5-2019

Estimating Suspended Solids and Phosphorus Loading in Urban Stormwater Systems Using High-Frequency, Continuous Data

Anthony A. Melcher
Utah State University

ESTIMATING SUSPENDED SOLIDS AND PHOSPHORUS LOADING IN URBAN STORMWATER SYSTEMS USING HIGH-FREQUENCY, CONTINUOUS DATA

by

Anthony A. Melcher

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

Approved:

_____	_____
Jeffery S. Horsburgh, Ph.D. Major Professor	R. Ryan Dupont, Ph.D. Committee Member
_____	_____
David K. Stevens, Ph.D. Committee Member	Michelle A. Baker, Ph.D. Committee Member
_____	_____
Bethany T. Neilson, Ph.D. Committee Member	Richard S. Inouye, Ph.D. Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2019

ABSTRACT

Estimating Suspended Solids and Phosphorus Loading in Urban Stormwater Systems Using High-Frequency, Continuous Data

by

Anthony A. Melcher, Doctor of Philosophy

Utah State University, 2019

Major Professor: Dr. Jeffery S. Horsburgh
Department: Civil and Environmental Engineering

Understanding the temporal and spatial variability of stormwater runoff and pollutant loading patterns is essential to managing stormwater in urban catchments. Recent advances in water monitoring technologies and wireless communication allow for data collection at much higher frequencies and at multiple locations than can be achieved using conventional methods. This research investigated methods for implementing modern stormwater monitoring technologies to quantify total suspended solids (TSS) and total phosphorus (TP) loads in an urban conveyance system. The research described in this dissertation includes the design and implementation of a novel stormwater observatory for collecting high-frequency data at multiple points within the Northwest Field Canal (NWFC), Logan, UT, USA, a comparison of statistical models that account for rapidly changing water quality conditions in stormwater conveyances, and an investigation of how high resolution monitoring data derived from the urban observatory can be used to improve the simulation of stormwater quantity. The principal findings of this research were that the urban observatory was able to capture and characterize short-duration storm events at

upstream and downstream ends of the NWFC and at multiple outfalls to the canal simultaneously without the use of field personnel. Additionally, we found that regression with categorical variables and mixed effects modeling were better suited than classical linear regression methods in developing surrogate relationships between suspended solids concentrations and in situ observations of turbidity to account for the dynamic nature of runoff events in an urban water conveyance. Finally, although data collected solely at the outlet of an urban drainage system can aid in the development of simulation models for predicting discharge values at the outlet, stormwater models calibrated using only data from the outlet location were unable to accurately predict discharge at interior, storm drain locations. Models calibrated using data collection from multiple sites within a wireless sensor network were able to better predict discharge values at interior points without compromising the accuracy of predictions at the model outlet. Results from this research are instructive for municipalities, water managers, and modelers for understanding what resources to dedicate to monitoring and modeling and what kind of benefits to expect.

(215 pages)

PUBLIC ABSTRACT

Estimating Suspended Solids and Phosphorus Loading in Urban Stormwater Systems Using High-Frequency, Continuous Data

Anthony A. Melcher

The introduction of pavement, buildings, and other impervious surfaces to urban landscapes greatly influences the quantity and quality of urban stormwater runoff. In this study, we designed and implemented modern stormwater monitoring technologies to establish a "smart" stormwater sensor network within the Northwest Field Canal (NWFC), an urban water conveyance located in Logan, Utah, USA. This network was designed to collect flow and water quality data at high frequencies and simultaneously at multiple locations. The observatory's innovative method of inter-site communication and changing sampling frequencies during storm events was able to capture short duration events at the upstream and downstream ends of the NWFC and at multiple outfalls in the canal simultaneously without human intervention. We then investigated statistical regression models between turbidity and TSS so as to predict TSS at high frequencies. Finally, the addition of the high-frequency discharge data in the calibration procedure for a stormwater simulation model developed using the Environmental Protection Agency's Stormwater Management Model did little to improve model performance at the downstream end of the canal, but did provide important insight into the overall contribution of discharge from individual stormwater outfalls to the NWFC. The results of this study inform water professionals on how to build and operate automated monitoring systems and how to create high-frequency estimates of TSS and TP loads in urban water systems.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Jeff Horsburgh, for his tutelage and patience as he helped me throughout my, at times painful, PhD journey. I have learned much from his work ethic and high academic standard. In addition to my advisor, I am grateful to all those that helped me carry out this research. Among those are Lance Houser and the City of Logan, Richard Boudrero and others at the Northwest Field Canal Company, as well as my extremely helpful and encouraging graduate committee for their willingness to provide guidance and advice through each phase of my PhD program.

I would also like to thank those that assisted me in so many facets of my research. Bryce Mihalevich and Phil Suiter for their advice and assistance with field and lab work. Amber Jones and Caleb Buahin for their high quality advice and encouragement. I am also grateful for the generous funding offered by the Utah Water Research Laboratory and the U.S. National Science Foundation under Grant IIA-1208732 innovative Urban Transitions and Aridregion Hydro-Sustainability (iUTAH), which made this research possible. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Last, I am in debt to my wife Jamie and our four kids. They are and always will be my motivation, and there is no way I could have done this without their loving support.

Anthony A. Melcher

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Urban stormwater runoff has been recognized as a major contributor to degraded water quality in many water bodies in the United States. The accumulation of pollutants and sediments can cause oxygen depletion in receiving waters, reduce reservoir capacities, degrade drinking water sources, and render water bodies unusable for recreational purposes (Chapra, 2008; Khaba and Griffiths, 2017; National Research Council, 2009; Sawyer, 1966). As a result, the United States Environmental Protection Agency's (USEPA) stormwater regulations for municipal separate storm sewer systems (MS4) have been promulgated to mitigate some of these impacts (Federal Register, 1999). These regulations require MS4s to develop a stormwater management program that highlights the stormwater control measures (SCM) to be implemented to meet downstream water quality standards. In order to accurately estimate constituent loads, select SCMs, and identify optimal locations for SCMs in the watershed, knowledge of the temporal and spatial constituent loading patterns in stormwater runoff must be obtained. This can be quite challenging due to the highly dynamic nature of loading events in urban catchments.

Changes to natural landscapes, such as the introduction of impervious surfaces, can greatly affect the size and shape of stormwater hydrographs and pollutographs (Hvitved-Jacobsen et al., 2010; Wanielista and Yousef, 1993). Shear stresses that mobilize urban sediments are directly related to runoff velocities (Berenbrock and Tranmer, 2008), which are increased due to increased imperviousness. Water quality monitoring programs that implement sampling plans that do not consider the effects of urban development on

constituent loading events are insufficient (Kirchner et al., 2004; National Research Council, 2009). Infrequent grab sampling methods that often include sampling at weekly or monthly intervals can bias load estimates and do little to characterize loading patterns in short-duration events (Harmel and King, 2005; Horsburgh et al., 2010; Jones et al., 2012; National Research Council, 2009; Rode et al., 2016). Additionally, sampling at a single location (e.g., at the outlet of a watershed) lacks the spatial resolution to draw any conclusions on loading patterns or identify locations within the watershed that could benefit from management practices like green infrastructure and stormwater treatment. While some studies of urban stormwater have used high-frequency sampling (Ackerman et al., 2011; Halliday et al., 2015; Viviano et al., 2014), much is still unknown about how high-frequency data collected simultaneously at multiple locations can be used to resolve temporal and spatial heterogeneity and help to better quantify the contribution of pollutants from urban stormwater runoff.

An emerging field of research is the use of automated technology with knowledge of engineered and natural systems to develop "smart" monitoring infrastructure that is able to collect high-frequency data at multiple locations and adapt according to changing conditions (Kerkez et al., 2016; Mullapudi et al., 2017; Wong and Kerkez, 2016). Because stormwater runoff and constituent loading events in urban catchments are highly dynamic, stormwater monitoring programs stand to benefit greatly from new techniques for detecting runoff events, adapting sampling frequencies based on predefined criteria, and communicating among monitoring sites to anticipate loading events at downstream locations. This type of monitoring infrastructure can improve the ability to estimate pollutant loads resulting from urban stormwater runoff by: 1) improving the understanding

of hydrologic processes, flows, and constituent loads in stormwater conveyances, 2) enhancing the ability to derive surrogate relationships between constituent concentrations and parameters measured *in situ*, and 3) increasing our ability to simulate urban stormwater systems.

Common in the western United States are water conveyances that serve multiple purposes. Base flows in urban streams and waterways may represent flows influenced primarily by groundwater and snow melt runoff, while flows during storm events might represent runoff from urban surfaces. Additionally, water diverted for agricultural purposes and return flows may be present in combined conveyances during irrigation seasons (City of Grand Junction, 2016; City of Logan, 2010; City of Sequim, 2016). The challenge of determining what fraction of the total constituent load measured at a catchment outlet can be attributed to stormwater runoff is not trivial. One widely accepted practice for making continuous estimates of constituent concentrations and loads is using high-frequency, *in situ* data as surrogates for concentration values typically obtained in a laboratory (e.g., turbidity as a surrogate for TSS) (Jones et al., 2011; Rasmussen et al., 2011; Ryberg, 2006; Viviano et al., 2014). This involves fitting a regression model with the *in situ* parameter(s) as the explanatory variable and the constituent concentration of interest as the response variable. Fitting a surrogate relationship with a single regression model, however, makes the assumption that the surrogate relationship is constant under all conditions. This has been found to not always be the case (Grayson et al., 1996; Jones et al., 2011; Ryberg, 2006), especially in water bodies that receive constituent loads from multiple sources, as is the case within combined urban/irrigation/stormwater conveyances.

High-frequency data collected via “smart” monitoring protocols are unique in that they represent multiple catchments’ simultaneous responses to a storm event. Typical urban stormwater modeling procedure is to calibrate and validate a model based on hydrographs and pollutographs observed at a catchment outlet. However, calibrating a model based on data collected at a single point is at risk of high model uncertainty and misrepresentation of the hydrologic processes at other locations within the catchment (Beven and Binley, 1992; Chiang et al., 2014; Neilson et al., 2010; Sun et al., 2013). Yet to be determined is how data collected at high temporal and spatial resolutions can affect stormwater model calibration and performance. While some research exists that considers the advantages and disadvantages of using multi-site data to calibrate and validate models (Lerat et al., 2012; Leta et al., 2017; Wang et al., 2012), much is still unknown about the degree to which high-resolution stormwater data collection using in situ sensors within an urban observatory can improve modeling procedures.

As many studies have attributed the degradation of water quality in receiving water bodies to urban stormwater runoff, the focus of the research described in this dissertation is on demonstrating the necessary monitoring and modeling efforts to substantiate those claims. It is hypothesized that high-frequency data collection at multiple monitoring sites within an urban catchment is necessary for identifying some of the processes and spatial heterogeneities that govern constituent loading events, as well as for advancing methods for making more accurate estimates of constituent loads in stormwater runoff. This hypothesis is tested in this dissertation by the design and deployment of an urban stormwater sensor network, or urban observatory. The high-frequency data collected by the urban observatory was then used to explore statistical and numerical modeling

procedures to estimate stormwater contributions of phosphorus and suspended solids to an urban receiving water body. The following research objectives were chosen to test the above hypothesis:

- *Objective 1: Design and establish an urban observatory for studying the effects of stormwater inputs on urban water systems.* Understanding the spatial and temporal variability in the fluxes of constituents in an urban stormwater system requires a coordinated plan for sampling and instrumentation. Under this objective, we designed and deployed a multi-node environmental sensor network capable of generating the high-frequency and high-resolution data required to better estimate urban runoff quantity and quality. *In situ* measurements of water quality (e.g., turbidity, specific conductance, dissolved oxygen, pH, and water temperature) and water quantity parameters (e.g., stage and precipitation) were coupled with periodic grab samples and in-stream discharge measurements to derive high-frequency constituent concentrations and load estimates. Due to the short duration of stormwater runoff events we developed algorithms that enabled upstream monitoring sites to detect stormwater runoff events in real-time and message downstream monitoring sites to better anticipate and ensure that loading events were monitored at adequate frequencies.
- *Objective 2: Investigate methods for quantifying suspended solids loads within urban water systems.* Making accurate constituent load estimates in a combined irrigation/stormwater conveyance requires a method that accounts for loads during base flow conditions as well as short duration storm event conditions. Under this objective, we used high-frequency data from *in situ* sensors and the collection of

periodic and event-based water quality samples to derive surrogate relationships to estimate TSS concentrations and loads in a combined irrigation/stormwater conveyance. Three regression methods were then evaluated, namely classical linear regression, linear regression with categorical variables, and linear mixed effects modeling for their ability to accurately predict base flow and storm event TSS concentrations and estimate TSS loads for two irrigation seasons.

- *Objective 3: Quantify the contributions of stormwater to suspended solids and phosphorus loading to the urban water system.* Information from field data collection campaigns and sensor deployments can be used to populate, calibrate, and validate rainfall-runoff models. However, with the recent availability of high resolution datasets, we are just now beginning to test how these datasets can be used to better inform and drive the models we use to simulate environmental systems. Under this objective, we built a U.S. Environmental Protection Agency's (USEPA) Storm Water Management Model (SWMM) of the NWFC drainage area. In addition to calibration at the outlet of the drainage area, we calibrated separate models of the drainage areas for the six monitored subcatchments. We then subsequently inserted each calibrated subcatchment model into the larger NWFC model to assess how predictions at the model outlet were affected by the availability of high resolution monitoring data.

These objectives were chosen to address the difficulties related to estimating constituent concentrations and loads in relatively small, urban catchments. By accomplishing these objectives, we created valuable information about the spatial and temporal loading patterns of water quality constituents with urban stormwater runoff and

we demonstrated valuable new techniques that can be used by other researchers and developers of urban water quality and stormwater sampling programs. Each of the above objectives is addressed within one chapter of this dissertation as follows.

Chapter 2 addresses the first objective by presenting the design and development of an urban environmental observatory within the Logan River watershed in northern Utah, USA. The purpose of the observatory was to capture stormwater runoff events at multiple locations and to synchronize the collection of water quality (*in situ* and grab samples) and quantity data so that better estimates of TSS and total phosphorus (TP) loads could be made. This was accomplished by detecting stormwater runoff events in real-time, using radio telemetry to communicate event detections between monitoring sites, and adapting both *in situ* and automated sample collection frequencies according to changing environmental conditions. Chapter 2 describes the monitoring infrastructure required, as well as the data collection and algorithms required to deploy an urban observatory and collect high-frequency water quality and quantity data. Results are presented demonstrating the observatory's capabilities to capture the spatial variability of TSS and TP event mean concentrations (EMC) between multiple monitoring sites, and to capture temporal variabilities such as the short duration, first flush phenomenon.

Chapter 3 addresses the second objective and compares multiple regression methods for deriving continuous estimates of TSS concentrations and loads for the duration of the study period in an urban catchment subject to stormwater runoff. In this chapter, we compare classical linear regression models to linear regression with categorical variables and linear mixed effects models, which use additional explanatory variables related to storm event characteristics to account for changes in the surrogate relationship and

undersampled event conditions. While common practice is to assume that surrogate relationships are unchanging (i.e., the classical linear regression method), we report how alternative regression methods that consider episodic constituent source changes can affect TSS load estimates.

Chapter 4 addresses the third research objective and describes techniques for incorporating high resolution monitoring data from multiple monitoring locations into urban stormwater modeling efforts to better estimate TSS and TP loads generated from urban stormwater runoff. SWMM was selected for this research because it is one of the most widely used stormwater models and currently represents the state of the practice in stormwater modeling (Niazi et al., 2017). Our work was focused on how the state of the practice could be advanced by combining urban stormwater modeling with high resolution data. We assessed how a semi-distributed stormwater model was affected by using additional time series datasets to calibrate subcatchments of the larger model domain. Specifically, we looked at how the addition of calibration datasets at the subcatchment scale affected the uncertainty of water quantity and quality predictions at the model outlet, or the outlet of the NWFC drainage area. Incorporating observational data from multiple locations within the watershed helped to justify monitoring efforts via an urban observatory and provided valuable information that stormwater managers may use in informing monitoring site selection based on stormwater modeling benefits.

The urban observatory along with the statistical methods and numerical modeling presented in this dissertation take advantage of technological advancements in the fields of stormwater monitoring and water resource management. The description of the observatory's infrastructure and sampling logic can be of benefit to water managers and

municipalities wanting to design a next-generation monitoring program and researchers seeking a greater understanding of processes governing pollutant loading in the built environment. The statistical and numerical modeling methods presented in this dissertation reveal valuable techniques for estimating TSS and TP loads to a receiving water body, as well as for providing insight into identifying optimal monitoring site locations. These techniques can be implemented in water quality and total maximum daily load (TMDL) studies.

References

Ackerman, D., Stein, E.D., Ritter, K.J., 2011. Evaluating performance of stormwater sampling approaches using a dynamic watershed model. Environ. Monit. Assess. 180, 283–302. https://doi.org/10.1007/s10661-010-1788-6

Berenbrock, C., Tranmer, A.W., 2008. Simulation of Flow, Sediment Transport, and Sediment Mobility of the Lower Coeur d'Alene River, Idaho.

Beven, K., Binley, A., 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. Hydrol. Process. 6, 279–298. https://doi.org/10.1002/hyp.3360060305

Chapra, S.C., 2008. Surface Water-Quality Modeling. Waveland Press, Inc., Long Grove.

Chiang, L.-C., Yuan, Y., Mehaffey, M., Jackson, M., Chaubey, I., 2014. Assessing SWAT's performance in the Kaskaskia River watershed as influenced by the number of calibration stations used. Hydrol. Process. 28, 676–687. https://doi.org/10.1002/hyp.9589

City of Grand Junction, 2016. Grand Junction Stormwater Management Manual [WWW Document]. Gd. Junction City Munic. Code Vol. II Dev. Regul. URL http://www.codepublishing.com/CO/GrandJunction/html2/GrandJunction28/GrandJunction2852.html (accessed 1.1.16).

City of Logan, 2010. 2010 Logan City Storm Water Management Plan. Logan.

City of Sequim, 2016. Storm and Surface Water Master Plan [WWW Document]. URL http://www.sequimwa.gov/DocumentCenter/View/7735 (accessed 1.1.16).

Federal Register, 1999. National Pollutant Discharge Elimination System - Regulations for Revision of the Water Pollution Control Program Addressing Storm Water Discharges. USA.

Grayson, R.B., Finlayson, B.L., Gippel, C.J., Hart, B.T., 1996. The Potential of Field Turbidity Measurements for the Computation of Total Phosphorus and Suspended Solids Loads. J. Environ. Manage. 47, 257–267. https://doi.org/10.1006/jema.1996.0051

Halliday, S.J., Skeffington, R. a., Wade, A.J., Bowes, M.J., Gozzard, E., Newman, J.R., Loewenthal, M., Palmer-Felgate, E.J., Jarvie, H.P., 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. Hydrol. Process. https://doi.org/10.1002/hyp.10453

Harmel, R.D., King, K.W., 2005. Uncertainty In Measured Sediment and Nutrient Flux In Runoff From Small Agricultural Watersheds. Am. Soc. Agric. Eng. 48, 1713–1722.

Horsburgh, J.S., Spackman Jones, A., Stevens, D.K., Tarboton, D.G., Mesner, N.O., 2010. A Sensor Network for High Frequency Estimation of Water Quality Constituent Fluxes Using Surrogates. Environ. Model. Softw. 25, 1031–1044. https://doi.org/10.1016/j.envsoft.2009.10.012

Hvitved-Jacobsen, T., Vollertsen, J., Nielsen, A.H., 2010. Urban and Highway Stormwater Pollution: Concepts and Engineering. CRC Press.

Jones, A.S., Horsburgh, J.S., Mesner, N.O., Ryel, R.J., Stevens, D.K., 2012. Influence of Sampling Frequency on Estimation of Annual Total Phosphorus and Total Suspended Solids Loads. J. Am. Water Resour. Assoc. 48, 1258–1275. https://doi.org/10.1111/j.1752-1688.2012.00684.x

Jones, A.S., Stevens, D.K., Horsburgh, J.S., Mesner, N.O., 2011. Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and Total Phosphorus Concentrations. JAWRA J. Am. Water Resour. Assoc. 47, 239–253. https://doi.org/10.1111/j.1752-1688.2010.00505.x

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., Bedig, A., Kertesz, R., Braun, T., Cadwalader, O., Poresky, A., Pak, C., 2016. Smarter Stormwater Systems. Environ. Sci. Technol. 50, 7267–7273. https://doi.org/10.1021/acs.est.5b05870

Khaba, L., Griffiths, J.A., 2017. Calculation of reservoir capacity loss due to sediment deposition in the `Muela reservoir, Northern Lesotho. Int. Soil Water Conserv. Res. 5, 130–140. https://doi.org/10.1016/j.iswcr.2017.05.005

Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The Fine Structure of water-Quality Dynamics: The (High-Frequency) Wave of the Future. Hydrol. Process. 18, 1353–1359. https://doi.org/10.1002/hyp.5537

Lerat, J., Andréassian, V., Perrin, C., Vaze, J., Perraud, J.M., Ribstein, P., Loumagne, C., 2012. Do internal flow measurements improve the calibration of rainfall-runoff models? Water Resour. Res. 48, W02511. https://doi.org/10.1029/2010WR010179

Leta, O.T., van Griensven, A., Bauwens, W., 2017. Effect of Single and Multisite Calibration Techniques on the Parameter Estimation, Performance, and Output of a SWAT Model of a Spatially Heterogeneous Catchment. J. Hydrol. Eng. 22, 05016036. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001471

Mullapudi, A., Wong, B.P., Kerkez, B., 2017. Emerging investigators series: building a theory for smart stormwater systems. Environ. Sci. Water Res. Technol. 3, 66–77.

https://doi.org/10.1039/C6EW00211K

National Research Council, 2009. Urban Stormwater Management in the United States.

Neilson, B.T., Chapra, S.C., Stevens, D.K., Bandaragoda, C., 2010. Two-zone transient storage modeling using temperature and solute data with multiobjective calibration: 1. Temperature. Water Resour. Res. 46. https://doi.org/10.1029/2009WR008756

Niazi, M., Nietch, C., Maghrebi, M., Jackson, N., Bennett, B.R., Tryby, M., Massoudieh, A., 2017. Storm Water Management Model: Performance Review and Gap Analysis. J. Sustain. Water Built Environ. 1–32. https://doi.org/10.1061/JSWBAY.0000817

Rasmussen, P.P., Gray, J.R., Glysson, G., Ziegler, A.C., 2011. Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data, in: Applications of Hydraulics Section C, Sediment and Erosion Techniques. U.S. Geological Survey, Reston, p. 52.

Rode, M., Wade, A.J., Cohen, M.J., Hensley, R.T., Bowes, M.J., Kirchner, J.W., Arhonditsis, G.B., Jordan, P., Kronvang, B., Halliday, S.J., Skeffington, R.A., Rozemeijer, J.C., Aubert, A.H., Rinke, K., Jomaa, S., 2016. Sensors in the Stream: The High-Frequency Wave of the Present. Environ. Sci. Technol. 50, 10297–10307. https://doi.org/10.1021/acs.est.6b02155

Ryberg, K.R., 2006. Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05. Reston.

Sawyer, C.N., 1966. Basic Concepts of Eutrophication. Water Pollut. Control Fed. 38, 737–744.

Sun, N., Hong, B., Hall, M., 2013. Assessment of the SWMM Model Uncertainties Within the Generalized Likelihood Uncertainty Estimation (GLUE) Framework for a High-Resolution Urban Sewershed. Hydrol. Process. 28, 3018–3034. https://doi.org/10.1002/hyp.9869

Viviano, G., Salerno, F., Manfredi, E.C., Polesello, S., Valsecchi, S., Tartari, G., 2014. Surrogate measures for providing high frequency estimates of total phosphorus concentrations in urban watersheds. Water Res. 64, 265–277. https://doi.org/10.1016/j.watres.2014.07.009

Wang, S., Zhang, Z., Sun, G., Strauss, P., Guo, J., Tang, Y., Yao, A., 2012. Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. Hydrol. Earth Syst. Sci. Katlenburg-Lindau 16, 4621. https://doi.org/http://dx.doi.org/10.5194/hess-16-4621-2012

Wanielista, M.P., Yousef, Y.A., 1993. Stormwater Management. John Wiley and Sons, Inc.

Wong, B.P., Kerkez, B., 2016. Adaptive Measurements of Urban Runoff Quality. Water Resour. Res. https://doi.org/10.1002/2015WR018013

CHAPTER 2

AN URBAN OBSERVATORY FOR QUANTIFYING PHOSPHORUS AND SUSPENDED SOLIDS LOADS IN COMBINED NATURAL AND STORMWATER CONVEYANCES[1]

Abstract

Water quality in urban streams and stormwater systems is highly dynamic, both spatially and temporally, and can change drastically during storm events. Infrequent grab samples commonly collected for estimating pollutant loadings are insufficient to characterize water quality in many urban water systems. *In-situ* water quality measurements are being used as surrogates for continuous pollutant load estimates; however, relatively few studies have tested the validity of surrogate indicators in urban stormwater conveyances. In this paper we describe an observatory aimed at demonstrating the infrastructure required for surrogate monitoring in urban water systems and for capturing the dynamic behavior of stormwater driven pollutant loads. We describe the instrumentation of multiple, autonomous water quality and quantity monitoring sites within an urban observatory. We also describe smart and adaptive sampling procedures implemented to improve data collection for developing surrogate relationships and for capturing the temporal and spatial variability of pollutant loading events in urban watersheds. Results show that the observatory is able to capture short-duration storm events within multiple catchments and, through inter-site communication, sampling efforts can be synchronized across multiple monitoring sites.

---

[1]Melcher, A.A. and J.S. Horsburgh, 2017. An Urban Observatory for Quantifying

Phosphorus and Suspended Solids Loads in Combined Natural and Stormwater Conveyances. Environmental Monitoring and Assessment June 2017, 189:285. The final publication is available at Springer via https://doi.org/10.1007/s10661-017-5974-7

2.1. Introduction

Urban stormwater runoff has proven to be a major contributor of sediment and nutrients to receiving water bodies (Sawyer, 1966; National Research Council, 2009; Utah Division of Water Quality, 2010). In many cases, this results in oxygen depletion and cultural eutrophication (Roy-Poirier et al., 2010). Stormwater regulations promulgated by the United States Environmental Protection Agency (USEPA) have charged municipal separate storm sewer systems (MS4s) with the responsibility of developing a stormwater management program that highlights the stormwater control measures (SCM) planned for mitigating loads of pollutant such as sediments and nutrients in stormwater runoff (Federal Register, 1999). In order to develop effective SCM plans, managers of MS4s need information on the quantity and quality of stormwater runoff to enable them to assess the impact of urban growth and land use changes and identify locations that will benefit the most from SCM implementation. However, monitoring for the estimation of pollutant loads is often one of the weakest parts of stormwater management programs and is often excluded or goes unreported (National Research Council, 2009; Aguilar and Dymond, 2015). Aguilar and Dymond (2015) state that of the 90 MS4s they surveyed in Virginia, USA, none reported the measurement of any water quality or water quantity parameters.

Common in urbanized areas of the western United States are drainage pipes, canals, and stream conveyances that serve multiple purposes (Douglas County, 2011; City of

Grand Junction, 2016; City of Sequim, 2016). Often, natural stream flows are combined with stormwater runoff and irrigation return flows within a single conveyance. Combined flows from multiple sources can make the quantification of pollutant loads from each specific source difficult. The overall pollutant concentration determined by sampling combined flows must be disaggregated, or sources must be monitored separately in order to attribute the pollutant contribution to a particular source. Fingerprinting or source tracking methods, which attempt to identify the chemical signature of a pollutant or another water constituent derived from each potential source (e.g., different land uses), have been used successfully to track sources of suspended sediment or microbial loading in rural, agriculturally dominated watersheds (Walling et al., 1999; Walling, 2005; Poleto et al., 2009). These methods have been less successful when applied within urban catchments and for a broad range of pollutants (Poleto et al., 2009). The unreliability of chemical fingerprinting and traditional sampling methods has emphasized the need for alternative methods for quantifying pollutant loads within urban catchments and conveyances.

Another challenge in quantifying pollutant loads resulting from stormwater runoff is characterizing the temporal and spatial scales of pollutant loading events (Tiefenthaler et al., 2001; Goonetilleke et al., 2005). Spatial and temporal variability are caused by variability in rainfall, which is especially important in urban settings where the varying hydrologic response from urban land uses depends on differing levels of impervious surface and the type and characteristics of urban stormwater infrastructure. Recent advances in water quality and quantity monitoring technology have produced in-situ sensors capable of deployment for long periods of time at a relatively low cost due to the

low maintenance required (Kirchner et al., 2004). The high frequency measurements made possible by in-situ sensors have allowed for a greater understanding of the temporal and spatial variability within storm events as well as seasonal and year-to-year trends (Kirchner et al., 2004). However, urban water systems present some unique challenges for automated, continuous monitoring.

First, the spatial and temporal scales at which important processes occur vary greatly. For example, monitoring at a lower frequency (e.g., every 30 minutes or even hourly) may be adequate for capturing seasonal water quality trends in urban conveyances with natural streamflow, but is insufficient to characterize the impact of short duration, high intensity pollutant loading events from storms that may last only minutes. In-situ monitoring at multiple locations provides comparative time-series datasets that allow for a greater understanding of the spatial variability of pollutant loading events and has been successfully used in hydrologic and environmental observatories (e.g., instrumented watersheds) for quantifying loads and identifying pollutant sources from short duration, high intensity events (Horsburgh et al., 2010; Cassidy and Jordan, 2011; Jones et al., 2011). These wireless sensor networks (WSN) are able to collect data via in-situ sensors at multiple monitoring nodes and then transmit them to a centralized location for storage, post-processing, and analysis. While the architecture and infrastructure of WSNs for environmental observatories have been defined (Corke et al., 2010; Horsburgh et al., 2010), there are fewer examples of using WSNs to characterize the quantity and quality of stormwater runoff within urban water systems (Wong and Kerkez, 2016).

Several recent studies have successfully used in-situ measurements as surrogates for pollutant concentrations and loads (Ryberg, 2006; Settle et al., 2007; Jones et al., 2011; Viviano et al., 2014; Hannouche et al., 2016; Nasrabadi et al., 2016). However, key to developing strong surrogate relationships (e.g., using turbidity to predict TSS or TP) is the collection of physical samples representative of the range of hydrologic conditions and constituent concentrations experienced in a catchment. This minimizes the potential for using derived relationships to extrapolate beyond the range of measured conditions. Manually collecting a large number of samples over time can accomplish this; however, the cost, effort, and timing required for field crews to collect and process a large number of samples can be prohibitive for many MS4s. There is an opportunity, therefore, to combine in-situ sensors and adaptive sampling logic to not only adjust sensor measurements to rapidly changing hydrologic conditions but to also trigger collection of physical samples to strategically capture important characteristics of storm events, reduce the number of samples required to capture a broader range of conditions, and better support development of surrogate relationships.

In this paper we describe the infrastructure, adaptive sampling logic, and communication requirements for collecting high frequency stormwater quantity and quality data in urban catchments. We then describe a case study in which an urban observatory, or an environmental sensor network located in an urban stormwater drainage system, was constructed. Within this observatory, we implemented an adaptive sampling logic to collect both high frequency, in-situ sensor observations and automated, pumped samples for laboratory analysis in an effort to quantify TP and TSS loading from rainfall

events. Results of the case study demonstrate the value of an urban observatory for quantifying pollutant loads in stormwater conveyances.

The following section includes a description of the temporal and spatial requirements for monitoring stormwater quality. We then describe the infrastructure, programming logic, and communication requirements for an adaptive, urban observatory that monitors water quantity and quality in urban stormwater runoff. Following our description of requirements, we describe a case study in Logan, Utah, USA in which an urban observatory that uses an adaptive sampling algorithm was installed to monitor the flux of TSS and TP in an urban stormwater conveyance. In the final section we summarize our results.

2.2. Quantifying Pollutant Loads in Urban Stormwater Runoff

Within urban water systems, precipitation is the driving force for runoff generation and pollutant mobilization. Data collection efforts can reveal how stormwater quantity and quality vary both within storm events and on a longer-term scale (e.g., seasonally and yearly). Additionally, data collection that targets characterization of multiple land use areas and/or pervious/impervious areas can reveal important spatial patterns. In the following sections we provide requirements and considerations for effective design of monitoring systems aimed at better characterizing both temporal and spatial characteristics of urban stormwater.

2.2.1. Temporal Variability of Pollutant Loads in Stormwater Runoff and the Need for High Frequency Data

Quantifying pollutant loads in urban runoff requires observations at a high temporal resolution. Hydrographs from urbanized watersheds have a shorter time-to-peak and higher runoff volumes per unit area when compared to watersheds with more natural landscapes due to their larger percentage of impervious area (M. P. Wanielista and Yousef, 1993; M. Wanielista et al., 1997; Hvitved-Jacobsen et al., 2010). Kirchner et al. (2004) noted that water quality and quantity measurements should be made at a temporal resolution greater than or equal to that of a catchment's hydrologic response (e.g., the rate at which flow and water quality change), which may be on the order of minutes in urban catchments with high levels of imperviousness. Data obtained at a high frequency can provide better understanding of the processes that drive pollutant loading on the catchment and subcatchment scale than grab samples obtained at a lower frequency.

Pollutant loading from urban stormwater runoff is greatly affected by human behavior (National Research Council, 2009). Some examples of practices that can increase runoff loads include land development and the introduction of impervious surfaces, sediments and other pollutants generated from construction sites, and anthropogenic pollutant sources based on land use (Waschbusch et al., 1999; National Research Council, 2009). As stormwater collection systems are typically dry between storm events, loading to receiving water bodies occurs during a short time window when water is actually flowing. In the semi-arid, western region of the United States, it is not uncommon to have weeks of dry period between very short and intense storm events. Combined stormwater conveyances, such as canals and urban rivers, may experience long periods of relatively

constant flow and water quality, interrupted by storm events that cause rapid changes in both. Effective stormwater monitoring efforts must be able to adapt sampling frequencies in order to quantify these short duration, high intensity inputs of pollutant loads while also capturing longer term conditions in the receiving waterbodies for assessing longer term impacts.

2.2.1.1. Monitoring for Characterizing the First Flush

Sampling during storm events has been conducted by many researchers using multiple different methods (Ackerman et al., 2011; Leecaster et al., 2002). One approach is to collect a composite or multiple composite samples during a storm event. This includes the collection of multiple sample aliquots at time or flow-weighted intervals that are then combined into fewer composite samples. Often, these composite samples are then analyzed to determine the concentrations of constituents of interest, thus obtaining what are effectively event mean concentrations (EMCs) – or the average concentration of each constituent over the course of the entire event. Another approach to storm event sampling is to collect discrete samples throughout the event that characterize how constituent concentrations change over time. It has been found in many cases that a majority of pollutant loading occurs during the first part of a runoff event, before the peak of the hydrograph. This phenomenon has been termed the "first flush" of the storm event (Bertrand-Krajewski et al., 1998; Sansalone and Cristina, 2004). The intensity and duration of the first flush has been observed to be watershed specific and is affected by the drainage area, land use of the catchment, and the amount of time that has passed since the last storm event (Lee et al., 2002; National Research Council, 2009). The first flush can be observed

and characterized when discrete samples are collected during a storm event. Studies that have compared multiple sampling methods for pollutant load estimation have found that the collection of discrete samples throughout the duration of the event, rather than composite sampling, introduces the least bias to an estimate of the total pollutant load occurring within a storm event (Ackerman et al., 2011; Leecaster et al., 2002).

2.2.1.2. Monitoring for the Development of Surrogate Relationships

It is currently difficult and costly to create high frequency datasets for constituents like TSS and TP from grab sampling alone. While in-situ phosphorus analyzers exist, they are expensive, require maintenance to ensure proper operation, and have been found to be prone to malfunction (Cassidy and Jordan, 2011). Many examples exist where variables monitored in-situ and at high frequencies have been used as surrogates for TSS and TP (Christensen et al., 2002; Fisher et al., 2016; Jones et al., 2011; Nasrabadi et al., 2016; Rasmussen et al., 2011; Ryberg, 2006). The majority of these examples, however, developed surrogate relationships (e.g., regression models between a surrogate such as turbidity and TSS or TP concentrations) in natural or agriculturally dominated watersheds. Many of these studies found that turbidity and streamflow were good predictors of TSS and TP. Fewer studies have investigated the use of surrogate monitoring techniques in systems affected or dominated by urban stormwater runoff (Fisher et al., 2016; Miguntanna et al., 2010; Settle et al., 2007; Viviano et al., 2014).

To obtain reliable parameters in a surrogate regression model, in-situ observations and physical samples for laboratory analysis must be collected across the range of expected values for water quality constituent concentrations (Lewis, 1996; McKee and Gilbreath,

2015). For example, when high frequency measurements of in-situ turbidity are used as the independent variable and TP concentrations are used as the dependent variable, TP samples should be collected over as much of the range of observed turbidity values as possible. Once developed, a surrogate relationship could then be used to create continuous estimates of TP concentrations for the period in which in-situ and sample data were collected if it can be reasonably assumed that the relationships remained constant (Rasmussen et al., 2011).

The selection of appropriate in-situ sensors is essential in the development of surrogate relationships and in collecting continuous observations of surrogate values for estimating pollutant concentrations. In-situ sensors are able to measure parameters such as turbidity and conductivity at high frequencies (e.g., on the order of seconds to minutes) that can then be used with surrogate methods to create continuous estimates of the concentrations of pollutants such as TP and TSS. A "continuous" dataset is one that accounts for all changes and variations in the pollutograph. If additional measurements will not provide any additional information about these variations, one can verify that the data set is continuous (Kirchner et al., 2004). Given a continuous dataset, the total pollutant mass loading (M) during a loading event can then be calculated by:

$$M = \int C(t)Q(t)dt \tag{2.1}$$

where C(t) is the pollutant concentration as a function of time and Q(t) is the stormwater discharge as a function of time. High frequency measurements must be made for both water quality and quantity to obtain continuous estimates of mass loads during a storm event. Using surrogate methods, C(t) is calculated based on the surrogate relationship. Thus, in

order to obtain accurate estimates of pollutant mass loads from the calculated datasets, surrogate measurements (e.g., turbidity) must be correlated with the concentrations of the pollutants that they are serving as surrogates for and must be frequent enough to characterize local minima and maxima in the pollutograph (Lewis, 1996) to ensure that peaks or valleys are represented in the integral calculation.

2.2.2. Spatial Variability of Pollutant Loads in Stormwater Runoff

Many studies have attempted to characterize and predict loading based on land uses within drainage areas (USEPA, 1983; Ahearn et al., 2005; Goonetilleke et al., 2005; Gunaratne et al., 2014). One of the earliest efforts to incorporate land use into pollutant load modeling and prediction was the nationwide urban runoff program (NURP) completed by the USEPA in 1979-1983. This program included 28 separate monitoring programs across the United States. Unfortunately, correlations between land use and pollutant loads were weak (USEPA, 1983), which may have been due to the variability in sampling methods and frequencies carried out by each individual monitoring program. Since the completion of NURP, however, new advances have been made in stormwater monitoring technology that allow for multi-nodal, high frequency sampling (Kirchner et al., 2004; Corke et al., 2010; Horsburgh et al., 2010; Cassidy and Jordan, 2011).

Similar to the inability of infrequent grab sampling to characterize temporal dynamics, the collection of water quality and quantity data at a single location is insufficient for characterizing the spatial variability in runoff and quality. Sampling at a single catchment outlet might provide an end-of-pipe snapshot of the pollutant concentrations within that catchment, but does not provide all of the information needed to

understand how pollutant loading changes across a range of catchments and urban landscapes. Instead, sampling at the outlets of a variety of catchments strategically chosen to represent a range of both catchment size and land cover composition can provide better understanding of the temporal and spatial scales on which processes occur and can lead to more accurate estimates of pollutant loading.

While it is infeasible to simultaneously sample every outfall to a larger stormwater conveyance, it is advantageous to simultaneously sample outfalls from a small number of catchments of varying size, slope, land use, and land cover to observe multiple catchments' responses to multiple storm events with differing characteristics (i.e., intensity, total volume, antecedent dry period, etc.). Mobile sampling equipment designed to be moved from outfall to outfall can aid in characterizing catchment responses, such as pollutant buildup during dry periods, hydrograph and pollutograph characteristics, and pollutant loading patterns.

2.3. Infrastructure Required for Quantifying Pollutant Loads from Urban Stormwater Runoff

Robust and automated sensing and data management infrastructure are required for quantifying pollutant loads from urban stormwater runoff given: 1) the "flashy" nature of the system being observed; 2) the need to create continuous records with sampling intensities high enough to capture storm events; and 3) the large volume of data generated. Adaptive sampling (e.g., automated collection of physical samples based on flow or other water quality conditions, or adaptation of the frequency of sensor observations during events) is a strategy that can be used to better characterize these types of systems that can

exhibit relatively long periods of unchanging flow and water quality followed by rapid changes during storms. Adaptive sampling poses multiple challenges, including synchronization of physical sample collection with in-situ sensing at individual sites, and synchronization of sampling across monitoring sites. In the following section, we describe the requirements that have been identified for sensing and data management infrastructure that enables automated, adaptive sampling, storm event detection, and in-network communications to quantify pollutant loads in urban water systems.

2.3.1. Requirements for Event Detection, Adaptive Sampling, and Site Communications

Urban monitoring sites designed to quantify pollutant loads from flashy runoff events must be able to detect those events in near real-time. For best results, sensor scan rates, or the rates at which observations are made, must be high enough to characterize the highest rate of change in discharge and pollutant concentrations at each monitoring site. Monitoring stations located at the outlet of smaller catchments generally require a high scan rate, whereas stations located within larger urban streams and canals may be effective using a slower scan rate. Additionally, the suite of sensors used must be able to distinguish storm events from other events that may generate discharge (e.g., lawn irrigation and other outdoor water use). For example, velocity/flow sensors and precipitation gages can be combined to ensure that sampling occurs during storm events (i.e., there is flow in a storm drain and rainfall is occurring).

An urban observatory designed for the estimation of TSS and TP loads in stormwater runoff needs to be capable of making adaptive observations triggered by

changing environmental conditions, such as sudden changes in turbidity, flow, or other parameters measured in-situ that would signal the beginning of an event to be sampled. This functionality can be supported when in-situ sensors are paired with automated samplers and a programmable datalogger. Programmable dataloggers provide the benefit of onsite data storage and the development of customizable programming logic to control sensor scanning rates, data recording intervals, and the triggering of automated sample collection. For example, Lewis and Eads (2001) describe a sampling logic that allows for water samples to be collected by an automated sampler when turbidity values rise above or fall below a specified threshold. This turbidity threshold sampling (TTS), which has been used for monitoring sediment loading from catchments affected by logging and other forestry practices, was designed to enable development of a strong surrogate relationship between turbidity values and suspended sediment concentrations (SSC). Likewise, in urban water systems, interfacing in-situ water quality and quantity sensors with automated samplers using custom programming logic can be critical for strategically collecting physical samples required for establishing surrogate relationships for TSS and TP.

Wireless communication allows sensor nodes to routinely transmit data to a centralized location for quality control and further analysis. The addition of a telemetry system to urban sensor nodes also provides flexibility and functionality for inter-nodal communications (Corke et al., 2010; Horsburgh et al., 2010; Kerkez et al., 2016). Advanced urban stormwater monitoring applications go beyond simple “sample and send” functionality by enabling in-network processing such as event detection and actuation of physical sample collection or adaptation of in-situ measurement timing (Corke et al., 2010).

This includes the ability to detect precipitation and runoff events at upstream sites, and then communicate event detection to downstream monitoring sites. Downstream monitoring sites are then able to "anticipate" changing water quality and quantity conditions and modify their data collection and sampling procedures accordingly. Facilitating this type of inter-network communication and processing requires that each monitoring node be equipped with appropriate communication infrastructure.

An obvious challenge and tradeoff for this type of operation is the increase in power required to operate nodes that may be remote and must be autonomously powered. However, in-network processing and event detection can be extremely beneficial for monitoring stormwater quality and quantity. In locations where storm events occur sporadically, it is often difficult or impossible to mobilize field personnel in time to sample the first flush of storm events and maintain them in the field for the duration of a storm at multiple sites. Additionally, automated sampling logic based on data from integrated sensors can detect and respond to conditions that humans cannot easily observe (e.g., triggering sample collection when a particular flow or turbidity threshold is reached). Thus, it is important that sample collection be automated.

2.4. An Urban Observatory for Monitoring Suspended Solids and Phosphorus – A Case Study

An urban observatory was established in the Northwest Field Canal (NWFC) in Logan, Utah, USA that demonstrates a specific implementation of the infrastructure and data collection scheme described in the previous sections. Much of Logan City's stormwater is conveyed out of the city through canals that were originally designed to carry

water diverted from the Logan River for agricultural irrigation. These combined conveyances are still used for this purpose, but in addition to return flows from agricultural irrigation, they also receive stormwater runoff from Logan City during storm events. This configuration is common in many parts of the western U.S. where larger municipalities grew within areas that were originally used for agriculture (City of Grand Junction, 2016; City of Sequim, 2016). Our specific case study was designed to test the hypothesis that continuous monitoring and surrogate relationships developed at the upstream and downstream ends of a study reach could be used to quantify pollutant loads contributed by stormwater runoff between the two sites. This hypothesis is addressed more in-depth in Chapter 3.

Combined flows in the NWFC eventually discharge to Cutler Reservoir, which is an impoundment on the Bear River originally built for irrigation, flood control, and hydropower generation. According to the recent Middle Bear River and Cutler Reservoir Total Maximum Daily Load (TMDL) study (Utah Division of Water Quality, 2010), Cutler Reservoir is impaired for low dissolved oxygen and excessive TP concentrations, with pollution from canal discharges and urban stormwater runoff identified as being significant nonpoint sources. The urban observatory described in this case study was designed to collect high frequency data for generating continuous estimates of TSS, TP, and total dissolved phosphorus (TDP) loads from urban stormwater runoff to the NWFC and, ultimately, to Cutler Reservoir.

2.4.1. Monitoring Site Selection and Infrastructure

The NWFC flows continuously during the irrigation season, which typically extends from April to October. Baseflows in the canal are diverted from the nearby Logan River, which originates in the Bear River Mountain range east of Logan City. The NWFC receives stormwater runoff from much of Logan's residential and commercial districts (Figure 2.1). The transition from residential land uses near the upstream end of the canal to commercial land uses toward the downstream end of the canal made the NWFC especially interesting as an observatory for investigating the effects of land use and spatial variability on stormwater quality.

