74).

Grant, D.M. and Dawson, B.D. (2001). Isco open channel flow measurement handbook. 5th edition. Isco, Inc., Lincoln, NE.

Greenway, M. (2004). "Constructed wetlands for water pollution control processes, parameters, and performance." Dev. Chem. Eng. Miner. Process., 12(5–6), 491–504.

Guo, J.C.Y., 2002. Overflow risk analysis for stormwater quality control basins. Journal of Hydrologic Engineering, 7 (6), 428–434, doi:10.1061/(ASCE)1084-0699(2002)7:6(428).

Hancock, G. S., Holley, J. W., & Chambers, R. M. (2010). A field-based evaluation of wet retention ponds: How effective are ponds at water quantity control? 1. JAWRA Journal of the American Water Resources Association, 46(6), 1145-1158.

Hollis, G. E. (1975). The effect of urbanization on floods of different recurrence interval. Water Resources Research, 11(3), 431-435.

Jacopin, C., Lucas, E., Desbordes, M., & Bourgogne, P. (2001). Optimisation of operational management practices for the detention basins. Water science and technology, 44(2-3), 277-285.

Jennings, D. B., & Jarnagin, S. T. (2002). Changes in anthropogenic impervious surfaces, precipitation and daily streamflow discharge: a historical perspective in a mid-Atlantic subwatershed. Landscape Ecology, 17(5), 471-489.

Kendon, E. J., N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, and C. A. Senior (2014), Heavier summer downpours with climate change revealed by weather forecast resolution model, Nat. Clim. Change, 4, 570– 576.

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., ... & Poresky, A. (2016). Smarter stormwater systems.

Knight, R. L., Walton, W. E., O'Meara, G. F., Reisen, W. K., & Wass, R. (2003). Strategies for effective mosquito control in constructed treatment wetlands. Ecological Engineering, 21(4-5), 211-232.

Kunkel, K. E., Karl, T. R., Brooks, H., Kossin, J., Lawrimore, J. H., Arndt, D., ... & Emanuel, K. (2013). Monitoring and understanding trends in extreme storms: State of knowledge. Bulletin of the American Meteorological Society, 94(4), 499-514.

Lefkowitz, J. R., Sarmanian, A. K., & Quigley, M. (2016). Continuous monitoring and adaptive control—the internet of things transforms stormwater management. Journal of the New England Water Environment Association, 50(1), 44-51.

Lenhart, H. A., & Hunt, W. F. (2011). Evaluating four storm-water performance metrics with a North Carolina Coastal Plain storm-water wetland. Journal of Environmental Engineering, 137(2), 155–162. doi:10.1061/(asce)ee.1943-7870.0000307

Leopold, L. B. (1973). River channel change with time: an example. Geological Society of America Bulletin, 84(6), 1845-1860.

Leopold, Luna B. 1994. A View of the River. Cambridge, MA: Harvard University Press.

Lewellyn, C., & Wadzuk, B. (2017). Evaluating Green Infrastructure Performance Using Real-Time Control from a Risk Perspective. Proceedings of the Water Environment Federation, 2017(6), 4726-4733.

Line, D. E., White, N. M., Osmond, D. L., Jennings, G. D., & Mojonnier, C. B. (2002). Pollutant export from various land uses in the Upper Neuse River Basin. Water Environment Research, 74(1), 100-108.Litton, R. B. (1968). Forest landscape description and inventories-a basis for land planning and design. Res.Paper PSW-RP-09. Albany, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture; 88 P, 49

Line, D. E., Jennings, G. D., Shaffer, M. B., Calabria, J., & Hunt, W. F. (2008). Evaluating the effectiveness of two stormwater wetlands in North Carolina. In 2008 Providence, Rhode Island, June 29–July 2, 2008 (p. 1). American Society of Agricultural and Biological Engineers.

Line, D. E., Brown, R. A., Hunt, W. F., & Lord, W. G. (2012). Effectiveness of LID for commercial development in North Carolina. Journal of Environmental Engineering, 138(6), 680-688.

Maniquiz-Redillas, M. C., Mercado, J. M. R., & Kim, L. H. (2013). Determination of the number of storm events representing the pollutant mean concentration in urban runoff. Desalination and Water Treatment, 51(19-21), 4002-4009.

Mann, M. E., Rahmstorf, S., Kornhuber, K., Steinman, B. A., Miller, S. K., & Coumou, D. (2017). Influence of anthropogenic climate change on planetary wave resonance and extreme weather events. Scientific Reports, 7, 45242.

Marchese, D., Johnson, J., Akers, N., Huffman, M., & Hlas, V. (2018). Quantitative comparison of active and passive stormwater infrastructure: Case study in Beckley, West Virginia. Proceedings of the Water Environment Federation, 2018(9), 4298-4311.

Marsalek, J. (2005). Evolution of urban drainage: from cloaca maxima to environmental sustainability (pp. 28-30). Burlington, ON: National Water Research Institute.

Martin, E. H., and Smoot, J. L. (1986). ‘‘Constituent-load changes in urban storm water runoff routed through a detention pond-wetland system in central Florida.’’ Water Resour. Investigation Rep. 85-4310, U.S. Geological Survey.

May, D., & Sivakumar, M. (2009). Optimum number of storms required to derive site mean concentrations at urban catchments. Urban Water Journal, 6(2), 107-113.

MDE (Maryland Department of the Environment). (2009). Maryland stormwater design manual, Water Management Administration, Baltimore, MD.

Merriman, L. S., & Hunt III, W. F. (2014). Maintenance versus maturation: Constructed storm-water wetland’s fifth-year water quality and hydrologic assessment. Journal of Environmental Engineering, 140(10), 05014003.

Miao, Weiwen, Yulia R. Gel, and Joseph L. Gastwirth. 2006. "A New Test of Symmetry about an Unknown Median." Pp 199-214 in Random Walk, Sequential Analysis and Related Topics - A Festschrift in Honor of Yuan-Shih Chow. Singapore: World Scientific Publisher, 199-214.

Middleton, J. R., & Barrett, M. E. (2008). Water Quality Performance of a Batch-Type Stormwater Detention Basin. Water Environment Research, 80(2), 172-178.

Miller, J.D., Hutchins, M., 2017. The impacts of urbanisation and climate change on urban flooding and urban water quality: a review of the evidence concerning the United Kingdom. J. Hydrol. Reg. Stud. https://doi.org/10.1016/j.ejrh.2017.06.006.

Muschalla, D., Vallet, B., Anctil, F., Lessard, P., Pelletier, G., & Vanrolleghem, P. A. (2014). Ecohydraulic-driven real-time control of stormwater basins. Journal of hydrology, 511, 82-91.

NCDENR (North Carolina Department of Environment and Natural Resources). (2009). Stormwater best management practices manual, Div. of Water Quality, Raleigh, NC.

NCDEQ (2017). North Carolina Stormwater Control Measure Credit Document. North Carolina Department of Environmental Quality: Raleigh, NC, USA. Retrieved from https://files.nc.gov/ncdeq/Energy%20Mineral%20and%20Land%20Resources/Stormwater/BMP%20Manual/SSW-SCM-Credit-Doc-20170807.pdf

NCDEQ. (2018). Stormwater Design Manual. Retrieved from https://deq.nc.gov/sw-bmp-manual

PADEP (Pennsylvania Department of Environmental Protection). (2006). Pennsylvania stormwater best management practices manual, Bureau of Waterways Engineering and Wetlands, Harrisburg, PA.

Papa, F., Adams, B., and Guo, Y., 1999. Detention time selection for stormwater quality control ponds. Canadian Journal of Civil Engineering, 26 (1), 72–82, doi:10.1139/l98-046.

Pettersson, T. J. R. (2002). Characteristics of suspended particles in a small stormwater pond. 9th International Conference on Urban Drainage, Portland, Oregon, USA.

Pleau, M., Colas, H., Lavallée, P., Pelletier, G., & Bonin, R. (2005). Global optimal real-time control of the Quebec urban drainage system. Environmental Modelling & Software, 20(4), 401-413.

Quigley, M., Rangarajan, S., Pankani, D., & Henning, D. (2008). New directions in real-time and dynamic control for stormwater management and low impact development. In World Environmental and Water Resources Congress 2008: Ahupua'A (pp. 1-7).

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Roman, D., Braga, A., Shetty, N., & Culligan, P. (2017). Design and modeling of an adaptively controlled rainwater harvesting system. Water, 9(12), 974.

Rosenberg, E. A., Keys, P. W., Booth, D. B., Hartley, D., Burkey, J., Steinemann, A. C., & Lettenmaier, D. P. (2010). Precipitation extremes and the impacts of climate change on stormwater infrastructure in Washington State. Climatic change, 102(1-2), 319-349.

s::can Messtechnik GmbH .(2019). spectro::lyser datasheet. https://www.s-can.at/plugins/system/vm_multiupload_attachment/js/multiupload/server/uploads/spectrolyser_ww_EN.pdf

Schmitt, Z. K., Hodges, C. C., & Dymond, R. L. (2020). Simulation and assessment of long-term stormwater basin performance under real-time control retrofits. Urban Water Journal, 1-14.

Schueler, T. R., and Holland, H. K. 2000. "Article 65: Irreducible pollutant concentrations discharged from stormwater practices." The practice of watershed protection, The Center for Watershed Protection, Ellicott City, The Center for Watershed Protection, Ellicott City, Md., 377–380.

Schütze, M., Campisano, A., Colas, H., Schilling, W., & Vanrolleghem, P. A. (2004). Real-time control of urban wastewater systems-where do we stand today? Journal of Hydrology. 299:335-348

Schwartz, D., Sample, D. J., & Grizzard, T. J. (2017). Evaluating the performance of a retrofitted stormwater wet pond for treatment of urban runoff. Environmental monitoring and assessment, 189(6), 256.

