



RESEARCH & DEVELOPMENT

Pre-restoration flow and water quality monitoring of the ‘the canal’ at Claridge Nursery, Goldsboro, NC

**François Birgand, PhD et al.
Department Biological and Agricultural Engineering
North Carolina State University**

NCDOT Project 2013-37

FHWA/NC/2013-37

July 2016

Technical Report Documentation Page

1. Report No. <i>FHWA/NC/2013-37</i>	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Pre-restoration flow and water quality monitoring of the 'the canal' at Claridge Nursery, Goldsboro, NC		5. Report Date <i>July 2016</i>	
		6. Performing Organization Code	
7. Author(s) <i>François Birgand, Chiao-Wen Lin, Bryan Maxwell, Kyle Aveni-Deforge</i>		8. Performing Organization Report No.	
9. Performing Organization Name and Address <i>Department of Biological and Agricultural Engineering 3110 Faucette Drive Raleigh, NC, 27695-7625</i>		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency Name and Address <i>Research and Development Unit 104 Fayetteville Street Raleigh, North Carolina 27601</i>		13. Type of Report and Period Covered Report <i>2013-2016</i>	
		14. Sponsoring Agency Code <i>2013-37</i>	
Supplementary Notes:			
16. Abstract This report constitutes the first phase to document the water quality benefits of a stream restoration in the coastal plain of North Carolina. This phase establishes the pre-restoration state of the water quality and hydrochemical signature of 'the canal' at the Claridge nursery in Goldsboro, NC, for comparison purposes with the during and after restoration phases. Because classical indicators of water quality (i.e., concentrations and loads) are calculated from infrequent samples (e.g., often monthly), these indicators are stained with large uncertainties (20% -50% or even higher). Classical monitoring methods are thus ill suited to detect and quantify water quality benefits of stream restoration, expected to be at or less than 20%. To have a chance to capture water quality benefits of a stream restoration, we have proposed to use state of the art monitoring equipment able to measure flow and concentrations at a high resolution in time, hoping to reduce uncertainties enough to detect and quantify stream restoration effects. We have used UV-Vis spectrophotometers as index data source and created Water Quality Rating Curves to obtain concentration data on a high frequency basis to calculate robust N, P, C, and material fluxes. In articular, we have shown that in the 'Claridge Canal' agricultural stream of the coastal plain of North Carolina, it was possible to construct robust Water Quality Rating Curves to measure nitrate, TDN, TKN, DOC, and TP, using the absorbance data as index data and Partial Least Square Regression (PLSR) as a rating method. We have also shown that this method did not work well for NH4-N and PO4-P. We show that the uncertainty on annual cumulative loads was within $\pm 3\%$ for most parameters. This gives a chance to capture restoration effects that would be at least twice this value.			
17. Key Words <i>Stream restoration; water quality; monitoring; high frequency</i>		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages <i>.. fill in...</i>	22. Price

DISCLAIMER

The contents of this report reflect the views of the author(s) and not necessarily the views of the University. The author(s) are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of either the North Carolina Department of Transportation or the Federal Highway Administration at the time of publication. This report does not constitute a standard, specification, or regulation.

**NCDOT
Research Digest**

Research Project Title: Pre-restoration flow and water quality monitoring of the ‘the canal’ at Claridge Nursery, Goldsboro, NC

Researcher’s Name, affiliation, and contact information:

François Birgand, PhD et al.
Department Biological and Agricultural Engineering
North Carolina State University

Picture *(related to the project):*



Background: (Circumstances that led to research initiation – specific opportunities, challenges, problems, etc.)

There are no accepted metrics to quantify the water quality benefits of stream restoration. As a result, there is no accepted nutrient credit for this practice. This work reports the methods tested to establish the pre-restoration state of water quality of the ‘Claridge canal’ in Goldsboro, NC.

Specific Research Objective:

Establish the benchmark state of water quality in a stream prior to restoration so as to be able to measure the WQ benefits of the restoration.

What research work was done?

We installed state of the art monitoring equipment for flow and concentrations at three stations, upstream, at the middle and downstream of the section to be restored. We have explored and provided calibration methods to best use this equipment. We have provided

evidence that the monitoring methods used are well suited for the task of detecting water quality benefits of stream restoration.

Implementable Research Product:

The eventual goal is to be able to associate nutrient removal credit for stream restoration. This will result from monitoring of the post restoration. The implementable research product currently lies in guidelines derived to install and calibrate water quality sensors.

How will the research product be used:

See above.

Who within NCDOT will use the research product:

Staff in charge of monitoring.