Six monitoring sites were installed during the 2015 and the first half of the 2016 irrigation seasons in or near the NWFC to monitor stormwater quality and quantity. The monitoring sites fall into two site types: continuously monitored canal and continuously monitored stormwater outfall sites (Figure 2.1). The outfall sites were equipped with a sensor suite that allowed for the collection of flow via the area/velocity method, precipitation using a tipping bucket rain gage, and an automated sampler that allowed for physical sample collection during storm events (Table 2.1). Given the multifunction (e.g., irrigation, return flows, and stormwater) nature of many of the outfalls to the NWFC, storm event detection required both a rain gage to detect precipitation and a flow module to detect resulting discharge, as discharge measurements alone would have been insufficient for distinguishing between the multiple flow sources.

The two canal sites were located at the upstream and downstream ends of our study reach, which extended from the beginning of the NWFC to where it leaves Logan City's

boundary. Both sites were equipped with a suite of water quality and quantity monitoring sensors. Water quality parameters were measured in-situ via a multi-parameter sonde with DO, water temperature, pH, fluorescent dissolved organic matter (fDOM), and electrical conductivity sensors. The DO, pH, fDOM, and conductivity sensors required regular calibration (YSI EXO, 2012), and this was performed biweekly. A turbidimeter was also installed to collect measurements of water clarity in nephelometric turbidity units (NTU). The turbidimeters used are calibrated at the factory using a multi-point calibration as part of their regular maintenance and are not calibrated in the field. Similar to the outfall sites, an automated sampler was included for the collection of water samples during storm events.

All samples collected at both outfall and canal sites via automated samplers were split three ways and analyzed for TSS, TP and total dissolved phosphorus (TDP). Prior to analysis, TSS samples were refrigerated for no more than 7 days and TP and TDP samples were frozen. TSS analysis was performed according to Standard Method 2540 D using a 1.5 µm glass fiber filter (APHA, 2012). TP analyses were performed according to EPA Method 134-A Revision 4 using an acid-persulfate digestion and a discrete analyzer (AQ2, Seal Analytical, Mequon, Wisconsin, USA). Samples were analyzed for TDP using the same method after being filtered through a 0.45 µm nylon filter. The discrete analyzer used for TP and TDP analysis was calibrated before each use. For quality control of the phosphorus samples, laboratory blanks, laboratory blank spikes, sample spikes, and duplicates were analyzed, and the 0.01 mg/L method detection limit was verified using the USEPA's procedure for determining the method detection limit (USEPA, 2016). For

quality control, 10 percent of all TSS, TP, and TDP samples to be analyzed were selected and analyzed as duplicates.

Flow in the canal was estimated by converting depth measurements from a pressure transducer installed at each site to discharge via a site-specific rating curve. Backwater effects caused by the periodic installation of damming structures in the canal for the purpose of diverting water to irrigation head gates weakened the stage-discharge relationship at the upstream site (200 South 400 West, Logan, UT). Consequently, an acoustic Doppler velocity meter (ADVM) was installed at that site to obtain more accurate discharge measurements using the index velocity method (Levesque and Oberg, 2012). Precipitation was measured at the canal sites using a tipping bucket rain gage, and all data were transmitted to a centralized location via radio communications. See Table 2.1 for a more detailed description of equipment installed at all four monitoring sites.

Dataloggers, batteries, radios, flow modules, and automated samplers were housed in a 100 x 68 x 85 cm fiberglass storm box (Figure 2.2). At the canal sites and in cases where level ground or space for the enclosure was not located in close proximity to the monitored outfall, platforms were constructed for the enclosures. In the case of the canal sites, these platforms were also used for mounting the sensor housings and antenna masts (Panel b of Figure 2.2).

Outfall sites were installed within significant stormwater outfalls to the canal. The outfall sites were designed to be mobile so that we could move them to monitor as many outfalls as possible over the course of the project. The contributing area for the outfall, the distribution of land use-land cover within the catchment, and the availability of a

permanent structure where monitoring equipment could be installed were considered when selecting and prioritizing outfall locations to install the monitoring equipment.

Two outfall sites were initially installed at the outlet of a residential catchment (300 North 300 West, Logan, Utah) and a commercial catchment (1250 North 200 West, Logan, Utah). This first set of storm drains were monitored for the entire 2015 irrigation season (April to October) so that sampling procedures, datalogger programs, and adaptive sampling logic could be established. The outlet sites were then moved to the outlet of another predominately residential catchment (800 North 150 West, Logan, Utah) and a different commercial catchment (1300 North 200 West, Logan, Utah). This second set of storm drains was monitored for the first half of the 2016 irrigation season. Finally, the two outfall sites were moved to two additional locations (1000 North 225 West and 1400 North 200 West). This paper will address only the first four outfall sites monitored during the 2015 and first half of the 2016 irrigation seasons. See Table 2.2 for a more detailed description of the outfall catchments monitored during this study.

2.4.2. Procedure for Storm Event Detection and Communication

The NWFC monitoring sites were used for the detection and monitoring of pollutant loading from stormwater runoff events. The dataloggers at the outfall sites were programmed to scan the sensors at 1-minute intervals, obtaining discharge, water temperature, and precipitation values. Under non stormwater runoff normal conditions, instantaneous discharge and water temperature values for the current scan were recorded at 15-minute intervals. Precipitation values were written to a separate file at 5-minute intervals and represent a summation of precipitation for the preceding 5-minute period.

The canal sites were programmed to scan sensors at 5-minute intervals. During each scan, the multi-parameter water quality sonde measuring pH, DO, water temperature, conductivity, and fDOM made high frequency (~4 Hz) continuous measurements and calculated a moving average of a defined window of values (YSI EXO, 2012). The turbidity sensor made a burst of 100 readings in 5 seconds when the measurement command was initiated. The mean, median, and variance values from the 100 measurements were then recorded. While sensors were scanned at 5-minute intervals to detect changing conditions, current and instantaneous values from the pressure transducer, sonde, and turbidity sensors were recorded at 15-minute intervals under non-storm conditions. Precipitation values were recorded every 15 minutes and represent the summation of precipitation over the 15-minute interval. The short scan intervals allowed for near real-time detection of changes in flow and/or water quality representing the onset of an event, but had to be balanced with the response time of the sensors we used, the time required for sensors with integrated wipers to execute a wipe, and the power budget for each of our site types.

Variables measured in-situ at the outfall sites were used to indicate stormwater runoff events. However, before these events could be detected accurately, an initial period of monitoring discharge from storm drains was required. We discovered that aging drainage infrastructure produced baseflows in some storm drains caused by leaky pipes and groundwater infiltration. We also discovered that Logan's stormwater drainage system serves multiple purposes. Some Logan citizens have the option of irrigating their lawns with water extracted from one of the four major irrigation canals in the city. Irrigation

return flows drain through stormwater conveyances, thus creating a scenario where discharge could be measured from a storm drain without an actual storm event occurring. Thus, our system required us to define stormwater runoff events as discharge from the storm drain above an event flow threshold (e.g., greater than observed baseflow in the storm drain) during the occurrence of precipitation (See Figure 2.3, which shows the logic for event detection and adaptive sampling). The initial monitoring period at each storm drain allowed us to "train" our sampling logic and set an event flow threshold that distinguished between baseflows, return flows, and stormwater runoff events.

Event detection at the outfall sites was programmed to initiate an alternate monitoring and sampling scheme at both the outfall and canal sites. This alternative sampling logic (Figure 2.3) is described in detail in the following section. When precipitation and discharge above the minimum threshold were observed at an outfall site, a binary flag was sent from that site via radio communications to the canal sites. Because our sites were battery powered, and power consumption was a concern, we implemented a "need-based" communication scheme. The event flag was transmitted to canal sites only when an event was detected at the outfall sites. This reduced power consumption relative to a scheme that requires regular data retrieval commands initiated by the canal sites.

2.4.3. Adaptive Sampling Procedure for Outfall and Canal Sites

At the outfall sites, upon detection of discharge above the event threshold and precipitation greater than "0" within a sensor scan, the datalogger was programmed to initiate a first flush sampling regime. The data recording interval was modified to 1-minute (rather than the 15-minute intervals during nonevent conditions), and the automated

sampler was triggered by the datalogger to collect a sample every 3 minutes for the estimated duration of the first flush. A sampling interval of 3 minutes during the first flush was determined based on the time required to purge the suction line before and after sample collection. A shorter sampling interval would not have been feasible given the type of automated sampler we used and the length of suction hose required. The duration of the first flush was initially estimated to be the first 15 minutes of the storm event (Grisham, 1995). After examination of resulting data, this estimate was then determined to be acceptable for the majority of events at each outfall site. After the first flush, the datalogger was programmed to trigger the automated sampler to collect a sample every 15 minutes for the remainder of the storm event duration or until a maximum of 24 samples were collected per storm event.

Upon receipt of a stormwater runoff event flag from an outfall site, the datalogger at each canal site was programmed to initiate adaptive, event-based sampling. The data recording interval was modified to 5-minutes (rather than the 15-minute intervals during non-event conditions), and the collection of samples based on a TTS sampling scheme was initiated. During events, the datalogger was programmed to trigger automated samples as turbidity values rose above or fell below predefined thresholds. Based on the suggestions of Lewis (1996), thresholds were determined by evenly spaced square root transformed turbidity values. The thresholds were calculated by:

$$TT_i = \left({TT_1}^{0.5} + (i-1) * \frac{{TT_n}^{0.5} - {TT_1}^{0.5}}{n-1}\right)^2 \qquad i = 1 \ldots n \tag{2.2}$$

where TT_i = the i^{th} turbidity threshold; TT_1 = the initial turbidity threshold calculated as 1.05 times the first turbidity value after an event is detected; TT_n = the maximum turbidity

threshold calculated as TT_1 + 200 NTU; and n = the number of rising thresholds. Thresholds were calculated for a range of the initial turbidity threshold plus 200 NTU. The 200 NTU range was determined based on existing storm event data, which showed that turbidity values tended to fluctuate approximately 200 NTU during storm events at the downstream canal site.

2.5. Results and Discussion

2.5.1. Spatial TSS and TP Loading Variability in the Northwest Field Canal

The inclusion of upstream and downstream canal sites in the observatory allowed for data collection at locations in the canal with varying degrees of stormwater influence. Figure 2.4 shows a boxplot of the TSS concentrations at the two sites for the 2015 irrigation season and the first half of the 2016 irrigation season. A comparison is made between the upstream and downstream sites under baseflow and storm event conditions. It is apparent from the figure that TSS concentrations are typically higher at the downstream canal site. It is also apparent that TSS concentrations at the upstream site vary little between baseflow and storm event conditions, with a difference between median concentrations of 0.97 mg TSS/L. Concentrations at the downstream site are higher under both baseflow and storm event conditions. Under baseflow conditions, irrigation return flows received between the upstream and downstream sites and sediment resuspension are likely causes of this increase in TSS concentration. Under storm event conditions there is a considerable increase in TSS both from upstream to downstream (median concentration increase of 31.2 mg TSS/L) and relative to baseflow conditions at the downstream site (median concentration increase of

21 mg TSS/L). This increase can be attributed to stormwater inputs during storm events. Figure 2.5 shows TSS concentrations at the two canal monitoring sites during a storm event that occurred on May 10, 2016. The plot shows a typical pattern we observed throughout our study in which TSS peaks are much higher at the downstream canal monitoring site.

The outfall sites allowed for monitoring of runoff from multiple catchments simultaneously during multiple individual storm events. This allowed for comparisons to be made of how different catchments responded during individual storm events (comparing one event across multiple sites) and how those catchments responded across multiple storm events (comparing multiple events at the same site). Although precipitation was not exactly the same at each site during each event, our results showed that runoff varied greatly across catchments due to catchment land cover, imperviousness, and storm event characteristics. The land use/land cover description for each monitored catchment is listed in Table 2.2.

TSS, TP, and TDP event mean concentrations (EMC) for the two outfall sites monitored during the 2015 irrigation season are shown in Figure 2.6. EMCs are often used to predict pollutant concentrations and estimate mass loadings from various land uses and degrees of imperviousness (Charbeneau and Barrett, 1998; Lin, 2004). Our analysis found that EMCs for TSS at 300 North and 1250 North do not appear to follow any obvious trend (Panel b of Figure 2.6). For the 15 events monitored, TSS EMCs ranged from 3 – 640 mg/L with the median EMC of 88 mg/L. These values are similar to those found in the literature for residential and commercial land uses (Lin, 2004; Rodriguez-Hernandez et al., 2013). It was found, however, that EMCs from the 300 North catchment were higher for both TP and TDP than the 1250 North catchment for the 15 storm events monitored. This is evident

from Figure 2.6 (Panels c and d). At 300 North, EMCs for TP and TDP range from 0.120 – 1.04 mg/L and 0.0160 – 0.485 mg/L, respectively. At 1250 North, EMCs for TP and TDP range from 0.0591 – 0.825 mg/L and 0.0113 – 0.209 mg/L, respectively. These values are similar to those found in the literature (Brezonik and Stadelmann, 2002; Lin, 2004). The higher values in the 300 North catchment could be due to the higher percentage of residential land use in that catchment. Higher concentrations of phosphorus from residential land uses have been observed by others (Dennis, 1986; Waschbusch et al., 1999; Brezonik and Stadelmann, 2002).

The range and variability in EMCs calculated for storms during the 2015 irrigation season at these two outfall sites illustrates the inadequacy of relying on EMCs alone to estimate loads. In some cases, such as TSS at 1250 North, the EMC varies by as much as two orders of magnitude, indicating the potential of greatly overestimating the TSS load if estimates are based only on land use and EMCs. Based on these results, the magnitude of the EMC is dependent not only on land use, but also on the characteristics of the storm event (e.g., average and peak rainfall intensities, antecedent dry period, rainfall volume, and the duration of the event). Table 2.3 shows these characteristics for the storm events monitored in 2015. The June 10, 2015, July 27, 2015, and September 14, 2015 storm events, which correspond with the largest EMCs of TSS at 1250 North and EMCs of TP at 300 North, likewise correspond with the longest antecedent dry periods. These results emphasize the importance of sampling multiple storm events at multiple locations. These comparisons would not have been possible without the data created using the urban observatory's infrastructure.

2.5.2. Temporal Pollutant Loading Variability in the Northwest Field Canal

Higher concentrations and the majority of pollutant mass load occurring at the start of the storm event indicate the existence of the first flush phenomenon. Our adaptive sampling scheme, which collected multiple, discrete samples at outfall monitoring sites with more frequent samples toward the beginning of each storm allowed us to examine the variability of pollutant concentrations within any single storm event and characterize the first flush. According to Bertrand-Krajewski et al. (1998), the occurrence of a significant first flush can be defined when 80 percent of the mass load is transported in the first 30 percent of runoff volume. According to this definition, the first flush phenomenon for TSS was observed in only two of the 15 storms at the 300 North site. None of the other outfall sites had first flush events that met this criterion. Similarly, Wanielista and Yousef (1993) proposed that a significant first flush could be defined when 50 percent of the mass load is transported in the first 25 percent of runoff volume. According to this definition, the first flush phenomenon for TSS was observed in approximately 33 percent of the storms at 300 North, 11 percent of the storms at 1250 North, 31 percent of the storms at 800 North, and 46 percent of the storms at 1300 North (Table 2.4).

Based on the results shown in Table 2.4, the presence of a first flush is affected by catchment area. The two monitoring sites installed during the 2015 irrigation season, 300 North and 1250 North, have catchment areas of 0.041 km^2 and 0.205 km^2, respectively. The smaller catchment (300 North), experienced a first flush of TSS more often than the larger catchment (1250 North). Likewise, the smaller catchment monitored during the first half of the 2016 irrigation season (1300 North) experienced a first flush of TSS more often

than the larger catchment (800 North). This agrees with the findings in Lee et al. (2002) and the National Research Council's report on Urban Stormwater Management in the United States (National Research Council, 2009) that smaller catchments tend to be more prone to a first flush.

As a further illustration, Sansalone and Cristina (2004) described a concentration-based first flush (CBFF) as a high pollutant concentration occurring during the rising limb of the runoff hydrograph. Figure 2.7 shows a CBFF that was observed at the 800 North outfall site during a storm that occurred on April 10, 2016. The maximum TSS concentration observed during this storm event was 474.8 mg/L, occurring at 04:21 MST, just 7 minutes after the start of the storm event. Because samples were collected at 3-minute intervals at the start of the event, the adaptive sampling logic was able to capture the CBFF during this intense storm event of short duration.

A major objective in our infrastructure design was to enable the collection of data to support development of surrogate relationships between turbidity and TSS and TP at the continuous canal sites so that we could derive continuous estimates of TSS and TP concentrations at the upstream and downstream monitoring sites of the canal. To demonstrate the observatory's ability to adequately sample storm events for the derivation of surrogate relationships, Figures 2.8 and 2.9 show the distribution of measured turbidity values as percent exceedance for the upstream and downstream canal sites, respectively. Turbidity values observed during storm events in 2015 and the first half of 2016 are represented by the filled points on the plot. Physical samples collected to be used in development of surrogate regression models are represented by open circles in the plots

and were plotted using the turbidity value corresponding to the time at which each sample was collected. Figure 2.8 and Figure 2.9 show that our adaptive sampling procedure enabled the collection of samples across nearly the full range of observed turbidity values at the two sites. The plots show that fewer physical samples were collected at higher turbidity values; however, this is to be expected as the higher turbidity values tend to occur sporadically and represent only about 10 – 15 % of the observed turbidity values.

2.5.3. Event Detection, Adaptive Sampling, and Inter-site Communications

In our urban observatory, the outfall sites were responsible for event detection and messaging to the canal sites that storm events were occurring. Communication between sites was critical because, while it was relatively easy to detect a storm event at the outfall sites because they responded quickly to runoff from a storm event, it was much more difficult to accurately detect the beginning of a storm event in the canal. Water diverted from the Logan River for agricultural purposes muted the stormwater signal to some extent due to dilution, and there was also a travel time effect as stormwater flowed from outfall locations to the canal monitoring sites. Initiating storm event sampling at the canal sites, therefore, relied on messaging from the storm drains.

Figure 2.10 shows a plot of turbidity values at the downstream canal site (1800 North) during the storm event on May 10, 2016. The times at which automated samples were collected are indicated by an “X,” and each was plotted at the turbidity value corresponding to the time at which the sample was collected for visualization purposes. The TTS method ensured that samples were collected throughout the range of observed turbidity values and during points of inflection in the pollutograph, capturing the response

of the canal to the stormwater inputs while minimizing the number of samples collected. These data provide a good estimate of the continuous shape of the pollutograph.

As a more comprehensive example, Figure 2.11 shows the observatory's ability to detect storm events at the outfall sites, communicate that information to corresponding canal sites, and coordinate sampling across the sites. This figure depicts a sampling event that occurred during the same storm on May 10, 2016. The 800 North and 1300 North outfall monitoring sites detected the storm event as flows increased, and initiated sampling. The storm event flag was communicated to the upstream (200 South) and downstream (1800 North) canal sites, which then initiated TTS. In this event, as well as in others we observed, the turbidity pulse was not as pronounced at the upstream canal site as it was at the downstream canal site. Because of this, we added additional logic to the program at the upstream canal site to first look for turbidity increases after having received a message from an outfall site that a storm even was occurring, but if none are present samples are collected at 30-minute intervals.

2.5.4. Surrogate Relationships in the Northwest Field Canal

Simple linear surrogate relationships for TSS and turbidity are shown in Figure 2.12 for the upstream and downstream canal sites. The least-squares regression equation and R^2 values are also shown in their corresponding plots. The R^2 value at the downstream canal site was 0.868, which is greater than the R^2 value at the upstream canal site (0.725). This may be due to the lower range of turbidity values at the upstream site. The median turbidity value at which samples were collected and analyzed for TSS at the upstream site was less than 3 NTU. Evidence of this can be seen in Panel a of Figure 2.12, where a cluster of

points on the graph are located near 3 NTU. At these lower turbidity values it becomes more difficult to distinguish actual changes in water quality conditions given the accuracy of the turbidity sensors we used.

Figure 2.13 shows the predicted TSS concentrations at the upstream canal site for the 2015 irrigation season derived using the surrogate relationship shown in Figure 2.12. These predictions were made under the assumption that the TSS-turbidity relationship remained constant for the duration of the study period. The 95% prediction intervals for the estimated TSS concentrations are represented by the shaded gray region. The red points on the plot represent actual measured TSS values from our samples. The gaps in the predicted concentration values correspond with times where the canal's headgate was shut to prevent flooding. During this time, the canal still received inflow from stormwater outfalls, but the flows were low and intermittent enough that we were unable to maintain our in-situ water quality sensors in the canal. For much of the month of May, the headgate was shut, restricting us from sampling during one of the wettest months of the year.

The predicted TSS concentrations shown in Figure 2.13 appear to follow a seasonal trend, decreasing through the Spring and early Summer months. This could be due to a combination of factors, including spring snowmelt, which results in elevated turbidity and suspended sediment in the Logan River. While the NWFC does not receive higher flows during Spring snowmelt due to manually operated hydraulic controls (diverted flows are driven by water rights and irrigation demands), it does appear to receive higher suspended solids. This can be significant as pollutants that sorb onto particulates (e.g., TP) may be found in higher concentrations during the first half of the irrigation season. Additionally,

as the vegetation on the banks and bed of the canal grows throughout the irrigation season, there is less opportunity for mobilization or resuspension of sediment from erosion within the channel, which likely also contributes to the overall decline in observed turbidity and TSS concentrations estimated via the surrogate relationship.

2.6. Conclusions

In this paper we have presented the required infrastructure and monitoring methods needed for capturing the spatial and temporal variability of TSS and TP loads in an urban stormwater system. Our case study from the NWFC shows how the collection of high frequency data in multiple catchments can aid in better understanding the processes that control pollutant load variability. For example, the data we present show the dynamic response of each of the different catchments to rainfall inputs, and the observatory's ability to capture high frequency data and coordinated samples enabled us to characterize the first flush phenomenon in multiple catchments. Additionally, the ability to compare catchment responses for a given storm event allowed us to better understand the influence drainage area has on the presence of the first flush and on pollutant EMC. It was found that smaller drainage areas are more prone to the first flush phenomenon. We also found that certain EMCs varied by two orders of magnitude within a single catchment across different storm events, indicating that simple EMCs are inadequate for estimating pollutant concentrations across a range of storms. Additional analyses could potentially relate EMCs to not only catchment characteristics (e.g., land use), but also storm event characteristics such as the antecedent dry period. However, without the ability to trigger event-based samples at

multiple catchment outlets, the observation of storm event temporal variability and first flush analysis would not have been possible.

The combination of adaptive sampling and inter-site communication proved to be critical in the observatory's ability to detect events and trigger samples at times of interest and at multiple locations in the urban drainage system. The example of event detection and inter-site communications represented by data presented from the storm event on May 10, 2016 demonstrates the degree of coordination needed to synchronize sampling efforts during runoff events across monitoring sites. This effectively reduced the number of samples collected, reduced field crew costs, and the autonomous nature of the observatory ensured that no storm events were missed due to the time of day at which the event occurred. This level of coordination would not have been possible without the observatory's ability to detect events at the outfall sites and communicate that detection to the canal sites. The TTS sampling scheme implemented at the canal sites also ensured that samples were collected throughout the entire range of observed turbidity values, which means that surrogate relationships derived from these datasets will not extrapolate beyond the range of observed turbidity values.

There are many advantages and some disadvantages to the use of an urban observatory similar to the one installed in the NWFC drainage area. Advantages include the ability to synchronize the monitoring of multiple catchments, detect storm events in real-time and adapt sampling frequencies accordingly, and trigger the collection of samples based on changes in environmental conditions that would be undetectable without the use of in-situ sensors. These advantages allow for better characterization of the spatial and

temporal variability of pollutant loading events. The high-resolution data obtained using this approach can also provide necessary information for testing and improving stormwater models. Disadvantages include the cost of installing and operating observatory equipment. While the observatory reduced likely personnel costs required for sampling across multiple sites, the instrumentation described in Table 2.1 is not inexpensive. However, it would be logistically challenging for personnel to perform the same function as automated equipment. Additionally, regular maintenance becomes a liability. While regular required maintenance reduces the chance of equipment malfunction, it does not remove the possibility entirely and also does not guard against potential theft and vandalism.

As sensor network technologies continue to improve, so will our ability to monitor pollutant fluxes in both rural and urban watersheds and derive surrogate relationships for pollutant concentrations. An example of such improvements might include more robust event detection and communication performed by Internet connected microcontrollers and dataloggers. The use of cellular phone modems for Internet downloads of flow, snowpack, precipitation, meteorological, or other data from other Internet connected devices could allow for communication between individual monitoring sites and data available for other reaches of an urban water system/watershed/river basin. This functionality could help in better predicting the onset of storm events, adapting physical sampling and sensor observation frequencies at urban observatory sites during flashy events that only occur at upstream reaches of the watershed, and in identifying processes that control the baseflow signal in combined urban conveyances. Another potential enhancement would be to include additional surrogate measurements to the sensor suite. For example, some studies

have emphasized the significance of particle size in the use of turbidity to predict TSS and TP. Monitoring sites with in-situ particle size analyzers might aid in deriving stronger surrogate relationships in both rural and urban environments (Landers and Sturm, 2013). This information may provide data more representative of physical conditions within a catchment for more accurate predictive modeling at the watershed scale and improved water resources management.

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. doi:10.1007/s10661-010-1788-6

Ahearn, D. S., Sheibley, R. W., Dahlgren, R. a., Anderson, M., Johnson, J., & Tate, K. W. (2005). Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California. Journal of Hydrology, 313(3–4), 234–247. doi:10.1016/j.jhydrol.2005.02.038

APHA. (2012). Standard Methods for the Examination of Water and Wastewater. (E. W. Rice, R. B. Baird, A. D. Eaton, & L. S. Clesceri, Eds.) (22nd ed.). Washington, D.C.: American Public Health Association, American Water Works Association, Water Environment Federation.

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. doi:10.1016/S0043-1354(97)00420-X

Brezonik, P. L., & Stadelmann, T. H. (2002). Analysis and Predictive Models of Stormwater Runoff Volumes, Loads, and Pollutant Concentrations from Watersheds in the Twin Cities Metropolitan Area, Minnesota, USA. Water research, 36(7), 1743–1757. http://www.ncbi.nlm.nih.gov/pubmed/12044074

Cassidy, R., & Jordan, P. (2011). Limitations of Instantaneous Water Quality Sampling in Surface-Water Catchments: Comparison with Near-Continuous Phosphorus Time-Series Data. Journal of Hydrology, 405(1–2), 182–193. doi:10.1016/j.jhydrol.2011.05.020

Charbeneau, R. J., & Barrett, M. E. (1998). Evaluation of methods for estimating stormwater pollutant loads. Water Environment Research, 70(7), 1295–1302.

Christensen, V. G., Rasmussen, P. P., & Ziegler, A. C. (2002). Real-time water-quality monitoring and regression analysis to estimate nutrient bacteria in Kansas Streams. http://ks.water.usgs.gov/pubs/reports/vgc.0610.html

City of Grand Junction. (2016). Grand Junction Stormwater Management Manual. Grand Junction City Municipal Code Volume II: Development and Regulations. http://www.codepublishing.com/CO/GrandJunction/html2/GrandJunction28/GrandJunction2852.html. Accessed 1 January 2016

City of Sequim. (2016). Storm and Surface Water Master Plan. http://www.sequimwa.gov/DocumentCenter/View/7735. Accessed 1 January 2016

Corke, P., Wark, T., Jurdak, R., Hu, W., Valencia, P., & Moore, D. (2010). Environmental Wireless Sensor Networks. In Proceedings of the IEEE (Vol. 98, pp. 1903–1917). doi:10.1109/JPROC.2010.2068530

Deletic, A. B., & Maksimovic, C. T. (1998). Evaluation of Water Quality Factors in Storm Runoff from Paved Areas. Journal of Environmental Engineering, 124(September), 869–879.

Deletic, A., Maksimovic, [Cbreve]edo, & Ivetic, M. (1997). Modelling of storm wash-off of suspended solids from impervious surfaces. Journal of Hydraulic Research, 35(1), 99–118. doi:10.1080/00221689709498646

Dennis, J. (1986). Phosphorus Export from a Low-Density Residential Watershed and an Adjacent Forested Watershed. Lake and Reservoir Management, 2(1), 401–407. doi:10.1080/07438148609354665

Fisher, J. R., Dvorak, B. I., & Admiraal, D. M. (2016). Pollutant Load Estimates Using Regression Models with In-Stream Measurements. Journal of Environmental Engineering, 142(3), 4015081. doi:10.1061/(ASCE)EE.1943-7870.0001049

Goonetilleke, A., Thomas, E., Ginn, S., & Gilbert, D. (2005). Understanding the role of land use in urban stormwater quality management. Journal of environmental management, 74(1), 31–42. doi:10.1016/j.jenvman.2004.08.006

Grisham, D. M. (1995). Designing for the first flush.pdf. Civil Engineering, 65(11), 67–69.

Gunaratne, G. L., Vogwill, R., & Hipsey, M. R. (2014). Effects of Changing Land Use on Seasonal Nutrient Wash-off Processes in an Urbanising Coastal Catchment in a Sub-tropical Climate. In 13th International Conference on Urban Drainage (pp. 1–8). Sarawak, Malaysia.

Horsburgh, J. S., Spackman Jones, A., Stevens, D. K., Tarboton, D. G., & Mesner, N. O. (2010). A Sensor Network for High Frequency Estimation of Water Quality Constituent Fluxes Using Surrogates. Environmental Modelling & Software, 25(9), 1031–1044. doi:10.1016/j.envsoft.2009.10.012

Hannouche, A., Joannis, C., & Chebbo, G. (2016). Assessment of Total Suspended Solids (TSS) Event Load and its Uncertainties in Combined Sewer System from Continuous Turbidity Measurements. Urban Water Journal, 1–8. doi:10.1080/1573062X.2016.1254256

Jones, A. S., Stevens, D. K., Horsburgh, J. S., & Mesner, N. O. (2011). Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and Total Phosphorus Concentrations. JAWRA Journal of the American Water Resources Association, 47(2), 239–253. doi:10.1111/j.1752-1688.2010.00505.x

Kerkez, B., Gruden, C., Lewis, M., Montestruque, L., Quigley, M., Wong, B., et al. (2016). Smarter Stormwater Systems. Environmental Science & Technology, 50(14), 7267–7273. doi:10.1021/acs.est.5b05870

Kirchner, J. W., Feng, X., Neal, C., & Robson, A. J. (2004). The Fine Structure of water-Quality Dynamics: The (High-Frequency) Wave of the Future. Hydrological Processes, 18(7), 1353–1359. doi:10.1002/hyp.5537

Landers, M. N., & Sturm, T. W. (2013). Hysteresis in Suspended Sediment to Turbidity Relations Due to Changing Particle Size Distributions. Water Resources Research, 49(9), 5487–5500. doi:10.1002/wrcr.20394

Lee, J. H., Bang, K. W., Ketchum, J. H., Choe, J. S., & Yu, M. J. (2002). First flush analysis of urban storm runoff. Science of the Total Environment, 293, 163–175. doi:10.1016/S0048-9697(02)00006-2

Leecaster, M. K., Schiff, K., & Tiefenthaler, L. L. (2002). Assessment of efficient sampling designs for urban stormwater monitoring. Water research, 36(6), 1556–64. http://www.ncbi.nlm.nih.gov/pubmed/11996344

Levesque, V. A., & Oberg, K. A. (2012). Computing Discharge Using the Index Velocity Method. Reston.

Lewis, J. (1996). Turbidity-Controlled Suspended Sediment Sampling for Runoff-Event Load Estimation. Water Resources Research, 32(7), 2299–2310.

Lewis, J., & Eads, R. (2001). Turbidity Threshold Sampling For Suspended Sediment Load Estimation. In Proceedings of the Seventh Federal Interagency Sedimentation Conference, March 25 to 29, 2001, Reno, Nevada (pp. 110–117).

Lin, J. P. (2004). Review of Published Export Coefficient and Event Mean Concentration (EMC) Data. Vicksburg, MS. www.wes.army.mil/el/wrap

McKee, L. J., & Gilbreath, A. N. (2015). Concentrations and loads of suspended sediment and trace element pollutants in a small semi-arid urban tributary, San Francisco Bay, California. Environmental Monitoring and Assessment, 187(8), 499. doi:10.1007/s10661-015-4710-4

Miguntanna, N. S., Egodawatta, P., Kokot, S., & Goonetilleke, A. (2010). Determination of a Set of Surrogate Parameters to Assess Urban Stormwater Quality. The Science of the total environment, 408(24), 6251–9. doi:10.1016/j.scitotenv.2010.09.015

Nasrabadi, T., Ruegner, H., Sirdari, Z. Z., Schwientek, M., & Grathwohl, P. (2016). Using Total Suspended Solids (TSS) and Turbidity as Proxies for Evaluation of Metal Transport in River Water. Applied Geochemistry, 68, 1–9. doi:10.1016/j.apgeochem.2016.03.003

National Research Council. (2009). Urban Stormwater Management in the United States.

Rasmussen, P. P., Gray, J. J. R., Glysson, G., & Ziegler, A. C. (2011). Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data. In Applications of Hydraulics Section C, Sediment and Erosion Techniques (p. 52). Reston: U.S. Geological Survey.

Rodriguez-Hernandez, J., Fernández-Barrera, A. H., Andrés-Valeri, V. C. A., Vega-Zamanillo, A., & Castro-Fresno, D. (2013). Relationship between Urban Runoff Pollutant and Catchment Characteristics. Journal of Irrigation and Drainage Engineering, 139(10), 833–840. doi:10.1061/(ASCE)IR.1943-4774.0000617

Ryberg, K. R. (2006). Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05. Reston.

Sansalone, J. J., & Cristina, C. M. (2004). First Flush Concepts for Suspended and Dissolved Solids in Small Impervious Watersheds, (November), 1301–1314.

Settle, S., Goonetilleke, A., & Ayoko, G. a. (2007). Determination of Surrogate Indicators for Phosphorus and Solids in Urban Stormwater: Application of Multivariate Data Analysis Techniques. Water, Air, and Soil Pollution, 182(1–4), 149–161. doi:10.1007/s11270-006-9328-2

Tang, Z., Engel, B. a, Pijanowski, B. C., & Lim, K. J. (2005). Forecasting land use change and its environmental impact at a watershed scale. Journal of environmental management, 76(1), 35–45. doi:10.1016/j.jenvman.2005.01.006

Tiefenthaler, L. L., Schiff, K. C., & Leecaster, M. K. (2001). Temporal Variability Patterns of Stormwater Concentrations in Urban Stormwater Runoff. Westminster, CA.

USEPA. (1983). Results of the Nationwide Urban Runoff Program: Volume I - Final Report. Washington, D.C.

USEPA. (2016). Definition and Procedure for the Determination of the Method Detection Limit, Revision 2. EPA 821-R-16-006. https://www.epa.gov/sites/production/files/2016-12/documents/mdl-procedure_rev2_12-13-2016.pdf

Utah Division of Water Quality. (2010). Middle Bear River and Cutler Reservoir Total Maximum Daily Load (TMDL). Salt Lake City.

Viviano, G., Salerno, F., Manfredi, E. C., Polesello, S., Valsecchi, S., & Tartari, G. (2014). Surrogate measures for providing high frequency estimates of total

phosphorus concentrations in urban watersheds. Water Research, 64, 265–277. doi:10.1016/j.watres.2014.07.009

Wanielista, M. P., & Yousef, Y. A. (1993). Stormwater Management. John Wiley and Sons, Inc.

Waschbusch, B. R. J., Selbig, W. R., & Bannerman, R. T. (1999). Sources of Phosphorus in Stormwater and Street Dirt from Two Urban Residential Basins In Madison , Wisconsin , 1994 – 95. Middleton.

Wong, B. P., & Kerkez, B. (2016). Adaptive Measurements of Urban Runoff Quality. Water Resources Research. doi:10.1002/2015WR018013

Yao, L., Wei, W., & Chen, L. (2016). How does imperviousness impact the urban rainfall-runoff process under various storm cases? Ecological Indicators, 60, 893–905. doi:10.1016/j.ecolind.2015.08.041

YSI EXO. (2012). EXO-User Manual.

Table 2.1. List of equipment deployed at the two types of monitoring sites

Site Type	Item	Measurement Frequency	Model
Continuously Monitored Stormwater Outfall Site	Area velocity flow module	1 minute	Teledyne ISCO 2150
	Automated sampler with 24,1-liter sample bottles	Adaptive (first flush)	Teledyne ISCO 3700
	Tipping bucket rain gage	5 minute	Campbell Scientific TE525WS
	Programmable datalogger	-	Campbell Scientific CR800
	Radio	-	Campbell Scientific RF450
	Antenna	-	PCTEL yagi
	Solar Panel	-	BP Solar 10W
	12V dc deep cycle marine battery	-	Super Start Marine
Continuously Monitored Canal Site	Multiparameter water quality sonde	5 minute (event) 15 minute (baseflow)	YSI EXO2
	Turbidity sensor	5 minute (event) 15 minute (baseflow)	FTS DTS-12
	Pressure transducer	5 minute (event) 15 minute (baseflow)	Campbell Scientific CS451
	Automated sampler with 24,1-liter sample bottles	Adaptive (TTS)	Teledyne ISCO 3700
	Tipping bucket rain gage	15 minute	Campbell Scientific TE525WS
	Programmable datalogger	-	Campbell Scientific CR800

	Side looking ADVM (200 South)	15 minute	Sontek SL3000
	Radio	-	Campbell Scientific RF450
	Antenna	-	PCTEL yagi
	Solar Panel	-	BP Solar 10W
	12V dc deep cycle marine battery	-	Super Start Marine

Table 2.2. Characteristics of catchments monitored

Catchment	Catchment Area (km^2)	Land Use (percent coverage)			Percent Impervious
		Residential	Commercial	Street	
300 North	0.041	29.5	39.0	31.5	62.3
1250 North	0.205	11.1	76.3	12.5	69.5
800 North	0.359	45.5	36.1	18.4	43.5
1300 North	0.065	18.0	78.9	3.1	84.6

Table 2.3. Storm event characteristics for the 2015 storm events at the two continuously monitored outfall sites

Site Name	Storm Event	Average Intensity (mm/hr)	Peak Intensity (mm/hr)	Antecedent Dry Period (days)	Duration (hours)	Rainfall Depth (mm)
300 North	May 6, 2015	-	-	-	-	-
	May 18, 2015	-	-	-	-	-
	May 20, 2015	-	-	-	-	-
	May 23, 2015	1.571	3.048	1.315	8.083	12.7
	May 27, 2015	2.814	9.144	0.731	1.083	3.048
	June 10, 2015	7.662	51.82	6.942	2.917	22.35
	July 5, 2015	1.321	9.144	24.32	2.5	3.302
	July 8, 2015	2.477	6.096	3.099	1.333	3.302
	July 27, 2015	2.111	9.144	18.09	1.083	2.286
	August 3, 2015	1.524	18.29	6.872	8.5	12.95
	August 7, 2015	2.032	6.096	4.274	1.5	3.048
	September 14, 2015	1.463	21.34	34.05	2.083	3.048
	September 15, 2015	1.159	12.19	0.273	5.917	6.858
	September 16, 2015	2.629	21.34	0.494	11.5	30.23
	October 3, 2015	1.806	15.24	17.18	2.25	4.064
1250 North	May 6, 2015	0.733	3.048	1.833	6.583	4.826
	May 18, 2015	4.572	9.144	1.758	1.5	6.858
	May 20, 2015	4.222	15.24	0.692	1.083	4.572
	May 23, 2015	1.541	6.096	1.356	7.75	11.94
	May 27, 2015	4.570	9.144	0.676	0.667	3.048
	June 10, 2015	8.467	42.67	6.92	3	25.4
	July 5, 2015	0.776	6.096	24.27	4.583	3.556
	July 8, 2015	1.972	9.144	1.858	1.417	2.794
	July 27, 2015	4.572	9.144	8.503	0.5	2.286

August 3, 2015	1.451	9.144	6.878	8.583	12.45
August 7, 2015	0.703	6.096	4.088	4.333	3.048
September 14, 2015	5.588	24.38	34.06	0.5	2.794
September 15, 2015	1.078	9.144	0.145	5.417	5.842
September 16, 2015	2.888	24.38	0.275	11.08	32
October 3, 2015	0.363	3.048	1.541	3.5	1.27

Table 2.4 First flush analysis at the four outfall sites monitored during the duration of the study

Catchment	Monitoring Period	Number of Storms	Catchment Area (km^2)	Percent of Storms with TSS First Flush (∑Mass/∑Volume)	
				30/80	25/50
300 North	April – October 2015	15	0.041	13.3%	33.3%
1250 North	April – October 2015	18	0.205	0%	11.1%
800 North	March – May 2016	13	0.359	0%	30.8%
1300 North	March – May 2016	13	0.065	0%	46.2%

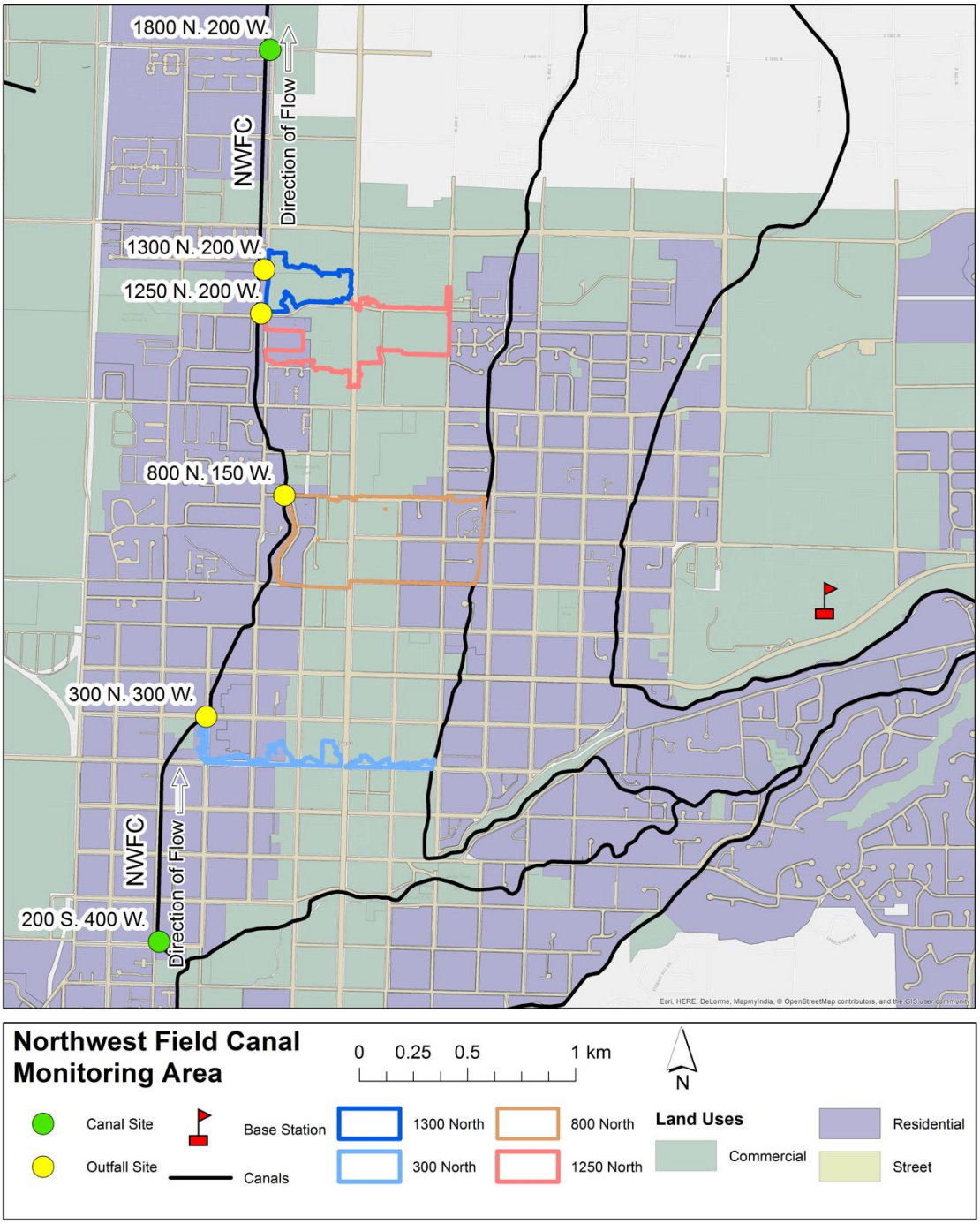
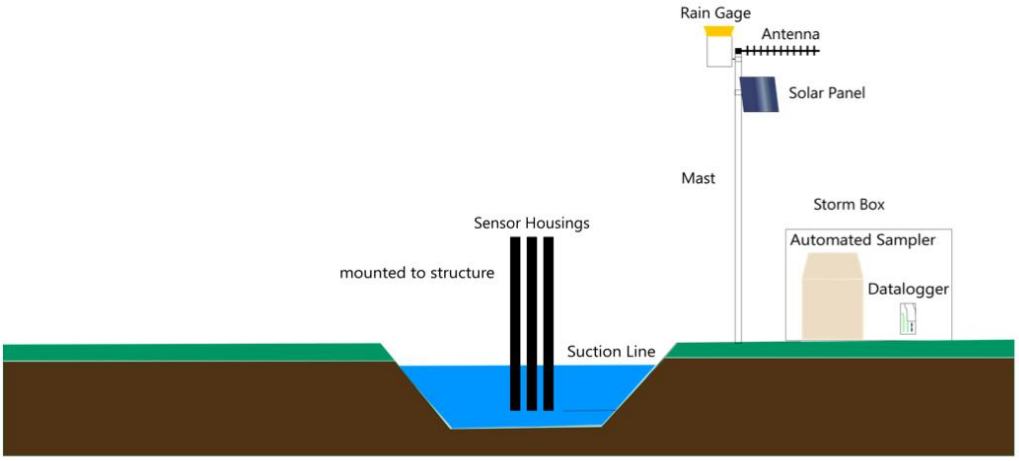


Fig 2.1. Northwest Field Canal monitoring area and urban observatory site locations



(a)

(b)

Fig 2.2. Example of Continuously Monitored Canal Site. (a) Diagram of typical canal site; (b) Photo of deployment at the upstream canal site