Shishegar, S., Duchesne, S., & Pelletier, G. (2018). Optimization methods applied to stormwater management problems: a review. Urban Water Journal, 15(3), 276-286.

Smolek, A. P., Hunt III, W. F., & Grabow, G. L. (2015). Influence of drawdown period on overflow volume and pollutant treatment for detention-based stormwater control measures in Raleigh, North Carolina. Journal of Sustainable Water in the Built Environment, 1(2), 05015001.

Stinson, M., & Vitasovic, Z. C. (2006). Real time control of sewers: US EPA manual. In World Environmental and Water Resource Congress 2006: Examining the Confluence of Environmental and Water Concerns (pp. 1-8).

Strauss, M., & Lefkowitz, J. (2017). Real-time Nitrogen and TSS Removal in an Active Control Wet Pond. Proceedings of the Water Environment Federation, 2017(6), 4734-4746.

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

Trimble, S. W. (1997). Contribution of stream channel erosion to sediment yield from an urbanizing watershed. Science, 278(5342), 1442-1444.

Vaze, J. and Chiew, F. H. S. (2004). Nutrient loads associated with different sediment sizes in urban stormwater and surface pollutants. Journal of Environmental Engineering-ASCE, 130(4), 391-396.

Vergeynst, L., Vallet, B., and Vanrolleghem, P.A., 2012. Modeling pathogen fate in stormwaters by a particle pathogen interaction model using population balances. Water Science and Technology, 65 (5), 823–832.

Walker, David J. 1998. "Modelling Residence Time in Stormwater Ponds." Ecological Engineering 10(1998):247–62.

Walsh, C. J. (2000). Urban impacts on the ecology of receiving waters: a framework for assessment, conservation and restoration. Hydrobiologia 431(2-3), 107-114.

Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., & Morgan, R. P. (2005). The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society, 24(3), 706-723.

Walsh, C. J., Fletcher, T. D., & Ladson, A. R. (2009). Retention capacity: A metric to link stream ecology and storm-water management. Journal of Hydrologic Engineering,14, 399–406.

Wuebbles, D., Fahey, D., Hibbard, K. A., Dokken, D. J., Stewart, B. C., & Maycock, T. K. (2017). US GLOBAL CHANGE RESEARCH PROGRAM CLIMATE SCIENCE SPECIAL REPORT (CSSR).

Zaremba, G., Bryant, S., Vacca, K., Wadzuk, B., (2019, May 19-23). Accounting for forecast uncertainty: using machine learning to optimize adaptive stormwater management [Conference presentation]. EWRI Congress, Pittsburgh, PA, United States. https://www.eventscribe.com/2019/ASCE-EWRI/fsPopup.asp?Mode=presInfo&PresentationID=481896

Supplemental material 3-A

As noted in the methods section, the outflow from the CSW was solely monitored by an AVM sensor post-retrofit. The calculated effluent from the CSW post-retrofit had cumulative vales higher than those of influent (effluent ≈ 1.06influent; *Figure 3.36*). RTC-estimated discharges had shown only slightly lower cumulative effluent volume (effluent ≈ 0.94influent; *Figure 3.36*). The high cumulative effluent volume could be (1) an effect of the retrofit, (2) overestimation due to a sensor issue (i.e., as noted in the methods section AVM sensors may not always capture stormwater fluctuations properly). Therefore, effluent volume was further investigated by comparing RTC-estimated effluent with observed values at the wet pond.

Wet pond RTC-estimated discharges had also shown cumulative effluent volumes higher than those of influent (effluent ≈ 1.13influent; *Figure 3.36*). However, the cumulative effluent volume calculated using the bubbler sensor (at the wet pond) was much lower (effluent ≈ 0.54influent; *Figure 3.36*). Therefore, on a cumulative basis, RTC-estimated discharge was overestimated 2.1 times (RTC-effluent ≈ 2.1effluent; *Figure 3.36*). Therefore, the CSW RTC-estimated effluent divided by the 2.1 correction factor was used as the effluent post-retrofit.

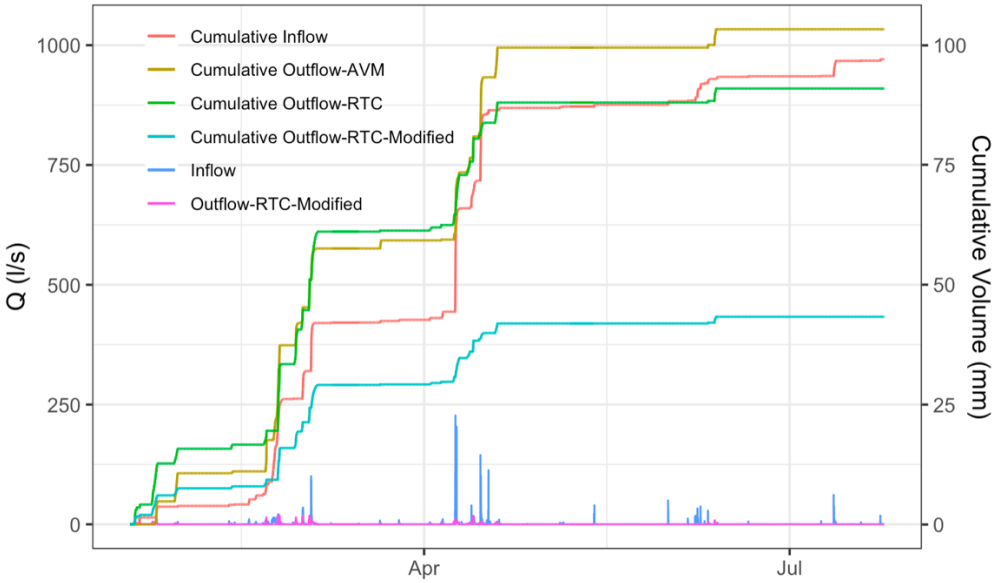


Figure 3.36 Post-retrofit CSW cumulative influent and effluent, calculated from the data collected by the AVM sensor, RTC estimate, and modified by a correction factor

Supplemental material 3-B

Two pre-retrofit events with extreme (perhaps odd) retention times at the CSW. These two observations were associated with (a) two events with less than 6 hours of dry weather in between, resulting in a negative retention time, or (b) continuous outflow that led to considering several inflows as pulses and thus a longer than usual retention time (*Figure 3.37*).

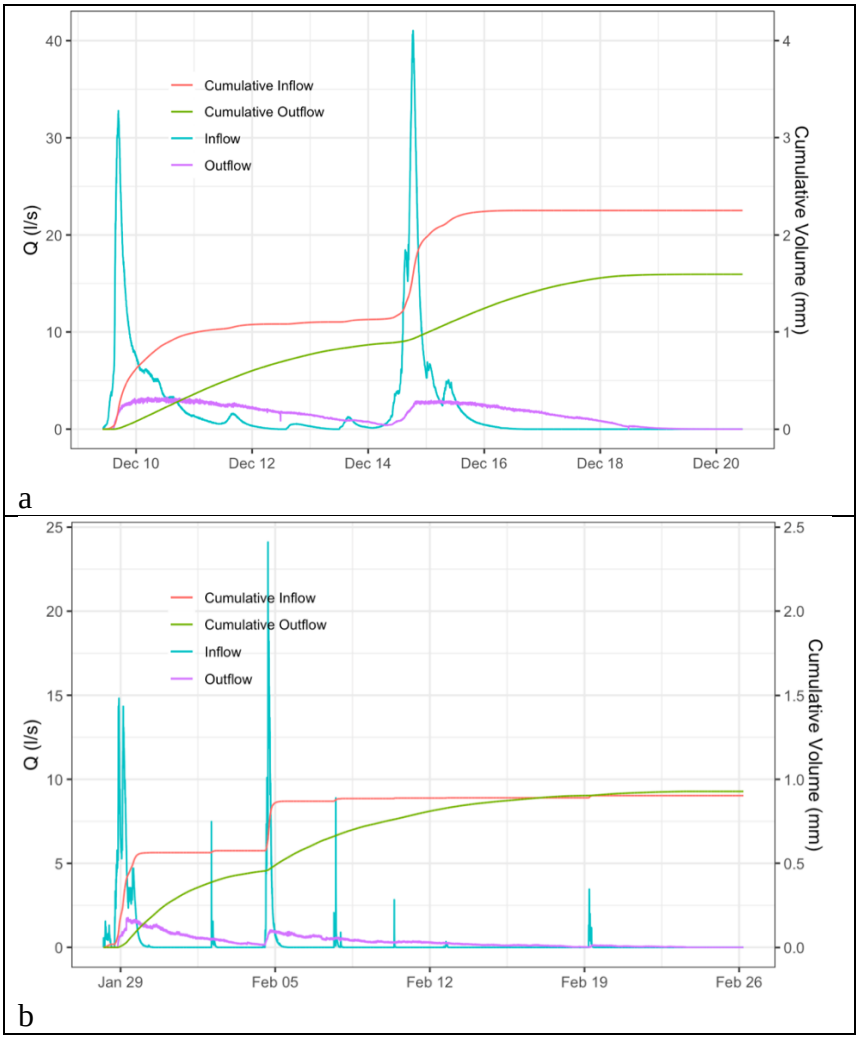


Figure 3.37 Pre-retrofit events with (a) negative retention time, and (b) extremely long retention time observed due to assumption of several inflow pulses as one event due to (a) less than 6-hour dry weather

Supplemental material 3-C

NCDEQ (2017) outlined a TSS removal performance standard for primary SCMs that is reflective of their field performance (*Table 3.27*). The standard removal efficiency varies based on influent TSS concentration. The removal efficiency test for influent EMC less than 20 mg/L is considered invalid. For higher influent concentrations, NCDEQ (2017) illustrates the standard effluent as a function of influent (illustrated as line *Figure 3.38*), rendering lower observations as meeting the requirement. *Figure 3.38* and *Figure 3.39* illustrate the effluent TSS EMCs against their corresponding influent values for all the paired samples at the CSW and the wet pond. Observations illustrated in gray ribbon indicate a test that is considered invalid. Observations below the standard line indicate a higher than standard removal efficiency and those above the line indicate that the SCM has not provided the required removal efficiency.