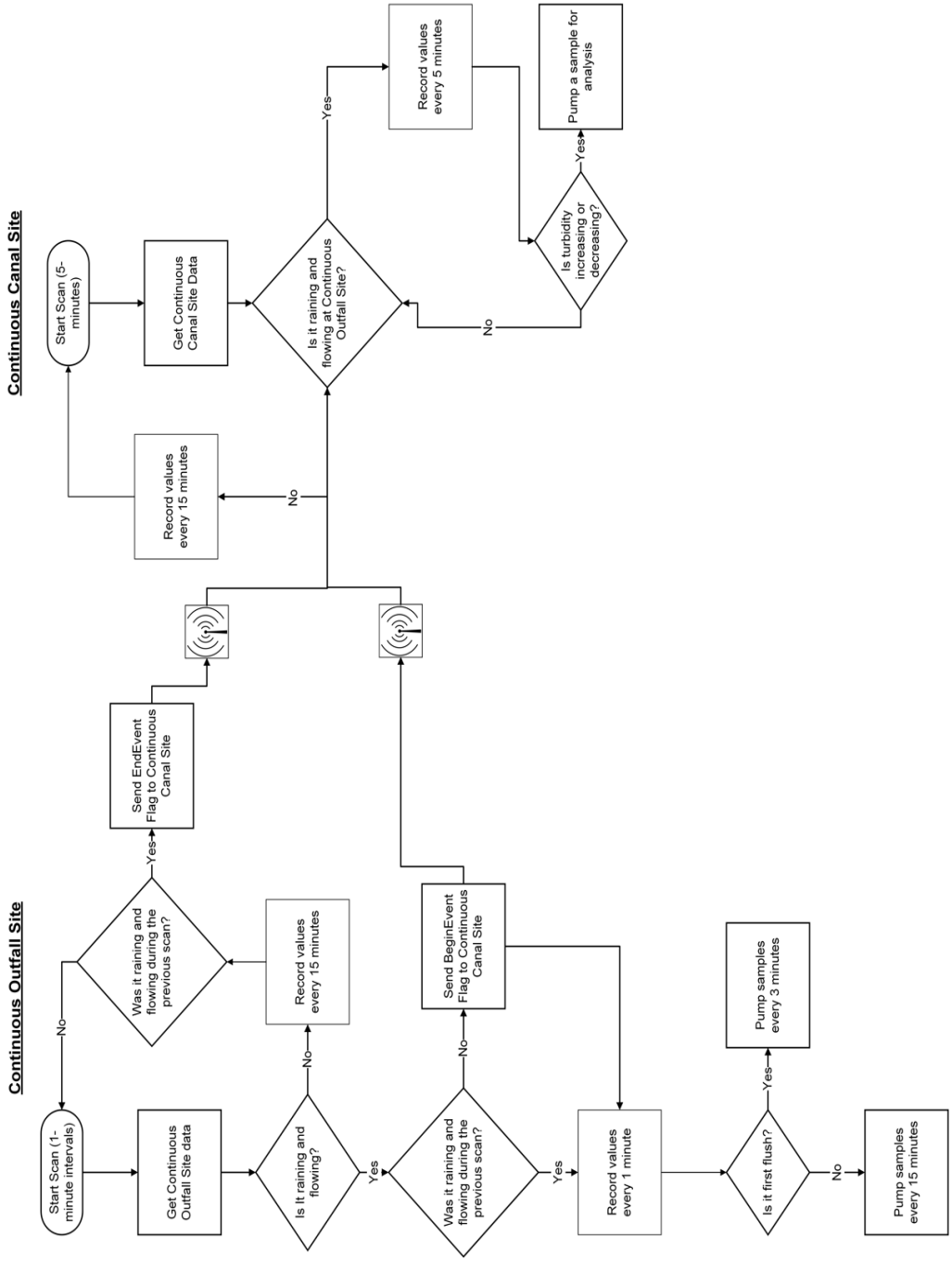


Fig 2.3. Flow chart of the urban observatory's programming logic

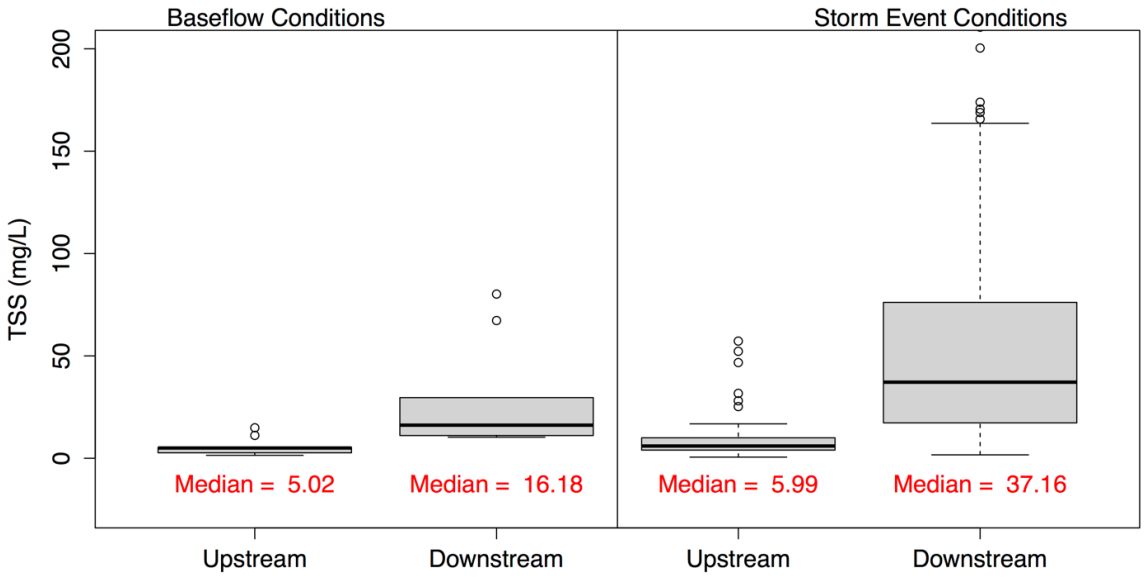


Fig 2.4. Boxplot of TSS concentrations collected the upstream and downstream continuously monitored canal sites

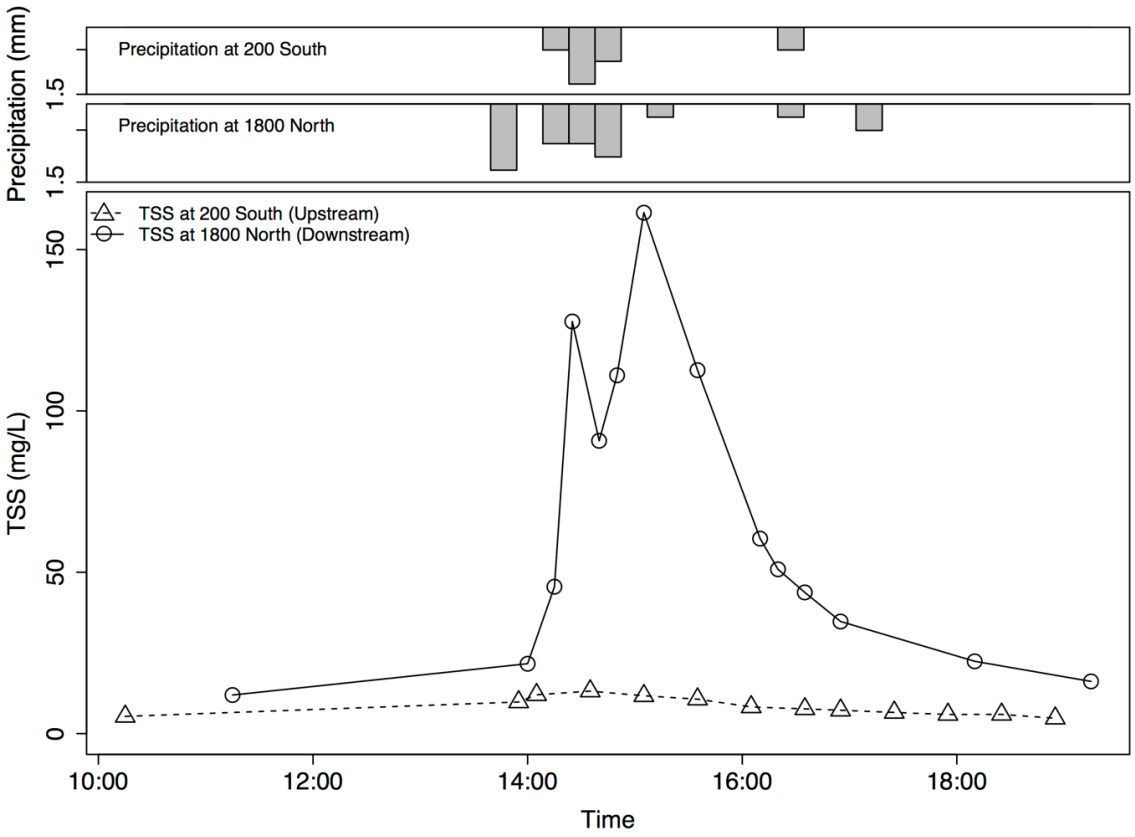


Fig 2.5. Comparison of TSS pollutographs for the upstream and downstream continuously monitored canal sites for the storm event on May 10, 2016

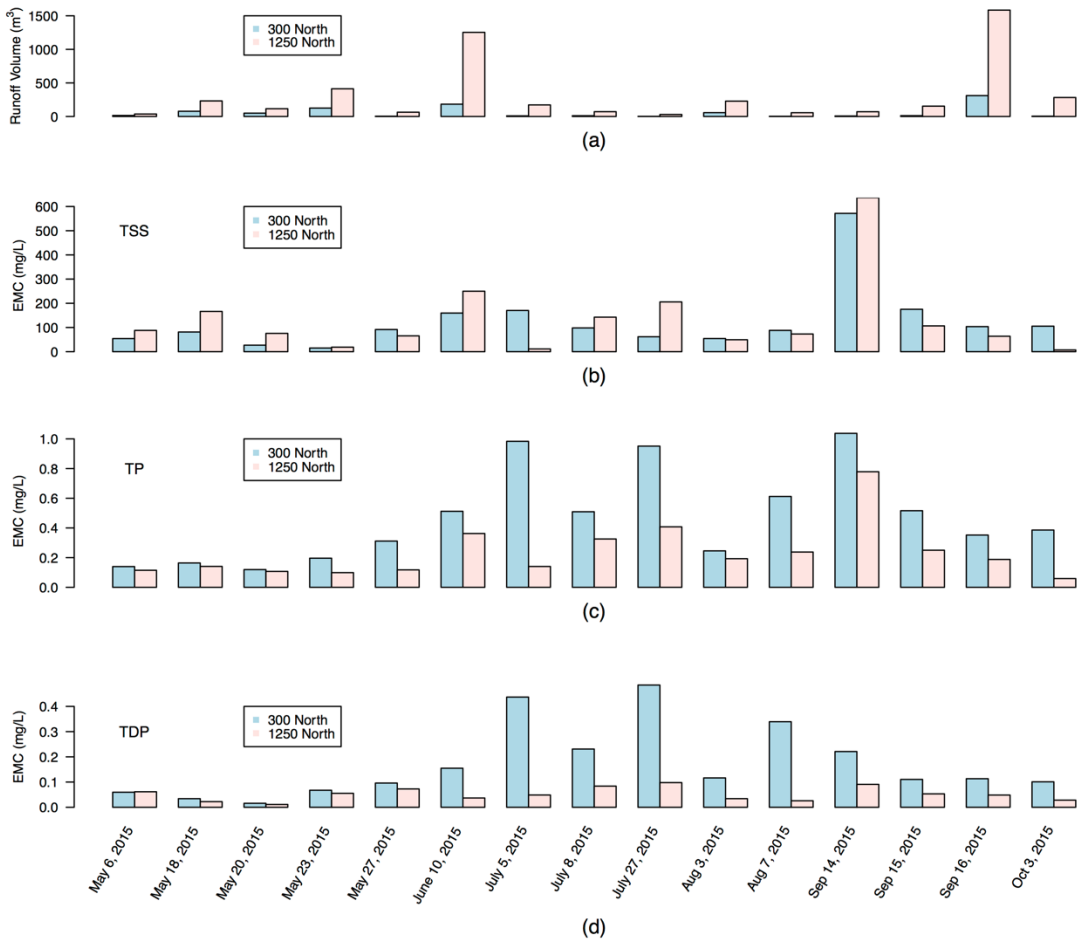


Fig 2.6. Comparison of Runoff Volumes and Pollutant EMCs at 300 North and 1250 North for 15 storm events during the 2015 irrigation season. (a) Runoff volumes for each event; (b) TSS EMCs; (c) TP EMCs; (d) TDP EMCs

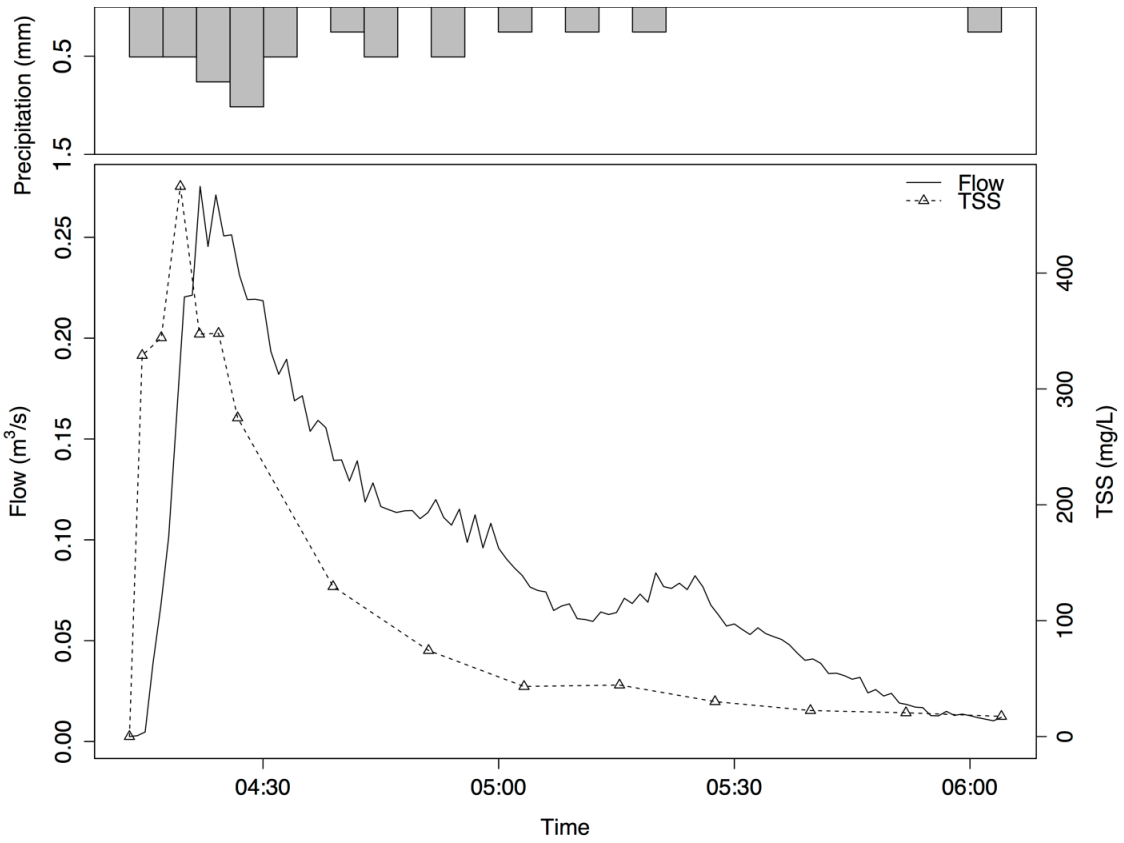


Fig 2.7. Example of a concentration-based first flush observed at the 800 North outfall site on April 10, 2016

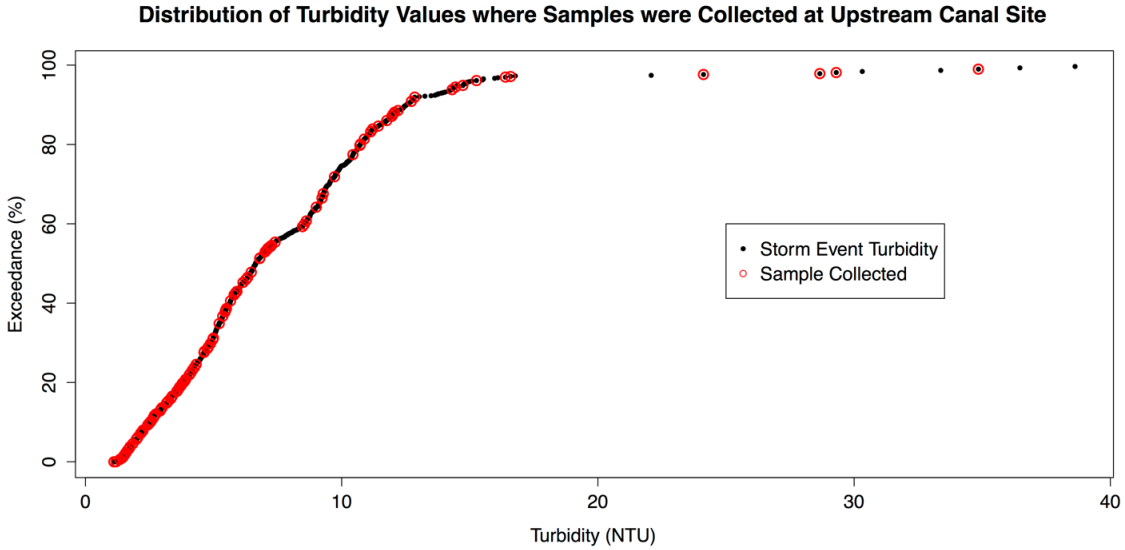


Fig 2.8. Distribution of turbidity values during storm events and samples collected at the upstream canal site (200 South)

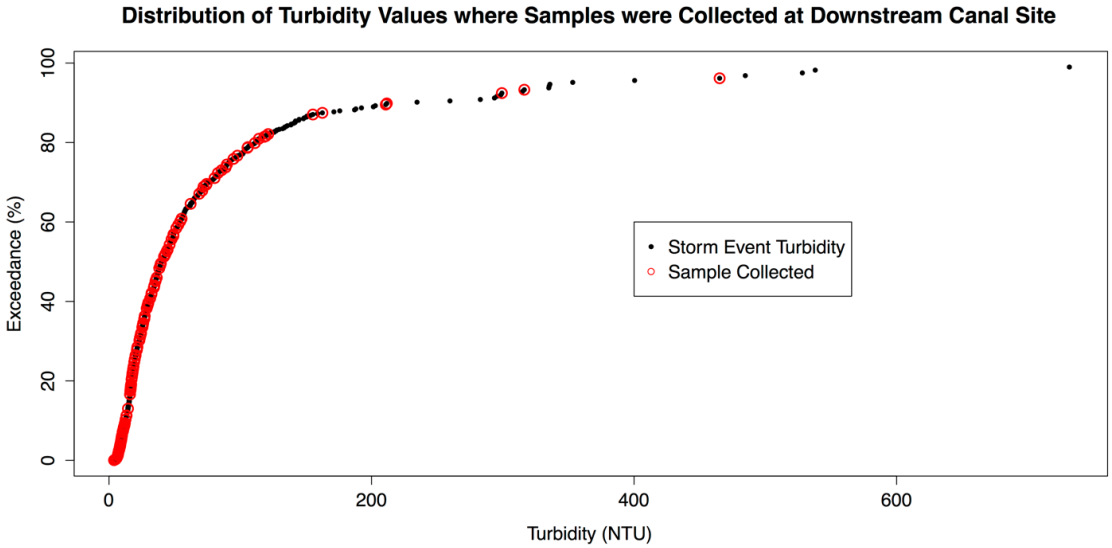


Fig 2.9. Distribution of turbidity values during storm events and samples collected at the downstream canal site (1800 North)

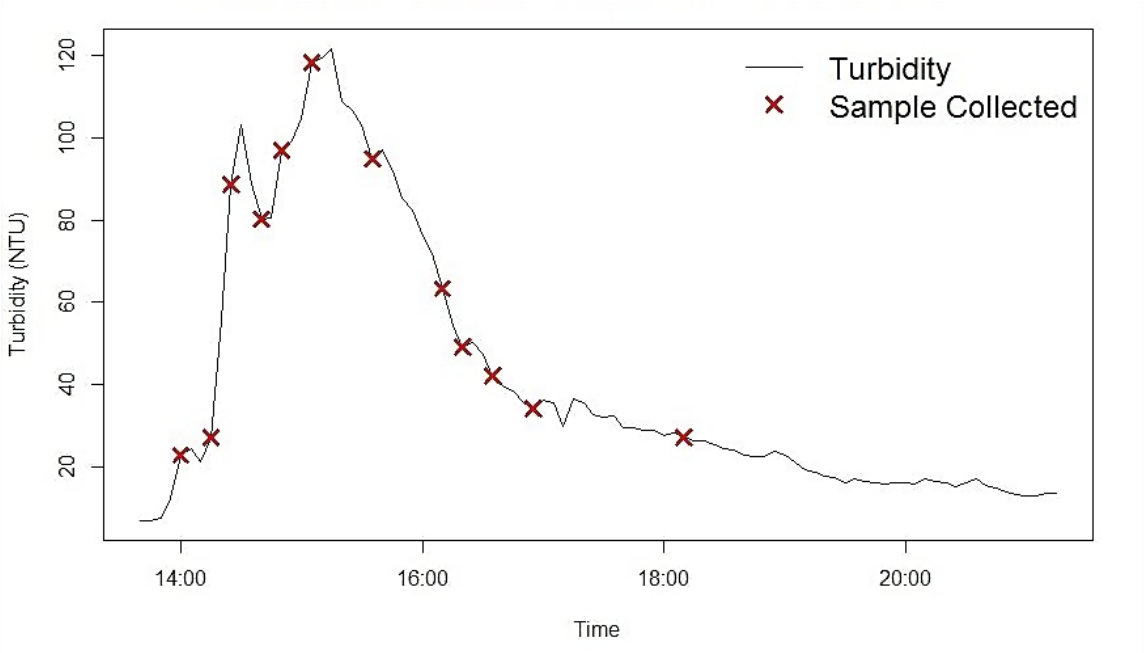


Fig 2.10. Example of the urban observatory's adaptive sampling at the 1800 North canal site based on the turbidity threshold sampling scheme

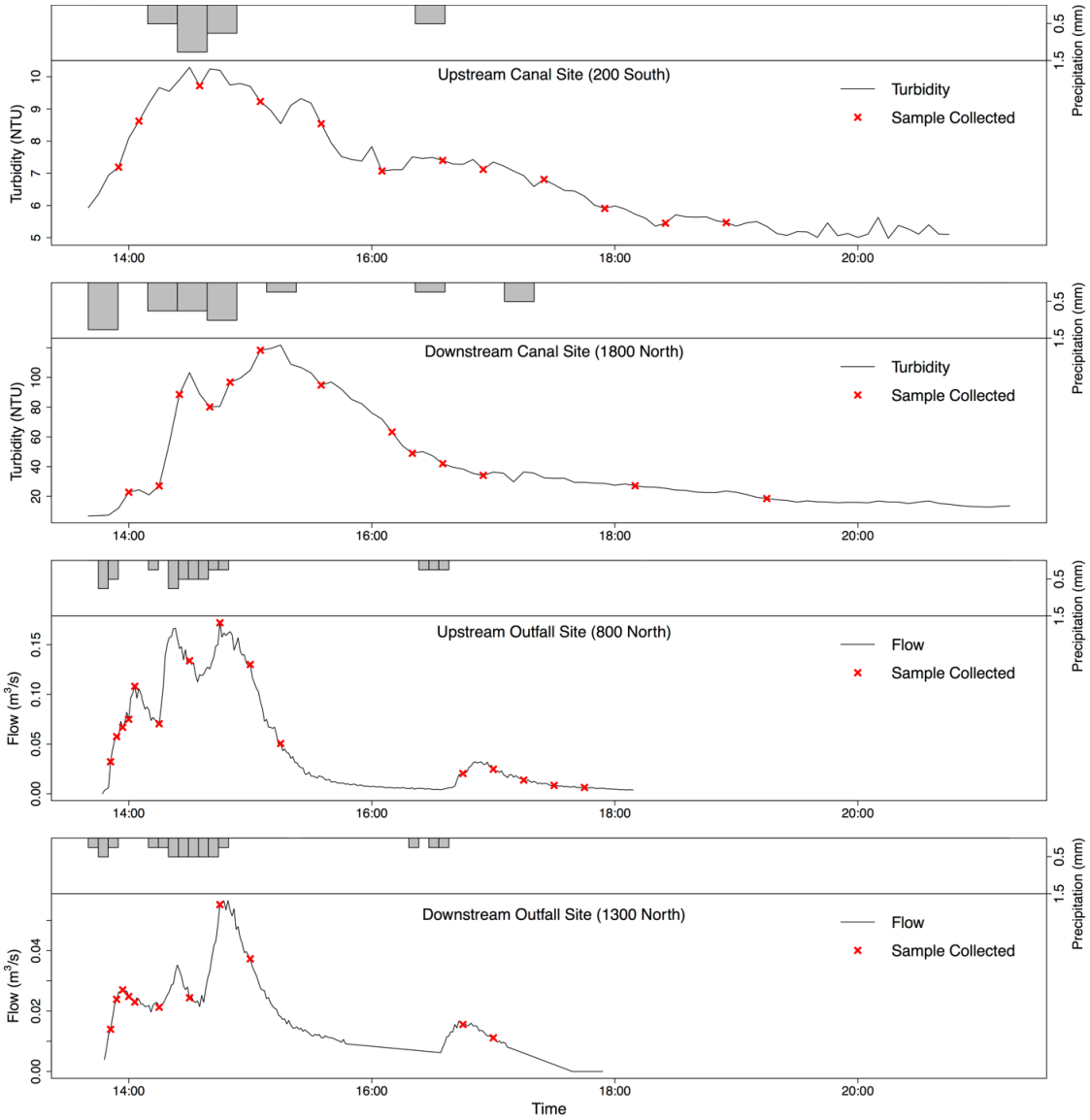


Fig 2.11. Example of the urban observatory's adaptive sampling, event detection, and inter-site communication

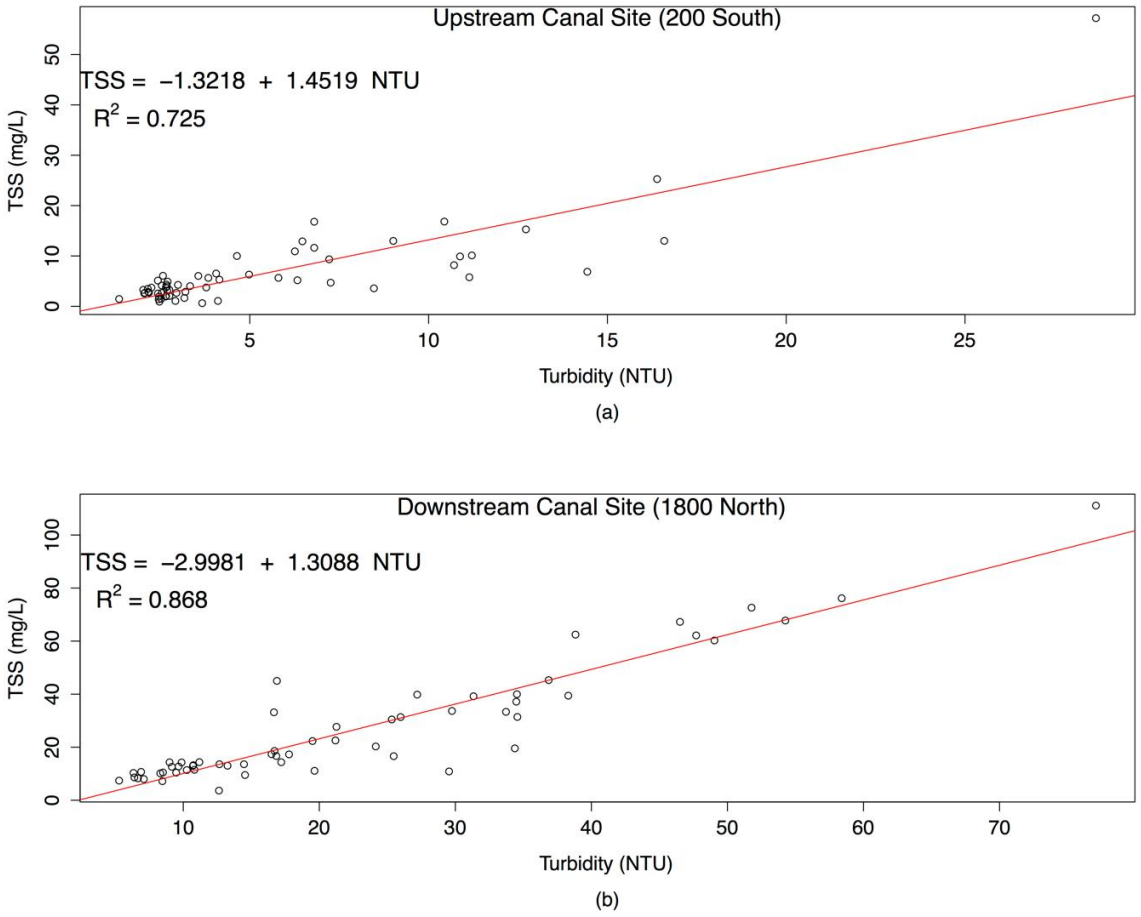


Fig 2.12. Examples of surrogate relationships. (a) Relationship between TSS and turbidity at upstream canal site; (b) Relationship between TSS and turbidity at downstream canal site

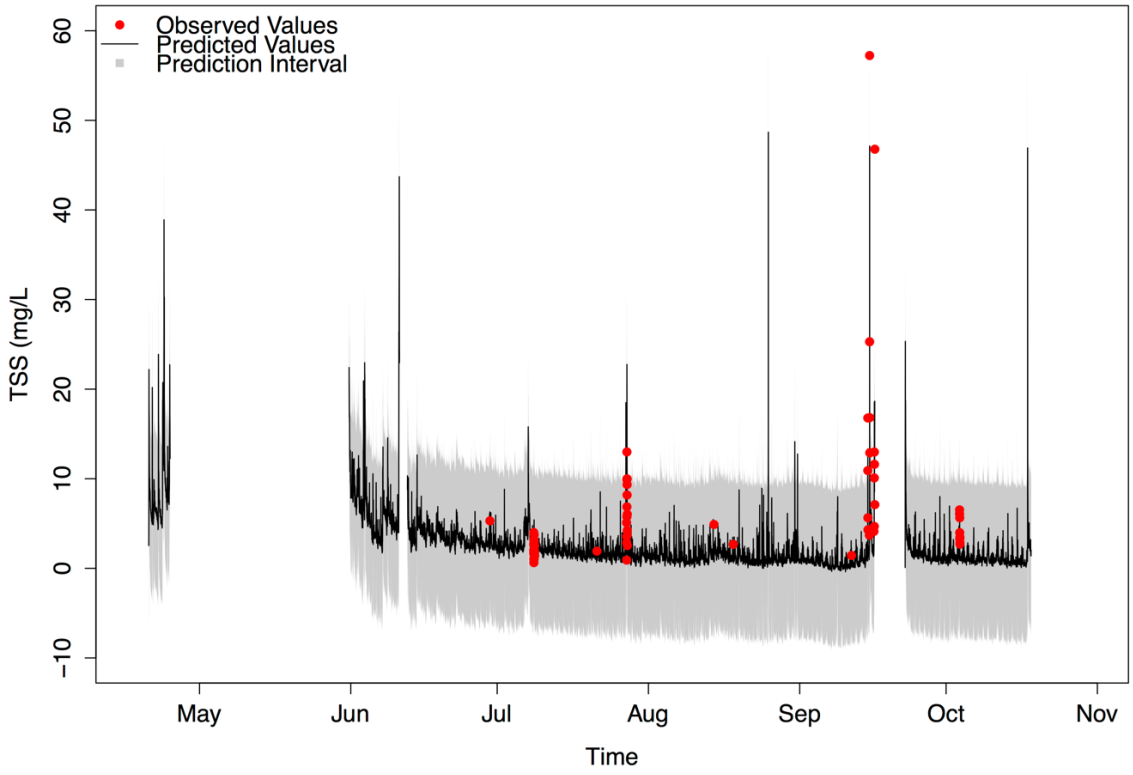


Fig 2.13. Total suspended solids concentrations predicted from turbidity at the upstream canal site during the 2015 irrigation season

CHAPTER 3

REGRESSION METHODS FOR ESTIMATING SUSPENDED SOLIDS CONCENTRATIONS WITHIN URBAN STORMWATER CONVEYANCES

Abstract

Linear regression methods have been used in water quality studies to estimate in-stream constituent concentrations from surrogate measurements made using in situ sensors. Linear regression models can be limited in their ability to account for conditions that may cause regression coefficients to differ or change, which may be particularly important in urban watersheds where short duration events can alter the source of a constituent load, changing the nature of the regression equation. This study compared three regression methods: classical linear regression, linear regression with categorical variables to distinguish events, and linear mixed effects (LME) models, which can account for changes in regressions based on conditions. We evaluated each method's predictions of total suspended solids (TSS) concentrations and loads at the upstream and downstream ends of an urban stormwater conveyance. Results show that turbidity and categorical variables representing the length of antecedent dry period and season were significant explanatory variables at the upstream monitoring site. Turbidity and categorical variables representing rainfall intensity and rising versus falling limbs of the pollutograph were significant explanatory variables at the downstream monitoring site. Based on statistical metrics and TSS load estimates, both LME and linear regression with categorical variables models were

superior to the classical linear regression method in their ability estimate TSS concentrations and account for undersampled events.

---

[1]Co-authored by Anthony A. Melcher, Jeffery S. Horsburgh, Amber S. Jones, and David K. Stevens

3.1. Introduction

The need for high frequency water quality data for constituent load estimation in both the natural and built environment and the inadequacy of infrequent grab sampling are well documented (Kirchner *et al.*, 2004; National Research Council, 2009; Horsburgh *et al.*, 2010; Wade *et al.*, 2012; Outram *et al.*, 2014; Rode *et al.*, 2016;). Instream diel cycles in constituent concentrations (Loperfido *et al.*, 2009; 2010), variable point source contributions, and seasonal concentration swings (Grayson et al., 1996; Jones et al., 2011) may go unobserved if samples are collected too infrequently (Brauer et al., 2009; Jones et al., 2012; Rode et al., 2016). Sporadic concentration values are traditionally used to obtain mass load estimates of constituents such as total suspended solids (TSS) with a significant margin of error, as infrequently sampled concentrations may not capture the variability of constituent concentrations (Tomlinson and De Carlo, 2003; Ryberg, 2006). High costs and logistical difficulties related to sample collection and laboratory analysis often prohibit water resource professionals and managing entities from collecting grab samples at a frequency required to capture changes in pollutant concentrations (Leecaster *et al.*, 2002; Coynel *et al.*, 2004; Fletcher and Deletic, 2007; Brauer *et al.*, 2009). One technique to

estimate constituent concentrations with higher frequency and determine mass loadings that capture a greater range of variability involves water quality surrogates. Surrogate methods rely on regression relationships (often linear) between water quality constituent concentrations derived from analysis of periodic grab samples (e.g., nutrients, suspended solids, E. coli, etc.) and parameters measured using *in situ* sensors at much higher frequencies (e.g., conductivity, turbidity, pH, etc.) (Jones *et al.,* 2011; Rasmussen *et al.,* 2011; Fisher *et al.*, 2016).

Particulates and other suspended solids can be harmful to stream biota and can be the means for the mobilization of other pollutants. Of assessed water bodies in the USA, 5.4 percent are impaired due to excess sediment (USEPA, 2017), making sediment one of the most common pollutants in aquatic systems. In efforts to better understand the timing, magnitude, and sources of suspended sediment, surrogate relationships have been used to obtain high frequency estimates of TSS concentrations and determine TSS load estimates (Irish Jr. *et al.*, 1998; Christensen, 2001; Tomlinson and De Carlo, 2003; Ryberg, 2006; Jones *et al.*, 2011; Rasmussen *et al.*, 2011). These studies have primarily used turbidity as a surrogate indicator for TSS across varying stream sizes, flow regimes, and surrounding land uses. While turbidity has become an accepted surrogate for TSS (Gippel, 1995; Grayson et al., 1996; Rasmussen et al., 2011), there are caveats and limitations related to using turbidity alone as a predictor. Pooling all samples into a classical linear regression model (referred to hereafter as CLR) tends to reveal systematic bias in the residual errors (Grayson et al., 1996; Jones et al., 2011) because the relationship between the surrogate and the constituent of interest may not be consistent across all conditions. For example,

Landers and Sturm (2013) found that turbidity-TSS relationships exhibited hysteresis, which they attributed to varying particle sizes between the rising and falling limbs of the pollutograph as turbidity readings are dependent on particle size and composition (Gippel, 1995; Patil *et al.*, 2011; Landers and Sturm, 2013).

Other documented surrogate regression models have found seasonal variability and storm event characteristics to be significant when estimating TSS and particulate pollutant loads, indicating that the relationship between the surrogate and the constituent of interest is not constant between these periods (Grayson *et al.*, 1996; Brezonik and Stadelmann, 2002; Ankcorn, 2003; Ryberg, 2006; Settle *et al.*, 2007; Rasmussen *et al.*, 2011; Fisher *et al.*, 2016). One modeling approach to address seasonal variability is to include sine and cosine functions of day of the year in the regression equation (Ryberg, 2006; Rasmussen *et al.*, 2011; Fisher *et al.*, 2016). However, seasonal changes do not always occur at the same time each year, and using this type of approach would require constant (i.e., annual) updates of the surrogate relationship, especially under changing climatic conditions. In another case, Jones *et al.* (2011) found that a binary categorical variable indicating whether or not samples were taken during spring snowmelt vs. base flow improved the quality of regressions between turbidity and total phosphorus (TP). In addition to turbidity, Kayhanian *et al.* (2007) found that stream flow and storm event characteristics were significant predictors of TSS event mean concentrations, and reported improved model performance after including these variables. A challenge with using categorical and seasonality variables is that runoff events and high-flow conditions need to be sufficiently sampled to representatively include them in regression relationships, as one group should

not outweigh another (e.g., runoff versus base flow or rising limb versus falling limb). However, sampling during these periods is often logistically difficult.

Given the variability of relationship types reported in the literature, it is not always clear which type of regression model should be used (e.g., CLR versus linear regression with categorical variables (referred to hereafter as LRCAT) or using transformed data versus untransformed data). These factors reflect the empirical nature of surrogate relationships. As they are developed distinctly for each site of interest (Miguntanna et al., 2010; Rasmussen et al., 2011; Viviano et al., 2014), surrogate relationships are not fully mechanistic and may not always capture the processes that drive concentrations of TSS. Furthermore, just as surrogate relationships vary from site-to-site, they may also vary between time periods.

While surrogate methods have been used in both rural and agricultural watersheds, fewer studies have examined their utility for quantifying pollutant concentrations within urban water systems (Settle *et al.*, 2007; Miguntanna *et al.*, 2010; Viviano *et al.*, 2014). Impervious drainage surfaces in urban systems create conditions by which relatively small precipitation events may produce disproportionately high runoff, which may cause intense loading events of short duration in urban streams and stormwater conveyances (Wanielista *et al.*, 1997; Maestre and Pitt, 2005; National Research Council, 2009; Hvitved-Jacobsen *et al.*, 2010a). Additionally, the complexity of pollutant sources in urban conveyances has the potential to affect the applicability of surrogate relationships. For example, Christensen (2001) found strong correlations between turbidity and TSS in rural streams while Miguntanna *et al.* (2010) observed evidence of weaker correlations in urban streams.

Adding to this complexity, some urban streams in the western U.S. receive both irrigation return flows and urban stormwater runoff (City of Logan, 2010; City of Grand Junction, 2016; City of Sequim, 2016).

Due to the spectrum of pollutant sources and flow conditions in urban conveyances and the likelihood that relationships may vary between seasons or flow conditions, surrogate relationships that use linear regression models may be inadequate for urban streams. This may be especially true when important hydrologic conditions (such as storms) are undersampled. One method that demonstrates potential for developing robust relationships while enabling categorical grouping of data and also accounting for undersampled groups is linear mixed effects (LME) modeling. Also called multilevel or hierarchical models, LME models are an alternative approach to model fitting and parameter estimation that attempt to explain some of the random and systematic error in regression models (Pinheiro and Bates, 2000; Gelman and Hill, 2007; Wu, 2010; Araujo *et al.*, 2012). In LME modeling, data points are assigned to groups, and each group is weighted based on its information content (Gelman and Hill, 2007). As a result, if fewer samples are collected within certain groups (e.g., rising limb of a storm event), that group has less influence on the overall model. Weighting each group according to the associated number of samples can effectively create a surrogate relationship that exhibits variability between different time periods without overemphasizing a condition that may have been undersampled. A primary motive for this research was to seek an option for developing surrogate relationships that would capture this variability.

The overall objective for this study was to evaluate the effectiveness of different regression techniques, including CLR, LRCAT, and LME modeling, for development of surrogate models for the estimation of TSS concentrations and loads in an urban stream. We evaluated models resulting from the different regression methods using multiple goodness-of-fit measures and examined the strengths and weaknesses of each modeling approach. In this paper, we first provide background on linear regression and LME modeling techniques for surrogate relationship development. We then describe methods for selection of CLR models, which are used as the base of the LRCAT and LME models, as well as the determination of categories and groups. Then, we compare the resulting models for their adequacy in estimating TSS concentrations within an urban water conveyance that aggregates snowmelt and groundwater, irrigation return flows, and short duration storm runoff from urban surfaces. We conclude with a discussion of the explanatory variables and factors that affect the variability in the goodness-of-fit values.

3.2. Theory and Background

A common approach for developing surrogate relationships for high-frequency estimates of water quality parameters is CLR (simple or multiple) as shown in Equation (3.1):

$$TSS_{\boldsymbol{i}} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \epsilon_i, \quad i = 1, \ldots, n \tag{3.1}$$

where $TSS_{\boldsymbol{i}}$ represents the response variable (in this case, TSS concentration) for the $\boldsymbol{i}$th observation, x_{ik} represents the k explanatory variables for the $\boldsymbol{i}$ th observation, β_0 represents the intercept value, β_1 to β_k represent the k regression (slope) coefficients, ϵ_i is

the residual error of the $\boldsymbol{i}$th observation, and n is the number of samples. The regression coefficients are most often estimated using ordinary least squares estimation (Berthouex and Brown, 2002; Helsel and Hirsch, 2002). This approach has been applied in many surrogate studies (Christensen, 2001; Ryberg, 2006; Jones *et al.*, 2011) and is recommended by the United States Geological Survey (USGS) (Rasmussen et al., 2011). This method assumes that residual errors (ϵ_i) are normally distributed, independent, and homoscedastic. If these assumptions are violated, the systematic error or correlation of residuals with another variable may indicate that the regression coefficients are not consistent across conditions (Helsel and Hirsch, 2002).

When developing surrogate relationships, variations of the CLR model may be used to account for changes in regression coefficients. One method for doing this is using categorical variables to indicate the occurrence of some phenomenon that changes the slope and intercept of the regression equation. Phenomena could include the occurrence of snowmelt runoff versus base flows, time trends in relationships, etc. (Berthouex and Brown, 2002; Jones et al., 2011). Equation (3.2) shows the general form of the regression equation with two categorical variables.

$$TSS_i = \beta_0 + \beta_1 x_i + Z_1(\alpha_0 + \alpha_1 x_i) + Z_2(\gamma_0 + \gamma_1 x_i) + \epsilon_i, \quad i = 1, \dots, n \qquad (3.2)$$

where α and γ represent parameters estimated by least squares. The categorical variables, Z_1 and Z_2, are variables that take on discrete values indicating factors or levels of the associated phenomenon. The accuracy of the estimates of α and γ depends on the number of samples collected for each value of Z. If Z represents an unpredictable event of short

duration or some phenomenon for which it is difficult to sample adequately, the lower number of samples will result in greater uncertainty in the associated estimates of α and γ.

The linear mixed effects (LME) modeling approach starts with the creation of a base model, similar to a CLR model, which uses *in situ* parameters (e.g., turbidity) and/or metavariables (e.g., rainfall intensity) as explanatory variables to predict TSS concentrations. This is referred to as the fixed effects portion of a mixed effects model. LME models also try to explain some of the systematic bias in the residual errors of the fixed effects model through the definition of groups caused by characteristic differences in the data under different conditions. This is referred to as the random effects portion of a mixed effects model. Gelman and Hill (2007) describe these groupings of data as categories between which regression coefficients are expected to vary. In the context of surrogate relationships, an example is a TSS-turbidity model with coefficients that vary between base flow and storm event conditions, across seasons, or between rising and falling limbs of a hydrograph/pollutograph. LME groups are similar to the categories used in a LRCAT model; however, the parameter estimates in the LME model are made by maximum likelihood or restricted maximum likelihood estimation methods (Pinheiro and Bates, 2000; Gelman and Hill, 2007; Zuur *et al.*, 2009), which consider all observations in each group simultaneously. This is in contrast to a LRCAT model, which fits a separate model to each category of observations using ordinary least squares and weights each group equally in the overall model (referred to by Gelman and Hill (2007) as the "no-pooling" estimate). In practice, LME models weight each group based on the information content (i.e., number of samples) and their overall influence on model precision.

LME models have proven powerful where observational data have a grouped, longitudinal, nested, or multilevel structure (Pinheiro and Bates, 2000; Gelman and Hill, 2007; Wu, 2010). This is particularly useful where repeated observations are made on subjects that belong to a certain group or class, such as geographic boundaries (Gelman, 2006; Wu, 2010). For example, if the objective is to assess the variability between base flow or storm runoff conditions, and if it can be assumed that regression coefficients (intercepts and slopes) vary between each, then base flow and storm runoff would be good candidates for groups in an LME model. LME models are common in life and social sciences to account for variability between individual observations and groups (Bagiella *et al.*, 2000; Gueorguieva and Krystal, 2004; Gelman, 2006; Gelman and Hill, 2007). Examples of the use of mixed-effects models in water resources applications, however, are fewer. Our review of the literature found only two studies that employed LME models in the development of surrogate relationships. Lessels and Bishop (2013) and Slaets *et al.* (2014) used mixed-effects modeling to account for auto-correlation in model residuals in developing surrogate relationships. Our literature search did not find any applications of mixed-effects modeling that considered groupings of the data in developing surrogate relationships. In this study, we directly explored the potential advantages of LME for estimating parameters in developing surrogate relationships.