Table 3.27 TSS removal performance standard for primary SCMs (NCDEQ, 2017)

Median Influent EMC	Applicable Performance Standard
< 20 mg/L	Invalid test
20 – 35 mg/L	≥ 29% removal
35 – 100 mg/L	≤ 25 mg/L
100 mg/L	≥ 75% removal

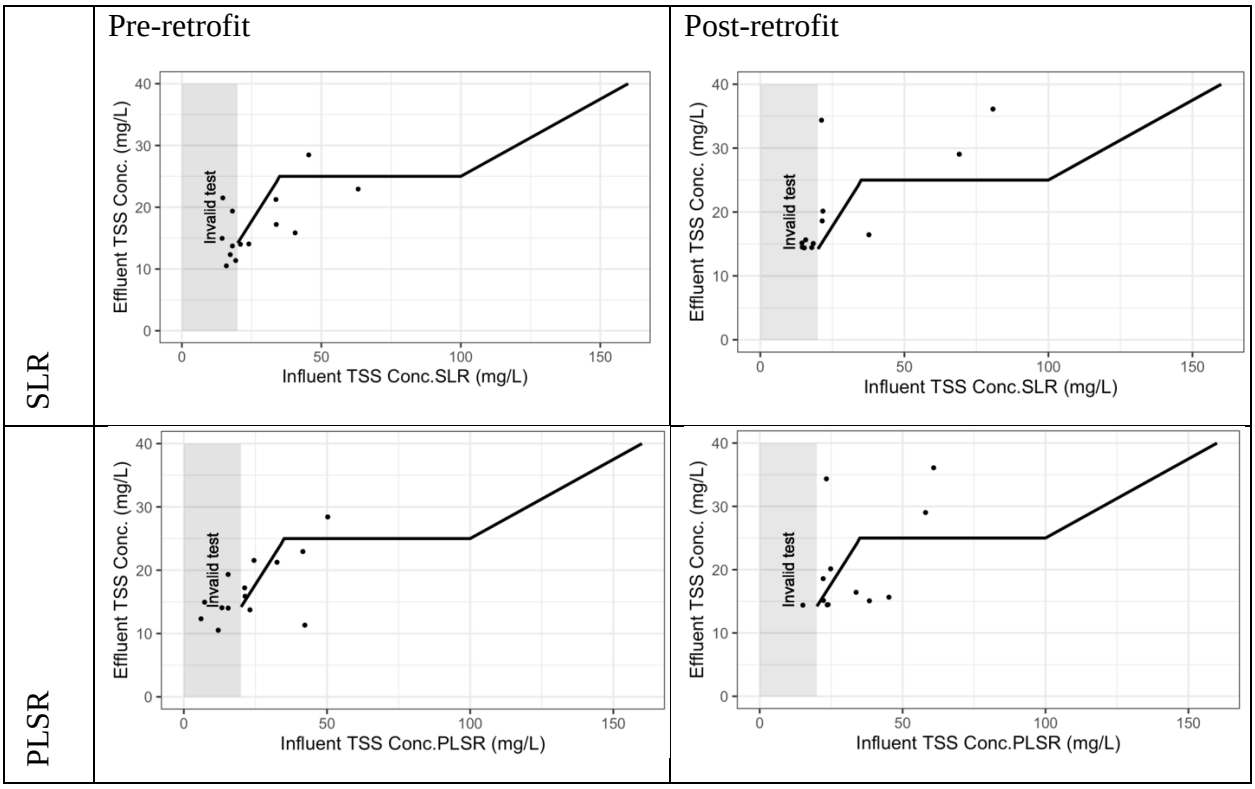


Figure 3.38 CSW event influent vs effluent TSS concentrations (dot) pre- and post-retrofit compared to NC required performance standard (line)

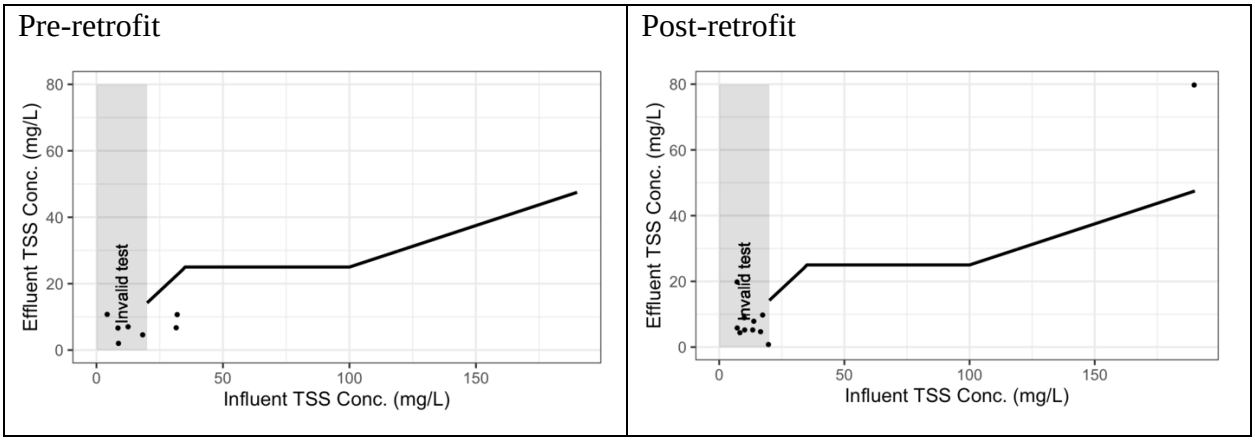


Figure 3.39 Wet pond event influent vs effluent TSS concentrations (dot) pre- and post-retrofit compared to NC required performance standard (line)

CHAPTER 4: EFFECT OF VISIBILITY ON MAINTENANCE INVESTMENT AND CONSEQUENT PERFOMANCE OF URBAN STORMWATER CONTROL MEASURES

Abstract

Studies on the performance of urban stormwater control measures (SCMs) mainly focus on hydrologic and biological factors. SCMs are located in an urban context and humans are part of this ecosystem, yet few studies have investigated the effect of human interaction on SCM performance. While SCM designs rarely encourage physical human interaction, their placement in the urban landscape does allow visual interaction. This study explores the impact of SCM visibility on the degree of maintenance received and, consequently, on the hydrologic performance of the system. Forty SCMs, including 20 bioretention cells and 20 wetlands or wet ponds, were assessed. Visibility was evaluated through SCM surveys to determine viewshed size, noticeability, and potential passer-by traffic. Hydrologic performance was evaluated through (1) visual inspection, (2) surveying vegetation health, (3) measuring drawdown rates, and (4) soil tests of bioretention media. As the degree of maintenance varied for each SCM, previous maintenance records, including cost data for the preceding year, were obtained and compared to visibility scores and hydrologic performance metrics. The results suggest that maintenance crews tend to prioritize more visible systems; however, based on the hydrologic performance, SCM priority did not significantly affect the quality of maintenance performed. Moreover, the SCMs examined tended to perform acceptably well. This finding is considered bias because of apparently competent and conscientious maintenance crews, as they hold a "SCM Inspection and Maintenance Certification". Further research is needed, using a greater number of maintenance crews, and controlling for crews with appropriate training. Ancillary

study findings (1) concluded that smaller practices (bioretention) were more expensive to maintain than larger practices (ponds and wetlands) on a per SCM-area basis and (2) findings underscore the importance of communication between the design community and the maintenance crew. As an example, because they are not aware that bioretention cell (BRCs) can drain too fast for effective nitrogen treatment, maintenance crews often assume a BRC with too a high drawdown rate is functioning well.

4.1 Introduction

Human-dominated landscapes are transformed due to increased impervious surfaces and decreased vegetation cover (Arnold and Gibbons, 1996; Brabec et al. 2002; Davis 2005; Pitt et al. 2008). Such transformations change the surface hydrology and water quality by increasing runoff volume and peak flow (Leopold 1973; Leopold 1994; Jennings and Jarnagin 2002), flooding (Leopold 1973; Hollis 1975) and pollutant loads (Trimble 1997; Line et al. 2002). Stormwater control measures (SCMs) are designed to mitigate the effect of anthropogenic alterations to the landscape by restoring (at least some of) the main characteristics of natural hydrology (Burns et al., 2012).

A key component to long-term hydrologic and water quality mitigation of SCMs, is ensuring their operability via regularly scheduled maintenance. Maintenance is critical to the functionality of SCMs as even the best-designed and best-constructed systems can fail without proper maintenance (Erickson et al. 2010; Hunt et al. 2011; Flynn et al. 2012). Consequently, research has expanded to the realms of maintenance to define key tasks, frequency, and costs associated with SCM maintenance (Kang et al. 2008; Asleson et al. 2009; Erickson et al. 2010; Wardynski & Hunt, 2012; Erickson et al. 2013; Houle et al. 2013; Johnson & Hunt, 2016; Al-Rubaei et al. 2017; Blecken et al. 2017). Key maintenance tasks ensure the functionality of SCMs. Some of these tasks include vegetation maintenance, pre-treatment (e.g., forebay) upkeep, outlet structure assessment, and media inspection (Blecken et al. 2017). Vegetation maintenance includes pruning of SCM-appropriate species and removing of other vegetation and dead plants to ensure plant health and prevent outlet clogging (Erickson et al. 2010, Blecken et al. 2017). The pre-treatment area should be inspected to avoid extensive sediment accumulation, clogging, and capacity reduction (Erickson et al. 2010; Johnson & Hunt, 2016; Al-Rubaei et al.