The literature provides multiple versions of the generic equations for mixed effects models. Equation (3.3) gives the simplest form as reported by Gelman and Hill (2007) with varying slopes and intercepts:

$$TSS_i = \beta_{0j[i]} + \beta_{1j[i]}x_{i1} + \beta_{2j[i]}x_{i2} + \cdots \beta_{kj[i]} + \epsilon_i \quad (3.3)$$

where the index $j[i]$ denotes the group (e.g., storm event, rising/falling limb of pollutograph, season) to which individual i (e.g., TSS observation) pertains, $\beta_{0j[i]}$ represents the varying-intercept value for each of J groups, and $\beta_{1j[i]}$ to $\beta_{kj[i]}$ represent varying-slope values for each of J groups. The maximum likelihood estimation of $\beta_{kj[i]}$ is made by maximizing the likelihood function or the product of Gaussian probability density functions by reducing the variance of the residual errors (Pinheiro and Bates, 2000; Gelman and Hill, 2007; Zuur *et al.*, 2009; Wu, 2010), which is typically done algorithmically using optimization algorithms (Powell, 2009; Bates *et al.*, 2015).

3.3. Study Area

Logan, Utah, USA, has a population of about 48,000, and is the largest city in Utah's Cache County. The city's primary surface water source is the Logan River, which enters city boundaries from the east at the mouth of Logan Canyon. Four agricultural irrigation canals are diverted from the Logan River, run north through the city, then to the west, eventually combining and emptying into Cutler Reservoir, which is also the receiving water body for the Logan River. These canals carry a large portion of the Logan River's flow during summer months when water is diverted from the main river for irrigation. These canals are also the primary recipients of stormwater runoff in Logan. This research focused on the Northwest Field Canal (NWFC), which is the farthest west of the four canals in Logan City. The NWFC was selected because it receives storm runoff from a variety of

land uses within its drainage area as well as the largest portion of Logan City's stormwater runoff relative to other canals (Figure 1).

The NWFC travels through residential and mixed residential/commercial neighborhoods and then through primarily commercial and mixed-use areas, receiving stormwater from much of Logan's city center and commercial zones. Drainage subcatchments in Logan City are bordered by the four irrigation canals, with stormwater traveling primarily from east to west. Irrigation within Logan City is accomplished by diverting water from the canal east (uphill) of each neighborhood and conveying it through city gutters and ditches to residential lawns and gardens. The gutters return unused water to the next canal to the west (downhill) either directly or after irrigation application.

To collect the data included in this analysis, monitoring sites were instrumented at the upstream and downstream ends of the NWFC. The upstream site (located at 200 South street) is located just downstream of the diversion from the Little Logan River that creates the canal. Thus, flows at this site are more characteristic of the Logan River (i.e., clear cool waters, limited algae growth). The downstream site (at 1800 North street) is located at the downstream end of the city after the NWFC has traveled through and received stormwater from residential, commercial, and industrial zones. Flows at the downstream site are more characteristic of the irrigation return flows and stormwater received by the canal.

3.4. Materials and Methods

3.4.1. Instrumentation and Monitoring

The monitoring sites from which data were obtained for this analysis are part of a larger urban observatory described by Melcher and Horsburgh (2017), which consists of two continuously-monitored canal sites (200 South, 1800 North) and two semi-mobile, continuously-monitored storm drain sites. Each monitoring site was equipped with a telemetry system for inter-site communication, a tipping-bucket rain gage (TE525, Campbell Scientific, Logan, Utah, USA), and an automated sampler (ISCO 3700, Teledyne ISCO, Lincoln, Nebraska, USA). The system was designed to detect stormwater runoff events at the storm drain sites and communicate alerts to the canal sites, upon receipt of which the canal sites would increase their data collection frequency and initiate collection of physical samples based on turbidity thresholds.

The continuously-monitored canal sites were each equipped with a suite of water quality and quantity monitoring equipment. Included in the water quality monitoring instrumentation were turbidity sensors (DTS-12, Forest Technology Systems, Victoria, BC, Canada) and multi-parameter water quality sondes (YSI EXO2, YSI Incorporated, Yellow Springs, Ohio, USA), which measured dissolved oxygen (DO), specific conductance (SC), pH, water temperature, and fluorescent dissolved organic matter (fDOM). All sensors recorded data at 15-minute intervals during base flow conditions and at 5-minute intervals when triggered during storm events.

Essential to pollutant load estimation is the collection of discharge data, and high frequency estimates of discharge were created at each continuously-monitored canal site. Both sites were equipped with a pressure transducer (CS451, Campbell Scientific, Logan, Utah, USA) to measure water depth, which we intended to use with a stage-discharge relationship to obtain high frequency estimates of discharge. However, characteristic of agricultural conveyances, the NWFC contains structures that change the hydraulic flow regime (e.g., drop structures, diversion gates, and other damming structures), which affected the development of valid stage-discharge relationships. The downstream site (1800 North) was located just above a large drop structure, which acted as a consistent hydraulic control, resulting in the derivation of a stage-discharge curve for that site (Melcher et al., 2018b). The relationship was developed by correlating stage measurements with periodic discharge measurements collected using the area-velocity method and an Acoustic Doppler Velocimeter (ADV) (Sontek FlowTracker, San Diego, California, USA). The upstream site (200 South) was located just above a location where water users dam the canal to raise the water level to facilitate diversion into lateral ditches. Thus, increases in water depth occurred when discharge remained constant. As an alternative to using stage to estimate discharge, a side-looking acoustic-Doppler velocity meter (ADVM) (Sontek SL3000, San Diego, California, USA) was installed at the upstream site, which uses the index velocity method to obtain reliable discharge estimates (Levesque and Oberg, 2012).

Water quality samples were collected at the upstream and downstream sites and analyzed for TSS. Automated samplers were used to collect storm event samples at both sites based on turbidity threshold sampling (TTS). The TTS scheme was originally

developed by the U.S. Forest Service and involves collecting physical samples when changes in turbidity values occur. As turbidity increased or decreased past a predefined threshold, a sample was collected, thus ensuring that samples were collected to represent entire storm periods and to capture the entire range of turbidity values (Lewis, 1996; Melcher and Horsburgh, 2017). If turbidity values didn't vary enough to trigger a sample collection during a storm event, as was often the case at the upstream site, a sample was collected at 30-minute intervals. Periodic grab samples were also collected during base flow conditions (i.e., when flow in the canal consisted of only diverted river water). Samples were refrigerated at 4 degrees Celsius and analyzed within 7 days. Laboratory analyses for TSS concentrations were performed at the Utah Water Research Laboratory (UWRL) using Standard Method 2540 D (APHA, 2012).

3.4.2. Data Quality Control

Prior to analysis, data review and quality control were performed on water quality and quantity data, as suggested by Zuur et al. (2009). Time series plots of all *in situ* variables were examined to identify anomalous values and irregular data value spikes, which were either discarded if they were clearly data errors, or interpolated based on field notes or knowledge of field conditions (Campbell et al., 2013; Horsburgh et al., 2010; Melcher et al., 2018a). This quality control post processing was performed using the ODM Tools Python software, Version 1.2.2 (Horsburgh et al., 2015). Plots of turbidity against TSS concentration (Melcher et al., 2018c) were created to visually detect anomalous values and potential outliers. Additionally, potential outliers in the TSS data were identified by

calculating the Cook's D value, which is a measurement of the overall influence that each point has on the position of the regression line, using Equations (3.4) and (3.5) (Helsel and Hirsch, 2002).

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{SS_x} \tag{3.4}$$

$$D_i = \frac{\epsilon_i^2 h_i}{k s^2} \tag{3.5}$$

where h_i is a measurement of the leverage of a point in the model and gives an indication of outliers along the x-axis, SS_x represents the sum of squares or the sum of the squared differences between x_i and $\bar{x}$, which represent the ith and the mean value of the explanatory variable x, s^2 represents the variance of the data, k represents the number of regression coefficients, and D_i is the Cook's distance for observation i. The n, ϵ, and x values are as previously defined. A critical value of $D_i >= 1.6$ obtained from the 10% F-distribution table (Helsel and Hirsch, 2002) was used to identify potential outliers. Values identified as outliers were then critically examined and removed if procedural errors or special sampling conditions could be verified (Helsel and Hirsch, 2002; Jones et al., 2011). The total number of outliers was never more than one percent of the total dataset for each site.

3.4.3. Statistical Regression Methods

Three regression methods were applied to the datasets collected by the urban observatory described in the previous sections: 1) CLR, 2) LRCAT, and 3) LME models. First, the explanatory variables to include in the CLR models for each site were determined. Those explanatory variables were then used as a base for the other two methods. The

strength of each model was evaluated by calculating multiple goodness-of-fit measures for comparison. Unless otherwise mentioned, all data analysis was performed using the R statistical computing software (Bates et al., 2015; R Core Team, 2016). LME models were created using the "lme4" statistical package for R (Bates et al., 2015).

A CLR model was selected for each site by analyzing the results from all possible models containing each combination of explanatory variables (2^k models with k representing the number of potential explanatory variables in the CLR model) (Helsel and Hirsch, 2002). Explanatory variables considered included discharge, turbidity, pH, DO, SC, and water temperature. In order to assess whether to include an explanatory variable, each model was tested to determine if the contribution of each variable offered a significant improvement in three statistics: 1) a reduction in the prediction error sum of squares ($PRESS$ - Equation (3.6)), 2) a reduction of the Mallow's C_p (Equation (3.7)), and 3) an increase in the adjusted coefficient of determination (R_a^2 - Equation (3.8)) (Helsel and Hirsch, 2002; Zuur et al., 2009):

$$PRESS = \sum_{i=1}^{n} \epsilon_i^2 \tag{3.6}$$

$$C_p = k + \frac{(n-k)(MSE_k - MSE_{min})}{MSE_{min}} \tag{3.7}$$

$$R_a^2 = 1 - \frac{(n-1)MSE_k}{SS_{TSS}} \tag{3.8}$$

$PRESS$ uses $n - 1$ observations to develop the linear model, which estimates the value of the one observation omitted from the model. This process is iterated through each observation, and the squared residuals are summed. For Mallow's C_p, the MSE_k represents the mean squared error or the average squared difference between the values predicted by

the model and the observed TSS values for each k set of parameters, and MSE_{min} represents the minimum mean squared error of all 2^k possible models. For the R_a^2, SS_{TSS} represents the sum of squares or the sum of the difference between the average predicted TSS concentration and the observed values, which is referred to as the "null" model (Berthouex and Brown, 2002).

To assess whether models met regression assumptions of normality and homoscedasticity in the residuals, Q-q plots and plots of residuals versus fitted values were examined. If these assumptions were violated, a Box-Cox transformation of the data was tested (Equation (3.9)) in an attempt to obtain near-normally distributed residuals with constant variance (Berthouex and Brown, 2002):

$$Y_i^{(\lambda)} = \frac{y_i^{\lambda}-1}{\lambda} \tag{3.9}$$

where $Y_i^{(\lambda)}$ represents the transformed value of the data series (either response or explanatory variable), λ represents the power of the Box-Cox transformation, and y_i represents the untransformed value of the data series. The power of the transformation (λ) typically takes on any value between -1 and 1. Values of -1, 0, 0.5, and 1 represent reciprocal, logarithmic, square-root, and no transformation of the data series respectively (Berthouex and Brown, 2002).

Using transformations requires retransformation back to the original units of analysis, which can introduce bias to the estimates of concentration (Berthouex and Brown, 2002). To overcome this bias, the Duan smearing estimator (DE) was used as a nonparametric estimate of the expected TSS concentration on the untransformed scale

(Duan, 1983; Helsel and Hirsch, 2002; Rasmussen et al., 2011). Equation (3.10) gives the retransformation of TSS including the smearing estimator:

$$TSS_{DE(0)} = \sum_{i=1}^{n} \frac{f^{-1}(TSS_0 + \epsilon_i)}{n} \tag{3.10}$$

where $TSS_{DE(0)}$ represents the untransformed TSS concentration for observation 0, f^{-1} represents the inverse function of the transformation performed on the data (e.g., taking the square of a dataset that was previously transformed to the square-root scale), TSS_0 represents the predicted TSS concentration on the transformed scale, and ϵ and n are as defined previously.

After using this procedure to select explanatory variables, the resulting CLR models for the upstream and downstream sites were used as the base models from which the LRCAT and LME models were created. Other than the categorical and grouping variables, each regression model for each site used the same explanatory variables. Grouping factors used in the LRCAT and LME models were selected graphically by creating multiple scatter plots of TSS vs turbidity. Each plot was analyzed for evidence of hysteresis in individual storm events, varying slopes and intercepts between base flow and storm events of varying size (e.g., small, medium, or large), and between spring and fall seasons, which could be used as grouping explanatory variables. Storm event size was further categorized based on storm event intensities (mm/hr), depths (mm), and antecedent dry periods (days). The storm event size thresholds were determined by dividing the range of storm event intensities, depths, and antecedent dry periods into three equal bins. Storm events that fell between these threshold values were then categorized as small, medium, large, or base flow. Thus, three explanatory variables were considered related to storm event characteristics (storm

event intensity, storm event depth, and antecedent dry period), each with four different levels (base flow, small, medium, and large). Additionally, a nested grouping structure of rising and falling pollutograph limbs within each storm event size was considered. All models were tested for normality, independence, and homoscedasticity of the residuals (Helsel and Hirsch, 2002; Zuur et al., 2009).

The three modeling techniques were compared by calculating multiple measurements of the goodness-of-fit for each selected model: $PRESS$, R_a^2, and $RMSE$ (Equation (3.11)):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(TSS_p(t_i) - TSS_o(t_i))^2} \tag{3.11}$$

where $TSS_p(t_i)$ is the predicted concentration from the surrogate relationship and $TSS_o(t_i)$ is the observed concentration at time t_i. $RMSE$ is an estimate of the quality of a model that takes into consideration lack of precision (Helsel and Hirsch, 2002). The $PRESS$ statistic (Equation (3.6)) is calculated by summing the squared errors if one observation was omitted from the model formulation and is an indicator of a model's ability to predict additional TSS values for conditions where an insufficient number of samples were collected. The R_a^2 is calculated by Equation (3.8) and indicates the ability of a model to improve upon the "null" model.

3.4.4. TSS Load Estimations

Once the regression models were created, estimations of TSS loads for the duration of the study were calculated. The estimation of TSS loads served two purposes: 1) to aid in explanatory variable selection by determining the impact of the inclusion of each

explanatory variable on the resulting TSS load estimation (i.e., we wanted to evaluate the models in terms of the practical endpoint for which they were intended and not just goodness of fit measures); and 2) to make estimates of runoff loads that enter the canal between the upstream and downstream monitoring sites as a practical application of these methods. Each regression model was used to generate estimates of TSS concentration from high frequency measurements of *in situ* and other metavariables, which were paired with continuous discharge values by matching time stamps. The mass loads were calculated by Equation (3.11):

$$M_{TSS} = \sum_{t=0}^{m} \left(\frac{TSS_t + TSS_{t+1}}{2}\right)\left(\frac{Q_t + Q_{t+1}}{2}\right)\Delta t \quad (3.11)$$

where M_{TSS} represents the estimated mass load of TSS for the duration of the study period (kg), TSS_t and TSS_{t+1} are the TSS concentrations (mg/L) at time t and $t+1$ respectively, Q_t and Q_{t+1} are the discharge values (m^3/s) at time t and $t+1$ respectively, a is a conversion factor to convert to kg per time period t, Δt is the length of the time interval between t and $t+1$, and m is the number of paired discharge and concentration estimates for the duration of the study period (Duvert et al., 2011; Jones et al., 2012; Phillips et al., 1999).

3.5. Results

3.5.1. Dataset Characterization

A total of 153 TSS samples were collected at the upstream site and 197 at the downstream site (Table 3.1). The difference between the number of samples collected at

the upstream and downstream sites can be explained by equipment malfunction and differences in sample collection logic (i.e., 30-minute intervals at upstream site in events where turbidity varied little and TTS sampling at the downstream site). The average number of storm event samples was determined by dividing the total number of storm event samples collected by the number of events monitored (17 at both sites) (Table 3.1). Additionally, the number of samples collected for storm event size (Table 3.2) and storm event (Table 3.3) are given.

Time sequence plots of turbidity at the upstream and downstream sites for the 2015 and 2016 irrigation seasons are shown in Figures 3.2 and 3.3, respectively. The gray shaded regions indicate the occurrence of a storm event. While the majority of storm events were sampled for TSS, some events were not sampled due to logistical constraints. Gaps in the turbidity data, such as in May 2015 at both sites, were caused by canal managers closing the canal diversion during large storm events or series of storm events. This is common practice in the NWFC to prevent flooding along the canal. Overall, turbidity values at the upstream site ranged between 0 and 35 NTU. Turbidity values at the downstream site ranged between 0 and 1500 NTU, but turbidity typically did not exceed 400 NTU other than during a few storm events.

3.5.2. Model and Variable Selection

The relationships between TSS and potential explanatory variables are shown for both sites (Figures 3.4 and 3.5). At both sites, there is strong correlation between turbidity and TSS concentration (Pearson's $r = 0.91$ upstream and 0.98 downstream). At the

upstream site, TSS is also correlated with SC and pH. The downstream site demonstrated correlations between TSS and discharge, turbidity, pH, DO, and SC. Worth noting is that the variables that correlate with TSS also correlate with each other, an indication of potential multi-collinearity of a CLR model that includes these explanatory variables. Thus, rather than incorporating every explanatory variable that has a significant correlation with TSS in the CLR models, additional analysis was required to obtain the most parsimonious model.

The CLR models were selected based on the criteria for optimizing model assessment statistics (Table 3.4). For almost all models examined, the assumptions of normality and homoscedasticity in the residuals were violated. In an attempt to achieve a model that met these assumptions, multiple values of λ in the Box-Cox transformation (Equation (9)) were examined. Values of λ near 0.5 in the transformation of the response variable (TSS) and turbidity obtained near-normally distributed residuals and constant variance for all the models tested with turbidity as an explanatory variable (Figure 3.6). A λ value of 0.5 is the equivalent of a square-root transformation, so a square-root transformation of TSS and turbidity was used in all cases.

In addition to turbidity, we demonstrate how the inclusion of other explanatory variables affects the quality of the models. Figure 3.7 shows each of the 2^5 models (the number of the models created using every combination of the 5 *in situ* explanatory variables) ranked by the calculated goodness-of-fit values is shown. Each panel has a point of discontinuity in the plot at model 32, indicated by a label and a vertical line in the plot, which corresponds to the improvement provided by including the square root

transformation of turbidity as an explanatory variable. Including the square root of turbidity explains over 80 percent of the total variability at the upstream site and over 96 percent at the downstream site. Any additional explanatory variable included in the models (all models to the right of the vertical line) provided limited improvement in terms of the C_p, R_a^2, and $PRESS$ statistics.

The TSS load estimates for the 32 models that include the square root of turbidity as an explanatory variable were calculated to better understand how each model and explanatory variable affects the associated estimate of TSS load for the study period. The TSS load estimate for the whole study period was used for this analysis because it offers a single numerical value of each model's predictive results that can be compared across all model realizations. The small range of estimated TSS loads for the 32 models at each monitoring site (59,500 – 62,100 kg at the upstream site and 114,100 – 120,600 kg at the downstream monitoring site) suggest that, for the purpose of this study, there is minimal effect of including explanatory variables other than the square root of turbidity in the CLR models (Table 3.4).

Multiple TSS-turbidity plots were examined to determine categories and groups for the LRCAT and LME models. In particular, we looked for cases of hysteresis and varying slopes and intercepts. Examples that motivated our selection of categories and groups to be included in the LRCAT and LME models are given (Figures 3.8, 3.9, and 3.10). The varying slopes and intercepts between TSS-turbidity plots of individual storm events were evident at both sites, but were especially prominent at the upstream site (Figure 3.8, downstream site not shown). Figure 3.8 demonstrates that further analysis was required to

determine how to best categorize the storm events. An analysis of the goodness-of-fit criteria for model realizations using combinations of each grouping factor (i.e., storm event intensity, storm event depth, and antecedent dry period) revealed that the storm events were best categorized (i.e., base flow, small, medium, or large) according to the length of the antecedent dry period at the upstream site and rainfall intensity at the downstream site. Additionally, at the upstream site, the season in which the sample was collected was found to be a significant categorical and grouping factor (Figure 3.9). Finally, the pollutographs for many of the storm events at the downstream site exhibited hysteresis between TSS and turbidity, indicating a change in the relationship from the rising to the falling limb of the pollutograph (e.g., Figure 3.10). Hysteresis was not observed at the upstream monitoring site, so this factor was not included in the models for that site.

3.5.3. Regression Type Comparison

Comparing the final three models (Table 3.4), the LRCAT model resulted in a slightly lower $RMSE$ than the LME and CLR models while the LME models demonstrated lower $PRESS$ statistics than the CLR and LRCAT models at both monitoring sites. R_a^2 values were similar between all models. At both monitoring sites, the estimates of TSS load vary between the three regression methods. At the upstream site, the difference in estimated TSS load between the CLR and LRCAT models is approximately 8,500 kg. At the downstream site, the difference in estimated TSS load between the CLR and LRCAT models is approximately 44,500 kg. These differences demonstrate how a model with only

marginal improvement in the goodness-of-fit criteria (i.e., between CLR and the LRCAT) can result in a large difference in TSS load estimate.

Given the magnitude of the observed differences in estimated TSS load, comparisons between load estimates resulting from storm event and base flow periods were made to further compare the regression methods in terms of load estimation (Table 3.5). At both the upstream and downstream sites, the estimated TSS loads during storm event conditions were similar for all three model types (approximately 6,400 kg at the upstream site and 43,800 kg at the downstream site). The base flow estimates, however, exhibited greater variability. Compared to the load estimates resulting from the CLR models during base flow conditions, the LRCAT load estimate is 16 percent higher at the upstream site and 60 percent higher at the downstream site. The LME load estimate during base flow conditions is similar to the corresponding CLR load estimate at the upstream site (53,620 kg and 53,350 kg respectively) and more similar to the LRCAT load estimate at the downstream site (110,680 kg and 118,200 kg respectively) (Table 3.5). Upon further examination, it was found that, at the upstream site, models like the LME and LRCAT are prone to artificial "steps" or points of discontinuity in the estimated TSS concentrations (Figure 3.11). This is due to the use of the grouping or categorical variables. Figure 3.11 shows the point at which the "season" categorical variable in the LME model transitions from "spring" to "fall." The discontinuity circled in red is an artifact of the LME model that resulted in a predicted TSS concentration that was less than the predicted concentrations from the CLR model during the spring months and greater during the fall

months. As a result the LRCAT model predicted the highest TSS loads and the CLR and LME models produced similar load estimates at the upstream site.

The predicted and observed TSS concentrations for base flow samples at the downstream site (Figure 3.12) show that the CLR models typically underestimate observed concentrations. Given this, and the fact that base flow conditions dominate at both sites, the CLR models are likely underestimating the load significantly as evidenced by the large differences between CLR predicted loads and the loads predicted by the other models. Additionally, the LME model appears to be slightly more robust as the restricted maximum likelihood method gives less emphasis to extreme values. This can be seen in Figure 3.12 when comparing the LRCAT predictions with the LME predictions (panels b-f). The extreme point in the upper right-hand corner of the plot is closer to the 1:1 line in the case of the LRCAT models, while the majority of the other points are above the 1:1 line. This is not the case for the LME models. This implies that the LRCAT models are less robust and more sensitive to extreme values than the LME models.

Variability between TSS concentrations estimated by the three methods are shown in plots of the predicted versus observed TSS concentrations on the square root scale for each method (Figure 3.13). At the upstream site, there is noticeably greater variance in the TSS concentrations predicted by the CLR model (Figure 3.13a) versus those predicted by the LRCAT and LME models (Figure 3.13b and c, respectively). A difference in variance also exists between the CLR and the LRCAT and LME models at the downstream site (Figure 3.13d, e, and f, respectively); however, the difference is less pronounced.

3.6. Discussion

3.6.1. Explanatory Variables

Of all the in situ parameters considered as potential explanatory variables in the CLR models, only turbidity was found to be significant based on our goodness-of-fit criteria and assessment of TSS load estimates. This is in agreement with the findings in other studies (Christensen, 2001; Jones et al., 2011; Fisher et al., 2016). Any additional explanatory variables resulted in marginal improvement in the goodness-of-fit criteria and little difference in the TSS load estimate. While others have found discharge to be a significant explanatory variable for TSS (Ryberg, 2006; Rasmussen et al., 2011), including discharge did not improve our models to the point that it met the criteria for inclusion as an explanatory variable, which we attribute to the hydraulic conditions that were controlled by the canal master in the NWFC.

The categorical and grouping factors included in the final models were found to be unique for each monitoring site. Models for both sites included grouping or categorical variables that describe qualities of storm events (antecedent dry period at the upstream site and rainfall intensity at the downstream site). At the upstream site, sediment source material is more likely to be a result of near-stream erosion and in-channel resuspension rather than from accumulation in storm drains, and the significance of the antecedent dry period may indicate the importance of time in the accumulation of these sediment supplies (Kayhanian et al., 2007). At the downstream site, the significance of storm magnitude, as represented by rainfall intensity, indicates that larger events are able to mobilize increased

volumes of sediment, which we hypothesize is associated with accumulated sediment within storm drains and stormwater conveyances (i.e., gutters). At both sites, sediment loading is associated with sediment buildup and the flushing of those sediments at the start of a storm event, but the time between storm events is influential for supply at the upper site while storm intensity is more important at the lower site.

At the upstream site, the season during which the sample was collected was also found to be a significant predictor of TSS concentrations. Seasonal trends have been found by others to affect sediment yield (Ryberg, 2006; Alberto et al., 2016). In addition to seasonal trends in storm event characteristics, this effect could be attributed to near-stream sediment sources and their susceptibility to erosion in the drier summer and fall months (Alberto et al., 2016). These effects were observed at the upstream site, where flows are less affected by urban runoff events, and represent the seasonality of Logan River water diverted into the canal.

At the downstream site, the limb of the pollutograph in which the sample was collected was found to be a significant grouping factor. Varying sediment sources and particle size distributions may result in unique characteristics between rising and falling pollutograph limbs (Patil et al., 2011; Landers and Sturm, 2013). For the majority of events monitored, a clockwise hysteresis pattern was observed. This indicates that the downstream site is affected by urban runoff events and TSS loadings from multiple sources more so than the upstream site. It is likely that the rising limb of the pollutograph is more influenced by urban stormwater runoff with particles that resemble road sediment deposits. After the runoff event flushes through the conveyance system, the falling limb may represent

sediments eroded farther upstream in the watershed or resuspended or eroded from sediment sources within the canal.

3.6.2. Comparison of Regression Methods

The range of load estimates during base flow conditions in contrast with the consistency of load estimates during storm events (Table 3.5) for the three different model types implies that the greatest source of variability stems from the estimated TSS concentrations during base flow conditions. This denotes the importance of regular base flow sampling to verify assumptions about the relatively constant nature of in-stream concentrations. The disparity between base flow and storm event loading also points to the importance of modeling techniques such as LRCAT and LME, which account for categorical variability between base flow and storm event conditions. Considering the relatively short duration of the storm events, the magnitude of the estimated event loads relative to the estimated loads during base flow conditions is significant.

Given that the estimated storm event loads determined by each model type were of approximately the same magnitude, we needed to directly use the statistical goodness-of-fit metrics and estimated loads during base flow conditions to determine the superior regression method. In terms of the RMSE, the LRCAT models outperformed CLR and LME models at both monitoring sites. This result was not unexpected, as the LRCAT models minimize the squared errors of TSS concentrations for each storm event category using ordinary least squares. Thus, model prediction errors for the entire set of TSS concentrations at both sites are minimized by using categorical variables. With regard to

the R_a^2, the LME and LRCAT models at both sites performed similarly, both slightly outperforming the CLR models.

For both monitoring sites, the PRESS statistic was greatly reduced using the LME and LRCAT methods to estimate TSS concentrations. In both cases, the LME method produced slightly better PRESS values than the LRCAT method. This metric is an indication of the model's ability to predict concentrations for samples that have been removed from the dataset used for model formulation. The PRESS statistic might be thought of as an indication of the regression method's ability to predict undersampled categories. This is important because certain storm event categories are represented by fewer samples than others (e.g., small events according to rainfall intensity at the upstream site (11 samples) and medium events according to antecedent dry period at the downstream site (13 samples) –Table 3.2). Additionally, the lowest PRESS value indicates the model that reduces the prediction errors of concentrations for storm events (for which the number of samples may be relatively low in comparison to the number of base flow samples (Table 3.3)) and for base flow conditions, yielding the most accurate results (Figure 3.12). For these reasons, the LME method appears to be a slightly more robust than the LRCAT method. However, as the LRCAT model produced superior RMSE values and the LME method produced superior PRESS values, both were considered superior and preferable to the CLR technique examined here, and we accept the resulting TSS load estimations to be more accurate for both monitoring sites (Tables 4 and 5).

3.6.3. Applicability

The premise of this study was to investigate multiple methods for creating surrogate relationships for estimating TSS concentrations and loads. The use of turbidity as a surrogate for TSS concentrations has been studied extensively and is a commonly accepted practice. This study, however, describes conditions where TSS-turbidity relationships vary based on storm event characteristics (antecedent dry period and rainfall intensity), season, and limb of the pollutograph. In urban streams and combined conveyances these conditions can be more prominent than in more natural stream settings. As a result, there is greater potential to bias the load estimate if a regression method that accounts for those conditions (e.g., LME or LRCAT) is not used. This study has shown LME and LRCAT modeling to be valuable tools for estimating TSS loads in urban streams and combined conveyances. Additionally, using both LME and LRCAT regression methods can provide insight on the uncertainty of the estimated TSS loads, as was seen with the variability of base load estimates at both the upstream and downstream sites.

3.7. Summary and Conclusions

The three regression methods tested in this study demonstrated varying results in terms of the estimated TSS loads for the upstream and downstream monitoring sites. While the CLR models at the upstream and downstream sites produced acceptable R_a^2 values, LME and LRCAT models resulted in significant improvements in terms of the models' PRESS statistic and their ability to predict TSS concentrations for conditions that may be undersampled. Both LME and LRCAT performed similarly and are more reliable

approaches for development of surrogate relationships that may account for limited sample collection during important periods (a typical challenge for water managing entities). Although LME methods can account for different numbers of samples collected in varying periods, the importance of both base flow and storm event sample collection should not be discounted.

While LRCAT and LME methods performed similarly based on the goodness-of-fit criteria and TSS load estimations, the LME method showed potential as a technique that is more robust to extreme values and undersampled categories (Figure 3.12). One potential limitation of our study is that data collection and regression model development took place at two sites in an irrigation/stormwater conveyance located in Logan, Utah, USA. While many studies in the literature report valid results with surrogate relationships for a single site or for a small number of sites in a single watershed (e.g., Christensen, 2001; Ryberg, 2006; Settle et al., 2007; Gao et al., 2008; Miguntanna et al., 2010; Jones et al., 2011; Landers and Sturm, 2013; Lessels and Bishop, 2013; Viviano et al., 2014; Fisher et al., 2016; Hannouche et al., 2016) a larger, synthesis study that integrates across many sites and watersheds could further investigate potential differences between the regression techniques we explored. This would aid in verifying the generality of our results.

The relationship between in situ variables and TSS concentrations can be complex and may vary between time periods and conditions such as seasons and storm events. Site specific and storm event specific characteristics were significant surrogate indicators of TSS concentrations at both sites in this study. Additional dedicated studies on the effects and interactions of each explanatory variable would be required to mechanistically

understand the processes that explain the need for site specific and storm event specific data. Additionally, many of these variables may have interaction terms and nonlinear relationships with TSS. Attempting to model all of these complexities with a single equation that meets the assumptions of regression would be infeasible.

Although we have demonstrated how to overcome these limitations to some degree and generate models that may be used for constituent estimates, further research might include investigation of alternative methods for obtaining constituent concentrations as a function of continuous in situ parameters. One promising approach that might be used for this analysis is that of random forests, which allow for regression equations on multiple partitions of the training dataset (Breiman, 2001). These methods allow for both linear and nonlinear relationships and the use of many explanatory variables without concern for linear regression assumptions or multi-collinearity and have demonstrated potential for making accurate predictions of TSS loads (Francke et al., 2008).

3.8. Acknowledgements

This work was supported by funding from the Utah Water Research Laboratory at Utah State University and from United States National Science Foundation award IIA-1208732. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Ahearn, D.S., Sheibley, R.W., Dahlgren, R. a., Anderson, M., Johnson, J., Tate, K.W., 2005. Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California. J. Hydrol. 313, 234–247. https://doi.org/10.1016/j.jhydrol.2005.02.038

Alberto, A., St-Hilaire, A., Courtenay, S.C., van den Heuvel, M.R., 2016. Monitoring stream sediment loads in response to agriculture in Prince Edward Island, Canada. Environ. Monit. Assess. 188, 415. https://doi.org/10.1007/s10661-016-5411-3

Ankcorn, P.D., 2003. Clarifying Turbidity - The Potential and Limitations of Turbidity as a Surrogate for Water-Quality Monitoring, in: Hatcher, K.J. (Ed.), Proceedings of the 2003 Georgia Water Resources Conference. Atlanta.

APHA, 2012. Standard Methods for the Examination of Water and Wastewater, 22nd ed. American Public Health Association, American Water Works Association, Water Environment Federation, Washington, D.C.

Araujo, H.A., Cooper, A.B., Hassan, M. a., Venditti, J., 2012. Estimating suspended sediment concentrations in areas with limited hydrological data using a mixed-effects model. Hydrol. Process. 26, 3678–3688. https://doi.org/10.1002/hyp.8462

Bagiella, E., Sloan, R.P., Heitjan, D.F., 2000. Mixed-effects models in psychophysiology. Psychophysiology 37, 13–20. https://doi.org/10.1017/S0048577200980648

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Berthouex, P. Mac, Brown, L.C., 2002. Statistics for Environmental Engineers, 2nd ed. CRC Press, Boca Raton.

Brauer, N., O'Geen, A.T., Dahlgren, R.A., 2009. Temporal Variability in Water Quality of Agricultural Tailwaters: Implications for Water Quality Monitoring. Agric. Water Manag. 96, 1001–1009. https://doi.org/10.1016/j.agwat.2009.01.011

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Brezonik, P.L., Stadelmann, T.H., 2002. Analysis and Predictive Models of Stormwater Runoff Volumes, Loads, and Pollutant Concentrations from Watersheds in the Twin Cities Metropolitan Area, Minnesota, USA. Water Res. 36, 1743–1757.

Campbell, J.L., Rustad, L.E., Porter, J.H., Taylor, J.R., Dereszynski, E.W., Shanley, J.B., Gries, C., Henshaw, D.L., Martin, M.E., Sheldon, W.M., Boose, E.R., 2013.

Quantity is Nothing without Quality. Bioscience 63, 574–585. https://doi.org/10.1525/bio.2013.63.7.10

Christensen, V.G., 2001. Characterization of Surface-Water Quality Based on Real-Time Monitoring and Regression Analysis, Quivira National Wildlife Refuge, South-Central Kansas, December 1998 Through June 2001.

City of Grand Junction, 2016. Grand Junction Stormwater Management Manual [WWW Document]. Gd. Junction City Munic. Code Vol. II Dev. Regul. URL http://www.codepublishing.com/CO/GrandJunction/html2/GrandJunction28/GrandJunction2852.html (accessed 1.1.16).

City of Logan, 2010. 2010 Logan City Storm Water Management Plan. Logan.

City of Sequim, 2016. Storm and Surface Water Master Plan [WWW Document]. URL http://www.sequimwa.gov/DocumentCenter/View/7735 (accessed 1.1.16).

Coynel, A., Schäfer, J., Hurtrez, J.-E., Dumas, J., Etcheber, H., Blanc, G., 2004. Sampling Frequency and Accuracy of SPM Flux Estimates in Two Contrasted Drainage Basins. Sci. Total Environ. 330, 233–247. https://doi.org/10.1016/j.scitotenv.2004.04.003

Duan, N., 1983. Smearing Estimate: A Nonparametric Retransformation Method. J. Am. Stat. Assoc. 78, 605–610. https://doi.org/10.2307/2288126

Duvert, C., Gratiot, N., Némery, J., Burgos, A., Navratil, O., 2011. Sub-daily Variability of Suspended Sediment Fluxes in Small Mountainous Catchments - Implications for Community-Based River Monitoring. Hydrol. Earth Syst. Sci. 15, 703–713. https://doi.org/10.5194/hess-15-703-2011

Fisher, J.R., Dvorak, B.I., Admiraal, D.M., 2016. Pollutant Load Estimates Using Regression Models with In-Stream Measurements. J. Environ. Eng. 142, 04015081. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001049

Fletcher, T.D., Deletic, A., 2007. Statistical Observations of a Stormwater Monitoring Programme ; Lessons for the Estimation of Pollutant Loads, in: Proceedings of Novatech. Lyon, France, pp. 1575–1582.

Francke, T., López-Tarazón, J.A., Schröder, B., 2008. Estimation of Suspended Sediment Concentration and Yield Using Linear Models, Random Forests and Quantile Regression Forests. Hydrol. Process. 22, 4892–4904. https://doi.org/10.1002/hyp.7110

Gao, P., Pasternack, G.B., Bali, K.M., Wallender, W.W., 2008. Estimating Suspended Sediment Concentration Using Turbidity in an Irrigation-Dominated Southeastern

California Watershed. J. Irrig. Drain. Eng. 134, 250–259. https://doi.org/10.1061/(ASCE)0733-9437(2008)134:2(250)

Gelman, A., 2006. Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. Technometrics 48, 432–435. https://doi.org/10.1198/004017005000000661

Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York.

Gippel, C.J., 1995. Potential of turbidity monitoring for measuring the transport of suspended solids in streams. Hydrol. Process. 9, 83–97. https://doi.org/10.1002/hyp.3360090108

Grayson, R.B., Finlayson, B.L., Gippel, C.J., Hart, B.T., 1996. The Potential of Field Turbidity Measurements for the Computation of Total Phosphorus and Suspended Solids Loads. J. Environ. Manage. 47, 257–267. https://doi.org/10.1006/jema.1996.0051

Gueorguieva, R., Krystal, J.H., 2004. Move Over ANOVA: Progress in Analyzing Repeated-Measures Data and Its Reflection in Papers Published in the Archives of General Psychiatry. Arch. Gen. Psychiatry 61, 310. https://doi.org/10.1001/archpsyc.61.3.310

Hannouche, A., Joannis, C., Chebbo, G., 2016. Assessment of Total Suspended Solids (TSS) Event Load and its Uncertainties in Combined Sewer System from Continuous Turbidity Measurements. Urban Water J. 1–8. https://doi.org/10.1080/1573062X.2016.1254256

Helsel, D.R., Hirsch, R.M., 2002. Statistical Methods in Water Resources, in: Hydrological Analysis and Interpretation. U.S. Geological Survey, p. 510. https://doi.org/10.2307/1269385

Horsburgh, J.S., Reeder, S.L., Jones, A.S., Meline, J., 2015. Open source software for visualization and quality control of continuous hydrologic and water quality sensor data. Environ. Model. Softw. 70, 32–44. https://doi.org/10.1016/j.envsoft.2015.04.002

Horsburgh, J.S., Spackman Jones, A., Stevens, D.K., Tarboton, D.G., Mesner, N.O., 2010. A Sensor Network for High Frequency Estimation of Water Quality Constituent Fluxes Using Surrogates. Environ. Model. Softw. 25, 1031–1044. https://doi.org/10.1016/j.envsoft.2009.10.012

Hvitved-Jacobsen, T., Vollertsen, J., Nielsen, A.H., 2010. Urban and Highway Stormwater Pollution: Concepts and Engineering, 1st ed. CRC Press, Boca Raton.

Irish Jr., L.B., Barrett, M.E., Malina Jr., J.E., Charbeneau, R.J., 1998. Use of Regression Models for Analyzing Highway Storm-Water Loads. J. Environ. Eng. 124, 987–993.

Jones, A.S., Horsburgh, J.S., Mesner, N.O., Ryel, R.J., Stevens, D.K., 2012. Influence of Sampling Frequency on Estimation of Annual Total Phosphorus and Total Suspended Solids Loads. J. Am. Water Resour. Assoc. 48, 1258–1275. https://doi.org/10.1111/j.1752-1688.2012.00684.x

Jones, A.S., Stevens, D.K., Horsburgh, J.S., Mesner, N.O., 2011. Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and Total Phosphorus Concentrations. JAWRA J. Am. Water Resour. Assoc. 47, 239–253. https://doi.org/10.1111/j.1752-1688.2010.00505.x

Kayhanian, M., Suverkropp, C., Ruby, a, Tsay, K., 2007. Characterization and prediction of highway runoff constituent event mean concentration. J. Environ. Manage. 85, 279–95. https://doi.org/10.1016/j.jenvman.2006.09.024

Kirchner, J.W., Feng, X., Neal, C., Robson, A.J., 2004. The Fine Structure of water-Quality Dynamics: The (High-Frequency) Wave of the Future. Hydrol. Process. 18, 1353–1359. https://doi.org/10.1002/hyp.5537

Landers, M.N., Sturm, T.W., 2013. Hysteresis in Suspended Sediment to Turbidity Relations Due to Changing Particle Size Distributions. Water Resour. Res. 49, 5487–5500. https://doi.org/10.1002/wrcr.20394

Leecaster, M.K., Schiff, K., Tiefenthaler, L.L., 2002. Assessment of efficient sampling designs for urban stormwater monitoring. Water Res. 36, 1556–64.

Lessels, J.S., Bishop, T.F.A., 2013. Estimating water quality using linear mixed models with stream discharge and turbidity. J. Hydrol. 498, 13–22. https://doi.org/10.1016/j.jhydrol.2013.06.006

Levesque, V.A., Oberg, K.A., 2012. Computing Discharge Using the Index Velocity Method. Reston.

Lewis, J., 1996. Turbidity-Controlled Suspended Sediment Sampling for Runoff-Event Load Estimation. Water Resour. Res. 32, 2299–2310.

Loperfido, J. V., Just, C.L., Papanicolaou, A.N., Schnoor, J.L., 2010. In situ sensing to understand diel turbidity cycles, suspended solids, and nutrient transport in Clear Creek, Iowa. Water Resour. Res. 46. https://doi.org/10.1029/2009WR008293

Loperfido, J. V., Just, C.L., Schnoor, J.L., 2009. High-Frequency Diel Dissolved Oxygen Stream Data Modeled for Variable Temperature and Scale. J. Environ. Eng. 135, 1250–1256. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000102

Maestre, A., Pitt, R., 2005. The National Stormwater Quality Database , Version 1 . 1: A Compilation and Analysis of NPDES Stormwater Monitoring Information. Washington, D.C.

Melcher, A.A., Horsburgh, J.S., 2017. An urban observatory for quantifying phosphorus and suspended solid loads in combined natural and stormwater conveyances. Environ. Monit. Assess. 189, 285. https://doi.org/10.1007/s10661-017-5974-7

Melcher, A.A., Horsburgh, J.S., Mihalevich, B.A., 2018a. Continuous In Situ Monitoring Data from the Northwest Field Canal in Logan, UT, USA [WWW Document]. Hydroshare. http://www.hydroshare.org/resource/fe05a4877eba46c182ea5ddc26bb140e

Melcher, A.A., J.S. Horsburgh, and B.A. Mihalevich, 2018b. Discharge Rating Curve at 1800 North and 200 West in the Northwest Field Canal in Logan, UT, USA. Hydroshare. http://www.hydroshare.org/resource/26aff3ccb4a7476fb191a0183cf98c46.

Melcher, A.A., Horsburgh, J.S., Mihalevich, B.A., 2018c. Grab Sample Data for the Northwest Field Canal in Logan, UT, USA [WWW Document]. Hydroshare. http://www.hydroshare.org/resource/05e3b46314be444ea89878ff3b0509f7

Miguntanna, N.S., Egodawatta, P., Kokot, S., Goonetilleke, A., 2010. Determination of a Set of Surrogate Parameters to Assess Urban Stormwater Quality. Sci. Total Environ. 408, 6251–9. https://doi.org/10.1016/j.scitotenv.2010.09.015

National Research Council, 2009. Urban Stormwater Management in the United States.

Outram, F.N., Lloyd, C.E.M., Jonczyk, J., Benskin, C.M.H., Grant, F., Perks, M.T., Deasy, C., Burke, S.P., Collins, a. L., Freer, J., Haygarth, P.M., Hiscock, K.M., Johnes, P.J., Lovett, a. L., 2014. High-Frequency Monitoring of Nitrogen and Phosphorus Response in Three Rural Catchments to the End of the 2011–2012 Drought in England. Hydrol. Earth Syst. Sci. 18, 3429–3448. https://doi.org/10.5194/hess-18-3429-2014

Patil, S.S., Barfield, B.J., Wilber, G.G., 2011. Turbidity Modeling Based on the Concentration of Total Suspended Solids for Stormwater Runoff from Construction and Development Sites, in: World Environmental and Water Resources Congress 2011: Bearing Knowledge for Sustainability. Reston, pp. 477–486.

Phillips, J.M., Webb, B.W., Walling, D.E., Leeks, G.J.L., 1999. Estimating the Suspended Sediment Loads of Rivers in the LOIS Study Area Using Infrequent Samples. Hydrol. Process. 13, 1035–1050. https://doi.org/10.1002/(SICI)1099-1085(199905)13:7<1035::AID-HYP788>3.0.CO;2-K

Pinheiro, J.C., Bates, D.M., 2000. Mixed Effects Models in S & S-Plus. Springer, Secaucus, NJ, USA.

Powell, M.J.D., 2009. The BOBYQA algorithm for bound constrained optimization without derivatives.

R Core Team, 2016. R: A language and environment for statistical computing.

Rasmussen, P.P., Gray, J.R., Glysson, G., Ziegler, A.C., 2011. Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data, in: Applications of Hydraulics Section C, Sediment and Erosion Techniques. U.S. Geological Survey, Reston, p. 52.

Rode, M., Wade, A.J., Cohen, M.J., Hensley, R.T., Bowes, M.J., Kirchner, J.W., Arhonditsis, G.B., Jordan, P., Kronvang, B., Halliday, S.J., Skeffington, R.A., Rozemeijer, J.C., Aubert, A.H., Rinke, K., Jomaa, S., 2016. Sensors in the Stream: The High-Frequency Wave of the Present. Environ. Sci. Technol. 50, 10297–10307. https://doi.org/10.1021/acs.est.6b02155

Ryberg, K.R., 2006. Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05. Reston.

Settle, S., Goonetilleke, A., Ayoko, G. a., 2007. Determination of Surrogate Indicators for Phosphorus and Solids in Urban Stormwater: Application of Multivariate Data Analysis Techniques. Water. Air. Soil Pollut. 182, 149–161. https://doi.org/10.1007/s11270-006-9328-2

Slaets, J.I.F., Schmitter, P., Hilger, T., Lamers, M., Piepho, H.-P., Vien, T.D., Cadisch, G., 2014. A turbidity-based method to continuously monitor sediment, carbon and nitrogen flows in mountainous watersheds. J. Hydrol. 513, 45–57. https://doi.org/10.1016/j.jhydrol.2014.03.034

Tomlinson, M.S., De Carlo, E.H., 2003. The Need for High Resolution Time Series Data to Characterize Hawaiian Streams. Journa Am. Water Resour. Assoc. 39, 113–123.