2017; Blecken et al. 2017). Inspection and unclogging of bioretention media facilitates infiltration and filtration, while preserving available storage (Johnson & Hunt, 2016; Blecken et al. 2017). Outlet structures should also be inspected to prevent leakage through a damaged structure (Al-Rubaei et al. 2017, Blecken et al. 2017).

Despite the emphasis on the importance of maintenance, SCMs are often neglected due to being dispersed over a wide area, budget limitations, or lack of knowledge (Blecken et al., 2017). Therefore, it is imperative that SCM maintenance professionals are educated on performance benchmarks that align with design and regulatory criteria. The North Carolina Stormwater Design Manual identifies maintenance as a "regulatory responsibility" and offers an "SCM Inspection and Maintenance Certification" program through N.C. State University (NCDEQ, 2018).

Blecken et al. (2017) suggested that future research should examine conditions that have led to proper (or lack of) maintenance, such as being visible to the public vs. being isolated systems. Isolated systems are considered to be SCMs that are either physically hidden (either by landscape or constructed objects) or that receive no passer-by traffic. Landscape visibility is defined by Ervin and Steinitz (2003) to be land surfaces that can be seen by an observer standing at a given point. In this definition, the observer and point of interest are interchangeable. As such, a point in an SCM also has a "viewshed" from which observers can see the practice.

Landscape visibility can be calculated by different software packages, such as geographic information systems (GIS), and is commonly used in urban planning for impact assessment and decision making (Fisher, 1996; Hernández et al. 2004; Yang et al. 2007).While the results of such computational evaluation reveal *potential* for visibility, a visible landscape is not necessarily "seen" or noticed (Ervin & Steinitz, 2003). Comparisons of computed versus field-

assessed visibility of lightly vegetated sites have shown an average of 50% accuracy for computed assessments (Maloy & Dean, 2001). Advances in spatial data manipulation likely improve this accuracy. However, the computational approach still insufficiently accounts for factors other than topographic changes, including atmospheric conditions (e.g., daylight, fog, clouds, Ogburn, 2006), speed of observer (e.g., walking, driving, Bell, 2019) and the observer's perception (e.g., culture, history, gender, psychology, (Lothian, 1999; Nassauer 1995; Ervin & Steinitz, 2003; Heft, 2013). Some visual features draw attention to a landscape: water bodies, elements of color or texture contrast, and informational signage (Ervin & Steinitz, 2003; Bell, 2019). The number of potential observers must also be accounted for (Litton, 1968). As humans are part of the watershed and ecosystem of an SCM, understanding the effect of human-SCM interactions (Grimm et al. 2000; Hull et al. 2015) can benefit the efficacy of SCMs. Current studies suggest that a positive experience of a landscape would encourage their maintenance (Tzoulas, 2007; Shandas 2015). The research herein explores whether SCM visibility impacts its maintenance and performance. As human-SCM interactions can be evaluated in different scales and types of interactions (e.g., Nassauer, 1995; Alberti and Marzluff, 2004; Tzoulas et al., 2007; Matsuoka and Kaplan, 2008; Green et al., 2012; Keeley et al., 2013), this study focuses specifically on the quantitative metric of visibility. Visibility of SCMs was characterized using the metrics of viewshed, noticeability, and traffic. Maintenance regimens for each SCM were assessed by interviewing maintenance professionals and examining records regarding (1) inspection and maintenance tasks, (2) performance/appearance benchmarks and (3) costs associated with maintaining an SCM. Lastly, visibility and maintenance of SCMs were compared to field evaluations of hydrologic performance. The results of this study can further inform the

design of SCMs and their placement in an urban environment, such that their effective lifetime is increased.

4.2 Methods

4.2.1 Site selection

Three types of SCMs were selected for examination: wetlands, wet ponds, and bioretention cells (BRCs). These practices are among the most common SCMs in North Carolina, due to design familiarity and each SCM's ability to remove pollutants. Forty individual SCMs were studied (twenty BRCs (in size range of 49-2479 m^2) and twenty wetlands (247-4193 m^2) or wet ponds (435-9145 m^2). All sites were either on North Carolina State University's campus or in the municipalities of Raleigh and adjacent Cary, NC, that receive 1101 mm average annual precipitation (NWS 2020). Two different maintenance crews, both of which employ many personnel who hold N.C. State University's "Stormwater Inspection and Maintenance" certification, were responsible for inspecting and maintaining these SCMs.

4.2.2 Performance evaluation

SCM performance was evaluated using four categories (*Table 4.1*; *Equation 4-1*; *Equation 4-2*): (1) visual inspection, (2) vegetation health, (3) drawdown rate and (4) hydric soil test (only BRCs). The selection of these assessment tools was based on the previous methods used for BRC performance evaluation (Asleson et al., 2009; Wardynski & Hunt, 2012) and for vegetation health assessment (Turk et al., 2014). Each SCM was evaluated using the aforementioned categories and given a 1-5 score. The final performance score of an SCM was the average of scores at each category, as shown in equations 1 and 2. Asleson et al. (2009) and Wardynski and Hunt (2012) also suggested assessing infiltration rate, which was tested herein

but not used to evaluate performance due to high variability of the results. Each assessment method is described in the following paragraphs.

Table 4.1 Categories used for performance evaluation at each SCM

Performance Evaluation Test	Bioretention	Wetland	Wet pond
Visual inspection	✓	✓	✓
Vegetation health	✓	✓	✓
Drawdown rate	✓	✓	✓
Hydric soil	✓	-	-

Equation 4-1

$$Performance\ Score(BRC) = Mean\ (Visual\ score, Vegetation\ score, Drawdown\ score, Soil\ score)$$

Equation 4-2

$$Performance\ Score(wetland\ or\ wet\ pond) = Mean\ (Visual\ score, Vegetation\ score, Drawdown\ score)$$

4.2.3 Visual inspection

Visual inspection includes checking for visual indicators of drainage problems as listed in *Table 4.2* (Asleson et al., 2009). Each of these visual indicators define the severity of impairment in 3 tiers (a, b, c; a = most severe impairment and c = least impaired) as it relates to the functionality of the SCM (*Table 4.2*). The final visual score was assigned based on the presence of the indicators and their tier of impairment. Finally, a synthesized visual inspection score accounting for all impairments was assigned (*Table 4.3*).

Each BRC was inspected for standing water 24 hours after a storm event of at least 20 mm (NCDEQ, 2018). Visual inspection entails checking for vegetation that is out-of-place (rather than vegetative health, which is described in a subsequent paragraph). For example,

BRCs are not intended to support obligate wetland species (e.g., cattail, arrowhead and marsh smartweed). Any BRC that contained these species was duly marked impaired.

Table 4.2 Valuation of visual indicators

Visual Indicator	Severity of Impairment*		
	Bioretention	Wetland	Wet pond
Standing Water	a	-	-
Sediment deposition	b	c	c
Clogged inlet	c	a	a
Clogged outlet	c	a	a
Erosion	b	c	c
No plants	b	a	-
Wetland Plants	a	-	-
Plant encroachment (outlet)	-	c	c

*a, b & c indicate the severity of impairment for the SCM with a = most severely impaired and c = least impaired.

Table 4.3 Overall impairment score for visual inspection

Indicator	Score
No indicators identified	5
One c tier impairment	4
Multiple c tier impairments	3
(At least) one b tier impairment	2
(At least) one a tier impairment	1

4.2.4 Vegetation health

Vegetation health was evaluated on a 1-5 scale, based on the presence of chlorosis or necrosis (Turk et al., 2014). Chlorosis is the discoloring of plant leaves due to lack of chlorophyll (Schuster, 2020), and necrosis is the death of plant tissue that reveals as dark/brown areas (Gunter, 2015).

Table 4.4 shows the indicators and scores used in this test, as taken from Turk et al. (2014). This test was performed in the months of July and August.

Table 4.4 Vegetation health score

Vegetation Test Scores	
Indicator	Score
No living material	1
Chlorosis/Necrosis > 50%	2
10% < Chlorosis/Necrosis < 50%	3
Chlorosis/Necrosis < 10%	4
Healthy Condition	5

(Turk et al., 2014).

4.2.5 Drawdown test

Based on NCDEQ (2018) regulations, SCMs should capture the design storm (25mm [1"] in the Piedmont and 38mm [1.5"] in the coastal counties). That volume is required to draw down at the rate of 25mm/h (1 in/hour) for bioretention and within 2-5 days for wetlands and wet ponds. Drawdown rate was measured once for each facility after a storm event of at least 20mm, using pressure transducers (HOBO U20L-04™). One pressure transducer was installed in the SCM to measure water elevation. By locating a second transducer to measure atmospheric pressure, the relative pressure of water was calculated. For BRCs, the sensor was placed in a shallow well such that the bottom of the sensor was at the same elevation as the top of the fill media. For wetlands and wet ponds, the sensor was placed near the outlet structure below the orifice elevation. The drawdown rate was calculated as the slope of the descending water level post-event (example in Supplemental material 4-A). The scoring index for drawdown rate is provided in

Table 4.5 and *Table 4.6*. Assigned scores were unacceptable (1), poor (2), good (4) and excellent (5). Because the authors decided that no moderately acceptable drawdown rate existed, no score of 3 was assigned. From a pollutant removal perspective, SCM drawdown rates can either be too fast or too slow; this is reflected in the assigned scores presented in and *Table 4.6*.

Table 4.5 Drawdown score for BRCs

Drawdown Score Index			
Drawdown rate (mm/h)	Drawdown rate (in/h)	Score*	Drawdown time
< 6.35	< 0.25	1	> 48 h
6.35 - 12.6	0.25 - 0.49	2	24-48 h
12.7 - 25.3	0.5 – 0.99	4	12-24 h
25.4 - 50.7	1.0 – 1.99	5	6-12 h
50.8 - 101.5	2.0 – 3.99	4	3-6 h
101.6 - 304.8	4.0 - 12	2	1-3 h
> 304.8	> 12	1	<1 h

*Assigned scores were unacceptable (1), poor (2), good (4) and excellent (5).