USEPA, 2017. National Summary of State Information [WWW Document]. URL https://iaspub.epa.gov/waters10/attains_nation_cy.control?p_report_type=T) (accessed 2.20.17).

Viviano, G., Salerno, F., Manfredi, E.C., Polesello, S., Valsecchi, S., Tartari, G., 2014. Surrogate measures for providing high frequency estimates of total phosphorus concentrations in urban watersheds. Water Res. 64, 265–277. https://doi.org/10.1016/j.watres.2014.07.009

Wade, A.J., Palmer-Felgate, E.J., Halliday, S.J., Skeffington, R.A., Loewenthal, M., Jarvie, H.P., Bowes, M.J., Greenway, G.M., Haswell, S.J., Bell, I.M., Joly, E., Fallatah, A., Neal, C., Williams, R.J., Gozzard, E., Newman, J.R., 2012. Hydrochemical Processes in Lowland Rivers: Insights from In Situ, High-Resolution Monitoring. Hydrol. Earth Syst. Sci. 16, 4323–4342. https://doi.org/10.5194/hess-16-4323-2012

Wanielista, M., Kersten, R., Eaglin, R., 1997. Hydrology: Water Quantity and Quality Control, 2nd ed. John Wiley and Sons, Inc., Hoboken.

Wu, L., 2010. Mixed Effects Models for Complex Data. CRC Press, Boca Raton.

Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M., 2009. Mixed effects models and extensions in ecology with R, Statistics for Biology and Health. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-87458-6

Table 3.1. Summary of the TSS sampling efforts in the NWFC study area

Site	Number of TSS Samples Collected				
	Base Flow	Average Storm Event	All Storm Events Monitored	Rising Limb	Falling Limb
Upstream (200 South)	10	8	143	63	80
Downstream (1800 North)	10	11	187	67	120

Table 3.2. Summary of sampling according to storm event size and antecedent dry period

Site	Storm Event Size[1]	Number of TSS Samples Collected		
		Antecedent Dry Period	Rainfall Intensity	Rainfall Depth
Upstream (200 South)	Small	59	11	87
	Medium	34	76	23
	Large	50	56	33
Downstream (1800 North)	Small	79	35	82
	Medium	13	75	53
	Large	95	77	52

[1]Small, medium, and large events were determined by taking the range of storm event values (antecedent dry period, rainfall intensity, and rainfall depth) and dividing them into three equal groups.

Table 3.3. TSS sample distribution for each storm event at both monitoring sites.

	Number of TSS Samples Collected	
Storm Event	Upstream Site (200 South)	Downstream Site (1800 North)
July 8, 2015	20	20
July 27, 2015	15	15
August 3, 2015	-	6
August 7, 2015	-	3
September 14, 2015	4	-
September 15, 2015	5	-
September 16, 2015	7	-
October 3, 2015	8	10
May 6, 2016	2	13
May 7, 2016	6	5
May 10, 2016	13	14
May 19, 2016	10	24
May 25, 2016	5	8
June 11, 2016	9	12
June 12, 2016	11	11
August 7, 2016	9	7
September 13, 2016 (a)	-	9
September 13, 2016 (b)	10	15
September 14, 2016	8	5
September 21, 2016	7	15

Table 3.4. Comparison of regression model results, where the base model (BM) represents the classical linear regression model at each site.

Site	Regression Method	Equation	C_p	R_a^2	$RMSE$	$PRESS$	TSS Load (kg)
Up-stream (200 South)	CLR (BM)	$TSS^{0.5} = 0.0730 + 1.067Turb^{0.5}$	27.90	0.810	0.476	36.3	59,900
	LRCAT	$TSS^{0.5} = BM + SizeADP + SizeADP * Turb^{0.5} + Season + Season * Turb^{0.5}$	-	0.875	0.376	26.0	68,400
	LME	$TSS^{0.5} = BM + SizeADP + SizeADP * Turb^{0.5} + Season + Season * Turb^{0.5}$	-	0.872	0.381	25.8	59,800
Down-stream (1800 North)	CLR (BM)	$TSS^{0.5} = -0.804 + 1.257Turb^{0.5}$	36.0	0.960	0.836	142	117,500
	LRCAT	$TSS^{0.5} = BM + Limb + Limb * Turb^{0.5} + SizeINT + SizeINT * Turb^{0.5}$	-	0.972	0.683	105	162,000
	LME	$TSS^{0.5} = BM + Limb + Limb * Turb^{0.5} + SizeINT + SizeINT * Turb^{0.5}$	-	0.972	0.686	103	154,400

Variable	Description
TSS	Total suspended solids (mg/L)
$Turb$	Turbidity (NTU)
$SizeADP$	Group and categorical variable for the length of the antecedent dry period (e.g., small, medium, large, or base flow)
$SizeINT$	Group and categorical variable for the size of the storm based on the rainfall intensity (e.g., small, medium, large, or base flow)
$Season$	Group and categorical variable for the season (e.g., Spring or Fall)
$Limb$	Group and categorical variable for the rising or falling limb of pollutograph

Table 3.5. Storm event and base flow predictions of total TSS load for each site and regression method.

Site	Regression Method	Storm Event TSS Load (kg)	Base Flow TSS Load (kg)
Upstream (200 South)	CLR	6,320	53,620
	LRCAT	6,490	61,960
	LME	6,470	53,350
Downstream (1800 North)	CLR	43,830	73,680
	LRCAT	43,750	118,200
	LME	43,720	110,680

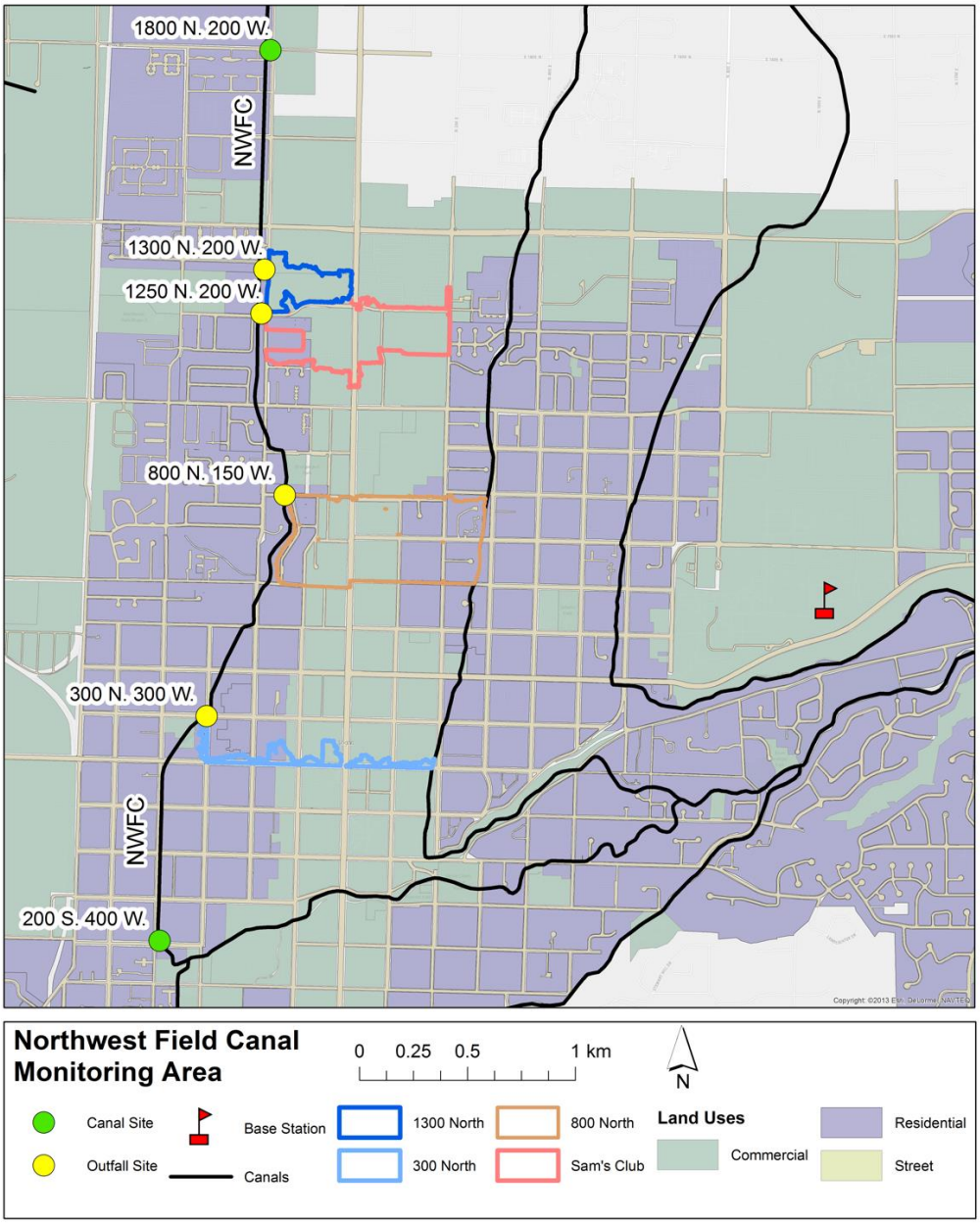


Fig 3.1. Map of the Northwest Field Canal study area.

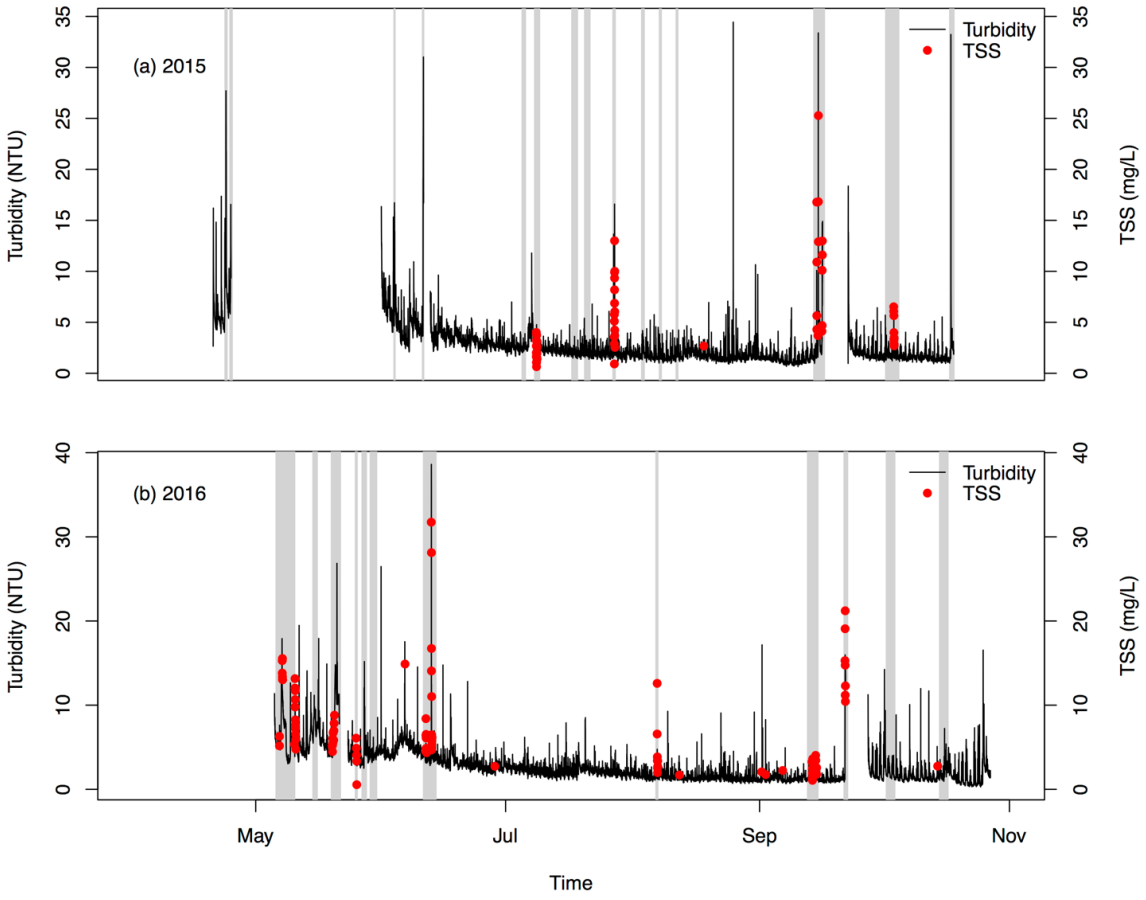


Fig 3.2. Plots of turbidity and TSS for the a) 2015 and b) 2016 irrigation seasons at the upstream site (200 South). Gray shaded areas indicate the occurrence of a storm event.

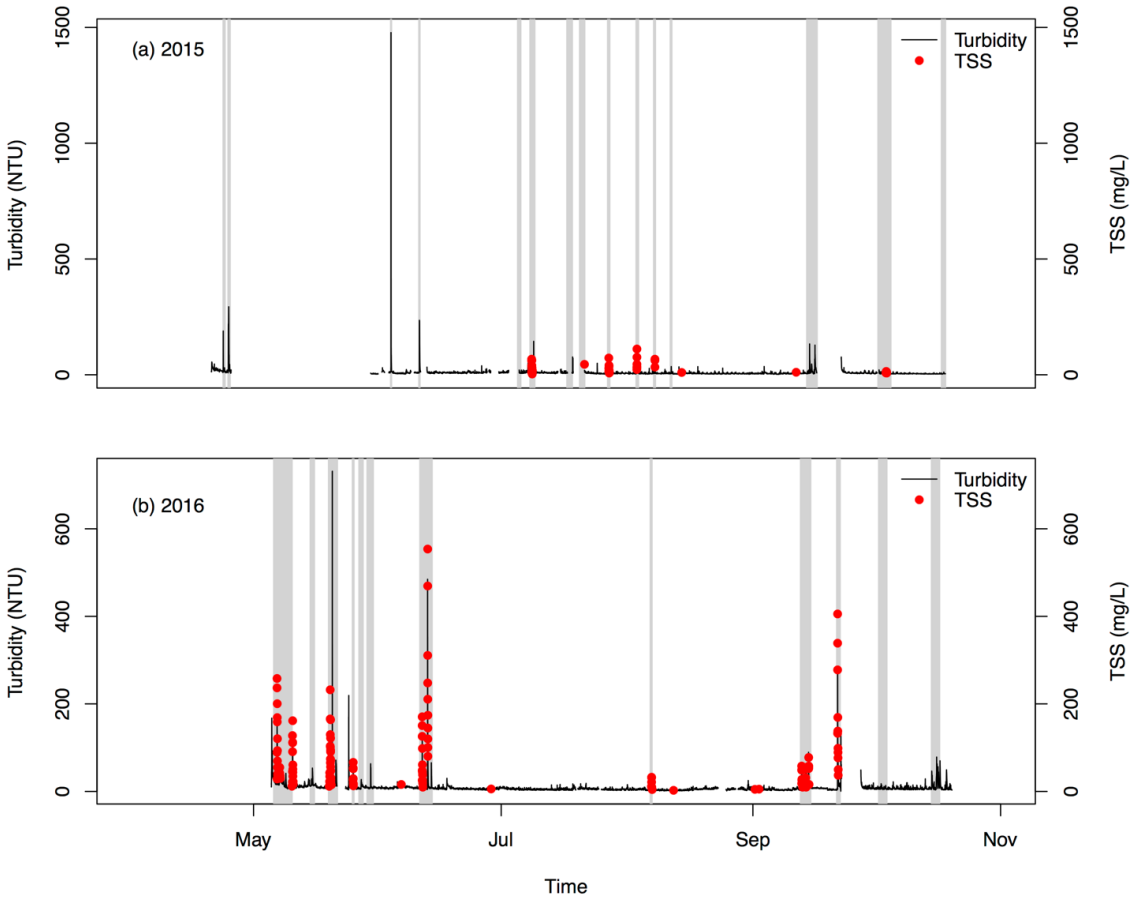


Fig 3.3. Plots of turbidity and TSS for the a) 2015 and b) 2016 irrigation seasons at the downstream site (1800 North). Gray shaded areas indicate the occurrence of a storm event.

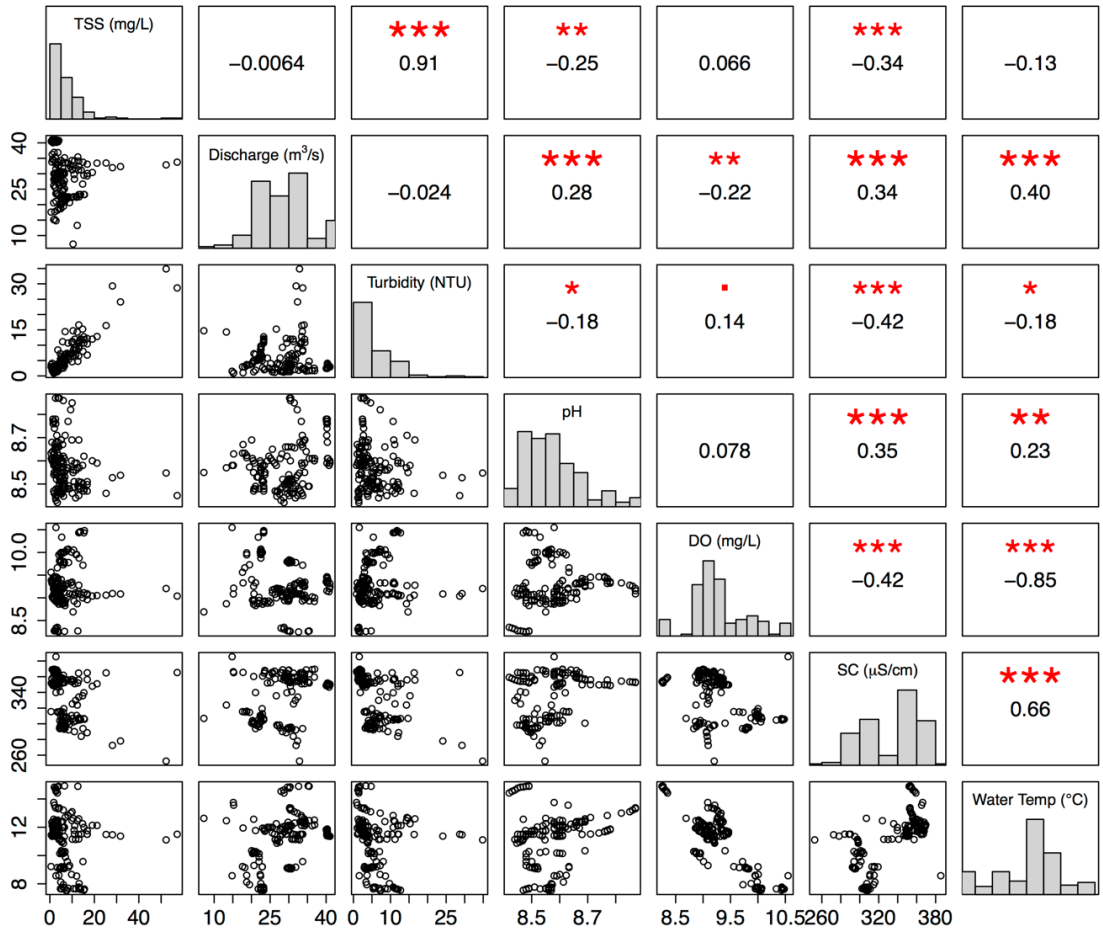


Fig 3.4. Correlation plots of each explanatory variable considered and TSS at the upstream site (200 South). Symbols in the upper panel indicate the significance of the Pearson's correlation coefficient (" *** ", 0.001;" ** ", 0.01;" * ", 0.05;" . ", 0.1).

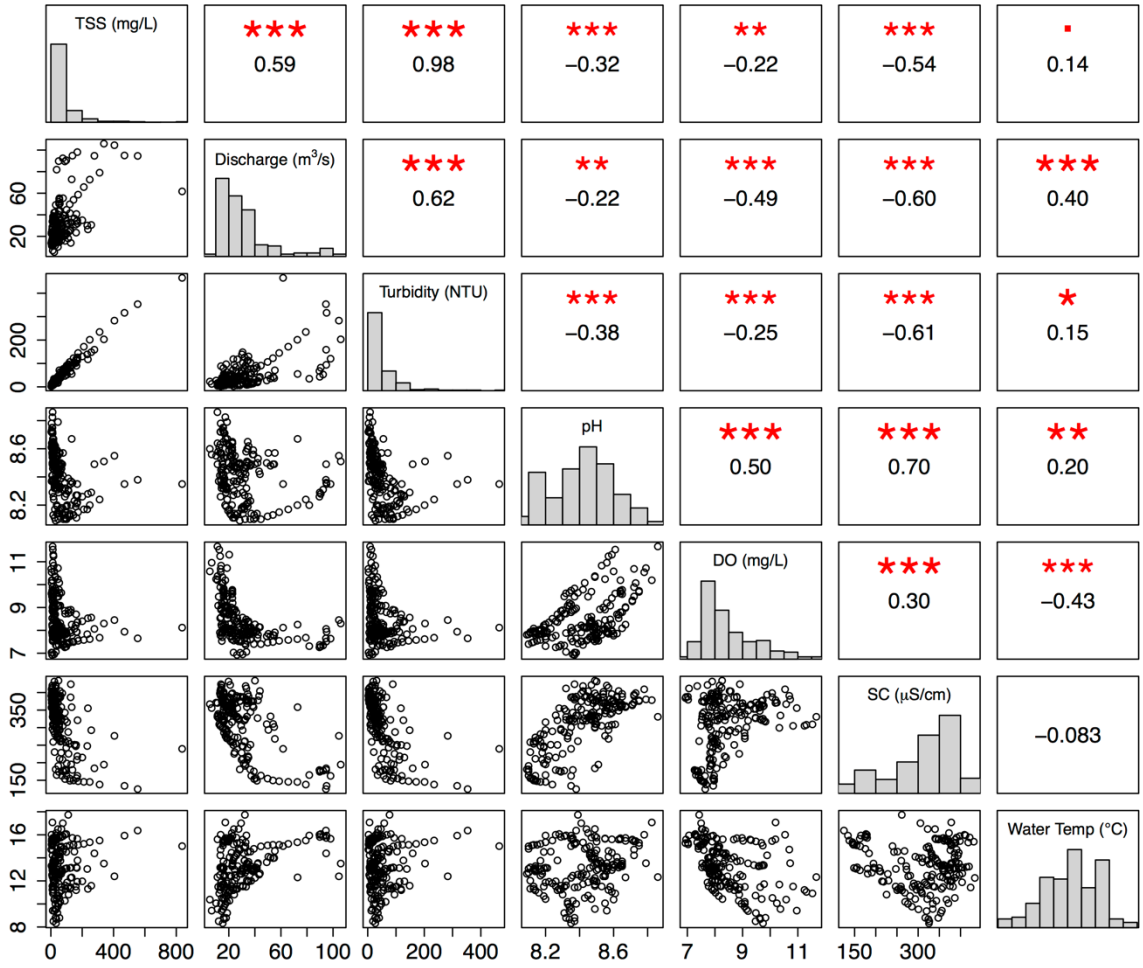


Fig 3.5. Correlation plots of each explanatory variable considered and TSS at the downstream site (1800 North). Symbols in the upper panel indicate the significance of the Pearson's correlation coefficient (" *** ", 0.001;" ** ", 0.01;" * ", 0.05;" . ", 0.1).

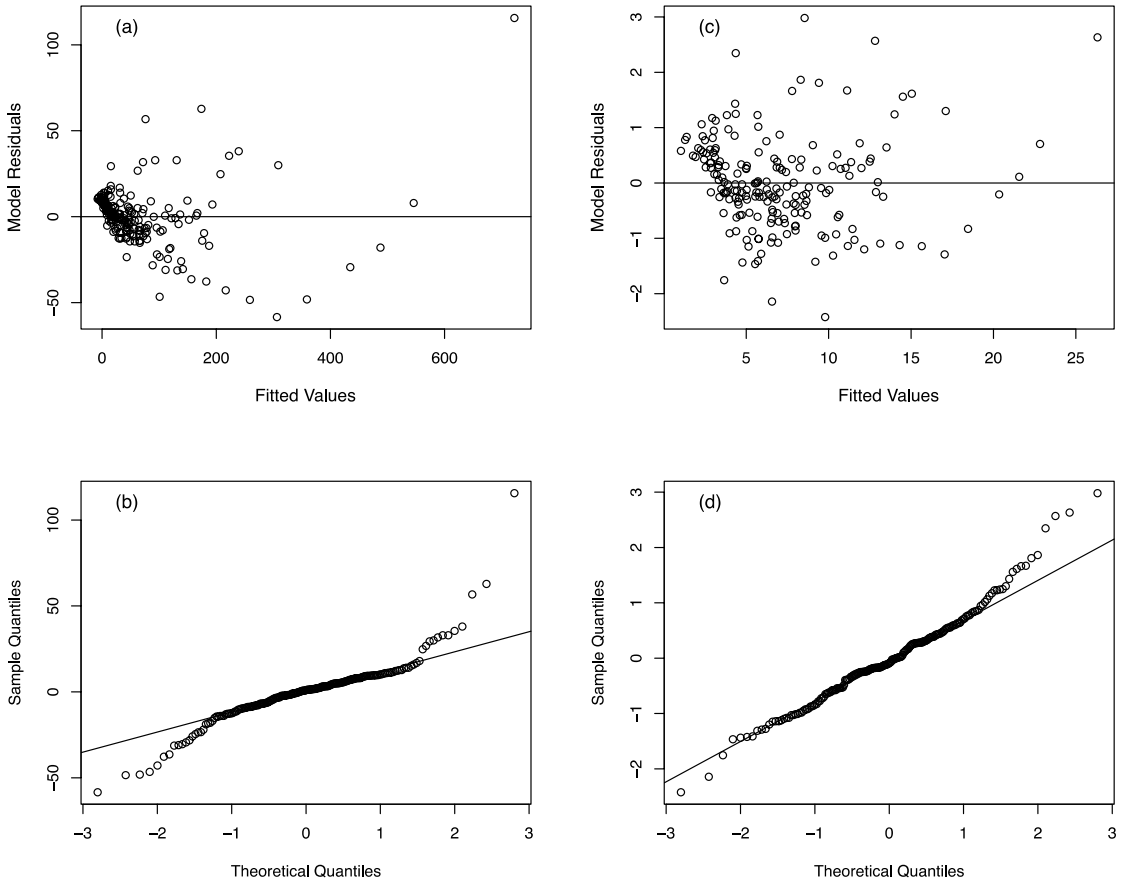


Fig 3.6. Residuals versus fitted plots and quantile-quantile plots for the TSS-turbidity model with no transformation (a, b) and with the square root transformation (c, d) for the downstream monitoring site.

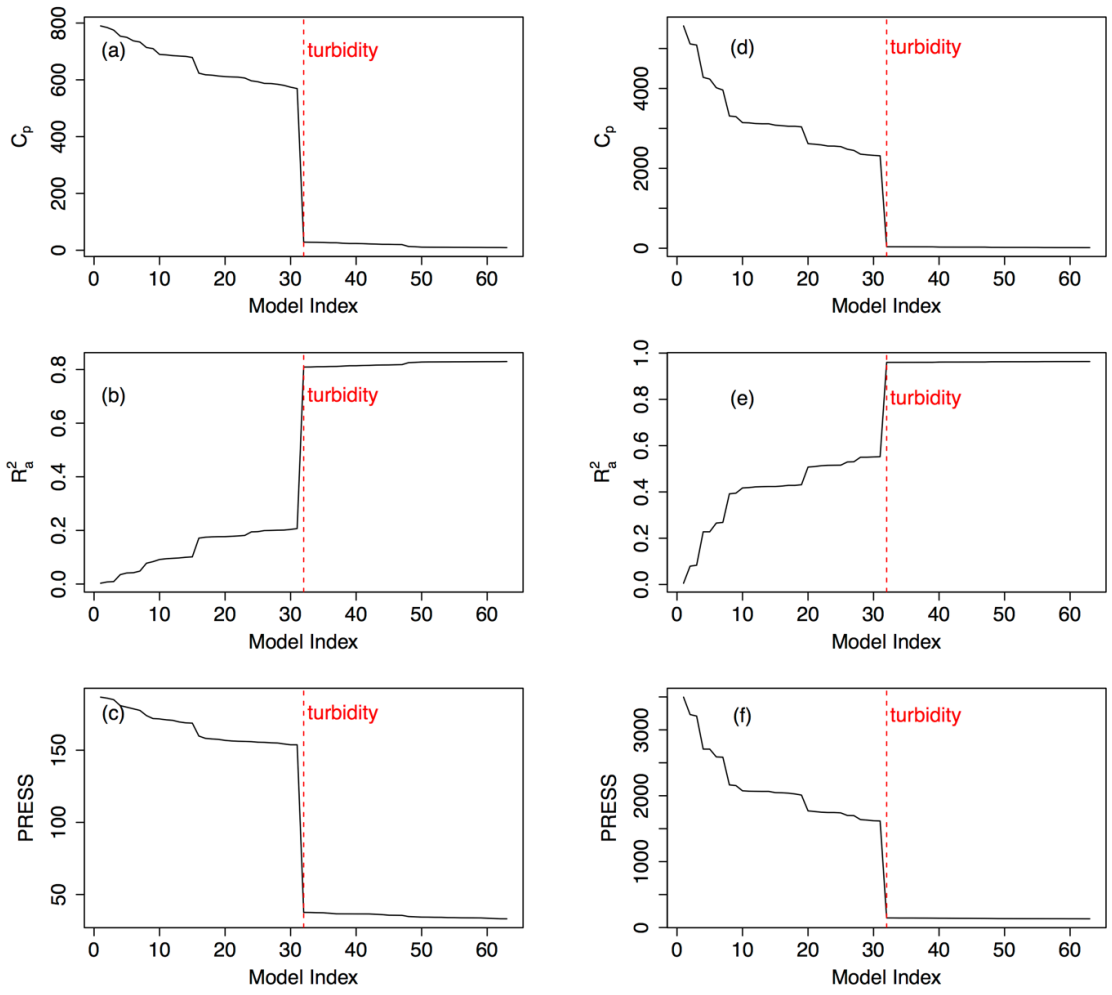


Fig 3.7. Plots of all of the 2^k possible CLR models against the model quality metrics (Malley's C_p, R_a^2, and the PRESS statistic) for the upstream site (200 South) (a, b, c) and downstream site (1800 North) (d, e, f). The vertical line with the "turbidity" label indicates the point at which the square root of turbidity was added to the model as an explanatory variable and shows the large increase in model quality that results when this term is included.

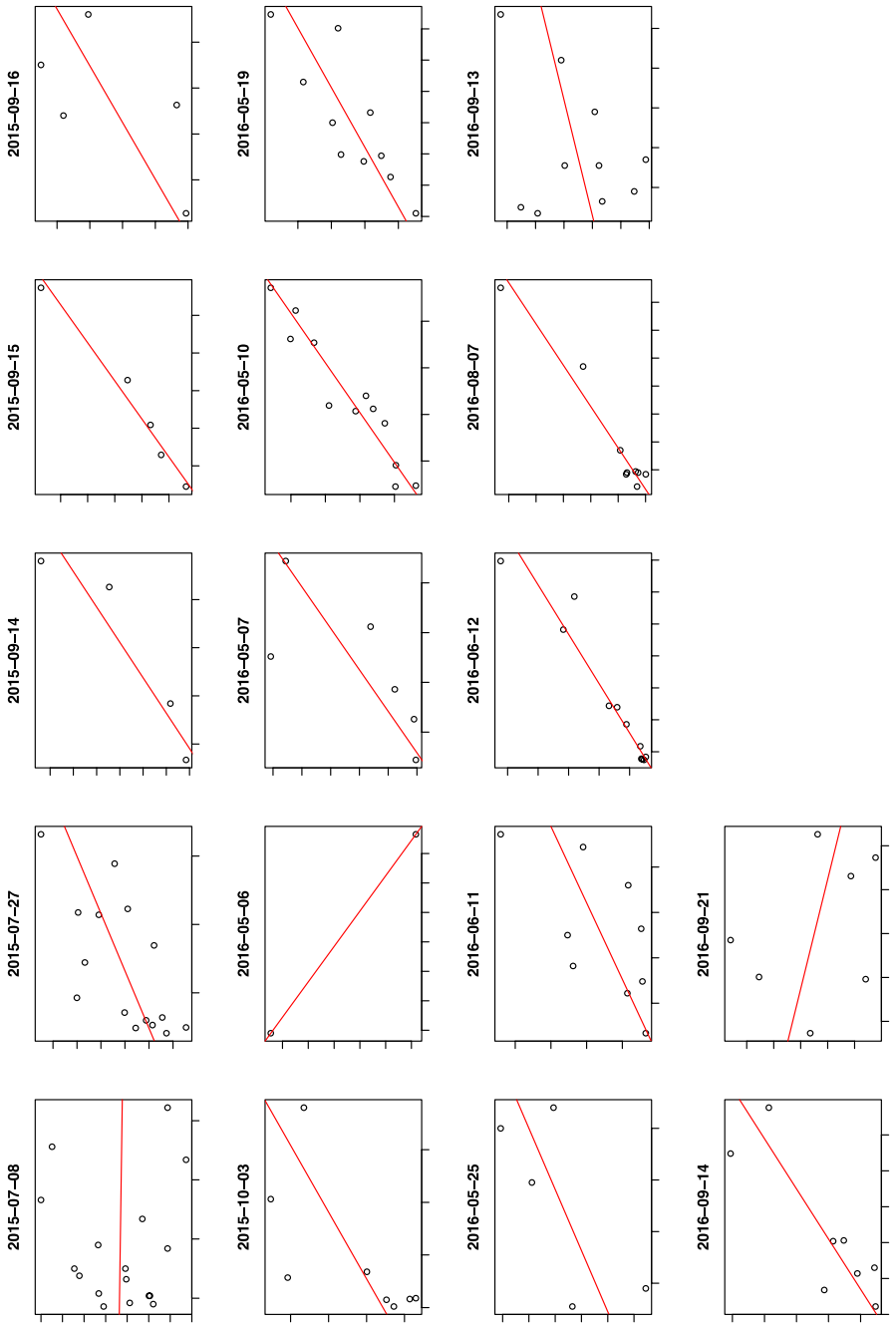


Fig 3.8. Plots of TSS versus turbidity for each storm event monitored at the upstream site (200 South) demonstrating the variability of the slopes and intercepts of the ordinary least squares regression line.

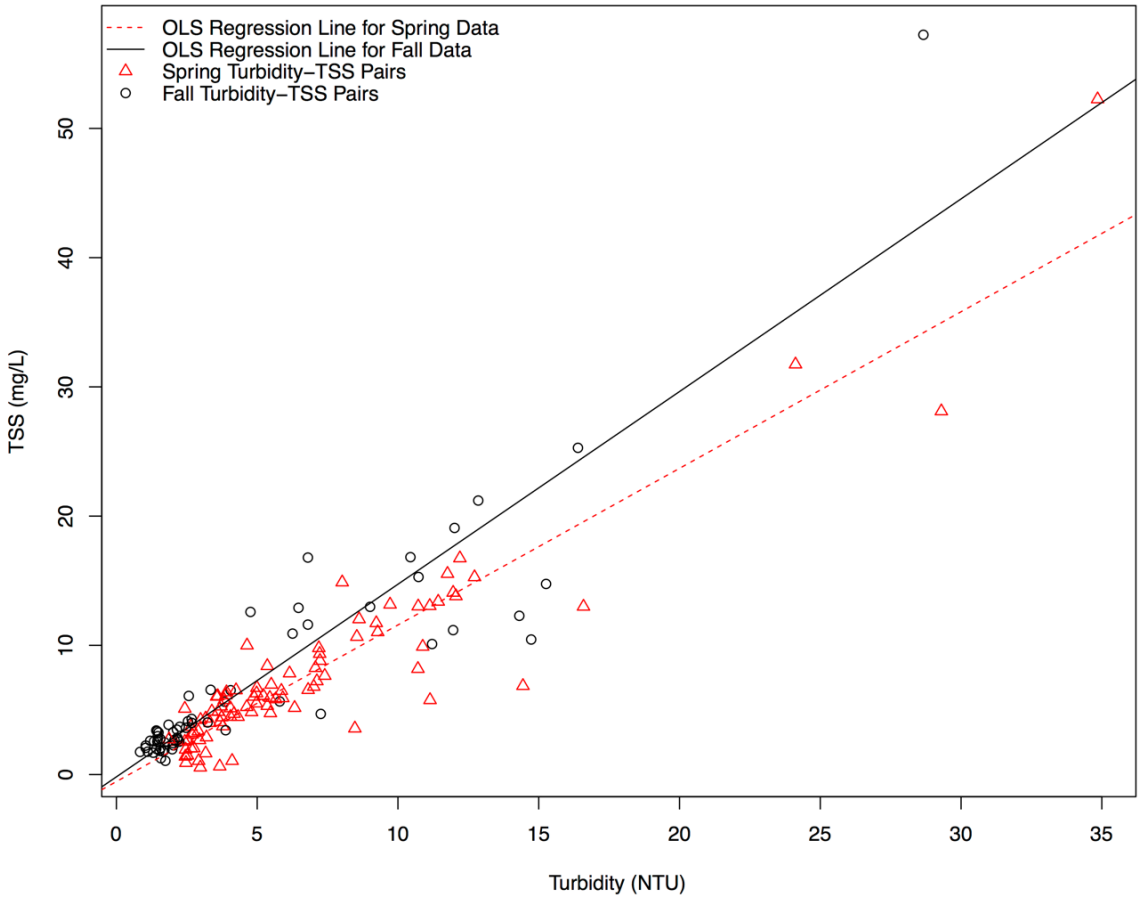


Fig 3.9. TSS versus turbidity plot with ordinary least squares lines fit to the Spring and Fall data.

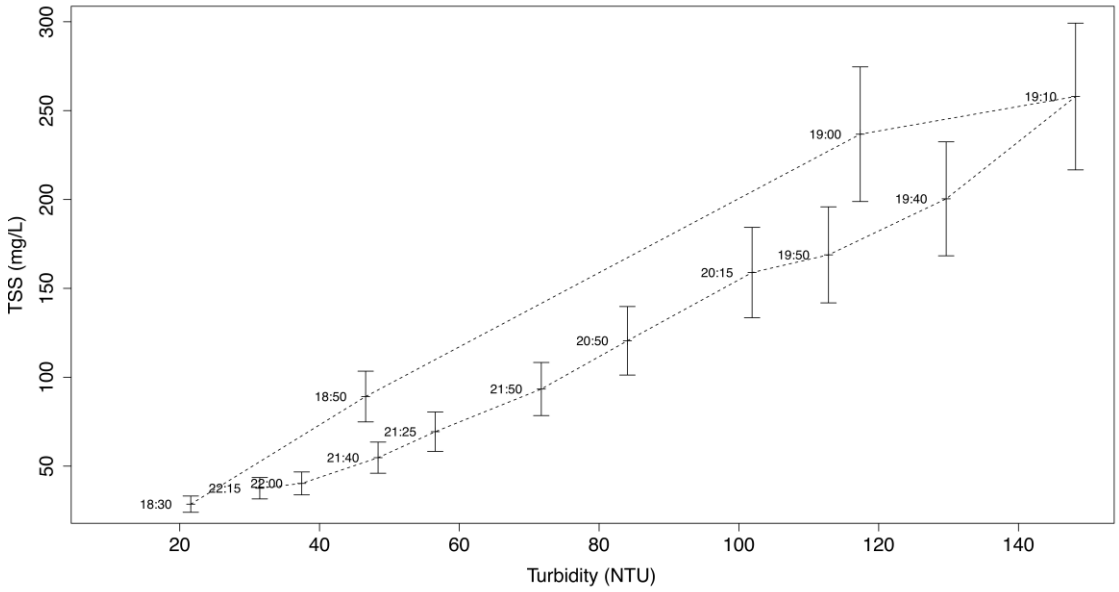


Fig 3.10. Plot of TSS versus turbidity for the storm event on May 6, 2016 at the downstream site (1800 North) showing an example of clockwise hysteresis. Times at which samples were collected are given in the plot as point labels. Error bars represent the plus or minus the standard deviation of the percent error (~16 percent) between sample duplicates

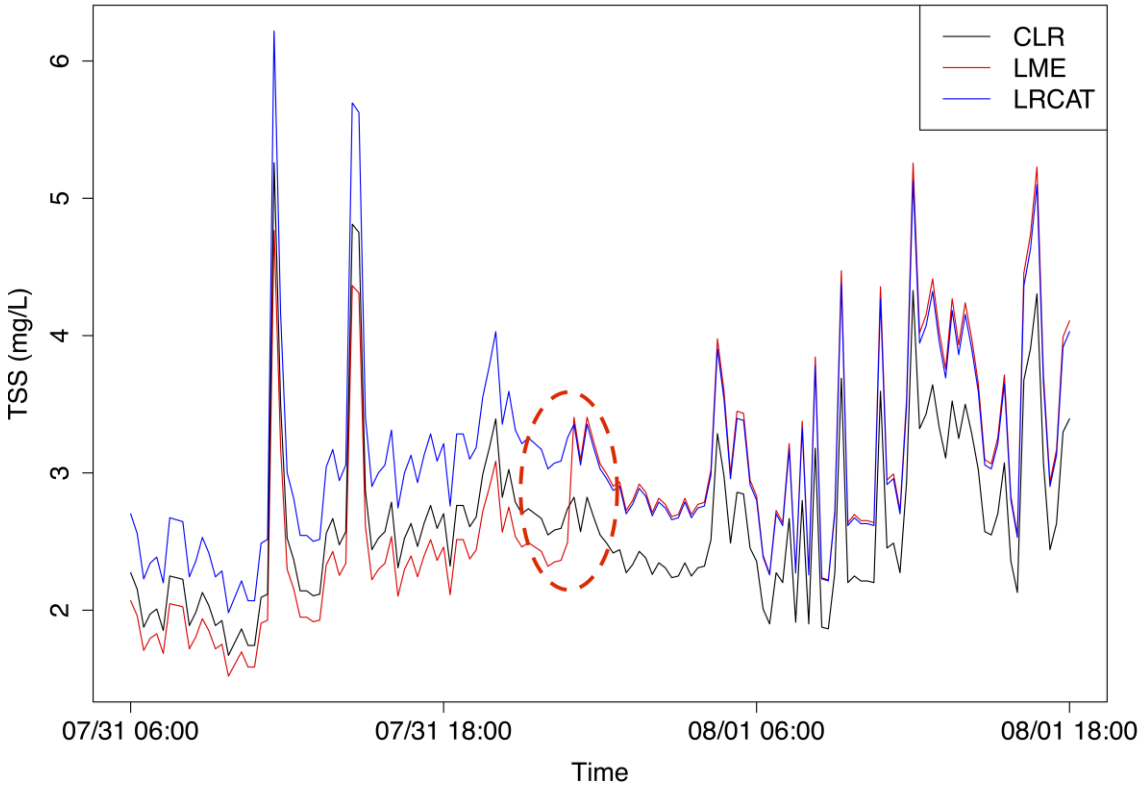


Fig 3.11. Predicted TSS concentrations at the upstream site for each of the linear regression methods. Circled is a discontinuity caused by the "season" categorical variable.

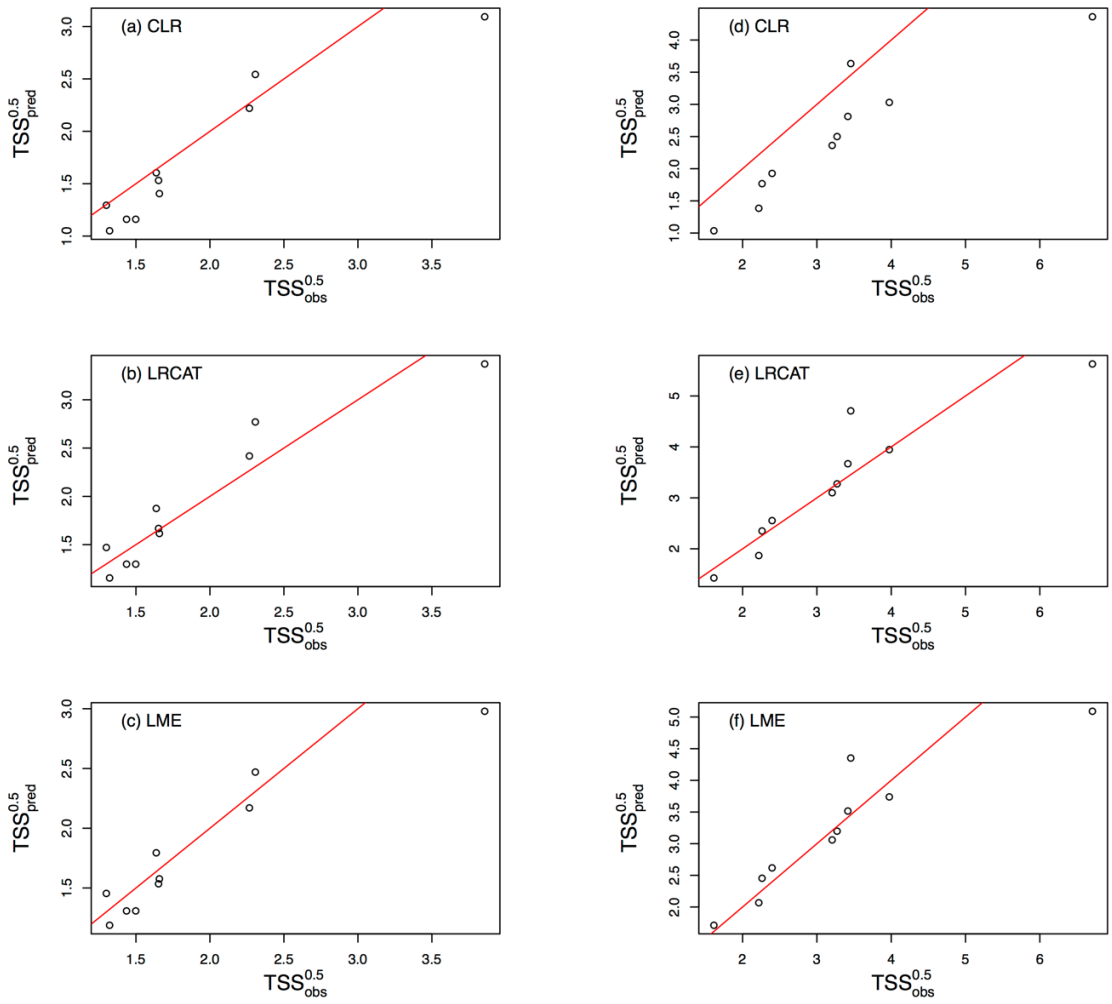


Fig 3.12. Predicted versus observed TSS base flow concentrations on the square root scale. Panels a, b, and c show the plots for the selected classical linear regression (CLR), linear regression with categorical variables (LRCAT), and linear mixed effects (LME) models at the upstream site (200 South) respectively. Plots d, e, and f show the plots for the selected CLR, LRCAT, and LME models at the downstream site (1800 North), respectively.