Table 4.6 Drawdown score for wetlands and wet ponds

Drawdown Score Index			
Drawdown rate (mm/h)	Drawdown rate (in/h)	*Score	Drawdown time
< 1.98	< 0.0781	1	> 8 days
1.98 - 3.17	0.0781 - 0.125	2	5-8 days
3.18 - 3.96	0.1251 - 0.1562	4	4-5 days
3.97 - 7.93	0.1563 - 0.3125	5	2-4 days
7.94 - 15.87	0.3126 - 0.625	2	1-2 days
> 15.87	> 0.625	1	< 1 day

* Assigned scores were unacceptable (1), poor (2), good (4) and excellent (5).

4.2.6 Soil test

The soil profiles were examined for each BRC to determine the existence of hydric conditions (*Table 4.7*). Samples were taken from the upper 30cm of soil, at three locations in the basin: (1) close to inlet, (2) close to the outlet and (3) in the middle of the BRC (example in Supplemental material 4-B). The texture of each soil layer was identified by the feel method (Thien, 1979), and layers were identified using Munsell Soil-color charts.

Table 4.7 Hydric soil scores

Soil Test Scores	
Number of samples with Hydric indication*	Score
0	5
1	3
2	2
3	1

* Out of 3 samples per basin at variable locations.

4.2.7 Maintenance data

Maintenance cost and time investment for the preceding year were obtained for each SCM from the staff that maintained that SCM. Some of the provided maintenance cost data also included the cost of equipment or occasional treatment based on the SCM's need (such as diverting construction runoff or replacing BRC media). The presented data herein only include the time investment (as measured in person-hours) of maintenance per square meter area of the SCM. Maintenance crews were also interviewed to note (1) the maintenance benchmarks they employed, (2) how crews evaluated SCM function and (3) whether they treated a high profile SCM differently. Maintenance crews considered some SCMs to be high profile because they were seen by influential people (e.g., mayor/university chancellor) or were part of high-profile landscapes (e.g., a student plaza). Hereon, this identified profile type is referred to as the maintenance priority of an SCM.

4.2.8 Visibility evaluation

Depending upon the practice, four or five stormwater professionals (engineers / trained technicians) visited each SCM and completed a survey to evaluate 3 tiers of visibility: (1) the size of viewshed, (2) noticeability of the SCM and (3) traffic (number of observers) for each SCM. These professionals were familiar with the sites and completed the survey based on the provided criteria by comparing the sites with each other, using their own observational judgment.

The survey was completed on-location for each SCM and included questions to encourage careful observation (*Table 4.8*). To evaluate the size of viewshed for an SCM, the survey takers were guided to identify the areas around the SCM that provided visual access with no obstruction. Each survey taker was then asked about the elements in or around the SCM that would suggest visual prominence or noticeability (such as trees, SCM size and proportion, a body of water, or an informational sign, *Table 4.8*). Finally, the potential numbers of passers-by to the SCM were estimated. Each of these estimations were scored by the survey takers on a 1-5 scale. For each SCM, each tier's score was determined by averaging that of all the survey participants. An SCM's final visibility score was the average score of the 3 tiers (*Equation 4-3*).

Equation 4-3

$$Visibility\ Score = Mean\ (Viewshed\ score, Noticeability\ score, Traffic\ score)$$

Table 4.8 Visibility evaluation questionnaire.

Question	Answer choices
Questions to encourage careful observation.	
Which elements exist in/around the SCM that would lead to its visual prominence (noticeability) against surrounding? (elements of color/contrast/texture)	Tree Vegetation Informational sign Water body Size/Proportion
Which features have visual access to the SCM?	Building Road Sidewalk/greenway
How do you rate the visual access to the SCM from each surrounding building? How do you rate the visual access to the SCM from each surrounding road? How do you rate the visual access to the SCM from each surrounding sidewalk?	No access Some access Moderate access Good access Excellent access
Question to evaluate each tier of visibility:	
How do you rate the size of viewshed for this SCM?	Very small (1) Small (2) Average (3) Large (4) Very large (5)
How do you rate the visual prominence (noticeability) of the SCM based on its elements?	Poor (1) Low (2) Moderate (3) High (4) Very high (5)
How do you rate the volume of traffic from each feature (building, road, sidewalk), with visual access to SCM? (consider direction and speed of observer)	No traffic (1) Some traffic (2) Moderate traffic (3) High traffic (4) Very high traffic (5)

4.2.9 Statistical analysis

The correlation between visibility and either maintenance cost or performance of SCMs was tested using a linear regression model (*Equation 4-4*).

Equation 4-4 $$\widehat{y_i} = \beta_0 + \beta_1 x_i + \varepsilon_i$$

where $\widehat{y_i}$: Performance score / Maintenance cost per square meter

x_i: Visibility score

β_0: intercept

β_1: regression coefficient

ε_i: residual error

For each regression model, the reported p-value and coefficient of determination (R^2) were used to evaluate the significance and degree of relationship between model parameters.

The Wilcoxon rank sum test was used to determine the effect of maintenance priority on performance and maintenance cost. This test was preferred over the two-sample t-test due to small number of observations (<30) in each maintenance priority group, which prevented the assumption of normality. Data analysis was performed using R statistical software and statistical significance was tested at $\alpha = 0.10$.

4.3 Results

The forty SCMs examined ranged from 2 to 27 years of age. Some systems had converted from one SCM type (e.g., a wet pond) to another (e.g., a wetland). "Converted" SCMs were evaluated based upon their current condition.

SCM performance score was the average score of visual assessment, vegetation health, soil test (BRCs only) and drawdown rates. Visibility score was calculated as the average scores for viewshed (1-5 scale), noticeability, and traffic. Maintenance cost was calculated on a per SCM surface area (square meters) basis. *Figure* 4.1 illustrates the relationship of visibility with performance (left) and maintenance cost (right). In general, practices examined were performing acceptably well (mean of 4.0 for high priority and 3.7 for low priority on a 1-5 scale). No strong relationship existed between visibility and performance (p-value = 0.51, R^2= 0.015) nor visibility

and maintenance cost (p-value = 0.17, R^2= 0.024). This might be due to bias introduced by the selection of study sites and/or the maintenance crews' level of training, as will be discussed later. On a per unit area of SCM basis, BRCs are more expensive to maintain than wetlands and wet ponds (p-value = 0.057); the maintenance cost of wetlands is not significantly different from that of wet ponds (p-value = 0.72) (*Figure* 4.1 right). Erickson et al. (2010) also found a higher annual maintenance cost associated with BRCs in comparison to that of wetland and wet ponds. Maintenance crews noted that based on the location prominence, some SCMs were considered high priority. *Figure* 4.2 shows the relationship of SCM priority with its visibility evaluation, maintenance cost, and performance. High priority SCMs were significantly more visible (p-value = 0.0033). High priority SCMs also tended to be more expensive to maintain (p-value = 0.20) and appeared to perform better (p-value = 0.23). Clearly, the latter two observations are not statistically supported. The lack of statistical significance for the latter two evaluations suggests that the competency and conscientiousness of the maintenance crews is important. Additionally, high priority SCMs had more variable performance.

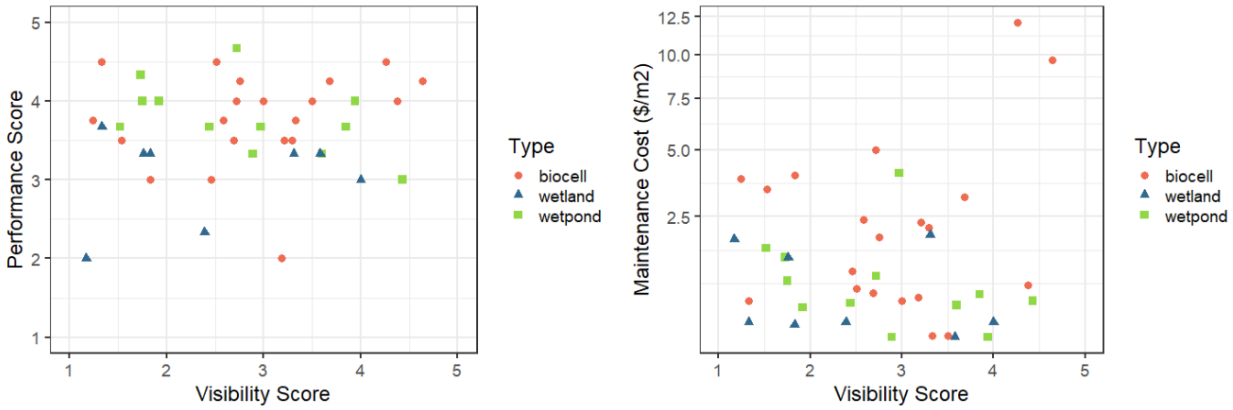


Figure 4.1 Visibility vs performance (left); visibility vs maintenance cost (right)

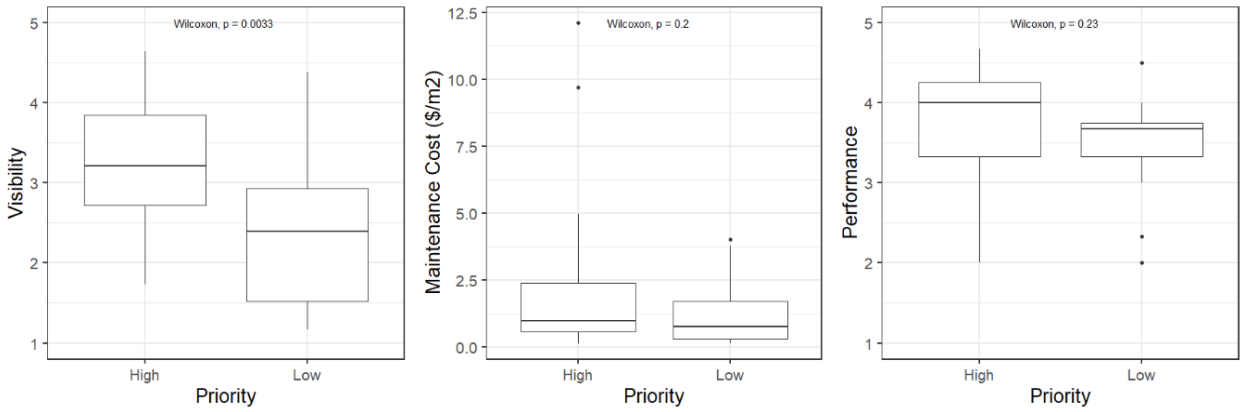


Figure 4.2 Relationship of maintenance priority with visibility(left), maintenance cost (middle) and performance (right).

Maintenance crews decide on tasks performed on each SCM based on their visual assessment of the device during site visits. Through visual assessment, impairments that can lead to hydrologic issues (such as flooding) are easier to identify than those leading to diminished treatment of water-borne pollutants (such as unintended leakage into a BRC outlet structure). The crews' maintenance benchmarks included tasks that would ensure both appearance of an SCM as a public amenity and hydrologic functioning of the SCM. However, hydrologic problems that did not cause flooding (such as too rapid a drawdown rate in BRCs which likely hampered water treatment) were often undetected. *Figure* 4.3 compares the drawdown rates of studied BRCs, related drawdown ratings and assigned priorities.

The background colors in *Figure 4.3* identify the score assigned to drawdown rates presented in

Table *4.5* and *Table 4.6* (orange(unacceptable)=1, yellow(poor)=2, green(good)=4, deep green (excellent)=5). Among the studied BRCs, 60% were considered high priority. About 16% of high priority BRCs had low drawdown rates and were thus assigned a low performance score of no more than 2 (*Figure 4.2*). The maintenance crews were able to (1) identify these low performing SCMs and (2) assign them to be renovated. In contrast, approximately 40% of high priority BRCs had too high of a drawdown rate and also received a low performance score due to

a likely inability to expose certain pollutants in runoff to their necessary removal mechanisms for a sufficient length of time (e.g., denitrification, thermal exchange, as per Hunt et al. (2012)). In these cases, maintenance crews assumed them to be well-functioning, and thus no restorative maintenance was assigned. The NC bioretention media "recipe" calls for moderate (50.8-101.6 mm/h or 2-4 in/h) drawdown rate, which threads the needle between providing sufficient contact time for pollutant treatment, while allowing bioretention cells to dry (NCDEQ, 2018; Hunt et al., 2012). Moreover, issues such as leakage through faulty outlet structures also lead to too rapid a drawdown rate. This emphasizes the importance of education and communication between SCM managers/designers and the maintenance crew about bioretention.

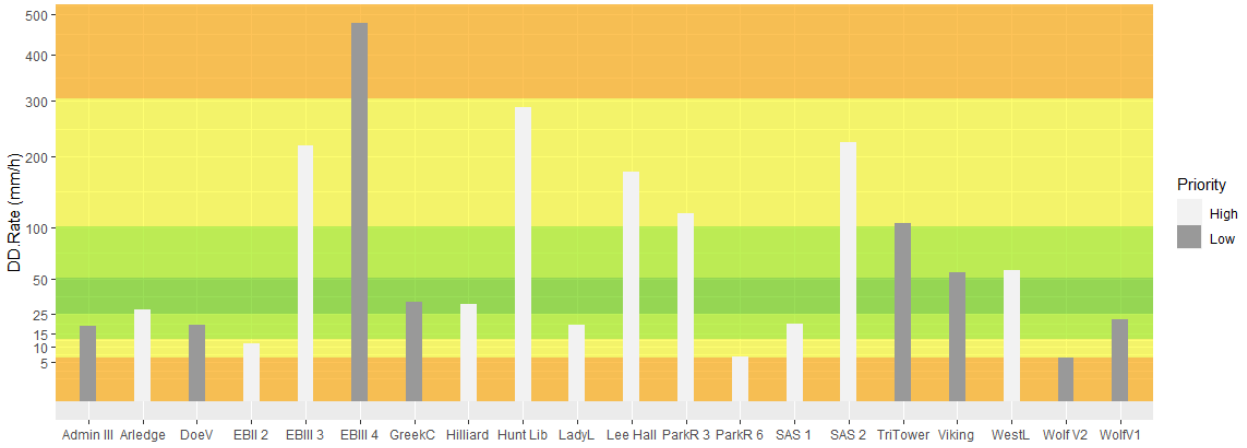


Figure 4.3 Drawdown rate (mm/h) of BRCs, (orange(unacceptable)=1, yellow(poor)=2, green(good)=4, deep green(excellent)=5).

Among the wetlands and wet ponds, 45% were assigned a high priority by maintenance crews. As shown in *Figure 4.4*, approximately 55% of high priority wetlands and wet ponds received low performance scores due to either very high or very low drawdown rates and therefore low scores. Both maintenance crews identified all but one of these SCMs as low functioning (the lone exception being an oversized wetland with low drawdown rate). A low drawdown rate score in the case of wetlands and wet ponds was given for either dewatering too

slowly (and thus reducing an SCM's capacity to capture runoff from subsequent events) or too quickly (as with bioretention, leading to potentially inadequate time allowed for treatment).

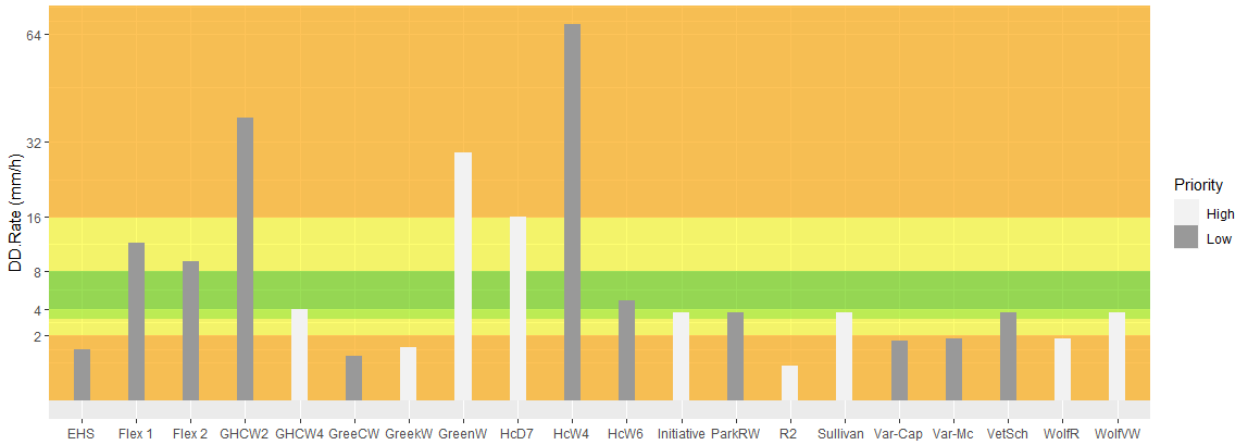


Figure 4.4 Drawdown rate (mm/h) of wetlands/wet ponds, (orange(unacceptable)=1, yellow(poor)=2, green(good)=4, deep green(excellent)=5).

Six of the 20 BRC's had at least one indicator of hydric soil (*Figure 4.5*). However, their drawdown rates were not uniformly critically low. These six BRCs appeared in 25% of high priority and in about 40% of low priority systems (*Figure 4.4*). This could be caused by either (1) leakage or (2) presence of heterogeneous media. Since there was no apparent leakage in outlet structure, the latter cause was assumed likely. Other studies have also reported heterogenous media in BRCs based on infiltration tests (Aslesson et al. 2009; Wardynski & Hunt, 2012). These systems were highly vegetated and surrounded by trees, with an age range of 5 to 14 years. A potential cause of soil profile variation is the decrease in media porosity due to accumulation of biofilm and shrinkage of organic matter (Rodgers et al. 2004).

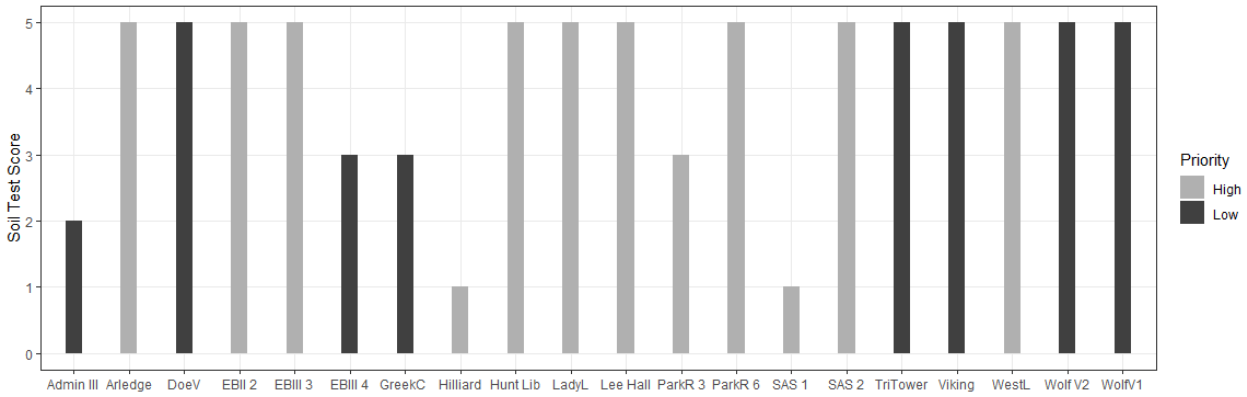


Figure 4.5 Soil test scores of BRCs and their maintenance priority (score1= 3 samples with hydric indicator, score2 = 2 samples with hydric indicator, score3 = 1 sample with hydric indicator, score5 = no hydric indicator).