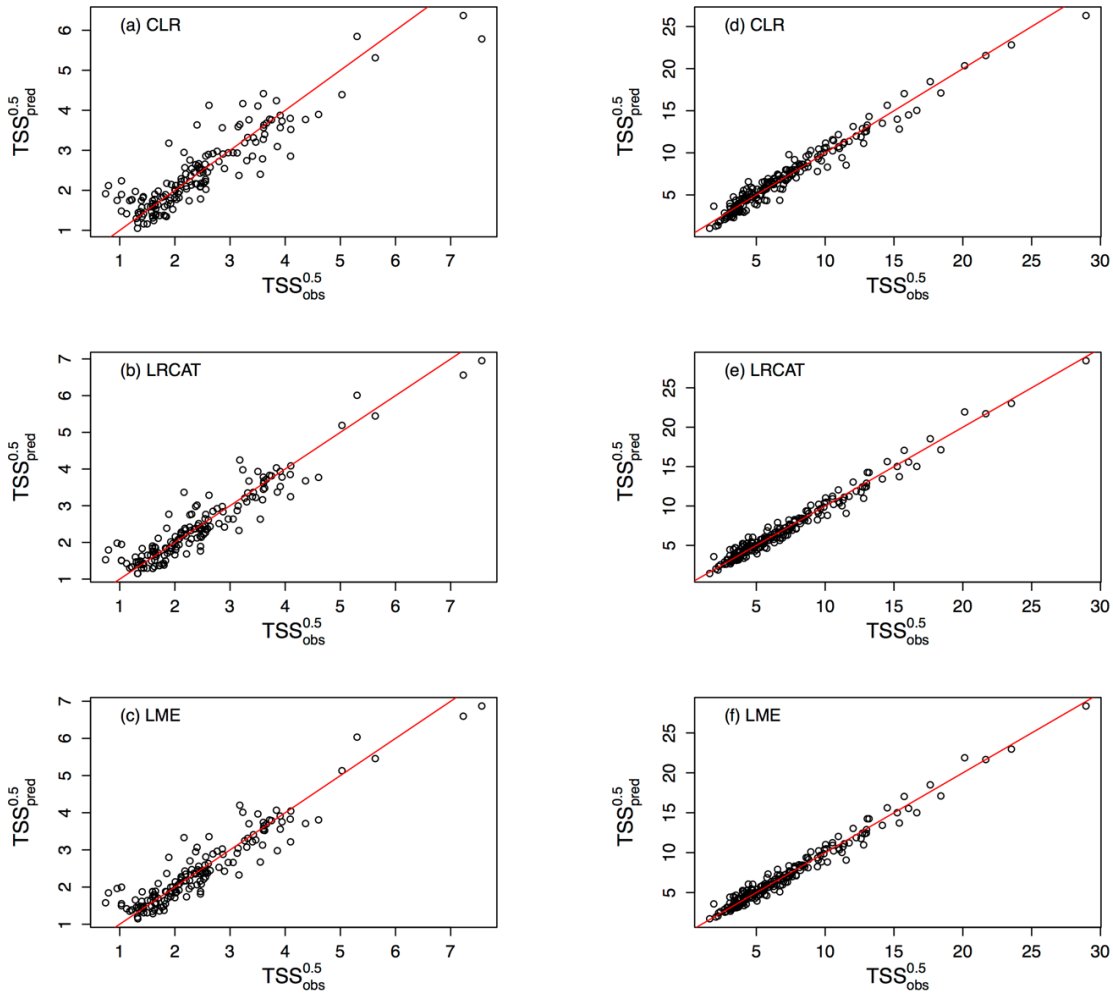


Fig 3.13. Predicted versus all observed TSS concentrations on the square root scale. Panels a, b, and c show the plots for the selected classical linear regression (CLR), linear regression with categorical variables (LRCAT), and linear mixed effects (LME) models at the upstream site (200 South) respectively. Plots d, e, and f show the plots for the selected CLR, LRCAT, and LME models at the downstream site (1800 North) respectively.

CHAPTER 4

MODELING RUNOFF IN AN URBAN STORMWATER CONVEYANCE USING HIGH FREQUENCY DATA

Abstract

Quantifying spatial and temporal variability in urban stormwater runoff is critical in identifying pollutant source areas and in evaluating the potential for management practices aimed at preventing pollution of downstream receiving waters. Simulation models can be used for this purpose, but must be driven and constrained by data. Wireless sensor networks (WSN) can be used to collect high-resolution data for modeling; however, operation of WSNs is logistically challenging, expensive, and the degree to which high-resolution data collected via WSNs can be used to improve ability to simulate stormwater discharge has not been well established. This study examined how using data from multiple continuous monitoring sites nested within an urban water system to calibrate a stormwater runoff model affected model performance in terms of simulating discharge when compared to the same model calibrated using only data collected at the system outlet. Multiple configurations of the United States Environmental Protection Agency's Storm Water Management Model (SWMM) were calibrated using a genetic algorithm for two summer irrigation seasons within an urban water system in Logan, Utah, USA that receives direct urban stormwater runoff. Model configurations represented inclusion or exclusion of data from different continuous monitoring sites in the calibration. Results showed that calibrating the model using data from nested outfall sites along with data from the model outlet predicted similar hydrographs at the model outlet when compared to the models

calibrated using only data from the model outlet. However, the models calibrated using only data from the model outlet were unable to accurately predict hydrographs at the nested outfall sites. Thus, ability to predict discharge at multiple locations within a study area can be enhanced by high-resolution data collection and can aid water managers in determining runoff contributions and selecting best management practices within an urban drainage.

[1]Co-authored by Anthony A. Melcher, Jeffery S. Horsburgh, Bethany T. Neilson, and Caleb A. Buahin

4.1. Introduction

Advances in environmental data collection technology have enabled better understanding of the characteristics of runoff events in both rural and urbanized watersheds (Horsburgh et al., 2010; Rode et al., 2016; Melcher and Horsburgh, 2017). Environmental wireless sensor networks (WSN) allow for multiple monitoring nodes equipped with in situ sensors that enable stormwater to be monitored at high, adaptive frequencies and to be synchronized across a larger geographic region (Corke et al., 2010; Kerkez et al., 2016; Wong and Kerkez, 2016; Melcher and Horsburgh, 2017). Such monitoring efforts are necessary, especially in urban watersheds where runoff events may occur on much smaller spatial and temporal scales than more rural watersheds (Deletic and Maksimovic, 1998; Waschbusch et al., 1999; Tomlinson and De Carlo, 2003; Kirchner et al., 2004; National Research Council, 2009; Gong et al., 2016). Understanding event dynamics is essential for characterizing runoff, predicting potential impacts of future events (Kirchner et al., 2004),

and can also allow best management practices (BMP) to be tailored to control runoff and pollution from specific source areas.

While many studies have employed high resolution data collection aimed at creating a better understanding of hydrologic processes, monitoring with high frequency at every location where data might be needed in managing a stormwater system is cost prohibitive. Simulation models can be developed to "fill in the gaps" and produce output data at locations that are not monitored. However, the authors were unable to find any studies that focused on how or if high resolution data collected across multiple monitoring nodes within an urban WSN enhances our ability to simulate stormwater runoff using current models. Stormwater data are used as inputs and to calibrate and validate stormwater models (Wanielista and Yousef, 1993; Mroczkowski et al., 1997). Not only do the frequency and volume of high resolution data pose a challenge for existing stormwater models, but the availability of high resolution data at multiple points within an urban water system poses a challenging question of how to approach simultaneous, multi-site calibration in an urban environment. An additional question is what benefits this type of calibration will have on the precision and accuracy of urban hydrology/hydraulic models such as the U.S. Environmental Protection Agency's (USEPA) Storm Water Management Model (SWMM).

Outside of the urban stormwater context, some authors have investigated the value of using multi-site, or multi-response, monitoring data to populate and calibrate runoff models (Lerat et al., 2012; Wang et al., 2012; Chiang et al., 2014; Leta et al., 2017). Their approaches and conclusions have varied. In their comprehensive study of streamflow from 187 catchments in France, Lerat et al. (2012) looked at benefits of multi-site versus single

site calibration of the GR4J model (in French, modèle du Génie Rural à 4 paramètres Journalier). Model performance metrics calculated in this study were indices derived from the root mean squared error (RMSE) and a total volume error. Those metrics were computed at the catchment outlet and up to two interior sites. It was found that multi-site calibration produced nearly identical results to the model calibrated using only data collected at the catchment outlet. Differing results between studies may be attributed to the underlying process equations unique to each model, or to each study's unique approach to model setup, calibration, and methods for quantifying model performance. Answers to the questions on the utility and how to approach multi-site model calibration continue to be pursued (Chiang et al., 2014; Lerat et al., 2012; Leta et al., 2017; Li et al., 2010; Shrestha et al., 2016; Wang et al., 2012).

The approach chosen for stormwater modeling is often determined by the motivation of the study (Kirchner, 2006; National Research Council, 2009; Niazi et al., 2017; Tsihrintzis and Hamid, 1997). From a management context, it might be of interest to model catchments to determine where a stormwater BMP might be best located. This context might include a coarser spatial discretization of the drainage area and larger modeling time steps, as the modeler is mostly concerned with total runoff volumes. A different approach might focus on better understanding the hydrologic response of subcatchments and distinguishing characteristics of the hydrograph. This category of model might be used to gain a greater understanding of the processes, unique to each modeled catchment, that drive runoff events. This latter approach provides information useful for determining which BMP would obtain the best hydrograph attenuation and would require a higher spatial and temporal resolution model.

With climate variability and the changing frequency of extreme storm events, the understanding of runoff processes and the ability to model those processes is increasingly pertinent. The purpose of this study was to investigate how high resolution stormwater runoff data collected via an environmental WSN can be used to improve the modeling of an urban water system. More specifically, we aimed to answer the following questions pertaining to stormwater modeling and its adequacy at estimating stormwater runoff quantities in an urban study area: 1) How can an urban stormwater model be effectively calibrated using high resolution data from boundary sites and multiple sites within the modeling domain?; and 2) What is the value of each additional monitoring site added to the calibration procedure in improving the accuracy of model predictions at model outlet? We chose to use the USEPA-SWMM stormwater model for this study primarily because it is one of the most widely used urban stormwater models (Niazi et al., 2017; Obropta and Kardos, 2007; Tsihrintzis and Hamid, 1997). While we anticipate that results may vary across models, SWMM provided us with the opportunity to explore these questions using a model that represents the current state of the practice in stormwater modeling (Niazi et al., 2017).

4.2. Background

SWMM is a semi-distributed, rainfall-runoff model used primarily to simulate water quantity and quality in urban water systems (Obropta and Kardos, 2007; Rossman and Huber, 2016). The components of SWMM's water quantity simulation include surface, sub-surface, and conveyance routing using precipitation and other meteorological values

as external forcing data (Rossman and Huber, 2016). Surface runoff is discretized into subcatchments, which are treated as nonlinear reservoirs where a depression storage must be overcome before runoff is generated (Rossman and Huber, 2016). Impervious and pervious areas can be defined within subcatchments, and the connectivity of impervious areas to the subcatchment outlet can be specified through subarea routing. Conveyances modeled in SWMM allow for both kinematic and dynamic wave routing. Kinematic wave routing solves a simplified version of the 1D Saint-Venant equations. Flows are assumed to be uniform, and water surfaces are parallel to conduit invert slopes. Kinematic wave routing does not account for pressurized flow or backwater effects and is more applicable to flows in steep-sloped conduits. Dynamic wave routing solves the complete forms of the Saint-Venant equations, including inertial and pressure terms, and allows for channel storage, unsteady, gradually varied, and pressurized flows (Niazi et al., 2017; Rossman, 2017; Sun et al., 2013).

Common applications of SWMM include municipal storm sewer system design and flood analysis. In both cases, common practice is to use previously designated “design storms” with a specified return period to determine whether infrastructure is sufficient to route the flows expected under worst-case scenarios (Niazi et al., 2017; Rossman and Huber, 2016). While useful for sizing infrastructure, these types of simulations do not provide information about how an urban water system might respond over the range of real storm events and conditions that might occur within the modeled catchment. Where the modeling objective is to examine the performance of the model with respect to observations of real conditions, design storms hold little value and emphasis should be on the use of measured data (Rossman and Huber, 2016). While others have looked at the effects of

varying temporal and spatial scales on SWMM water quantity results (Krebs et al., 2013; Niazi et al., 2017; Petrucci and Bonhomme, 2014; Sun et al., 2014), our search of the literature did not find any examples of the use of high frequency data from a synchronized, multi-node WSN to calibrate a SWMM model or evaluation of how the addition of data from multiple monitoring sites might affect water quantity predictions.

4.3. Study Area

The Northwest Field Canal (NWFC) is a primarily un-lined, open-channel, combined irrigation/stormwater conveyance that runs through the heart of the City of Logan, Utah, USA. Runoff from much of the commercial and residential zones of Logan is received by the NWFC and ultimately conveyed to Cutler Reservoir. In the mid-nineteenth century, when the canal was originally constructed, it conveyed solely irrigation water. However, with unregulated stormwater flows that were introduced as the City grew, the canal now frequently floods during larger storm events, causing damage to adjacent properties.

The upstream monitoring site (200 South) was located near the NWFC diversion from the Little Logan River (Figure 4.1). The Logan River-Little Logan River watercourse has very few stormwater inputs upstream of this point. The downstream monitoring site is located about 4.5 km downstream of the upstream monitoring site and is near where the canal leaves Logan City's boundary. The upstream and downstream monitoring sites bookend the model domain and provide water quantity boundary conditions in the canal. In addition to the upstream and downstream monitoring sites, six stormwater monitoring

sites were located in outfalls to the canal. Land uses within the NWFC drainage area are comprised primarily of low and high-density residential and commercial zones. Table 4.1 provides a description of the NWFC drainage area as well as the drainage areas for the outfall monitoring sites.

4.4. Methods

4.4.1. Data Collection

Water quantity data were obtained from six monitoring sites in the NWFC to provide the data needed to populate and calibrate the models that were developed for this study. Details of the data collection procedure are described in Melcher and Horsburgh (2017), but a brief summary is provided here for completeness. At any given time, data collection consisted of two continuous canal monitoring sites at the upstream and downstream ends of the study area and two semi-mobile monitoring sites located in stormwater outfalls to the canal (Figure 4.1). The two outfall sites were moved periodically during the study after enough storm events were monitored of varying sizes and intensities. All monitoring sites were equipped with a tipping-bucket rain gage (TE525, Campbell Scientific, Logan, Utah, USA), which summed precipitation depths over a 15-minute interval at the canal sites and at a 5-minute interval at the outfall sites. The total number of storm events monitored and other catchment characteristics are summarized in Table 4.1.

Discharge values at the canal sites were obtained by either the rating curve method (Rantz, 1982) or the index velocity method (Levesque and Oberg, 2012) using a side-looking acoustic Doppler velocity meter (ADVM) (Sontek SL3000, San Diego, California,

USA). The downstream canal site (1800 North, 200 West) was equipped with a pressure transducer for in-situ stage measurements. These measurements were then correlated to periodic discharge measurements made with an acoustic Doppler velocimeter (ADV) (Sontek FlowTracker, San Diego, California, USA) (Melcher et al., 2018a). As a result, a rating curve was derived, and discharge values were obtained at 15-minute intervals during base flow conditions and 5-minute intervals during storm events. The upstream canal site (200 South, 400 West) was equipped with an ADVM given that at that site there were multiple downstream hydraulic controls and water diversion structures that caused conditions where the same discharge value could correspond to multiple water depths. The ADVM made flow measurements as a function of channel geometry, water depth, and velocity, thus circumventing what would have been a poor stage-discharge relationship. Those measurements were then correlated with ADV measurements so as to ensure that upstream and downstream discharge values were derived from observations made with the same instrument. Discharge values at the outfall sites were obtained at 1-minute intervals during storm events using an area-velocity flow module (ISCO 2150, Teledyne ISCO, Lincoln, Nebraska, USA). All discharge data were then used to calibrate and validate the SWMM model.

In addition to water quantity data, geospatial datasets were obtained from Logan and North Logan Cities including a high-resolution LiDAR elevation dataset for the modeled area. Stormwater infrastructure datasets included the locations of drainage conveyances and nodes such as catch basins, junctions, closed conduits, canals, and curb and gutter. These stormwater infrastructure datasets were verified in the field, to the extent possible. Land use, land cover, and impervious area datasets, which were collected by

Logan and North Logan Cities, were used in calculating catchment characteristics required for the runoff blocks of the SWMM model (Rossman and Huber, 2016).

4.4.2. Catchment Delineations and SWMM Model Design

The SWMM modeling framework includes overland and conduit flow components. Overland flow is defined by subcatchment discretization and parameterization (Rossman and Huber, 2016). Subcatchments are further divided into three subareas: impervious with no depression storage, impervious with depression storage, and pervious area. Common practice is to use map overlays and manually digitize subcatchments based on landscape features such as parking lots, streets, and city parks such that each subcatchment contains primarily only impervious or pervious subareas (Krebs et al., 2013; Sun et al., 2014). In this study, the discretization process was expedited by using a high-resolution digital elevation model (DEM) derived from LiDAR data. LiDAR point data were interpolated to a 0.5 m resolution digital elevation model (DEM) raster using the ArcGIS Topo to Raster geoprocessing tool from ESRI's ArcGIS software. Subcatchment delineations were then performed using ESRI's ArcGIS 10.3.1 for Desktop software using the procedure described below.

The DEM was manipulated using the Arc Hydro toolbox extension for ArcGIS (ESRI, 2017). While the Arc Hydro toolbox was not developed primarily for watersheds that include built infrastructure, others have noted its utility in urbanized watersheds (Johnson, 2008). Arc Hydro tools were used to "burn-in" flow paths and slopes in locations of known curbs, ditches, and other drainage structures. Artificial walls in the DEM were also built using Arc Hydro tools in known locations of subcatchment divides such as

property line fences/walls and other structures. These artificial flow paths and walls serve to represent landscape features that are not well captured in the DEM and that help constrain and improve the accuracy of surface water flow.

The DEM was then hydrologically conditioned to remove non-draining pits, flow directions and accumulations were calculated, and subcatchments were delineated as per the suggestions in the Arc Hydro Tools tutorial (ESRI, 2017). This process required multiple iterations. Subcatchments were delineated and then visually inspected for feasibility and accuracy. Additional "streams" and "walls" were added to the input raster as needed, and the delineation was repeated until the delineation results resembled subcatchments that were observed during field visits to the study area. Initially, all storm drains were used as subcatchment outlets, and flows were then routed through the pipe network. Following this procedure, the first iteration of the model for the NWFC drainage area had a total of 869 subcatchments and 1,105 storm pipes/canal links.

This study aimed to assess the prediction of event hydrographs, requiring a high spatial and temporal resolution model. Thus, we sought to represent the subcatchments and drainage network with as much detail as possible given the input data we had. However, after a series of preliminary simulations using this detailed model, it was apparent that calibration for the chosen simulation period using the calibration procedure described below would not be feasible with the level of detail in our initial delineation because model run times were constrained by a 3-day time limit imposed by the managers of the high-performance computing resources used for this study. Given this, storm pipes were manually merged or removed from the model based on their length and whether they were directly connected to the outfall to the canal. Subcatchments were manually merged if

runoff from one subcatchment drained to another subcatchment with little or no storm pipe flow connecting the two subcatchments. As a result, the final iteration of the model of the NWFC drainage area model had a total of 803 subcatchments and 666 storm pipe/canal links.

Impervious areas, flow width estimates, and subcatchment slopes were calculated based on geospatial data provided by Logan City. The impervious area was treated as a calibration parameter, due to the coarseness of the impervious dataset. Set initially based on the geospatial data provided by the City, the parameter was allowed to vary plus or minus 10 percent during calibration. The flow width estimates were calculated using Equation (4.1) (Rossman and Huber, 2016):

$$W = A/(2 * MC) \tag{4.1}$$

where W represents the flow width, A represents the subcatchment area, and MC represents the main channel length. The flow width value was then allowed to vary plus or minus 30 percent during the calibration procedure due to uncertainties related to representing irregularly shaped subcatchments as rectangles (Guo and Urbonas, 2009; Rossman and Huber, 2016). The subcatchment slopes were calculated by taking the average slope in each subcatchment based on the DEM derived from LiDAR data. All subcatchment geometries, subcatchment slopes, subcatchment impervious areas, junction and outfall features and elevations, and conduit and conduit slopes were then imported into SWMM via a Python script that used ESRI's ArcPy module and the necessary geoprocessing tools from ArcGIS. The NWFC was represented by an open-channel, irregularly shaped conduit. The cross-sectional geometry and slope of the canal were defined by 21 cross sections that were

surveyed along the canal at locations that best captured changes in channel geometry and slope.

Within the SWMM model, each subcatchment requires the specification of a rain gauge to drive the runoff processes. As not all subcatchments in the NWFC model were equipped with a rain gauge, a Thiessen polygon analysis was performed to determine the nearest rain gauge to each subcatchment. Subcatchments were then assigned the observed precipitation values from the nearest rain gauge as driving data for the simulations.

Accurately modeling the NWFC required accounting for all gains and losses to the canal. A series of longitudinal discharge measurements on two different occasions were made to identify potential groundwater gains and losses along the canal. Each longitudinal measurement event involved six discharge measurements, working our way from upstream to downstream. This was performed once at a steady flow rate of ~0.4 m^3/s (about 40 percent capacity), once at a steady flow rate of ~0.5 m^3/s (about 50 percent capacity), and once at a steady flow rate of ~0.7 m^3/s (about 70 percent capacity). These longitudinal measurement events occurred during controlled conditions, which means flow rates were constant and maintained by the canal master, and no diverted water was being added to or removed from the canal. In both instances, while there were reaches of the canal that both gained and lost water, net losses from upstream to downstream were negligible and were determined to be within acceptable uncertainty (8 percent loss, 2.5 percent loss, and 1 percent loss respectively for the three events). For this reason, groundwater gains and losses along the canal were neglected in the model (see Appendix A).

An additional challenge to the water balance of the NWFC were irrigation diversions along the length of the canal. Diversions in this canal are handled in an unofficial

manner by several different water users, where a single water share equates to about 4,320 m3/yr. This associates with an allotted time where a water user can open a diversion headgate and extract water from the canal. Thus, only imprecise estimates of diversions were available due to the human factor of headgate operation and incomplete flow records. That being the case, a constant value of 0.22 m3/s was used as an estimate of diversion flows based on total water shares, which was able to close the daily and monthly flow balance within about two percent. This agreed quite well with the discrepancy in discharge values measured at upstream and downstream ends of the canal during base flow/nonevent conditions (Melcher et al., 2018b).

Three different time periods were simulated in this study, 1) the second half of the 2015 irrigation season (July 26, 2015 - September 17, 2015), 2) Spring 2016 (May 5, 2016 -May 26, 2016), and 3) Fall 2016 (August 5, 2016 – September 22, 2016). These periods correspond with the time periods that each stormwater outfall was monitored (Table 4.3). For our calibration procedure, we ran continuous SWMM simulations for the 3 monitoring periods in Table 4.3. Antecedent soil moisture conditions were accounted for during the continuous simulations via the Green-Ampt method (Rawls et al., 1983). One-dimensional hydraulics were modeled using the dynamic wave model as there were reaches of the canal where gradual and uphill slopes were measured.

4.4.3. Sensitivity Analysis

For this study we retrieved ranges for each of SWMM's important parameters from the literature. These parameters have been investigated extensively across many case studies and have been specified in the literature with a relatively high level of confidence

(Rossman and Huber, 2016; Rossman, 2017; Sun et al., 2013; Tsihrintzis and Hamid, 1998). While we relied on the results of prior studies and the parameter ranges extracted from the literature, we also performed a sensitivity analysis for the NWFC SWMM model to gain a better understanding of which parameters were more important in predicting discharge. Using an exploratory set of simulations, we manually perturbed SWMM's water quantity parameters (i.e., Manning's N for impervious surfaces, Manning's N for pervious surfaces, depression storage for impervious surfaces, depression storage for pervious surfaces, Manning's N for all closed conduits, Manning's N for the canal, subcatchment flow width, saturated hydraulic conductivity, suction head, percent of area routed to outlet, initial soil moisture deficit, and subarea routing configuration) individually and examined the degree to which these parameter changes affected discharge output by the model. These changes were examined by plotting hydrographs and visually inspecting the effects on hydrograph peak values, runoff volumes, and hydrograph shapes. This allowed us to isolate the effects of each parameter on discharge and remove the less sensitive parameters from the calibration procedure by assigning values that closely represent the physical system. Among the less sensitive parameters that were assigned constant values were the percent of area routed to outlet (assigned a value of 40 percent), the initial soil moisture deficit (assigned a value of 0.2), and the selection of impervious to pervious subarea routing. The 9 more sensitive parameters whose values were calibrated are listed in Table 4.2.

4.4.4. Model Calibration and Analysis

Single and multi-objective calibration was used for this study. Because we wanted to evaluate the potential value of each monitoring site in improving the performance of the

model, we calibrated the model multiple times, each time including data from different monitoring sites. We refer to each of the model calibrations as "calibration instances" in the text that follows.

When calibrating hydrologic models with observed data at multiple locations, or multi-response data, many approaches exist in the literature (e.g., Chiang *et al.*, 2014; Lerat *et al.*, 2012; Leta *et al.*, 2017; Li *et al.*, 2010; Shinma and Reis, 2014; Shrestha *et al.*, 2016; Wang *et al.*, 2012). Among those approaches is calibrating a set of parameter values, θ_1, which is applied uniformly across the entire model domain. Then, as data from interior locations are included in the calibration procedure, additional parameter sets: θ_2, θ_3, … θ_m where m is the total number of locations where model calibration is performed, are included in the model (Lerat et al., 2012; Shrestha et al., 2016). This approach, however, can lead to $m * n$ parameters to be calibrated, where m is the number of calibration locations and n is the number of parameters. For this reason, we decided to include n parameters (Table 4.2) for all calibration instances. This procedure maintains a constant number of model parameters between calibration instances (9, see Table 4.2), and the number of objective functions varied between one and two. This allowed us to build a more parsimonious model and make better comparisons among calibration results.

The objective functions selected for this study were the root-mean squared error for discharge at multiple sites ($RMSE_Q$) (Equation 4.2). The $RMSE_Q$ was selected because the minimization of the $RMSE_Q$ value best captured the shape of storm event hydrographs:

$$RMSE_Q = \sqrt{\frac{\sum(Q_o - Q_p)^2}{n_Q}} \tag{4.2}$$

where Q_o and Q_p represent the observed and modeled or predicted discharge (m^3/s) respectively, n_Q represents the number of observation/prediction pairs of discharge. The $RMSE_Q$ was calculated for each calibration instance and simulation period (Table 4.3).

The workflow for the calibration procedure carried out in this study was as follows:

1. Use an evolutionary algorithm (NSGA-II – discussed in more detail below, Deb et al., 2002) to calibrate the model of the entire NWFC drainage area using solely data collected at the downstream canal site (1800 North) to obtain an optimal parameter set. The optimal parameter set was determined to be the parameter set that yielded the minimized root-mean squared error at the downstream canal site. This step represents three calibration instances, one for each time period that was simulated (Table 4.3), which are referred to as the benchmark calibration instances.
2. Recalibrate the NWFC model considering the high-frequency data collected at the downstream canal site as well as at one of the outfall sites. Data from each outfall was used to calibrate a simulation of its corresponding simulation time period. For example, a calibration instance that considered data from the downstream canal site (1800 North) and the 300 North site was simulated for the 2015 time period as this corresponds with the time period in which both sites were actively collecting data. The calibration instances for this step were two-objective, the RMSE at the downstream canal site and the RMSE at the outfall site. The optimal parameter set was then determined to be the parameter set that gave the value closest to origin in the two-objective space based on the

normalized Euclidean distance. This step was repeated for each outfall, resulting in six calibration instances.

3. Calibrate models for the drainage areas of each of the six outfall sites, minimizing the RMSE at each outfall site (six different calibrated subcatchment models). Use these calibrated models as boundary conditions to the model of the entire NWFC drainage area and calibrate the rest of the drainage area, using the RMSE at the model outlet as a single objective function. This step represents six calibrated subcatchment models (one for each outfall model), which were used as boundary conditions to three calibration instances of the NWFC model – one for each simulation time period (Table 4.3).
4. Compare results of each of the six canal site plus single outfall calibration instances with the results of the calibration instance that only used data from the downstream canal site.

For the purposes of discussion in this paper, the calibration instances generated will be referred by the following nomenclature. The models calibrated solely using data from the downstream canal site, or at the model outlet, are referred to as the NWFC_2015, NWFC_spr2016, and NWFC_fal2016 models and serve as benchmark calibration instances for this study. The model calibrated using data from the downstream canal site and data from the outfall located at the street corner of 800 North is referred to as NWFC_800. This naming scheme applies to the five other calibration instances that include one of the six monitored outfalls (i.e., NWFC_300, NWFC_1000, NWFC_1250, NWFC_1300, NWFC_1400) where the numbers represent the street address at which the outfall is located. Finally, the calibration instances that were calibrated based on data

collected at the downstream monitoring site and use the output from the calibrated subcatchment models as boundary conditions (step 3 above) are referred to as NWFC_ALL_2015, NWFC_ALL_spr2016, and NWFC_ALL_fal2016. It was also determined that these calibration instances would allow us to evaluate the effects of each outfall dataset on the overall performance of the model and whether these effects were unique to a specific outfall.

4.4.5. SWMM Parameterization and Calibration

The calibration procedure that attempts to minimize the $RMSE_Q$ at multiple locations within a SWMM model needs the ability to find a set of solutions that are equally optimal with respect to objective function values. This set of solutions make up the Pareto front (Neilson et al., 2010; Vrugt et al., 2003). Multi-objective evolutionary algorithms (MOEA) have been found to converge on the Pareto front without getting "stuck" in suboptimal solutions (Deb et al., 2002). The Nondominated Sorting Genetic Algorithm (NSGA-II) (Deb et al., 2002) was selected as the MOEA for the calibration procedure in this study. NSGA-II was selected because it has been found to be effective at quickly converging at the Pareto front, thus requiring fewer SWMM simulations to converge at an optimal set of solutions.

As inputs, the NSGA-II algorithm receives a specified population size, number of generations, the probability of crossover, and the probability of mutation. The population size represents the number of SWMM parameter sets simulated at each generation or iteration of the calibration instance. The number of generations represents the number of

iterations in the calibration instance. The probability of crossover is a number from 0 to 1 that represents the probability that a solution will swap some of its SWMM parameter values with another solution for the next iteration of the calibration instance. The probability of mutation is a number from 0 to 1 that represents the probability that some of the SWMM parameter values of a solution will be allowed to vary slightly for the next iteration of the calibration instance. Both crossover and mutation are included to ensure that the calibration procedure finds globally optimal solutions.

As this study was focused on examining the modeling benefits of using multi-response data from a WSN rather than on examining the sensitivities of the genetic algorithm inputs, the default value of 0.7 for crossover probability was used. The mutation probability used was calculated to be 1/L where L is the number of parameters to be optimized in the calibration instance (Deb et al., 2002; Mala-Jetmarova et al., 2015). In our case, this value was calculated to be 1/9 = 0.111. Due to the cost of simulation time, a population size of 100 and 100 generations were used. This resulted in a total of 10,100 total model simulations for each calibration instance (100 by 100 simulations plus the initial 100 parent population simulations). The NSGA-II algorithm included in the MCO R package (Mersmann, 2014) and an R script that was originally written by Peter Steinberg for Herrera Environmental Consultants (Steinberg, 2014) were modified such that the NSGA-II algorithm could be parallelized for execution on a high performance computer.

Readers are referred to Deb et al. (2002) for a more in-depth description of the NSGA-II algorithm. Calibration instances were run in parallel on a high-performance computing resource consisting of multiple 2.8 GHz Intel Xeon processors. A flow chart of the calibration procedure is given in Figure 4.2.

4.4.6. Analysis of Calibration Results

The result of each calibration instance was a solution that minimizes the objective function value in the case of a single-objective calibration or a solution along the Pareto front that minimizes the normalized Euclidean distance to origin in the case of a two-objective calibration. In order to compare the quality of each calibration instance and make comparisons among instances, the difference in $RMSE_Q$ (Equation 4.2) values between calibration instances was calculated and compared to the corresponding benchmark calibration instance. The benchmark calibration instances are the instances that used solely the data from the model outlet to calibrate the model (i.e., NWFC_2015, NWFC_spr2016, and NWFC_fal2016). The other calibration instances considered were the instances with two calibration locations (e.g., NWFC_300, NWFC_800, etc.), and the calibration instances using all outfall models as boundary conditions to the model of the entire drainage area (i.e.,NWFC_ALL_2015, NWFC_ALL_spr2016, and NWFC_ALL_fal2016).

4.5. Results

4.5.1 Calibration Results at the Model Outlet

Assessing the differences in discharge values at the model outlet tells the story of how each calibration instance predicted the volume and timing of water leaving the NWFC drainage system. For the event on August 7, 2015, each of the calibration instances provided similar results (Figure 4.3a). Each instance provided an accurate prediction of the peak discharge; however, they each underestimated the base flow conditions prior and

subsequent to the event. This underestimation of pre and post base flow conditions is likely due to uncertainty in the operation of individual diversion headgates along the canal. For the event on September 16, 2015, each calibration instance overestimated the peak discharge by ~0.5 – 0.7 m^3/s (Figure 4.3b). While the shape of the predicted hydrographs is somewhat consistent, each of the four calibration instances estimated a peak discharge of nearly double the observed peak. The NWFC_300N calibration instance predicted the highest discharge values of the four calibration instances. Although the NWFC_300 calibration instance yielded the greatest $RMSE_Q$ value at the model outlet, the difference in $RMSE_Q$ between the NWFC_300 and the benchmark calibration instances was ~ 0.015 m^3/s or about one percent of the canal's total capacity (Figure 4.6a). The $RMSE_Q$ values for the NWFC_1250 and NWFC_ALL_2015 were more similar to the benchmark value with differences in $RMSE_Q$ values of 0.0009 and 0.0094 m^3/s respectively.

The Spring 2016 simulation period included results from the NWFC_spr2016 (benchmark), NWFC_800, NWFC_1300, and NWFC_ALL_spr2016 calibration instances. The NWFC_1300 calibration instance most greatly overestimates both observed hydrographs. This is most obvious in the May 19, 2016 storm event (Figure 4.4a), where the overestimation of the peak discharge was about 0.7 m^3/s. The other three calibration instances produced similar results, with the NWFC_spr2016 (benchmark) instance demonstrating slightly improved peak discharge predictions. The overestimation is less obvious in for the May 21, 2016 storm (Figure 4.4b). In both panels (Figure 4.4a and 4.4b), it appears as though the hydrograph peaks are best captured by the benchmark instance (NWFC_spr2016). These results are confirmed by Figure 4.6b, which shows similar

$RMSE_Q$ values for the Spring 2016 simulation period for all calibration instances, with the lowest value for the benchmark instance and the highest value for the NWFC_1300 instance. The difference in $RMSE_Q$ values between the NWFC_1300 and benchmark calibration instances was ~0.03 m^3/s or about 3 percent of canal capacity. The overestimation from the NWFC_1300 calibration instance may be explained by uncertainty in the observations made at the 1300 North outfall. This outfall pipe was prone to submerged/backflow conditions which added noise to the discharge data measured by the area-velocity flow module (see Melcher et al., 2018b).

The NWFC_fal2016 (benchmark), NWFC_1000, NWFC_1400, and the NWFC_ALL_fal2016 calibration instances were calibrated for the Fall 2016 simulation period. Each calibration instance yielded similar results for a series of storm events that occurred between September 12 and September 15, 2016 (Figure 4.5a). The general shape of the hydrograph was captured by all instances, with the NWFC_1000 instance overestimating some of the peaks. All instances slightly underestimated the peak near September 14 18:00. Again, the NWFC_1000 overestimated the peak discharge of the storm event that occurred on September 21, 2016 while the benchmark and NWFC_ALL_fal2016 instances yielded quite similar results (Figure 4.5b). The $RMSE_Q$ values for each calibration instance of the Fall 2016 simulation period were quite similar (Figure 4.6c)..

4.5.2. Calibration Results at Outfalls

Comparing simulated discharge values at each of the monitored outfalls to the observed values helps in determining how well each calibration instance was able to

represent the hydrology of specific sections of the NWFC drainage system. It is evident that the NWFC_300 and NWFC_ALL_2015 calibration instances yielded the results that best captured the hydrograph peaks at 300 North for the event on September 16, 2015, whereas the NWFC_2015 (benchmark) instance greatly underestimated the peak (Figure 4.7a). This is to be expected as both the NWFC_300 instance and the NWFC_ALL_2015 instance used data observed at the 300 North outfall to calibrate the model.

The observed hydrograph included peaks and valleys that were not represented by any of the calibration instances. This discrepancy may be, in part, due to uncertainty in rain gauge data. While precipitation data were quality controlled to ensure accuracy, often precipitation measurements at the outfall site, which were used as input for some of the modeled subcatchments, were not fully representative of precipitation patterns within other subcatchments. This may explain why the simulations for outfalls were unable to fully capture the observed hydrograph shape. The improvement in the simulated discharge values for the 300 North outfall can be seen by looking at the $RMSE_Q$ value at 300 North for the NWFC_2015, NWFC_300, NWFC_1250, and NWFC_ALL_2015 calibration instances (Figure 4.10a). It is evident that the NWFC_300 and NWFC_ALL_2015 instances provided better $RMSE_Q$ values than the benchmark and NWFC_1250 instances. The amount of improvement was about 36 and 44 percent for the NWFC_300 and NWFC_ALL_2015 instances respectively, relative to the benchmark instance.

Each calibration instance overestimated the observed hydrograph for the same event at the 1250 North outfall (Figure 4.7b). The NWFC_ALL_2015 calibration instance yielded the lowest $RMSE_Q$ value at 1250 North (Figure 4.10b); however, the improvement

in the $RMSE_Q$ value between the NWFC_ALL_2015 and benchmark calibration instances is not as pronounced (~0.000149 m^3/s). The calibrated parameter values for the benchmark calibration instance and the 1250N subcatchment model were quite similar (Table 4.4).

The NWFC_ALL_spr2016 calibration instance best captures the hydrograph peaks for the May 19, 2016 storm event at 800 North (Figure 4.8a); The other three instances (i.e., NWFC_800, NWFC_1300, and NWFC_spr2016 (benchmark)), however, were able to capture the overall shape of the observed hydrograph. These results are confirmed by looking at the $RMSE_Q$ values for the benchmark and NWFC_ALL_spr2016 instances (Figure 4.11a). The difference between the two $RMSE_Q$ is ~0.0064 m^3/s or about 50 percent. The NWFC_1300 and NWFC_ALL_spr2016 instances performed similarly for the May 19, 2016 storm event at 1300 North. These results are also confirmed by the similar $RMSE_Q$ values (Figure 4.11b).

At 1000 North, both the NWFC_1400 and benchmark calibration instances underestimated the hydrograph for the storm event on September 21, 2016 (Figure 4.9a). This is especially evident at the peak of the hydrograph, which occurred at about September 21 20:30. Again, the calibration instances that best captured hydrograph peaks were the instances that used data from the 1000 North outfall to calibrate the model (i.e., NWFC_1000, NWFC_ALL_fal2016). These results are confirmed by the $RMSE_Q$ values (Figure 4.12a). The difference between the NWFC_ALL_fal2016 and benchmark instances is ~0.00326 m^3/s or about 25 percent. At the 1400 North outfall, the NWFC_1000 instance appears to greatly overestimate the September 21, 2016 hydrograph peak, while the other instances perform similarly (Figure 4.9b). Additionally, the

NWFC_ALL_fal2016 instance reduced the $RMSE_Q$ value from the benchmark instance by about 0.0027 m^3/s or 16 percent (Figure 4.12b).

4.6. Discussion

Assessing the calibration results, the benchmark calibration instances, which used only data from the 1800 North outlet point for calibration, minimized the $RMSE_Q$ at the 1800 North outlet point for each simulation period. However, with the exception of the NWFC_1300 calibration instance, the other calibration instances produced similar results at the 1800 North outlet point in terms of capturing hydrograph peaks, shapes, and $RMSE_Q$ values. This is similar to the results reported by Lerat *et al.* (2012), which also demonstrated only minimal improvement in the simulation predictions at the model outlet when including interior, nested calibration points during the calibration procedure.

Results at the individual stormwater outfalls (Figures 4.7 – 4.12) showed that the use of additional calibration points yielded better $RMSE_Q$ values for the simulated flows at each of these interior points while still maintaining similar performance at the overall model outlet (1800 North). Again this is similar to the results of Lerat et al. (2012). However, Lerat et al. (2012) state that these results are intuitive and have very little significance in justifying the use of additional data sets to calibrate the model at interior locations. This may be true when the only objective is to examine flows at the most downstream outlet point. An example might be watershed modeling and management for determining how changes to landscapes and management procedures affect runoff volumes and hydrograph peaks at a watershed outlet. Effective urban stormwater management,

however, requires not only the ability to quantify runoff volumes leaving the drainage system, but also the runoff volumes and pollutant loads from sections of the larger drainage (i.e., at the neighborhood scale). Our results showed that the benchmark calibration instances did not do well at predicting flows from outfalls internal to the model. Thus, where modeling objectives include accurately predicting flow at the model outlet and at points internal to the model, inclusion of additional outfall monitoring stations in the calibration can improve results.

Additional insight may be gained through an analysis of the optimized parameter values for each calibration instance. As an example, the values of each of the nine calibrated parameters for the NWFC_800, NWFC_spr2016 (benchmark), and NWFC_ALL_spr2016 calibration instances are given in Figure 4.13 along with the parameter values for the calibrated 800 North subcatchment model. Panel a of the figure gives the values of one of the more sensitive SWMM parameters, N_{Imp}. Upon observation, the N_{Imp} values for the NWFC_800 calibration instance and the 800N subcatchment model are similar. In fact, for most SWMM parameters shown in Figure 4.13, the NWFC_800 calibration instance is more similar to the 800N subcatchment model than to the benchmark instance. This demonstrates that, in the compromise between the two objective functions (i.e., $RMSE_Q$ at the model outlet and $RMSE_Q$ at the 800 North outfall), the $RMSE_Q$ at 800 North had more influence on model calibration. Similar results can be seen with other sensitive parameters (e.g., S_{Imp}, S_{Perv}, N_{Cond}, and FW). This may imply that some parameters are more sensitive at different spatial scales or with different land uses and covers, or that parameters such as N_{Imp} and N_{Cond} are more sensitive at the subcatchment

model scale than at the scale of the entire NWFC drainage area. Krebs *et al.* (2014) also found that some SWMM parameters were sensitive to spatial scale and land cover, supporting our findings. The N_{Perv}, $N_{Channel}$, k_s, and ψ SWMM parameters differed between the NWFC_800 and 800N calibration instances. As the $N_{Channel}$ parameter was not calibrated in the 800N instance (the NWFC canal is not included in the subcatchment models), the higher $N_{Channel}$ value in the NWFC_800 instance than in the benchmark instance may be compensating for the lower N_{Imp} and N_{Perv} values.

Additionally, a comparison of the parameter values for the benchmark instance and the NWFC_ALL_spr2016 instance is given in Figure 4.13. In this case, the parameter values for the NWFC_ALL_spr2016 calibration instance are the parameter values for the portion of the model domain that received no boundary conditions from the outfall models. This ends up being about 60 percent of the total drainage area. The N_{Perv}, S_{Imp}, N_{Cond}, k_s, and ψ parameter values show differences between the NWFC_ALL_spr2016 and the benchmark calibration instance, and differences in parameter values were variable. Relative to the benchmark calibration instance, the N_{Perv} and ψ parameter values for the NWFC_ALL_spr2016 calibration instance were lower and would cause greater discharge values at the model outlet. Likewise, relative to the benchmark calibration instance, the S_{Imp}, N_{Cond}, and k_s parameter values for the NWFC_ALL_spr2016 calibration instance were higher and would cause reduced discharge values. The values of the five above mentioned calibrated parameters (N_{Perv}, S_{Imp}, N_{Cond}, k_s, and ψ) for the individual subcatchment models do not appear to follow any recognizable pattern relative to the benchmark values. This implies some of the equifinality of the SWMM parameters, but

also that similar discharge values at the model outlet can be obtained without much influence by the calibrated subcatchments.

Two of the underlying objectives of this study were to determine the value of monitoring efforts in the NWFC for improving a detailed stormwater model and to assess whether the modeling results could provide information for planning and executing future monitoring efforts in urban catchments. We observed similar simulated discharge values at the model outlet, regardless of whether we included interior monitoring/calibration points in the calibration and regardless of which interior points we included. This was the case in both the two-site calibration instances (NWFC_300, NWFC_800, NWFC_1000, NWFC_1250, NWFC_1300, and NWFC_1400) and the calibration instances that used all subcatchment models as boundary conditions (NWFC_ALL_2015, NWFC_ALL_spr2016, and NWFC_ALL_fal2016). In other words, when the only objective was to match flows at the model outlet, the choice of which interior sites at which to monitor/calibrate or whether to monitor at these sites at all mattered little.

The benefits of adding additional monitoring sites were only realized when examining and predicting flows at both the model outlet and at the interior calibration points. For studies that require quantifying outflows from the modeled domain as well as the relative contributions of individual subcatchments (e.g., where stormwater BMPs are being considered within specific subcatchments), collecting data at points interior to the model may be critical in ensuring good results throughout the modeled domain. In these cases, monitoring efforts should be guided by the judgement of stormwater managers and engineers. For example, additional data should be collected for subcatchments being investigated for stormwater BMPs, for subcatchments where flooding has occurred, or

where the subcatchment is the location of a potential contaminant source. Additionally, modelers may perform preliminary simulations of the model domain to identify important subcatchments (i.e., those that contribute large portions of the domain outflow) before initiating a data collection program. With data interior to the modeled domain, multiple scenarios can more accurately be modeled and effects at subcatchment outfalls and at the outlet of the modeled system could be assessed.

4.7. Summary and Conclusions

High resolution monitoring data of the type we used in this study provide a wealth of information about the behavior of urban water systems, but they present a challenge for use in current urban stormwater models. We sought to determine the degree to which inclusion of data from multiple monitoring sites internal to the modeled domain would improve the model's ability to predict stormwater discharge as a way of investigating the value of these monitoring sites in supporting modeling. This question is important because it is cost prohibitive to monitor continuously at every point within an urban stormwater system where discharge information may be needed. Stormwater models like SWMM represent an important tool for water managers to use in overcoming data shortcomings, understanding overall system behavior, assessing the effects of BMPs, and simulating management scenarios. However, effectively using models requires incorporation of data, and the process for using high resolution sensor datasets for calibrating urban stormwater models is not straightforward.

We investigated this problem in the context of a suite of model calibration instances representing multiple configurations of a SWMM model calibrated using data from one or more of our continuous monitoring sites. While most calibration instances considered in this study produced similar results in terms of predicting discharge at the model outlet (1800 North), results at the individual outfall sites showed much more variability. The calibration instances that used data from the model outlet and from one of the six monitored outfall sites tended to improve results at the outfall sites, showing that multiple parameter sets could yield similar results at the model outlet while producing much different results at points interior to the model (i.e., discharge from the individual outfalls).

Based on these results, water managers and modelers may consider the tradeoffs and costs associated with including additional monitoring sites for the purpose of improving modeling accuracy and precision. Where detailed knowledge of discharge from individual outfalls is needed for tracking runoff and drainage through the system, the availability of high resolution data at multiple locations within an urban water system (including outfalls and boundary conditions) is a tremendous asset. Where results are only needed at a downstream outlet point, the cost of installing and maintaining monitoring sites, along with processing/post-processing all the data may not be worth it. Finally, the simulation time associated with the calibration procedure described in this study was not trivial. Although access to HPC resources is increasing, many practicing engineers/hydrologists may not have access to these resources with the ability to run calibration instances in parallel across multiple computing nodes.

4.8. Acknowledgments

This work was supported by funding from the Utah Water Research Laboratory at Utah State University and from United States National Science Foundation award IIA-1208732. Any opinions, findings, and conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation. The support and resources from the Center for High Performance Computing at the University of Utah are gratefully acknowledged.

References

Chiang, L.-C., Y. Yuan, M. Mehaffey, M. Jackson, and I. Chaubey, 2014. Assessing SWAT's Performance in the Kaskaskia River Watershed as Influenced by the Number of Calibration Stations Used. Hydrological Processes 28:676–687.

Corke, P., T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, 2010. Environmental Wireless Sensor Networks. Proceedings of the IEEE., pp. 1903–1917.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan, 2002. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6:182–197.

Deletic, A.B. and C.T. Maksimovic, 1998. Evaluation of Water Quality Factors in Storm Runoff from Paved Areas. Journal of Environmental Engineering 124:869–879.

ESRI, 2017. Arc Hydro Tools V10.3.

Gong, Y., X. Liang, X. Li, J. Li, X. Fang, and R. Song, 2016. Influence of Rainfall Characteristics on Total Suspended Solids in Urban Runoff: A Case Study in Beijing, China. Water 8:23.

Guo, J.C. and B. Urbonas, 2009. Conversion of Natural Watershed to Kinematic Wave Cascading Plane. Journal of Hydrologic Engineering 14:839–846.

Horsburgh, J.S., A. Spackman Jones, D.K. Stevens, D.G. Tarboton, and N.O. Mesner, 2010. A Sensor Network for High Frequency Estimation of Water Quality Constituent Fluxes Using Surrogates. Environmental Modelling & Software 25:1031–1044.

Johnson, L.E., 2008. Geographic Information Systems in Water Resources Engineering.

Kerkez, B., C. Gruden, M. Lewis, L. Montestruque, M. Quigley, B. Wong, A. Bedig, R. Kertesz, T. Braun, O. Cadwalader, A. Poresky, and C. Pak, 2016. Smarter Stormwater Systems. Environmental Science & Technology 50:7267–7273.

Kirchner, J.W., 2006. Getting the Right Answers for the Right Reasons: Linking Measurements, Analyses, and Models to Advance the Science of Hydrology. Water Resources Research 42. doi:10.1029/2005WR004362.

Kirchner, J.W., X. Feng, C. Neal, and A.J. Robson, 2004. The Fine Structure of Water-Quality Dynamics: The (High-Frequency) Wave of the Future. Hydrological Processes 18:1353–1359.

Krebs, G., T. Kokkonen, M. Valtanen, H. Koivusalo, and H. Setälä, 2013. A High

Resolution Application of a Stormwater Management Model (SWMM) Using Genetic Parameter Optimization. Urban Water Journal 10:394–410.

Krebs, G., T. Kokkonen, M. Valtanen, H. Setälä, and H. Koivusalo, 2014. Spatial Resolution Considerations for Urban Hydrological Modelling. Journal of Hydrology 512:482–497.

Lerat, J., V. Andréassian, C. Perrin, J. Vaze, J.M. Perraud, P. Ribstein, and C. Loumagne, 2012. Do Internal Flow Measurements Improve the Calibration of Rainfall-Runoff Models? Water Resources Research 48:W02511.

Leta, O.T., A. van Griensven, and W. Bauwens, 2017. Effect of Single and Multisite Calibration Techniques on the Parameter Estimation, Performance, and Output of a SWAT Model of a Spatially Heterogeneous Catchment. Journal of Hydrologic Engineering 22:05016036.

Levesque, V.A. and K.A. Oberg, 2012. Computing Discharge Using the Index Velocity Method. Reston.

Li, X., D.E. Weller, and T.E. Jordan, 2010. Watershed Model Calibration Using Multi-Objective Optimization and Multi-Site Averaging. Journal of Hydrology 380:277–288.

Mala-Jetmarova, H., A. Barton, and A. Bagirov, 2015. Sensitivity of Algorithm Parameters and Objective Function Scaling in Multi-Objective Optimisation of Water Distribution Systems. Journal of Hydroinformatics 17. http://iwaponline.com/content/17/6/891. Accessed 4 Sep 2017.

Melcher, A.A. and J.S. Horsburgh, 2017. An Urban Observatory for Quantifying Phosphorus and Suspended Solid Loads in Combined Natural and Stormwater Conveyances. Environmental Monitoring and Assessment 189:285.

Melcher, A.A., J.S. Horsburgh, and B.A. Mihalevich, 2018a. Discharge Rating Curve at 1800 North and 200 West in the Northwest Field Canal in Logan, UT, USA. Hydroshare. http://www.hydroshare.org/resource/26aff3ccb4a7476fb191a0183cf98c46.

Melcher, A.A., J.S. Horsburgh, and B.A. Mihalevich, 2018b. Continuous In Situ Monitoring Data from the Northwest Field Canal in Logan, UT, USA. Hydroshare. http://www.hydroshare.org/resource/fe05a4877eba46c182ea5ddc26bb140e.

Mersmann, O., 2014. Multiple Criteria Optimization Algorithms and Related Functions.

Mroczkowski, M., P.G. Raper, and G. Kuczera, 1997. The Quest for More Powerful Validation of Conceptual Catchment Models. Water Resources Research 33:2325–2335.

National Research Council, 2009. Urban Stormwater Management in the United States.

Neilson, B.T., S.C. Chapra, D.K. Stevens, and C. Bandaragoda, 2010. Two-Zone Transient Storage Modeling Using Temperature and Solute Data with Multiobjective Calibration: 1. Temperature. Water Resources Research 46. doi:10.1029/2009WR008756.

Niazi, M., C. Nietch, M. Maghrebi, N. Jackson, B.R. Bennett, M. Tryby, and A. Massoudieh, 2017. Storm Water Management Model: Performance Review and Gap Analysis. Journal of Sustainable Water in the Built Environment:1–32.

Obropta, C.C. and J.S. Kardos, 2007. Review of Urban Stormwater Quality Models: Deterministic, Stochastic, and Hybrid Approaches. JAWRA Journal of the American Water Resources Association 43:1508–1523.

Petrucci, G. and C. Bonhomme, 2014. The Dilemma of Spatial Representation for Urban Hydrology Semi-Distributed Modelling: Trade-Offs Among Complexity, Calibration and Geographical Data. Journal of Hydrology 517:997–1007.

Rantz, S.E., 1982. Measurement and Computation of Streamflow: Volume 1. Measurement of Stage and Discharge. USGS, p. 284.

Rawls, W.J., D.L. Brakensiek, and N. Miller, 1983. Green-ampt Infiltration Parameters from Soils Data. Journal of Hydraulic Engineering 109:62–70.

Rode, M., A.J. Wade, M.J. Cohen, R.T. Hensley, M.J. Bowes, J.W. Kirchner, G.B. Arhonditsis, P. Jordan, B. Kronvang, S.J. Halliday, R.A. Skeffington, J.C. Rozemeijer, A.H. Aubert, K. Rinke, and S. Jomaa, 2016. Sensors in the Stream: The High-Frequency Wave of the Present. Environmental Science & Technology 50:10297–10307.

Rossman, L.A., 2017. Storm Water Management Model Reference Manual Volume II – Hydraulics. U.S. Environmental Protection Agency.

Rossman, L.A. and W.C. Huber, 2016. Storm Water Management Model Reference Manual Volume I – Hydrology. www2.epa.gov/water-research.

Shinma, T.A. and L.F.R. Reis, 2014. Incorporating Multi-Event and Multi-Site Data in the Calibration of SWMM. Procedia Engineering 70:75–84.

Shrestha, M.K., F. Recknagel, J. Frizenschaf, and W. Meyer, 2016. Assessing SWAT Models Based on Single and Multi-Site Calibration for the Simulation of Flow and Nutrient Loads in the Semi-Arid Onkaparinga Catchment in South Australia. Agricultural Water Management 175:61–71.

Steinberg, P.D., 2014. RSWMM. https://github.com/PeterDSteinberg/RSWMM.

Sun, N., M. Hall, B. Hong, and L. Zhang, 2014. Impact of SWMM Catchment Discretization: Case Study in Syracuse, New York. Journal of Hydrologic Engineering 19:223–234.

Sun, N., B. Hong, and M. Hall, 2013. Assessment of the SWMM Model Uncertainties Within the Generalized Likelihood Uncertainty Estimation (GLUE) Framework for a High-Resolution Urban Sewershed. Hydrological Processes 28:3018–3034.

Tomlinson, M.S. and E.H. De Carlo, 2003. The Need for High Resolution Time Series Data to Characterize Hawaiian Streams. Journa of the American Water Resources Association 39:113–123.

Tsihrintzis, V.A. and R. Hamid, 1997. Modeling and Management of Urban Stormwater Runoff Quality: A Review. Water Resources Management 11:137–164.

Tsihrintzis, V.A. and R. Hamid, 1998. Runoff Quality Prediction from Small Urban Catchments Using SWMM. Hydrological Processes 12:311–329.

Vrugt, J.A., H. V. Gupta, L.A. Bastidas, W. Bouten, and S. Sorooshian, 2003. Effective and Efficient Algorithm for Multiobjective Optimization of Hydrologic Models. Water Resources Research 39. doi:10.1029/2002WR001746.

Wang, S., Z. Zhang, G. Sun, P. Strauss, J. Guo, Y. Tang, and A. Yao, 2012. Multi-Site Calibration, Validation, and Sensitivity Analysis of the MIKE SHE Model for a Large Watershed in Northern China. Hydrology and Earth System Sciences; Katlenburg-Lindau 16:4621.

Wanielista, M.P. and Y.A. Yousef, 1993. Stormwater Management. John Wiley and Sons, Inc.

Waschbusch, B.R.J., W.R. Selbig, and R.T. Bannerman, 1999. Sources of Phosphorus in Stormwater and Street Dirt from Two Urban Residential Basins In Madison , Wisconsin , 1994 – 95. Middleton.

Wong, B.P. and B. Kerkez, 2016. Adaptive Measurements of Urban Runoff Quality. Water Resources Research. doi:10.1002/2015WR018013.

Table 4.1. Monitoring sites, sampling/monitoring coverage, and their catchment characteristics.

Outlet Location	Total Area (ha)	Imperviousness (%)	Average Slope (%)	Residential (%)	Commercial (%)	Street (%)
300 North	4.25	64.5	4.1	30.6	29.1	26.6
1250 North	19.8	74.4	3.7	19.0	64.8	14.6
800 North	35.3	52.2	3.9	32.1	25.5	17.7
1000 North	53.6	53.0	3.7	21.4	44.1	12.8
1300 North	5.57	87.1	4.0	0.00	75.8	5.27
1400 North	39.1	80.9	3.7	8.48	56.5	11.9
NWFC Drainage Area	389	60.7	4.0	21.6	39.2	14.8

Table 4.2. Calibration parameters used for the NSGA-II calibration procedure.

Parameter Name	Symbol	Modeling Component	Minimum Bound	Maximum Bound	Source
Manning's N for impervious surfaces	N_{imp}	Hydrologic	0.01	0.035	(Rossman and Huber, 2016)
Manning's N for pervious surfaces	N_{perv}	Hydrologic	0.02	0.15	(Rossman and Huber, 2016)
Depression Storage for impervious surfaces	S_{imp}	Hydrologic	0.1	12	(Rossman and Huber, 2016)
Depression storage for pervious surfaces	S_{perv}	Hydrologic	2	40	(Rossman and Huber, 2016)
Manning's N of conduit	N_{cond}	Hydraulic	0.01	0.03	(Rossman and Huber, 2016), corrugated materials
Manning's N of canal	$N_{channel}$	Hydraulic	0.02	0.08	(Rossman and Huber, 2016), weedy reaches
Flow width	FW	Hydrologic	-30 percent	+30 percent	Estimated based on geospatial data
Green-Ampt saturated hydraulic conductivity	k_s	Infiltration	1	8	(Rossman and Huber, 2016), silty clay loam/silt loam
Green-Ampt suction head	ψ	Infiltration	13	1315	(Rossman and Huber, 2016), silty clay loam

Table 4.3. Description of the time periods monitored at each site.

Monitoring Location	Time Periods Monitored		
	2015 (July 26-September 17)	Spring 2016 (May 5, 2016 -May 26, 2016)	Fall 2016 (August 5, 2016 – September 22, 2016)
200 South (Upstream Canal Site)	X	X	X
300 North	X		
1250 North	X		
800 North		X	
1300 North		X	
1000 North			X
1400 North			X
1800 North (Downstream Canal Site)	X	X	X

Table 4.4. Calibrated SWMM parameter values for each calibration instance.

Calibration Instances	Calibration Parameters								
	N_{imp}	N_{perv}	S_{imp}	S_{perv}	N_{cond}	$N_{channel}$	FW	k_s	ψ
NWFC_2015	0.0350	0.1495	11.99	39.94	0.0111	0.0276	-29.94	7.999	778.5
NWFC_spr2016	0.0350	0.1471	0.1065	39.99	0.0100	0.0239	-29.97	4.346	1289
NWFC_fal2016	0.0109	0.1490	2.301	23.97	0.0208	0.0341	-27.96	2.811	760.1
NWFC_300	0.0138	0.0607	11.52	2.252	0.0101	0.0795	30.00	1.969	1273
NWFC_1250	0.0348	0.1496	11.96	39.80	0.0101	0.0263	-29.87	3.315	1312
NWFC_800	0.0100	0.0884	0.1333	6.219	0.0100	0.0479	29.64	5.378	1298
NWFC_1300	0.0192	0.1456	0.1012	10.47	0.0100	0.0333	-28.06	1.006	75.42
NWFC_1000	0.0100	0.1488	0.5541	39.89	0.0100	0.0367	29.26	1.042	88.20
NWFC_1400	0.0100	0.1498	0.1880	39.60	0.0214	0.0551	29.95	7.997	1311
NWFC_ALL_2015	0.0350	0.1500	12.00	39.99	0.0102	0.0759	-29.83	7.997	1275
NWFC_ALL_spr2016	0.0346	0.0582	3.527	39.99	0.0228	0.0299	-29.82	7.883	847.2
NWFC_ALL_fal2016	0.0100	0.1499	6.002	39.57	0.0100	0.0429	28.76	2.721	1227
300N	0.0100	0.0426	1.491	2.002	0.0113	0.0213	29.83	1.912	1298
1250N	0.0350	0.1451	12.00	40.00	0.0100	0.0646	-29.99	3.321	1315
800N	0.0100	0.0200	0.1035	2.877	0.0100	0.0655	30.00	1.687	203.0
1300N	0.0200	0.1494	0.1121	8.402	0.0100	0.0456	-29.43	1.000	123.0
1000N	0.0100	0.1497	1.808	15.76	0.0100	0.0799	27.53	1.028	671.3
1400N	0.0100	0.1479	0.1153	40.00	0.0211	0.0668	29.90	7.998	1313

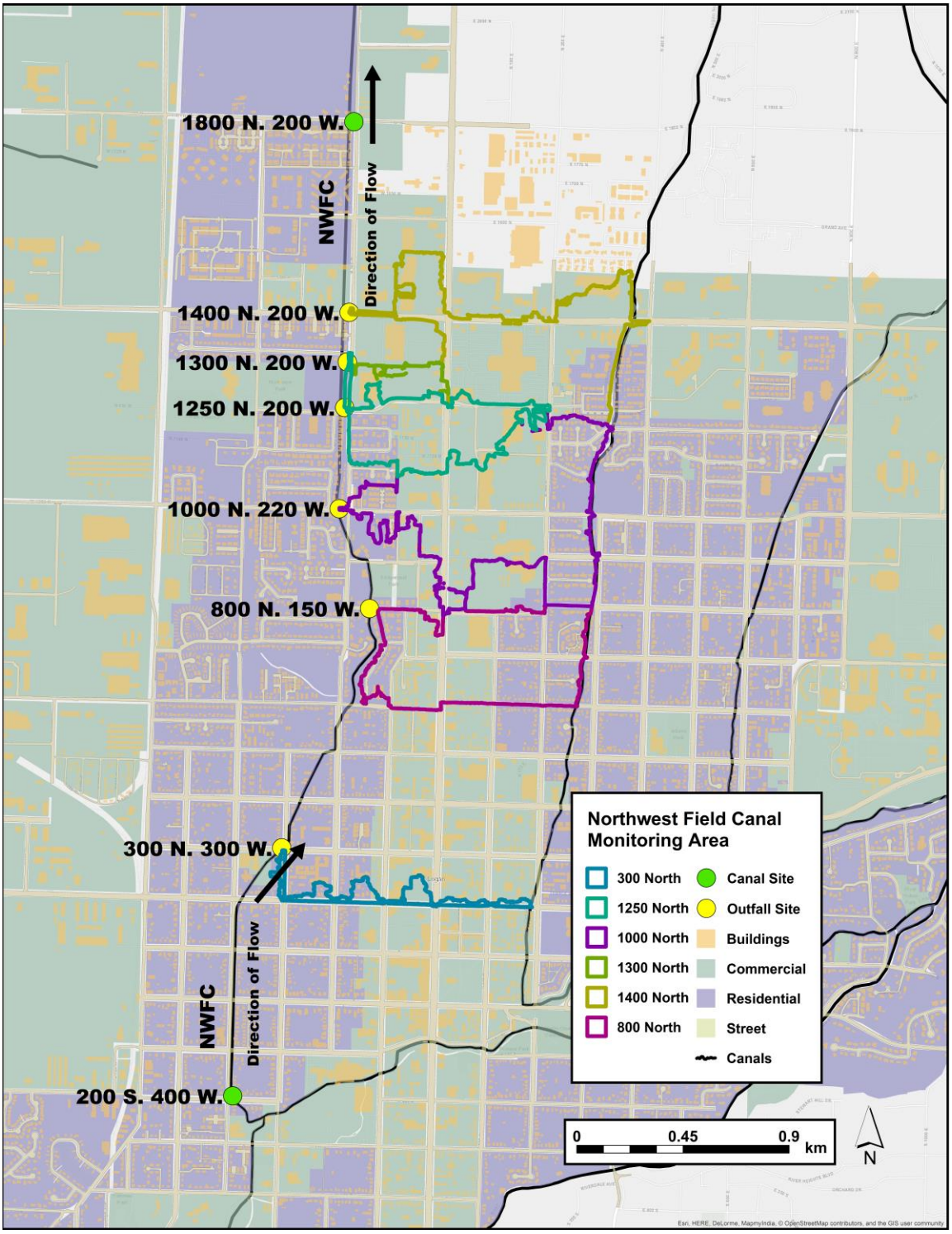


Figure 4.1. Northwest Field Canal Study Area

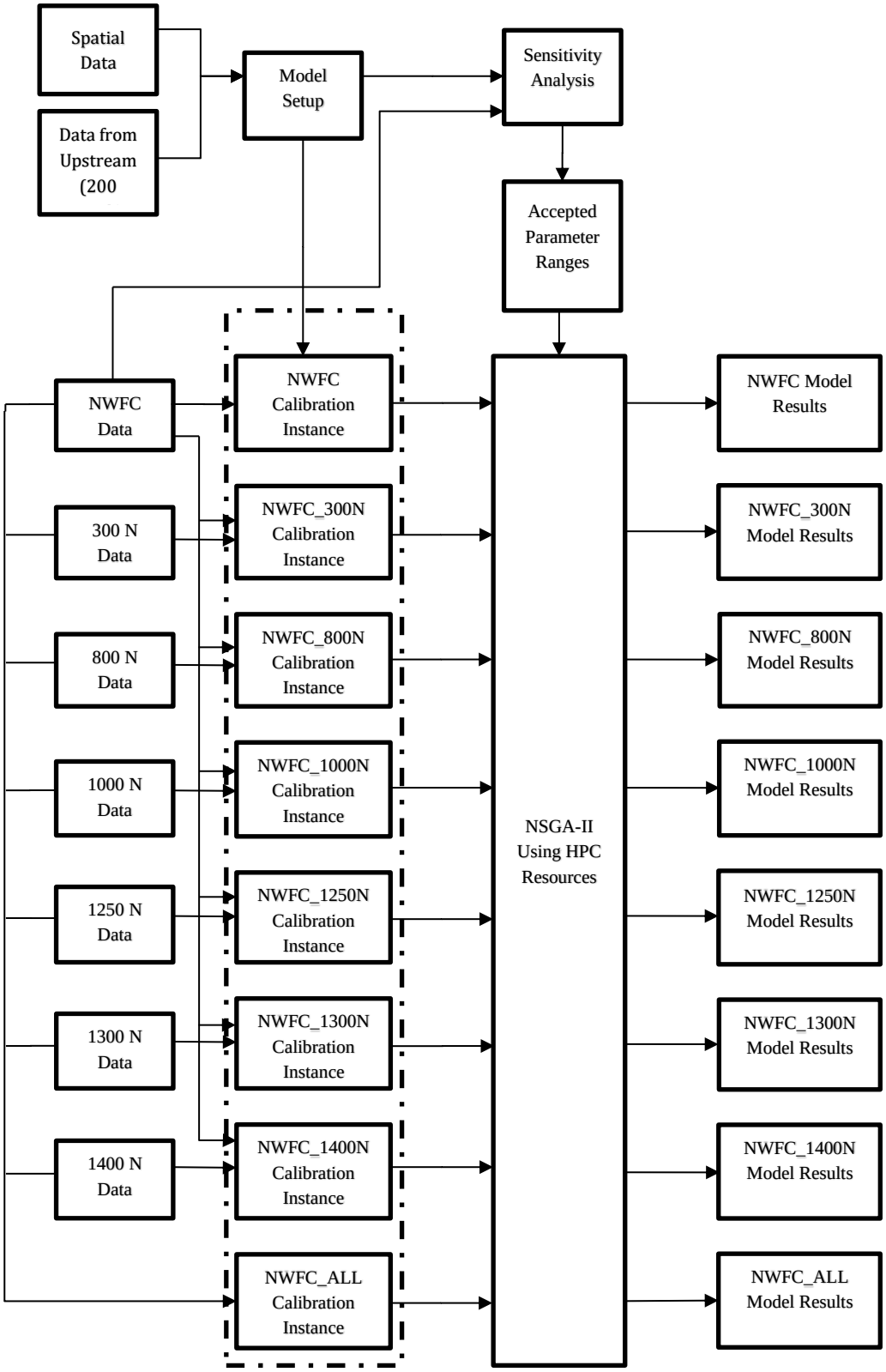


Fig 4.2. Flow chart of calibration procedure for study

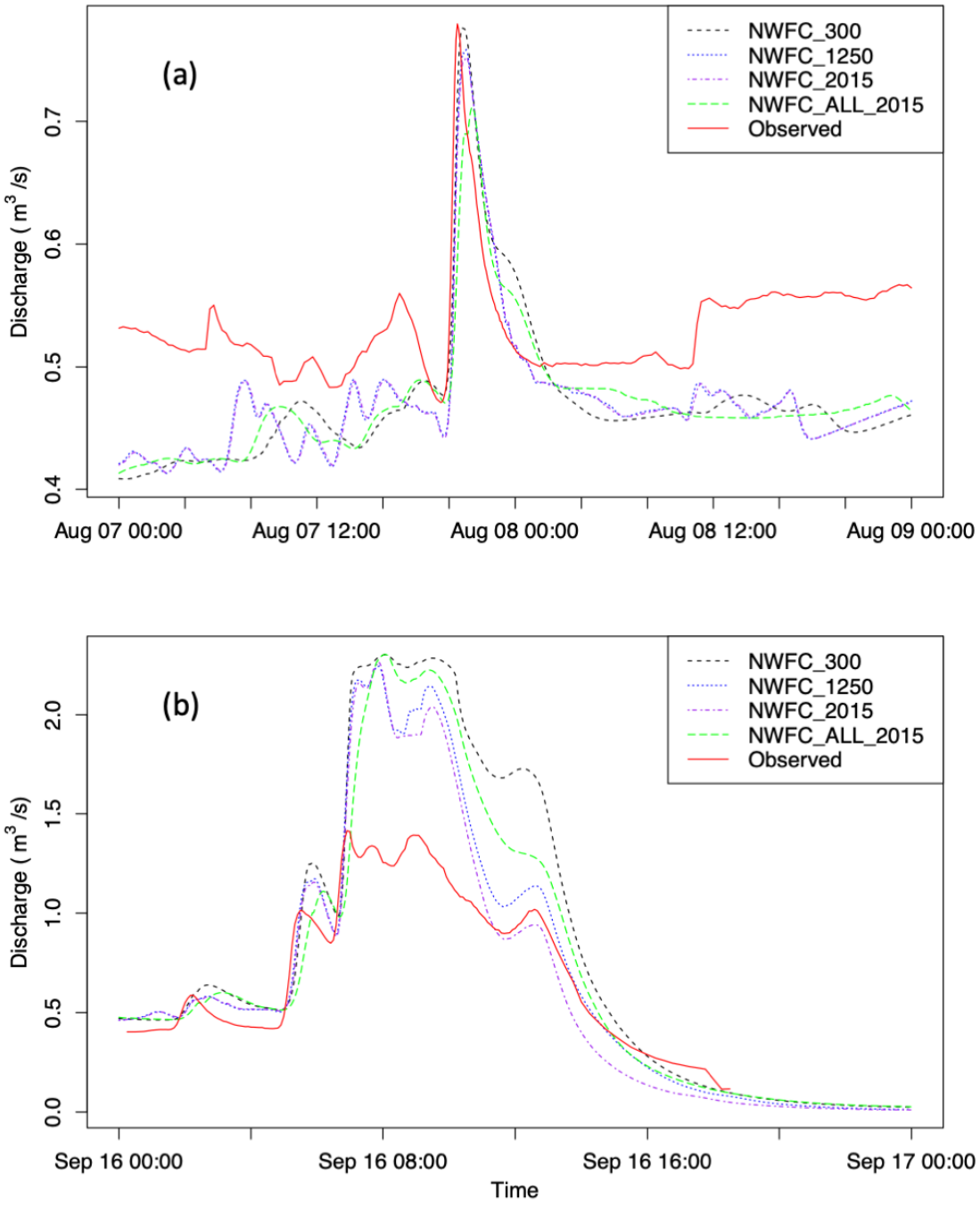


Fig 4.3. Calibrated hydrographs at the model outlet (1800 North) for the NWFC_300, NWFC_1250, NWFC_2015 (benchmark), and NWFC_ALL_2015 calibration instances for storm events on August 7, 2015 (a) and September 16, 2015 (b). The observed hydrograph is shown in red.

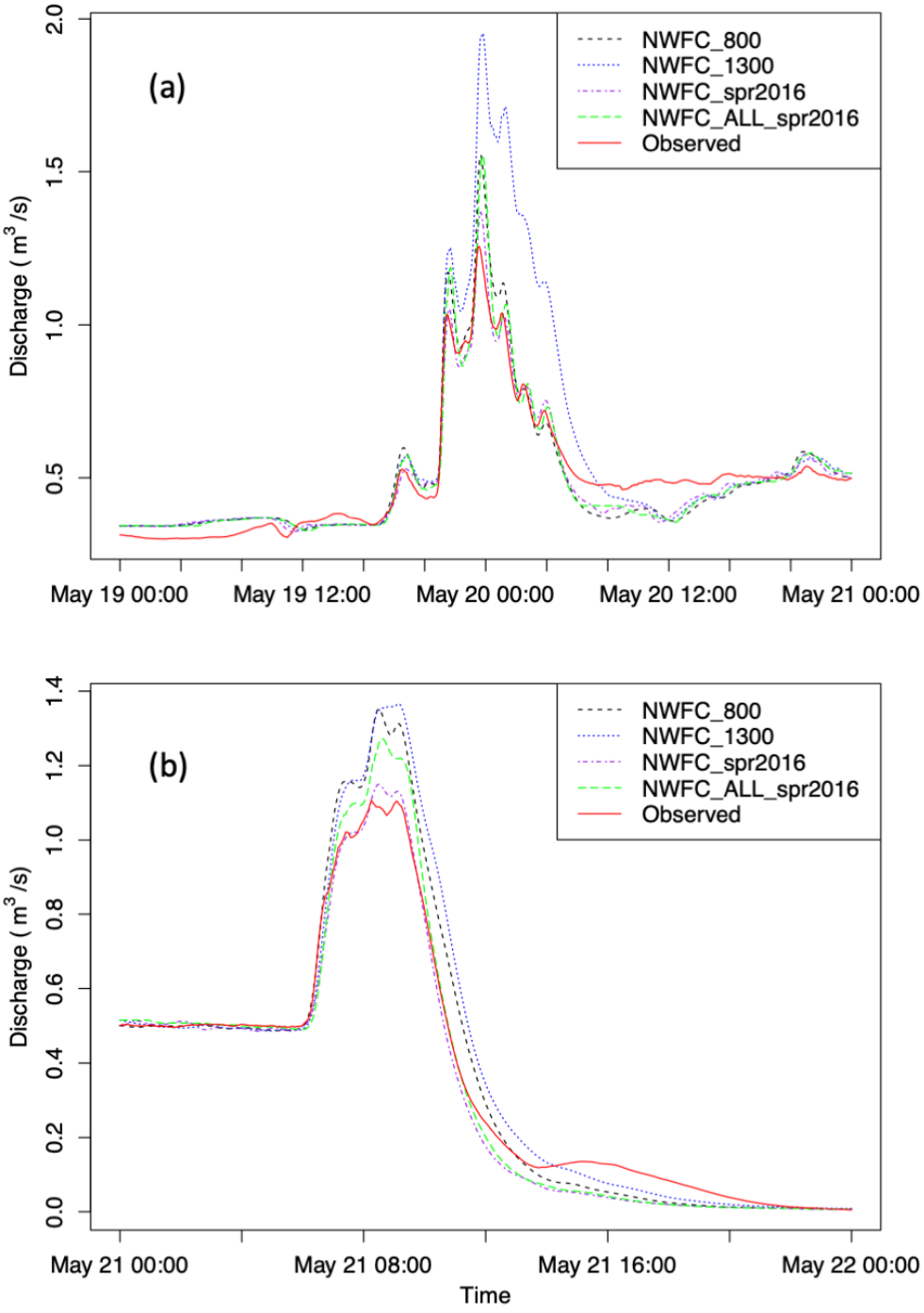


Fig 4.4. Calibrated hydrographs at the model outlet (1800 North) for the NWFC_800, NWFC_1300, NWFC_spr2016 (benchmark), and NWFC_ALL_spr2016 calibration instances for storm events on May 19, 2016 (a) and May 21, 2016 (b). The observed hydrograph is shown in red.

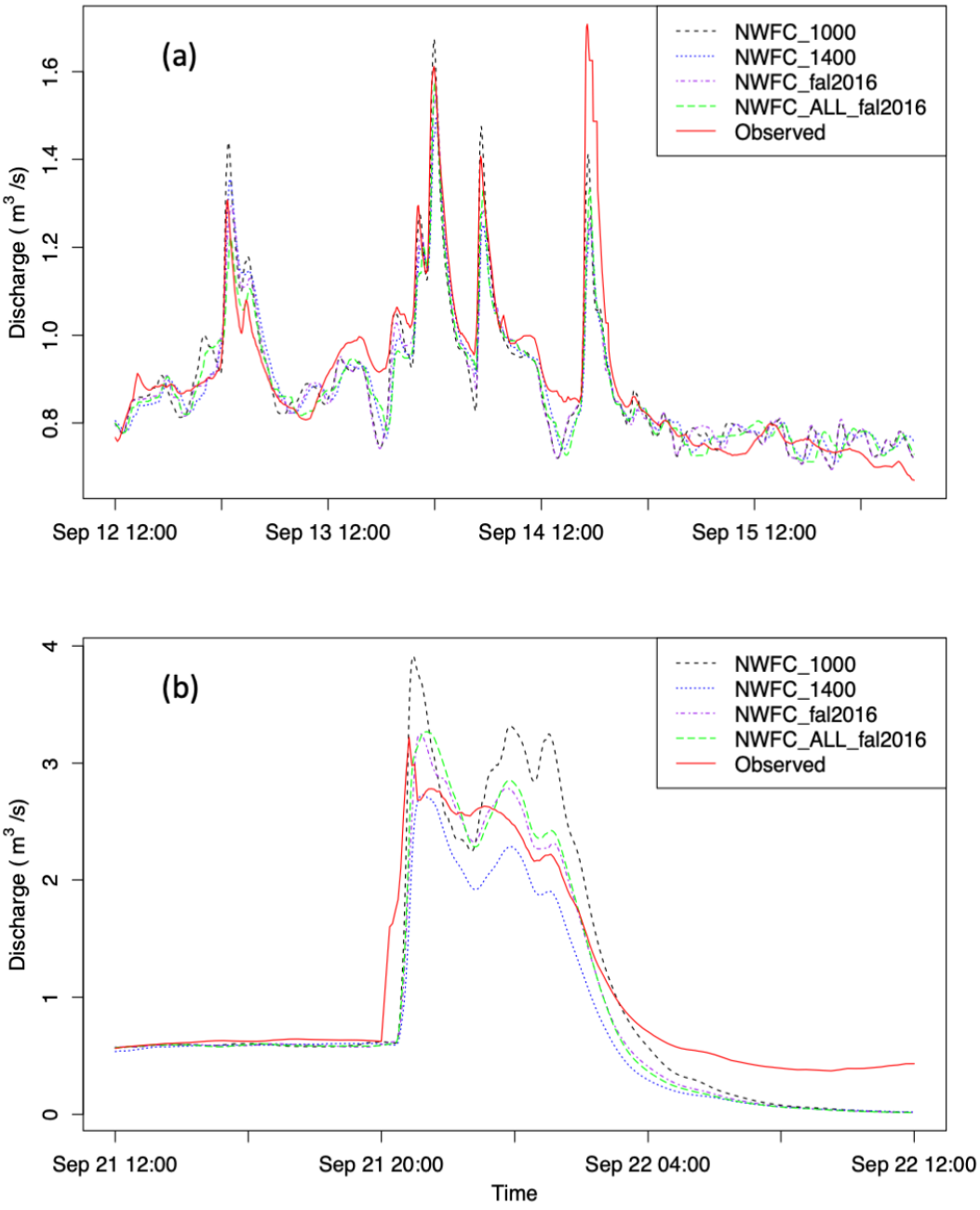


Fig 4.5. Calibrated hydrographs at the model outlet (1800 North) for the NWFC_1000, NWFC_1400, NWFC_fal2016 (benchmark), and NWFC_ALL_fal2016 calibration instances for storm events on September 12 – September 16, 2016 (a) and September 21, 2016 (b). The observed hydrograph is shown in red.

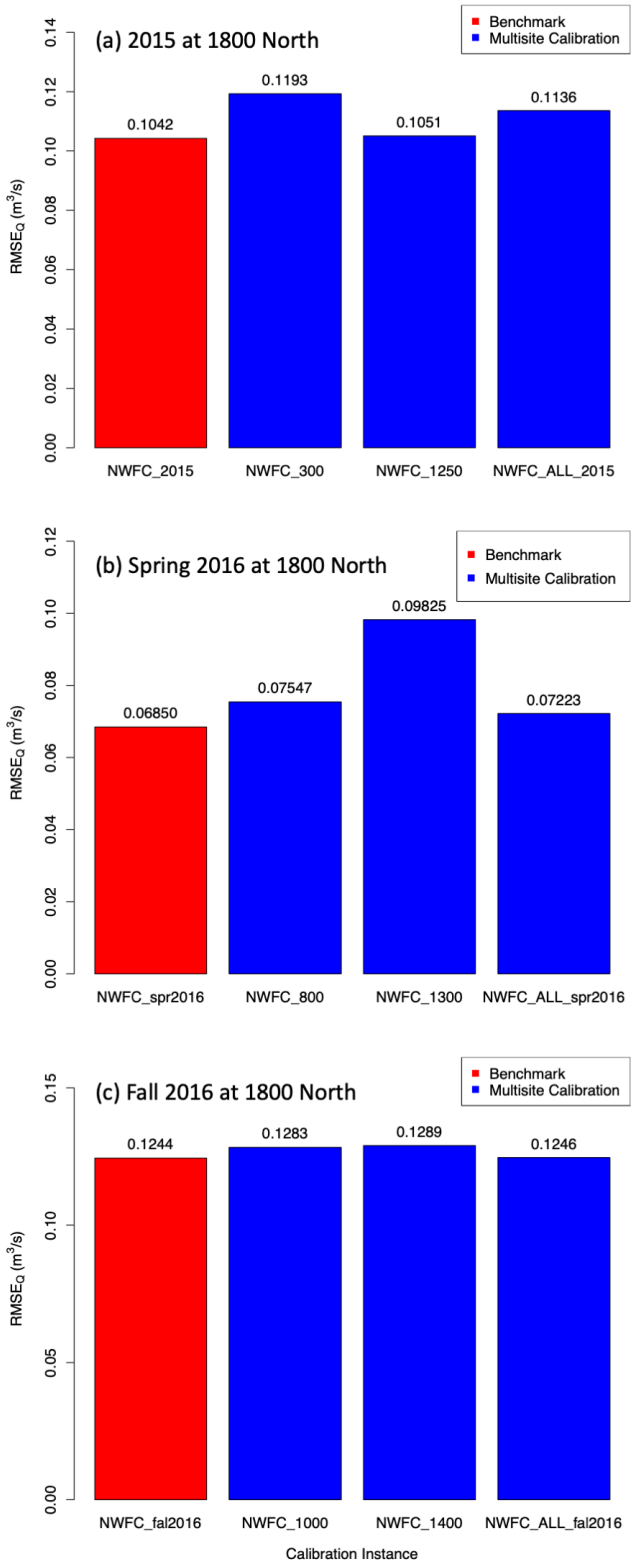


Fig 4.6. $RMSE_Q$ values at the model outlet for the three simulation periods – 2015 (a), Spring 2016 (b), and Fall 2016 (c).

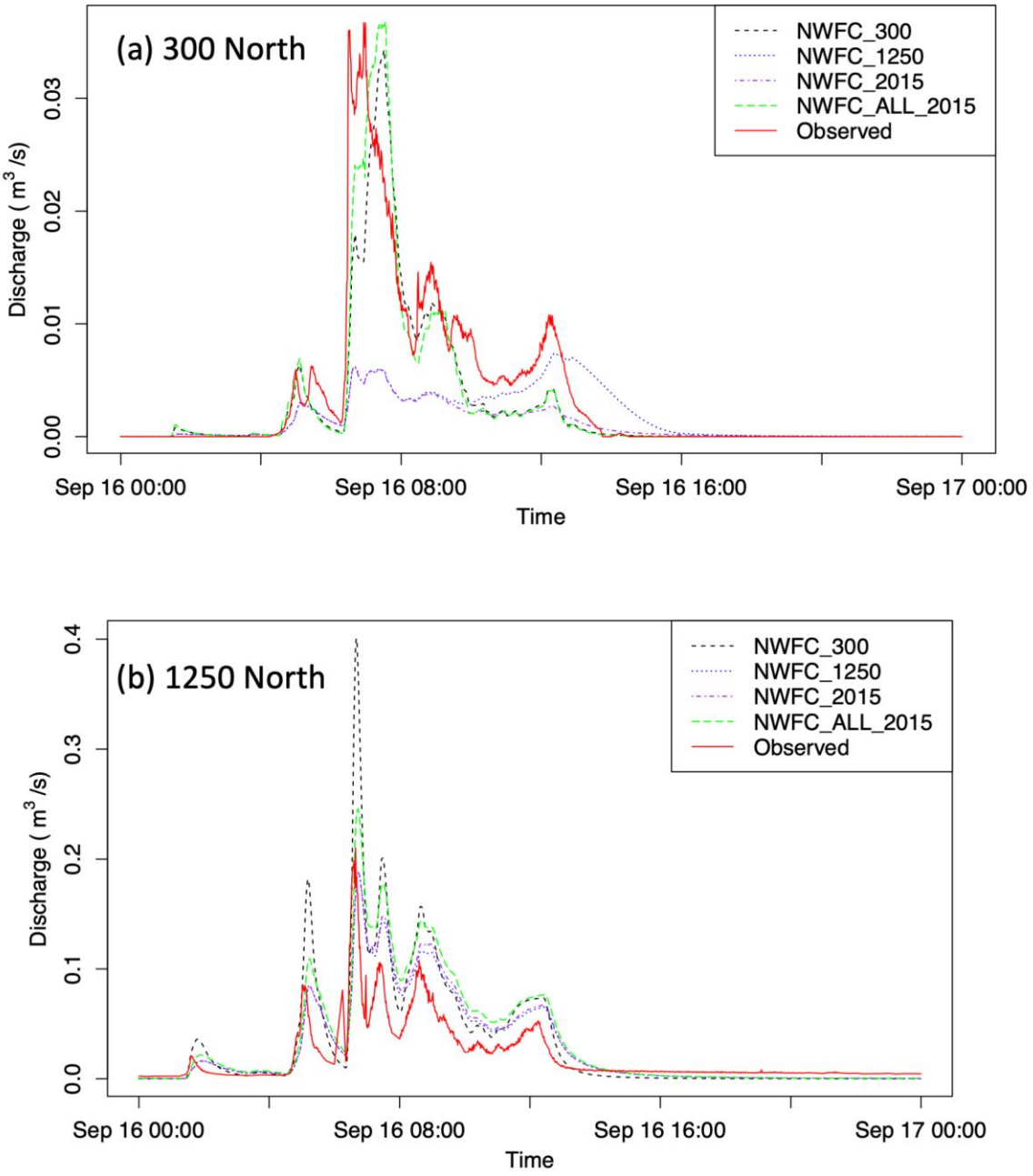


Fig 4.7. Calibrated hydrographs at the 300 North outfall (a) and at the 1250 North outfall (b) for the NWFC_300, NWFC_1250, NWFC_2015 (benchmark), and NWFC_ALL_2015 calibration instances for the storm event on September 16, 2015. The observed hydrograph is shown in red.

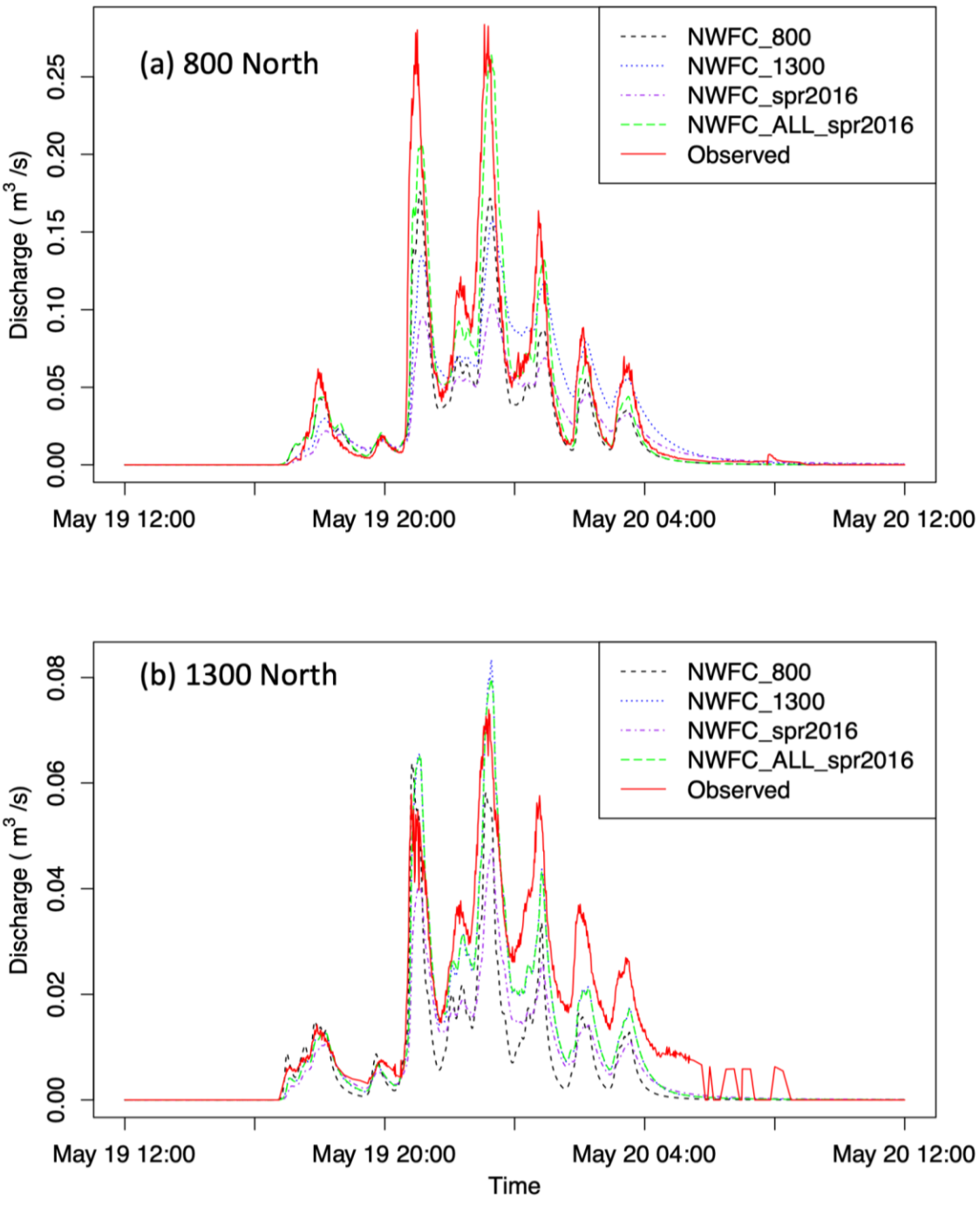


Fig 4.8. Calibrated hydrographs at the 800 North outfall (a) and at the 1300 North outfall (b) for the NWFC_800, NWFC_1300, NWFC_spr2016 (benchmark), and NWFC_ALL_spr2016 calibration instances for the storm event on May 19, 2016. The observed hydrograph is shown in red.