4.4 Conclusions

Twenty BRCs and twenty wet ponds/wetlands were studied at NC State University; Raleigh, NC; and Cary, NC. The purpose of this examination was to determine if SCM visibility lead to a prioritization in maintenance, and if so, did these SCMs function better than their lower priority counterparts. The research focused on SCMs that were maintained by two well-trained maintenance crews and demonstrated that these crews do tend to prioritize highly visible SCMs, which translates to more money invested in the upkeep of high priority SCMs. Despite this additional maintenance investment, the performance (as based on hydrology, soil morphology and plant health) of the high priority (also highly visible) SCMs herein were not significantly different from that of those considered low priority. This suggests that both maintenance crews were competent and provided sufficient functional maintenance no matter the priority and visibility. Other studies have suggested lack of knowledge and unclear communication amongst stakeholders as a potential cause of improper maintenance (Blecken et al. 2017). This hypothesis was not supported on NC State University campus. But our findings reflect the situation at NC State University, and not necessarily of SCMs found elsewhere. The authors therefore conclude that this exercise should be repeated using a larger pool of maintenance firms and SCM owners.

Other important study findings include that smaller scale SCMs (bioretention, herein) are more costly to maintain on a per unit of SCM area basis than larger scale SCMs (wetlands and wet ponds). This finding supports those of earlier studies (e.g., Erickson et al. 2010). There was no difference in maintenance costs between ponds and wetlands. The typical SCM was generally performing acceptably well. High priority practices had a mean performance score of 4.0 (high of 5), while low priority SCMs' mean score was 3.7, which are quite similar and show that regardless of priority the SCMs in this study performed well and therefore were well maintenance.

Education of and communication with the maintenance crew, played an important role in how SCMs received maintenance. Both maintenance crews in this study hold the state's SCM maintenance certificate and their SCMs were generally well-functioning. However, SCM impairments that hamper water quality treatment proved harder to detect and in specific cases herein were missed by these crews. In this study, the maintenance crews assumed a BRC with too high a drawdown rate was functioning well, while it hindered effective nitrogen treatment. Keeping SCMs appropriately maintained is critical for long-term performance, because once the designer and contractor construct an SCM, the onus lies squarely on (informed) maintenance crews. Going forward it will be critical that maintenance benchmarks align with design and regulatory goals.

4.5 Acknowledgments

This project was made possible with the assistance of many NCSU faculty, staff and students, stormwater division in the cities of Raleigh and Cary, NC. Special thanks to the following individuals for their notable support:

William G. Lord and Sarah E. Waickowski for help with the visibility evaluation survey. Mick Ribault of Dragonfly Pond Works for providing research sites off-campus and financial support of the Dragonfly Pond Works Scholars program. Jacqueline Welles (Dragonfly Pond Works Scholar) for her help with off-campus data collection.

REFERENCES

Al-Rubaei, A. M., Merriman, L. S., Hunt III, W. F., Viklander, M., Marsalek, J., & Blecken, G. T. (2017). Survey of the operational status of 25 Swedish municipal Stormwater management ponds. Journal of Environmental Engineering, 143(6), 05017001.

Alberti, M., & Marzluff, J. M. (2004). Ecological resilience in urban ecosystems: linking urban patterns to human and ecological functions. Urban ecosystems, 7(3), 241-265.

Arnold Jr, C. L., & Gibbons, C. J. (1996). Impervious surface coverage: the emergence of a key environmental indicator. Journal of the American planning Association, 62(2), 243-258.

Asleson, B. C., Nestingen, R. N., Gulliver, J. S., Hozalski, R. M., & Nieber, J. L. (2009). Performance assessment of rain gardens. Herndon, VA: Assoc.

Bell, S. (2019). Elements of visual design in the landscape. Routledge.

Blecken, G., Hunt, W. F., Al-Rubaei, A. M., Viklander, M., & Lord, W. G. (2017). Stormwater control measure (SCM) maintenance considerations to ensure designed functionality. Urban Water Journal, 14(3), 278-290. doi:10.1080/1573062X.2015.1111913

Brabec, E., Schulte, S., & Richards, P. L. (2002). Impervious surfaces and water quality: a review of current literature and its implications for watershed planning. Journal of planning literature, 16(4), 499-514.

Burns, M. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., & Hatt, B. E. (2012). Hydrologic shortcomings of conventional urban stormwater management and opportunities for reform. Landscape and urban planning, 105(3), 230-240.

Davis, A.P. 2005. Green Engineering Principles: Promote Low-Impact Development. Environmental Science & Technology, 39(16), 338A-344A.

Erickson, A. J., Gulliver, J. S., Kang, J., Weiss, P. T., & Wilson, C. B. (2010). Maintenance for stormwater treatment practices. Journal of Contemporary Water Research & Education, 146(1), 75-82.

Erickson, A. J., Weiss, P. T., & Gulliver, J. S. (2013). Optimizing stormwater treatment practices. A Handbook of Assessment and Maintenance, 1(1), 1.

Ervin, S., & Steinitz, C. (2003). Landscape visibility computation: Necessary, but not sufficient. Environment and Planning B: Planning and Design, 30(5), 757-766.

Fisher, P. F. (1996). Extending the applicability of viewsheds in landscape planning. Photogrammetric engineering and remote sensing, 62(11), 1297-1302.

Flynn, K. M., Linkous, B. W., & Buechter, M. T. (2012). Operation and maintenance assessment for structural stormwater BMPs. Paper presented at the World Environmental and Water Resources Congress 2012: Crossing Boundaries, 3662-3673.

Gobster, P.H., Nassauer, J.I., Daniel, T.C. et al. (2007). The shared landscape: What does aesthetics have to do with ecology? Landscape Ecol.

Green, O. O., Shuster, W. D., Rhea, L. K., Garmestani, A. S., & Thurston, H. W. (2012). Identification and induction of human, social, and cultural capitals through an experimental approach to stormwater management. Sustainability, 4(8), 1669-1682.

Grimm NB, Grove MJ, Pickett STA, Redman CL. 2000. Integrated approaches to long-term studies of urban ecological systems. Bioscience 50:571–84

Gunter, C. C. (2015). Illustrated definitions of plant problems. Retrieved from https://extension.purdue.edu/extmedia/ID/ID-319-W.pdf

Heft, H. (2013). Environment, cognition, and culture: Reconsidering the cognitive map. Journal of Environmental Psychology, 33, 14-25.

Hernández, J., García, L., & Ayuga, F. (2004). Assessment of the visual impact made on the landscape by new buildings: a methodology for site selection. Landscape and Urban Planning, 68(1), 15-28.

Hollis, G. E. (1975). The effect of urbanization on floods of different recurrence interval. Water Resources Research, 11(3), 431-435.

Houle, J. J., Roseen, R. M., Ballestero, T. P., Puls, T. A., & Sherrard Jr, J. (2013). Comparison of maintenance cost, labor demands, and system performance for LID and conventional stormwater management. Journal of environmental engineering, 139(7), 932-938.

Hull, V., Tuanmu, M. N., & Liu, J. (2015). Synthesis of human-nature feedbacks. Ecology and Society, 20(3).

Hunt, W. F., Greenway, M., Moore, T. C., Brown, R. A., Kennedy, S. G., Line, D. E., & Lord, W. G. (2011). Constructed storm-water wetland installation and maintenance: Are we getting it right?. Journal of irrigation and drainage engineering, 137(8), 469-474.

Jennings, D. B., & Jarnagin, S. T. (2002). Changes in anthropogenic impervious surfaces, precipitation and daily streamflow discharge: a historical perspective in a mid-Atlantic subwatershed. Landscape Ecology, 17(5), 471-489.

Johnson, J. P., & Hunt, W. F. (2016). Evaluating the spatial distribution of pollutants and associated maintenance requirements in an 11 year-old bioretention cell in urban Charlotte, NC. Journal of environmental management, 184, 363-370.

Kang, J. H., Weiss, P. T., Gulliver, J. S., & Wilson, B. C. (2008). Maintenance of Stormwater BMPs Frequency, effort, and cost. Stormwater, 8, 18-28.

Keeley, M., Koburger, A., Dolowitz, D. P., Medearis, D., Nickel, D., & Shuster, W. (2013). Perspectives on the use of green infrastructure for stormwater management in Cleveland and Milwaukee. Environmental management, 51(6), 1093-1108.

Leopold, L. B. (1973). River channel change with time: an example. Geological Society of America Bulletin, 84(6), 1845-1860.

Leopold, Luna B. 1994. A View of the River. Cambridge, MA: Harvard University Press.

Line, D. E., White, N. M., Osmond, D. L., Jennings, G. D., & Mojonnier, C. B. (2002). Pollutant export from various land uses in the Upper Neuse River Basin. Water Environment Research, 74(1), 100-108.Litton, R. B. (1968). Forest landscape description and inventories-a basis for land planning and design. Res.Paper PSW-RP-09. Albany, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture; 88 P, 49

Lothian, A. (1999). Landscape and the philosophy of aesthetics: is landscape quality inherent in the landscape or in the eye of the beholder?. Landscape and urban planning, 44(4), 177-198.

Maloy, M. A., & Dean, D. J. (2001). An accuracy assessment of various GIS-based viewshed delineation techniques. Photogrammetric Engineering and Remote Sensing, 67(11), 1293-1298.

Matsuoka, R. H., & Kaplan, R. (2008). People needs in the urban landscape: analysis of landscape and urban planning contributions. Landscape and urban planning, 84(1), 7-19.

Middleton, J. R., & Barrett, M. E. (2008). Water quality performance of a Batch-Type stormwater detention basin. Water Environment Research, 80(2), 172-178.

Nassauer, J. I. (1995). Culture and changing landscape structure. Landscape ecology, 10(4), 229-237.

NCDEQ. (2018). Stormwater Design Manual. Retrieved from https://deq.nc.gov/sw-bmp-manual

National Weather Service. (2020). Climatological Report. Retrieved from https://forecast.weather.gov/product.php?site=NWS&issuedby=RDU&product=CLA&format=CI&version=1&glossary=0&highlight=off

Ogburn, D. E. (2006). Assessing the level of visibility of cultural objects in past landscapes. Journal of Archaeological Science, 33(3), 405-413.

Pitt, R., Chen, S. E., Clark, S. E., Swenson, J., & Ong, C. K. (2008). Compaction's impacts on urban storm-water infiltration. Journal of Irrigation and Drainage Engineering, 134(5), 652-658.

Schuster, J. (2020). Focus on plant problems. Retrieved from https://web.extension.illinois.edu/focus/index.cfm?problem=chlorosis

Shandas, V. (2015). Neighborhood change and the role of environmental stewardship: a case study of green infrastructure for stormwater in the City of Portland, Oregon, USA. Ecology and Society, 20(3).

Trimble, S. W. (1997). Contribution of stream channel erosion to sediment yield from an urbanizing watershed. Science, 278(5342), 1442-1444.

Turk, R. L., Kraus, H. T., Bilderback, T. E., Hunt William, F., & Fonteno William, C. (2014). Rain garden filter bed substrates affect stormwater nutrient remediation. Hort Science, 49(5), 645-652.

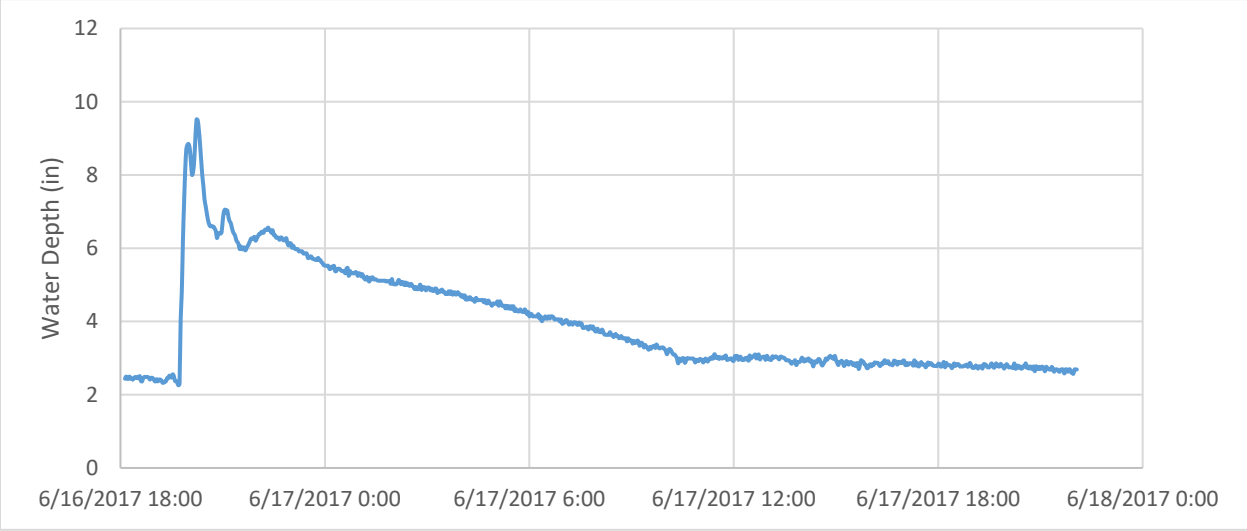
Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., & James, P. (2007). Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. Landscape and urban planning, 81(3), 167-178.

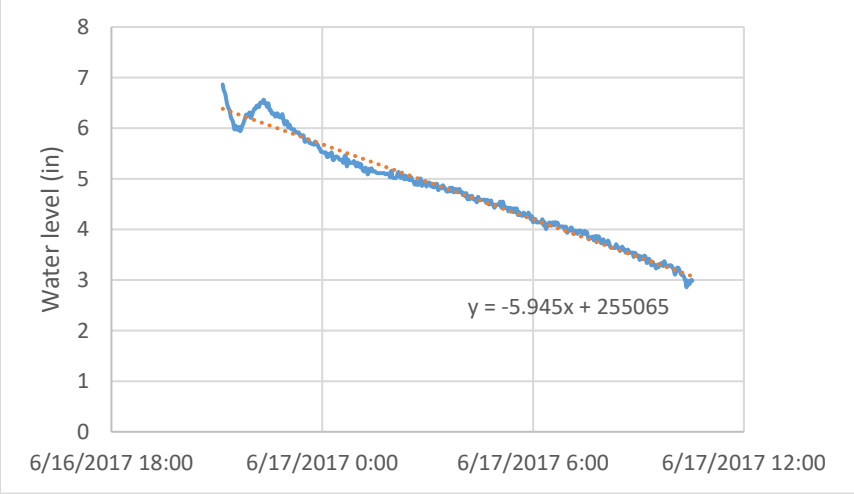
Wardynski, B. J., & Hunt, W. F. (2012). Are bioretention cells being installed per design standards in north carolina? A field study. Journal of Environmental Engineering, 138(12), 1210-1217. doi:10.1061/(ASCE)EE.1943-7870.0000575

Yang, P. P. J., Putra, S. Y., & Li, W. (2007). Viewsphere: a GIS-based 3D visibility analysis for urban design evaluation. Environment and Planning B: Planning and Design, 34(6), 971-992.

Supplemental material 4-A

The following figures provide an example of drawdown rate calculation of a BRC. Figure 1 shows water level inside the well containing the pressure sensor. Figure 2 shows the part of this graph related to descending water level after a storm event. The drawdown rate is the slope of this line, herein calculated in inch per day.





Drawdown Rate: -5.945 in/day = - 0.248 in/hr.

Supplemental material 4-B

The following figures provide an example of locations for soil test sampling.

CHAPTER 5: DESIGN CONSIDERATION

Stormwater control measures (SCMs) are one of the main strategies used to mitigate the negative effects of stormwater runoff by restoring some of the characteristics of natural hydrologic cycle (Burns et al., 2012). Current efforts in stormwater management focus on optimizing existing SCMs' hydrologic and water quality performance. This study investigated three main areas: (1) pollutant dynamics of stormwater runoff at near-continuous resolution, (2) use of real-time control (RTC) technology, (3) and SCM maintenance. Based on the research findings, the following design considerations can offer the potential to increase SCM performance.

Chapter 1 presented the challenges and the advantages associated with the use of UV-vis high-frequency water quality sensors in stormwater settings. The results presented in Chapter 2 illustrated the pollutant dynamics of the studied watershed (9.12 ha with 30% impervious surface) collected at near-continuous resolution (4 minutes). The findings suggested that total suspended solids (TSS) concentration peaks occurred close to the flow peak while those for nitrate and dissolved organic carbon (DOC) often occurred after the peak flow. Consequently, the majority of TSS load was carried by the first 40% of the event's cumulative volume and those of nitrate and DOC were transported by the last 40 percentile (60-100% of cumulative volume). Therefore, to treat TSS and particulate bound pollutants, capturing and treatment of the first half of inflow was more critical than the entire event, while to treat nitrate and DOC, that of second half was more important. Based on the watershed characteristics, designers can optimize the SCM performance and potentially reduce their size by selecting and placing SCMs for targeted pollutant treatment (i.e., in this watershed, capturing the first portion of volume with sediment based SCMs and divert the rest for nitrate treatment).

Chapter 3 quantified the hydrologic and water quality benefits of using real-time control (RTC) in a stormwater wetland and wet pond. RTC implementation increased the number of no outflow events for smaller storms and increased the number of higher volume events. RTC did not affect the cumulative volume reduction and peak flow mitigation. However, it increased the retention time at wet pond by a factor of four. Consequently, the TSS, DOC, and phosphate removal improved by 29%, 17%, and 7% respectively. Both SCMs had higher than average hydrologic and water quality performance both pre and post RTC retrofit. Therefore, SCMs designed based on the guidelines meet the NC required removal rates (NCDEQ, 2017) and may not substantially benefit from an RTC retrofit. Although with mentioned increased in retention time at wet pond, RTC offers the potential to substantially reduce the SCM size while maintaining water quality performance. Therefore, designers can consider use of RTC on wet pond design for projects with real estate restrictions and reduce the wet pond size. It is possible that the relatively large size of the CSW diminished the overall benefit of the RTC and that for much smaller SCM, benefits of RTC would have been more obvious.

Chapter 4 examined the effect of SCM visibility on the maintenance received and consequent performance of it. The results suggest that maintenance teams tend to prioritize SCMs with higher visibility. Visibility was defined in terms of viewshed size, traffic, and noticeability. Therefore, a visible SCM is one with a combined high score of all three parameter. To improve chances of SCM maintenance, designers should (1) decide on SCM location based on not only the watershed but also the viewshed, (2) consider placing SCM in a location where it would be visually accessible to more people (foot traffic/low speed vehicle), (3) consult with a landscape architect on plant selection and arrangement to create a more noticeable SCM that also harmonize with the surrounding landscape.

REFERENCES

Burns, M. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., & Hatt, B. E. (2012). Hydrologic shortcomings of conventional urban stormwater management and opportunities for reform. Landscape and urban planning, 105(3), 230-240.

NCDEQ (2017). North Carolina Stormwater Control Measure Credit Document. North Carolina Department of Environmental Quality: Raleigh, NC, USA. Retrieved from https://files.nc.gov/ncdeq/Energy%20Mineral%20and%20Land%20Resources/Stormwater/BMP%20Manual/SSW-SCM-Credit-Doc-20170807.pdf