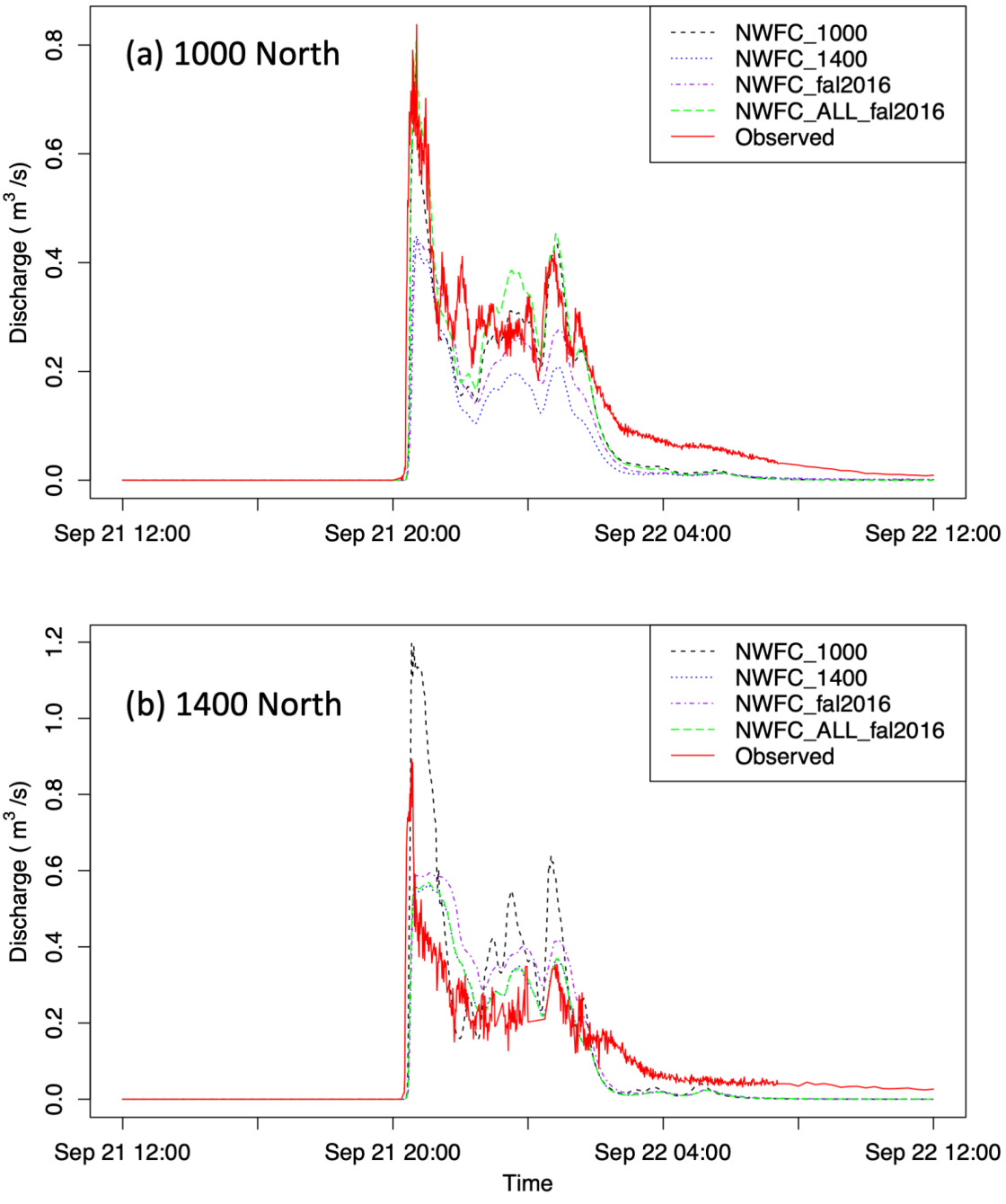


Fig 4.9. Calibrated hydrographs at the 1000 North outfall (a) and at the 1400 North outfall (b) for the NWFC_1000, NWFC_1400, NWFC_fal2016 (benchmark), and NWFC_ALL_fal2016 calibration instances for the storm event on September 21, 2016. The observed hydrograph is shown in red.

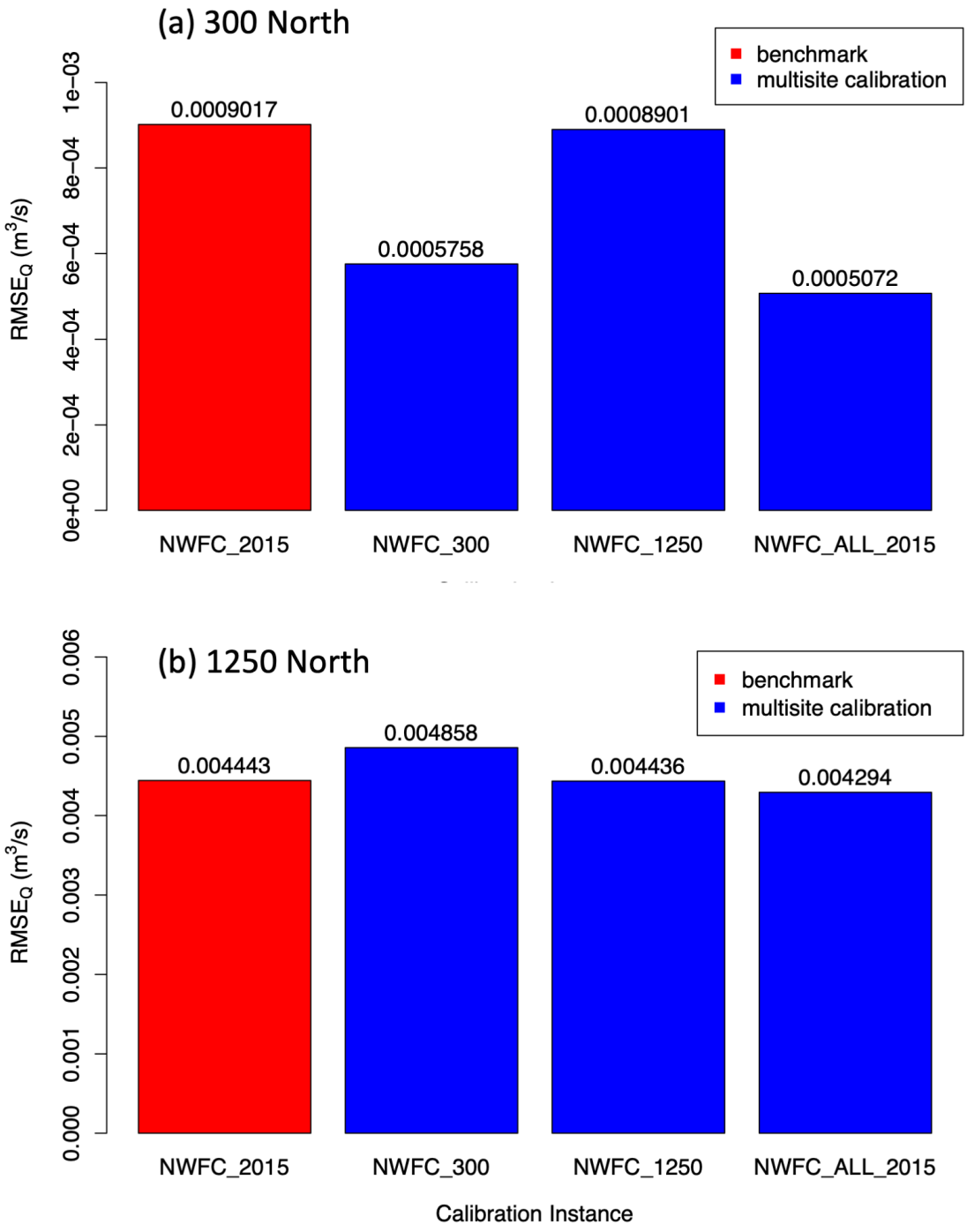


Fig 4.10. $RMSE_Q$ values at the 300 North outfall (a) and at the 1250 North outfall (b).

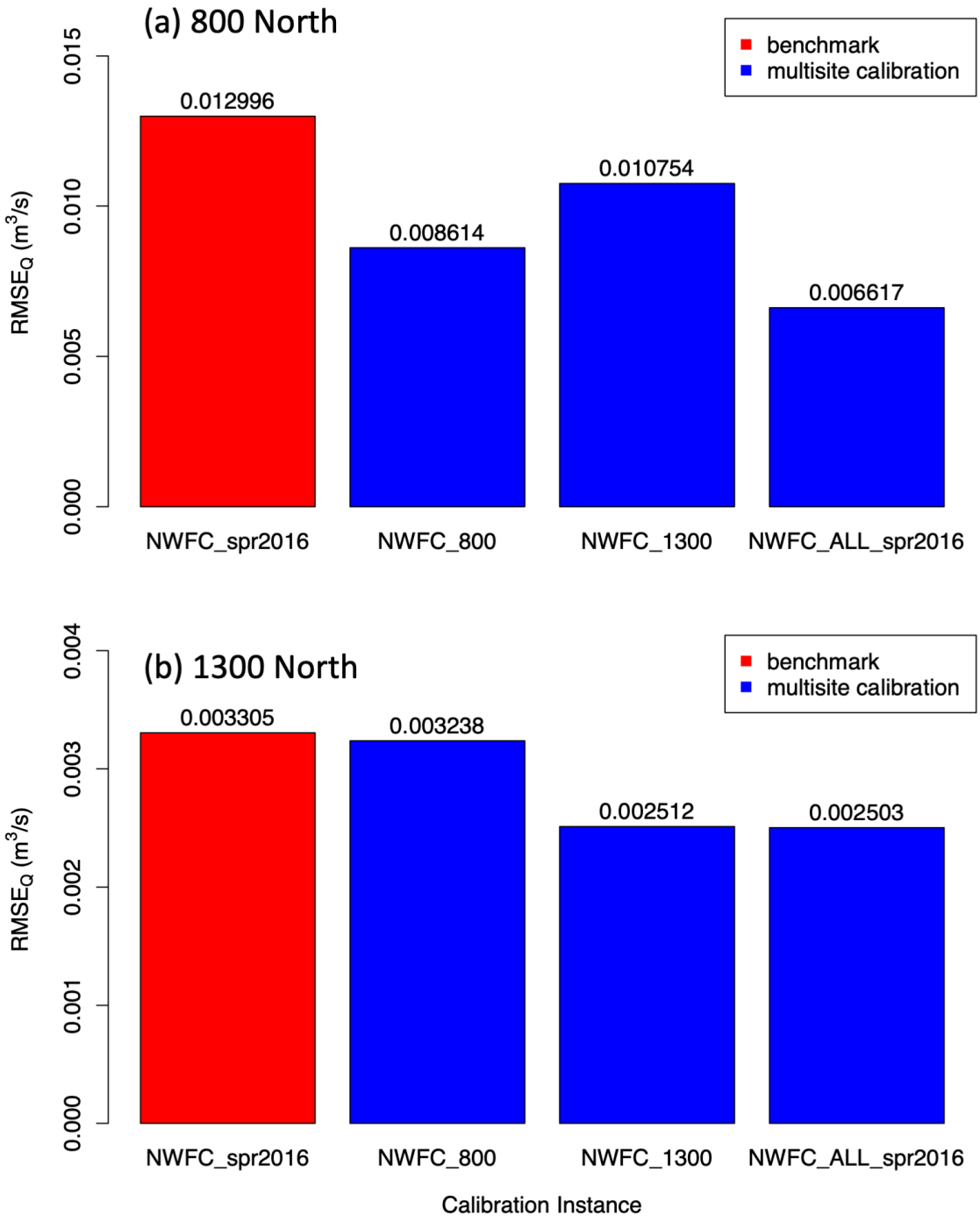


Fig 4.11. $RMSE_Q$ values at the 800 North outfall (a) and at the 1300 North outfall (b).

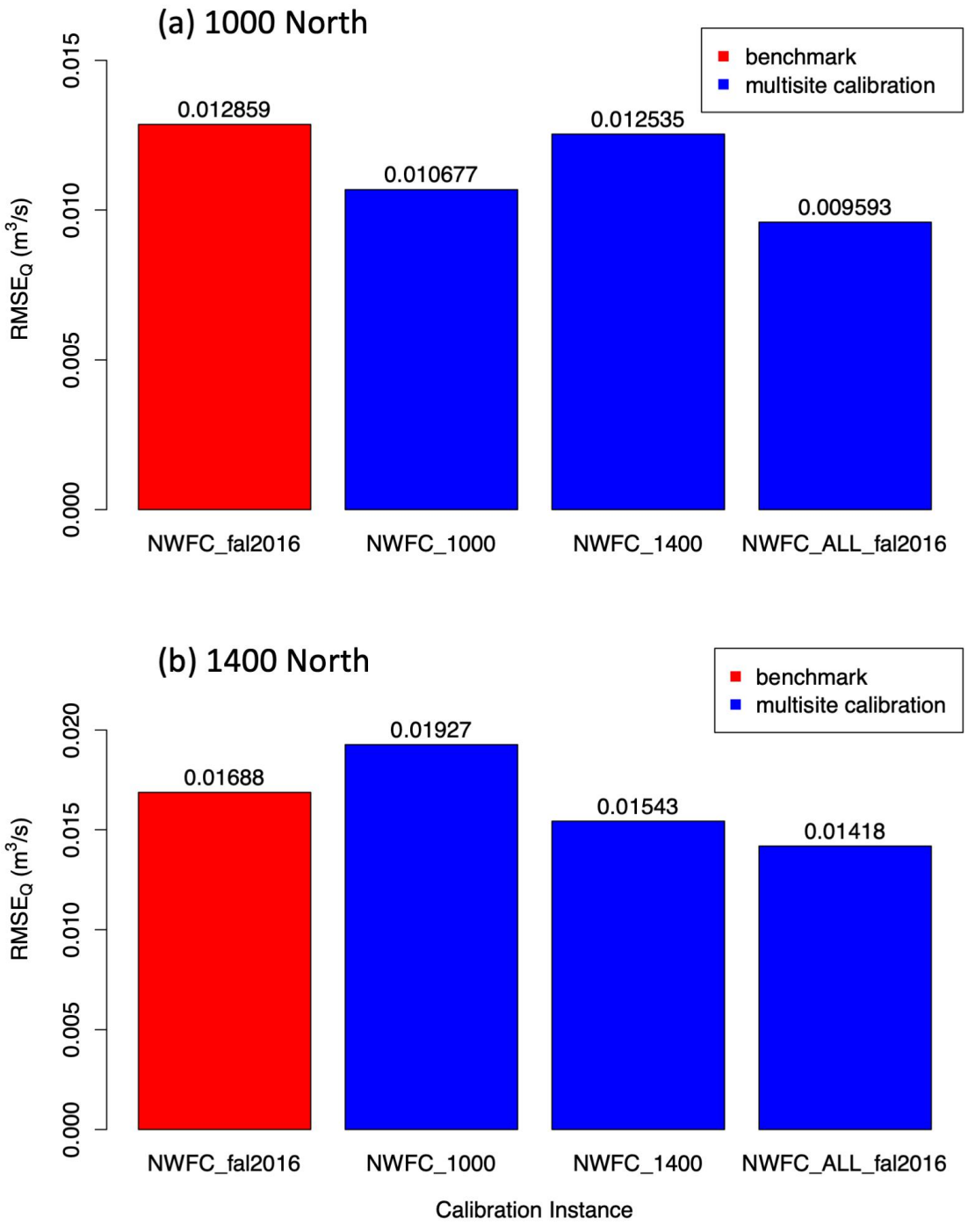


Fig 4.12. $RMSE_Q$ values at the 1000 North outfall (a) and at the 1400 North outfall (b).

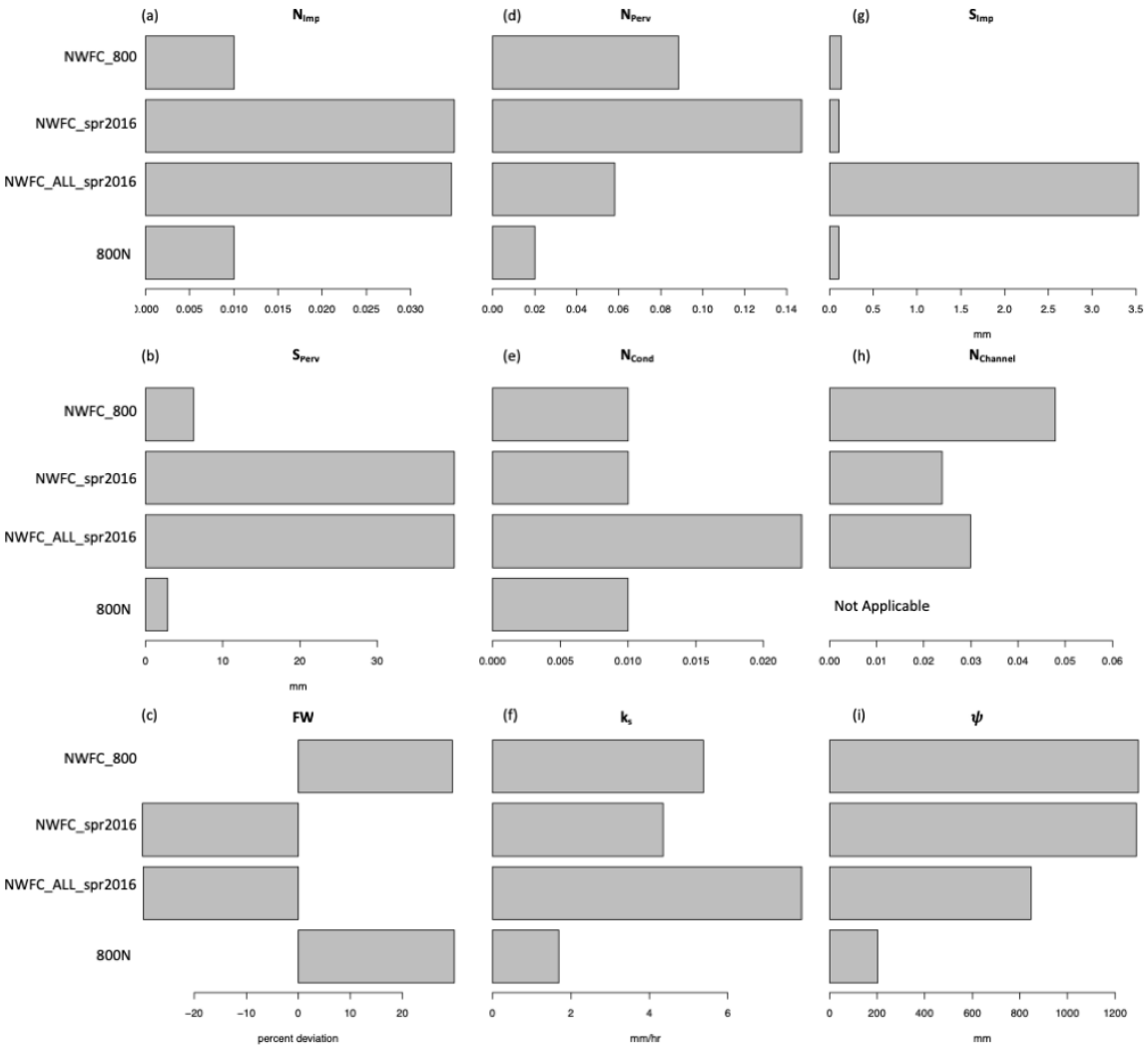


Fig 4.13. Comparison of calibrated parameter values for the NWFC_800, NWFC_spr2016, and NWFC_ALL_spr2016 calibration instances, as well as the calibrated values for the 800N subcatchment model

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The research described in this dissertation aims to address the need within the engineering and water management communities to understand some of the temporal and spatial loading patterns of total suspended solids (TSS) and total phosphorus (TP) loads resulting from stormwater runoff in combined urban water conveyances. This need exists, in part, as a result of the United States Environmental Protection Agency's (USEPA) release of the Phase II Stormwater Rules (Federal Register, 1999). As engineers and water managers are required develop a stormwater management plan and select stormwater pollution mitigation measures under the Phase II Stormwater Rules, this research emphasizes the monitoring and modeling required for water managers to make accurate estimates of TSS and TP loads. Emphasis in this dissertation was given to the need for high-frequency data collection at multiple locations across the study area, the required monitoring and telemetry infrastructure, a comparison of statistical methods for deriving continuous estimates of TSS concentrations from the high-frequency data, and stormwater modeling techniques that take advantage of the high-frequency data collected across multiple monitoring sites.

The results from the urban observatory case study demonstrate the need for synchronized and adaptive sampling, or sampling that is coordinated across multiple sites and at varying frequencies, for characterizing constituent loading patterns within and between multiple storm events. This novel approach to adapting sampling frequencies based on whether or not stormwater is present in the canal and real-time detection of runoff

events takes much of the guesswork out of when to collect grab samples and how to coordinate simultaneous collection of samples at multiple locations. Additionally, the collection of grab samples initiated by water quality parameters measured in situ ensured that loading events were sampled that may have been logistically difficult to sample (e.g., during storm events occurring in the middle of the night) or undetectable to the naked eye (e.g., continuation of storm sampling until the influence of the storm has passed through the system even after it has stopped raining). This form of "smart" sampling promotes data collection that can better enable the derivation of continuous constituent concentration estimates.

The statistical analysis and results obtained from this research build upon the many studies already performed that use turbidity as a surrogate for TSS and suspended sediment concentrations (Christensen et al., 2002; Lewis and Eads, 2001; Rasmussen et al., 2008; Ryberg, 2006). However, this study is one of the first to investigate the use of linear mixed effects (LME) modeling in developing surrogate regression equations in urban water systems. The use of LME models to estimate TSS concentrations and loads allows for a model that accounts for changes in the relationship between turbidity and TSS within short duration loading events and other phenomena characteristic of small urban catchments. During these investigations, it was observed that the LME and LRCAT models performed better than the CLR model tested for the sites at which data were collected. As the effects of constituents in stormwater runoff are being recognized in receiving water bodies (National Research Council, 2009) and progress is made toward a better quantification and understanding of constituent loading, examination of multiple statistical modeling techniques, including linear mixed effects modeling, is necessary to account for changing

and undersampled environmental conditions within regression equations and to ensure the quality of regression results.

The numerical modeling performed using USEPA's Storm Water Management Model (SWMM) was used to assess the degree to which the model's ability to simulate stormwater runoff and conveyance could be improved using high-resolution data collected via the urban observatory. Through the use of data from multiple continuous monitoring sites and an analysis of multiple goodness-of-fit measures, we were able to identify some of the challenges related to multi-site/multi-objective calibration procedures. The simulations performed demonstrated that the use of calibration data at stormwater outfalls could be used to improve predictions at those locations without compromising prediction accuracy at the model outlet This type of information can be beneficial to water managers and engineers as they determine where to make investments in collecting data through grab sampling or installation of continuous monitoring equipment in efforts to isolate runoff volumes and, subsequently, constituent loads.

Chapters 2 through 4 of this dissertation present the main results of this research and were focused on three research objectives: 1) design and establish an urban observatory for studying the effects of stormwater inputs on urban water systems, 2) investigate methods for quantifying suspended solids loads within urban water systems, and 3) quantify the contributions of stormwater runoff to urban water systems. These research objectives were chosen to address the challenges of monitoring and quantifying TSS and TP loads in urban water systems having dynamic, precipitation-driven stormwater runoff.

In Chapter 2, we describe the design of an urban environmental wireless sensor network installed along the Northwest Field Canal (NWFC) in Logan, Utah, USA. We

describe the programming logic for each continuously monitored canal and outfall site, the adaptive sampling protocol, and inter-site communications. This research demonstrated how automation of in situ observations, event-based physical sample collection, and synchronization of sampling efforts across multiple locations along a receiving water body can provide valuable information for municipal separate storm sewers (MS4s) on the temporal and spatial variability of TSS and TP loading events. As a result of this adaptive sampling and smart stormwater monitoring effort, we were able to extract valuable information related to the spatial and temporal variability of TSS and TP loading events.

Using the multi-site configuration and synchronized sampling scheme, we found that event mean concentrations (EMC) for TSS, TP, and total dissolved phosphorus (TDP) varied greatly between storm events and between monitoring sites. It was also found that while TSS EMCs were similar at both outfall sites (300 North and 1250 North), the TP and TDP EMCs were much higher at a site that drained more residential neighborhoods (300 North). Furthermore, we found that TSS EMCs for each storm event varied by as much as two orders of magnitude, providing evidence that estimating TSS loads for unmonitored events by using a single, previously derived EMC has potential to greatly bias those estimates given that our per-storm EMCs were highly variable.

Adaptive sampling and inter-site communication allowed for real-time event detection and the capture of temporal loading characteristics within storm events. As a result, it was found that constituent concentrations varied greatly, even within the short-duration of most of the urban runoff events we monitored. It was found that the first flush phenomenon was more prevalent at outfalls that drained catchments with smaller surface area. The detection of the presence of a first flush, especially for the smaller monitored

catchments, was made possible by the high-frequency physical samples collected at the onset of the storm event and the lower-frequency sample collection for the remainder of the event. This method of adaptive sampling ensured that the initial peak, which often occurred within the first 10 minutes of the event, and the receding limb of the pollutograph were adequately sampled. The ability to do this type of sampling is critical in adequately quantifying discharge and loading from stormwater outfalls in urban water systems of the western United States like the one we studied. These systems are characterized by brief, but intense rainstorms, many of which last on the order of hours. The time required to mobilize sampling crews to manually sample these events would preclude our ability to collect many of the samples we were able to get in an automated way, undermining ability to examine first flush effects and how concentrations vary over the course of individual storms.

The implementation of the turbidity threshold sampling (TTS) scheme at the continuously monitored canal sites ensured that physical samples were collected for the entire range of observed turbidity values during storm events and enabled us to better track pulses of stormwater traveling through the canal system – even after rainfall had ended. The TTS scheme enhanced our ability to generate surrogate relationships by ensuring that extreme turbidity values were adequately sampled, that the regression equations accounted for these extremes, and that extrapolation beyond those extreme values was not necessary. The TTS and constant-time interval sampling schemes were initiated by rainfall-runoff detection at outfall sites. Communicating event detections to upstream and downstream canal sites enabled us to subsequently estimate TSS loads during base flow and storm event conditions by means of surrogate relationships.

The urban observatory described in Chapter 2 had the specific purpose of understanding TSS and TP loading event processes and making observations at a high enough frequency to capture the characteristics of small urban catchment responses. To reduce data storage requirements, personnel and laboratory analysis time, and power demands at monitoring sites, in situ observations and periodic base flow sampling occurred at a much lower frequency under the assumption that constituent concentrations remained relatively constant. Future work, however, could include a sampling scheme that describes more of the variability during base flow conditions. This could be accomplished in part by installing Internet connected dataloggers that make regular queries of other Internet connected devices upstream of the observatory. Examples could include an urban observatory that makes regular queries of snowpack data during spring and summer months to anticipate snowmelt events, or data downloads and communication with automated dam spillways and diversion gates upstream to assess how control structures impact ephemeral in-stream water quality. These types of advancements would allow for more accurate constituent mass budgets to downstream impaired water bodies and provide water managers with the required information for implementing a more precise approach to stormwater treatment and constituent removal.

Deriving continuous time series of constituent concentrations by means of surrogate relationships aims at understanding in-stream water quality constituent fluxes as well as making more accurate estimates of constituent loads to a receiving water body. Using turbidity as a surrogate for TSS has been investigated and used in numerous studies in the last few decades. The work reported in Chapter 3 explored multiple regression methods for estimating TSS concentrations and loads and an analysis of potential

explanatory variables. Results of this analysis showed that turbidity and categorical variables related to the season that a sample was collected in (upstream), the antecedent dry period (upstream), rising vs. falling limb (downstream), and storm event intensity (downstream) were found to be significant predictor variables. It was also found that linear regression with categorical variables and linear mixed effects models outperformed classical linear regression based on multiple goodness-of-fit criteria and that this was reflected in TSS load estimates for the duration of the study. It was also found that linear mixed effects modeling was a more robust method for estimating TSS loads during undersampled conditions (e.g., undersampled storm events or base flow conditions) as the maximum likelihood parameter estimation was less affected by extreme values. These results represent the value of applying linear mixed effects modeling for estimating TSS concentrations in environments where constituent source materials may vary, causing surrogate relationships vary. These statistical modeling methods and the results of this study can greatly benefit Total Maximum Daily Load (TMDL) or other studies concerned with estimating TSS loads to receiving water bodies in highly urbanized environments.

Using the high frequency turbidity data and other metavariables we were able to derive continuous TSS estimates with satisfactory accuracy. One challenge with using surrogate indicators is the often-violated assumptions of heteroscedasticity and independence in residuals. High-frequency time series datasets often display high degrees of autocorrelation, which often results in interdependence in residual values. One direction that future research could lead is the exploration of alternative methods for estimating TSS concentrations as a function of continuous in situ parameters that are less sensitive to nonlinearity and explanatory variable interactions. One example is random forests, which

allow for high numbers of explanatory variables without being bound by assumptions of normality, independence, and homoscedasticity.

This surrogate analysis could then be expanded to include TP. One challenge presented with TP was that there were numerous samples that were below the detection limit. Large percentages of censored data prohibit the use of classical linear regression methods. While other methods exist to account for censored data, such as maximum likelihood estimation methods, additional research is required to understand which method is most applicable in water bodies where surrogate relationships are expected to change, as is the case in the NWFC.

Finally, in Chapter 4, we used data from multiple continuous monitoring sites to calibrate a SWMM model for the NWFC drainage area in an effort to determine how high-resolution data can improve our ability to simulate stormwater quantity. It was found that the additional calibration data were essential for predicting runoff events at both the watershed outlet and at outfall locations interior to the modeled domain. While this result can be useful to stormwater managers/planners, we do not feel that the procedure we used in Chapter 4 exhausts potential options for addressing the question of how high-resolution data from multiple monitoring sites can improve our ability to simulate stormwater. For example, there remain alternative approaches to multisite calibration that may produce different results. As this study used high-frequency data at its full temporal resolution to calibrate the model, another alternative approach might be to aggregate the high-frequency data to a larger time step and consider total runoff volumes and TSS or TP loads for the study period. This would allow engineers and water managers to identify subcatchments that would most benefit from a BMP in order to reduce seasonal or annual loads.

The case study of the urban observatory in the NWFC and the subsequent statistical and numerical modeling efforts address many of the requirements for estimating TSS and TP loads in an urban combined conveyance. One trajectory for future research is to explore how this research benefits water management and decision making. For example, the drainage system in Logan, UT is comprised of a network of canals that were originally constructed for irrigation purposes, but have later been dedicated for stormwater conveyance. This type of system is not unique, especially in the western U.S. The added stress to these canals caused by the reception of stormwater inputs results in frequent flooding events in which canal banks are overtopped. Often it is logistically difficult for canal masters to adjust diversion structures in time to avoid flooding events. Future research could include the investigation of algorithms and infrastructure required for using high-frequency data as well as weather forecasts for real-time stormwater model calibration, scenario prediction, and potentially even automated operation of the stormwater system (e.g., automate closing diversion gates during storm events). Modeling results could then be used to alert water managers of expected runoff rates and management decisions could be made accordingly in a manual or automated way. This could be a relatively low-cost alternative to increasing canal capacity and would save the cost of repairing property damaged by flooding events.

As the research performed in this dissertation progressed, the focus evolved from TSS, TP, and TDP to an emphasis on only particulates (TSS and TP). Additional research is required to understand the dynamics of dissolved phosphorus as this is the form that is most bioavailable. The results of our case study in Chapter 2 found that higher EMCs for TDP were observed at 300 North, the smaller, residential subcatchment. This may be due

to the residential land use, or this could be because it is a smaller drainage area with more directly connected impervious area, which can reduce the residence time of TDP. Future work could include the installation of monitoring equipment within storm drains and conduits to better grasp the age and flow paths of dissolved nutrients. The genesis and residence time of TDP would need to be investigated in order to understand which catchments could benefit from green infrastructure and other nonstructural SCMs. While TP, DTP, and TSS are important in the context of the TMDL for Cutler Reservoir, which is the downstream receiving water body for Logan City's stormwater runoff, other water quality constituents not included in this study (e.g., pesticides, metals, etc.) may also be important in managing the multiple uses for this urban water course (irrigation flows, return flows, stormwater flows, etc.).

The SWMM modeling conducted in Chapter 4 was able to shine light on some of the challenges encountered when attempting to develop an accurate water budget for a naturally-lined urban conveyance that is highly influenced by human behavior and operational strategies. While this study found that net losses in the naturally-lined canal were negligible, this is not always the case (Molina, 2008). While longitudinal flow measurements along the canal can be used to get an idea of the potential magnitude of surface water/groundwater interactions, it is imprecise as it is difficult to ensure that flowrates are constant during flow measurements and that the same section of water is measured at each location. Recent investigations have shown the utility of mobile platforms that are able to collect high-frequency water quality data along an urban surface water conveyance (Mihalevich et al., 2017). Future research could include the adaptation of a mobile platform to also collect water quantity data and allow for identification of reaches

of canals that were susceptible to groundwater losses so that costly repairs to the canal could be localized to problem areas. A mobile platform for making longitudinal flow measurements would greatly improve our water balance estimates for the canal. The operation of head gates by irrigation water users was also found to be unpredictable. In the case of the NWFC where water users were allowed to operate flow diversion head gates in an unregulated fashion, human behavior during storm events was found to be a significant unknown and potential area of uncertainty when modeling the canal. Future research could include surveys and data collection from water users to better understand trends and behaviors, especially during storm events.

The focus area of this research was the NWFC drainage area, a subcatchment of the Logan City urban water system. SWMM parameters that were obtained using the multi-objective evolutionary algorithm we used for calibrating the model were specific for the NWFC drainage area and monitored subcatchments. Techniques, however, need to be derived for using monitoring data and existing models to scale up from the small area we were able to intensively monitor and model to the larger Logan City urban water system. This would include an investigation of the portability of calibrated parameters to ungauged subcatchments and an investigation of how current model results might be transferred to other canals and sections of the drainage system. This would allow for more accurate constituent load predictions for the entire Logan City MS4, and could influence NPDES permitting and the enforcement of water quality standard infractions or lack thereof. Such an investigation could also serve as a model for how this type of work could be effectively carried out in other urban water systems.

References

Christensen, V.G., Rasmussen, P.P., Ziegler, A.C., 2002. Real-time water-quality monitoring and regression analysis to estimate nutrient bacteria in Kansas Streams [WWW Document]. URL http://ks.water.usgs.gov/pubs/reports/vgc.0610.html

Federal Register, 1999. National Pollutant Discharge Elimination System - Regulations for Revision of the Water Pollution Control Program Addressing Storm Water Discharges. USA.

Lewis, J., Eads, R., 2001. Turbidity Threshold Sampling For Suspended Sediment Load Estimation, in: Proceedings of the Seventh Federal Interagency Sedimentation Conference, March 25 to 29, 2001, Reno, Nevada. pp. 110–117.

Mihalevich, B.A., Horsburgh, J.S., Melcher, A.A., 2017. High-frequency measurements reveal spatial and temporal patterns of dissolved organic matter in an urban water conveyance. Environ. Monit. Assess. 189, 593. https://doi.org/10.1007/s10661-017-6310-y

Molina, K.N., 2008. Seepage Evaluations in Cache Valley Irrigation Canals. Utah State University.

National Research Council, 2009. Urban Stormwater Management in the United States.

Rasmussen, T.J., Lee, C.J., Ziegler, A.C., 2008. Estimation of Constituent Concentrations, Loads, and Yields in Streams of Johnson County, Northeast Kansas, Using Continuous Water-Quality Monitoring and Regression Models, October 2002 through December 2006.

Ryberg, K.R., 2006. Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05. Reston.

APPENDICES

APPENDIX A

Results from Longitudinal Flow Measurements

Below are the results from the longitudinal flow measurement events. Flow measurements were made at six different locations.

Table A1. Results from the longitudinal flow measurement event on October 13, 2016

Longitudinal NWF Flow on 10-13-2016						
Measurement Location	200 South	100 South	100 North	500 North	1000 North	1800 North
Discharge (cfs)	**13.8795**	**13.9834**	**15.264**	**14.4137**	**15.7428**	**12.7985**
Time (MST)	0900-0935	0900-0935	1007-1037	1000-1037	1115-1158	1108-1145
Stage @ 200 South (ft)	1.11					1.16

Table A2. Results from the two longitudinal flow measurement events on October 26, 2016

Longitudinal NWF Flow on 10-26-2016						
Event 1						
Measurement Location	200 South	100 South	100 North	500 North	1000 North	1800 North
Discharge (cfs)	**24.1573**	**24.7921**	**24.5958**	**23.3852**	**25.8404**	**23.9317**
Time (MST)	0840-0906	0830-0910	0930-1000	0930-1040	1106-1200	1100-1150
Stage @ 200 South (ft)	1.5-1.42	1.5-1.42	1.41-1.4	1.41-1.4	1.4	1.4
Event 2						
Measurement Location	200 South	100 South	100 North	500 North	1000 North	1800 North
Discharge (cfs)	**17.6624**	**18.0219**	**18.3222**	**16.4077**	**18.183**	**17.2137**
Time (MST)	1320-1355	1320-1352	1400-1430	1405-1455	1515-1555	1515-1545
Stage @ 200 South (ft)	1.17-1.16	1.17-1.16	1.18	1.18	1.18	1.18

APPENDIX B

Coauthor Approval Letters

UtahState University
Department of Civil and Environmental Engineering
4110 Old Main Hill
Logan, UT 84322-4110
Telephone: (435) 797-2932
Fax: (435) 797-1185

2-25-2019

Amber Spackman Jones
Utah Water Research Laboratory
Utah State University
8200 Old Main Hill
Logan, UT 84322-8200

Dear Amber,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in March of 2019.

I am requesting your permission to include the attached paper, of which you are a coauthor, as a chapter in my dissertation. I will include acknowledgments to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Anthony A. Melcher

I hereby give permission to Anthony A. Melcher to use and reprint all of the material that I have contributed to Chapter 3 of his dissertation.

_____

Amber Spackman Jones

UtahState University
Department of Civil and Environmental Engineering
4110 Old Main Hill
Logan, UT 84322-4110
Telephone: (435) 797-2932
Fax: (435) 797-1185

2-25-2019

Caleb Amoa Buahin
Utah Water Research Laboratory
Utah State University
8200 Old Main Hill
Logan, UT 84322-8200

Dear Caleb,

I am in the process of preparing my dissertation in the Civil and Environmental Engineering Department at Utah State University. I hope to complete my degree in March of 2019.

I am requesting your permission to include the attached paper, of which you are a coauthor, as a chapter in my dissertation. I will include acknowledgments to your contributions as indicated. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please contact me.

Thank you,

Anthony A. Melcher

I hereby give permission to Anthony A. Melcher to use and reprint all of the material that I have contributed to Chapter 4 of his dissertation.

_____
Caleb Amoa Buahin

CURRICULUM VITAE

ANTHONY A. MELCHER
Ph.D. Candidate – Utah Water Research Laboratory, Department of Civil and Environmental Engineering, Utah Water Research Laboratory, Utah State University
Email: tmelcher@aggiemail.usu.edu
Phone: (801) 602-6897

Education

Doctor of Philosophy in Civil and Environmental Engineering **Spring 2019**
Utah State University, Logan, UT
Dissertation: "Estimating Suspended Solids and Phosphorus Loading and Stormwater Impacts in Urban Water Systems Using Continuous Data"
Advisor: Dr. Jeffery S. Horsburgh

M.S. Civil and Environmental Engineering **April 2013**
Brigham Young University, Provo, UT
Water Resources Emphasis
Advisor: A. Woodruff Miller

B.S. Civil and Environmental Engineering **April 2012**
Brigham Young University, Provo, UT

Expertise

My research and training are in the fields of watershed hydrology, hydrodynamics, surface and groundwater quality, contaminant transport, and data analysis. I focus on the development of new technology and modeling procedures to increase the understanding of environmental processes and promote intelligent systems and infrastructure. My current research is focused on the design and instrumentation of aquatic wireless sensor networks for the collection of high frequency water quality and quantity data and the development of modeling and analysis techniques for pollutant load estimation in urban environments.

Research Experience

Graduate Research Assistant **August 2013 – Present**
Utah Water Research Laboratory, Utah State University

- Designed an urban wireless sensor network which linked multiple water quality and quantity monitoring sites capable of adaptive sampling and inter-site communication
- Derived linear mixed effects regression models using R statistical software that relate total suspended solids and total phosphorus to parameters measured *in-situ* at high frequencies
- Developed high-resolution SWMM model for the Northwest Field Canal catchment in Logan, UT using high-frequency water quality and quantity data as model inputs and for model calibration. Simulations were parallelized and executed on high-performance computers
- Managed and maintained 9-10 water quality/quantity monitoring sites located in the Little Bear River Watershed, Cache County, UT

Graduate Research Assistant **April 2012 – April 2013**
Brigham Young University

- Analyzed 46 lakes in Yellowstone National Park to determine their trophic state classification

- Performed principal component analysis (PCA) and cluster analysis on water quality data for Yellowstone lakes

Publications and Presentations

1. **Melcher, A. A.** and J.S. Horsburgh, 2017. An Urban Observatory for Quantifying Phosphorus and Suspended Solids Loads in Combined Natural and Stormwater Conveyances. *Environmental Monitoring and Assessment*. http://dx.doi.org/10.1007/s10661-017-5974-7

2. Mihalevich, B.A., J.S. Horsburgh, **A.A. Melcher**, 2017. High Frequency Measurements Reveal Spatial and Temporal Patterns of Dissolved Organic Matter in an Urban Water Conveyance. *Environmental Monitoring and Assessment.* https://doi.org/10.1007/s10661-017-6310-y

3. **Melcher, A. A.** and J.S. Horsburgh. Regression Methods for Estimating Suspended Solids Concentrations within Urban Stormwater Conveyances. (In Review)

4. Mihalevich, B.A., J.S. Horsburgh, **A.A. Melcher.** Investigating Stormwater Impacts on Dissolved Organic Matter in an Urban Water System. *Urban Ecosystems* (In Review)

5. **Melcher, A. A.** and J.S. Horsburgh. Modeling Phosphorus and Suspended Solids Loads in an Urban Stormwater Conveyance Using High Frequency Data. (to be submitted)

6. **Melcher, A. A.** and J.S. Horsburgh, 2016. Continuous Surrogate Monitoring for Pollutant Load Estimation in Urban Water Systems. The National Nonpoint Source Monitoring Workshop. Salt Lake City, UT. https://water-meetings.tetratech.com/NPS/Content/Docs/32AMelcher.pdf

7. **Melcher, A. A.,** J.S. Horsburgh, B.A. Mihalevich, 2016. Continuous Monitoring for Pollutant Load Estimation in Combined Conveyances. Spring Runoff Conference. Utah State University. http://conference.usu.edu/springrunoff/schedule/griddetails.cfm?aid=675&type=poster

Professional Experience

Water Resources Engineer **January 2018 – January 2019**
Utah Division of Water Resources, Salt Lake City, UT

- Managed the Division's RiverWare model of northern Utah's Weber River basin.
- Analyzed RiverWare model simulations to assess Weber Basin Water Conservancy District's resiliency to drought under various climate change scenarios.
- Developed standalone desktop applications in Python, VB.NET, and R to analyze stream gauge data, create graphics for displaying data to the public, and format model inputs.

Water Resources Software Support Representative **November 2011 – August 2013**
Aquaveo, LLC, Provo, UT

- Provided daily aid to an average of 15 engineers/hydrologists with groundwater, watershed, and surface water models
- Supervised 6 technical support representatives on modeling and customer service techniques
- Offered basic consulting engineering services to software users

Design Engineering Intern **May 2010 – August 2010**
CH2M-WG-Idaho (CWI), Idaho Falls, ID

- Aided in the design of air-tight containment structures to be built over radioactive waste excavation sites

- Verified that design of structures complied with Department of Energy regulations

Teaching and Mentoring

Graduate Teaching Assistant, Hydrology for Engineers **August 2012 – December 2012**
Civil and Environmental Engineering, Brigham Young University

- Assisted 20 senior-level undergraduate engineering students with homework and laboratory report questions
- Graded homework and laboratory report assignments for technical merit and presentation

Vice President/Treasurer **August 2011 – April 2012**
American Society of Civil Engineers BYU Chapter

- Scheduled guest speakers for weekly seminar for all Civil Engineering undergraduate students
- Planned Civil Engineering related service projects and activities for ASCE BYU Chapter
- Maintained the Chapter's budget and costs incurred from activities

Technical Skills

- Programming languages and mathematical packages: Python, MATLAB, R, Basic, C#, Mathcad, JavaScript, C, Bash, HPC
- Software Packages: ArcGIS, HEC-RAS, HEC-HMS, SWMM, WMS, GMS, SMS, ArcHydro, ArcHydro Groundwater

Honors and Awards

Outstanding Poster Presentation **2016**
Utah State University Spring Runoff Conference

- Poster Title: "Continuous Monitoring for Pollutant Load Estimation in Combined Conveyances"

Academic Scholarship **2009, 2012, 2013**
Brigham Young University

- Awarded for academic excellence in both B.S. (3.6/4.0 G.P.A) and M.S. (3.8/4.0 G.P.A.) programs

Languages

English: Native
Italian: Fluent
Spanish: limited working proficiency

Professional Affiliations and Activities

Member of the American Society of Civil Engineers
Member of the Environmental Water Resources Institute