How the use of the research product will benefit the Department:

The department will be the first one to have reliable numbers on stream restoration benefits.

Estimated dollar value of research product (if it can be readily estimated):

NA.

Any other comment:

Executive summary

This report constitutes the first phase to document the water quality benefits of a stream restoration in the coastal plain of North Carolina. This phase establishes the pre-restoration state of the water quality and hydrochemical signature of ‘the canal’ at the Claridge nursery in Goldsboro, NC, for comparison purposes with the during and after restoration phases. Classical indicators of water quality include concentrations and loads, both calculated from samples generally collected infrequently (e.g., monthly). Because they cannot capture the large variability in flow and concentrations these indicators are stained with large uncertainties. They are thought to be suited to detect large effects (e.g., >20%) of a treatment, but not when smaller impacts are expected (e.g., <20%), as surmised in the case of stream restoration. Large uncertainties tend to result from classical sampling and monitoring approaches. Not surprisingly these techniques have failed to provide reliable numbers.

We thus proposed to use different approaches using state of the art high frequency monitoring technology to lower the uncertainty as much as possible. In particular, for this project, we have hypothesized that we could largely reduce uncertainties on nutrient and pollutant loads, which we used to quantify water quality benefits, by using novel continuous water quality sensors. We have proposed to use UV-Vis spectrophotometers, which measure light absorbance from the UV to the visible range. The absorbance data serve as index data source and the creation of Water Quality Rating Curves to obtain concentration data on a high frequency basis to calculate robust N, P, C, and material fluxes.

We have shown that in an agricultural stream of the coastal plain of North Carolina, it was possible to construct robust Water Quality Rating Curves to measure nitrate (NO₃-N), Total Dissolved Nitrogen (TDN), Total Kjeldahl Nitrogen (TKN), Dissolved Organic Carbon (DOC), and Total Phosphorus (TP), using the absorbance data as index data and PLSR as a rating method. We have also shown that this method did not work well for ammonium (NH₄-N) and phosphate (PO₄-P), although it is possible that this was the result of not optimal calibration points pool.

Our results suggest that it possible to reduce uncertainties within $\pm 3\%$ for most constituents for the annual load, which opens the possibility of capturing effects of at least twice the uncertainty level, which corresponds to approximately a 10-fold improvement on current practices. We have shown that to obtain robust Water Quality Rating Curves, it is necessary to have as large as possible a calibration concentration range, where calibration concentrations are as stratified as possible. Extrapolating beyond the calibration range is inherently risky and may result in errors, and in the case of parameters that exhibit concentration effect during events, large overestimation as we showed for TSS predicted using PLSR. For Total Suspended Solids (TSS), we suggest that it is preferable to use the more conservative turbidity based method.

We thus believe that the methods developed during the pre-restoration period of this stream restoration are very promising to detect the effects of water quality of such a practice in the Coastal Plain of North Carolina.

CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER 1: Introduction.....	xii
CHAPTER 2: Application of ultraviolet-visual spectrophotometers and water quality rating curves to monitor concentrations and loads on a continuous basis in the ‘Claridge canal’, an agricultural coastal plain stream	17
Background.....	17
Objectives of this chapter.....	17
2.2 Methods.....	18
Site description.....	18
Flow measurements	19
Water samples collections and measurements	21
Installation of the monitoring system	22
Equipment maintenance.....	23
Laboratory analyses	24
Cumulative load calculations.....	25
Calibration methods for concentrations from spectral data	26
Methods for correcting erroneous data and filling missing data.....	27
2.3 Results and Discussion	28
Developing index velocity ratings to calculate flow.....	28
Degradation sample study results	28
Discrete sample results	29
Global Calibration versus discrete sample results	30
Large improvement provided by the PLSR method on nitrate and DOC concentrations	33
It is possible to go beyond the instrument advertised capabilities and create Water Quality Rating Curves for most parameters tested	34
TSS Predictions from PLSR and Turbidity	36
Discussion: can we reliably measure nutrient concentrations and loads with Water Quality Rating Curves?.....	37
Beyond water quality rating curves: additional local concentration corrections	38
Results for filling missing flow and concentration data	40

Challenges and maintenance for the application of UV-Vis spectrophotometers <i>in situ</i>	42
2.4 Conclusion	43
2.5 References.....	45
CHAPTER 3: Uncertainties assessment for continuous and infrequent sampling in measuring nutrient loads with the application of UV-Visual spectrophotometers in an agricultural coastal plain stream	53
3.1 Introduction.....	53
Water quality monitoring <i>in situ</i> probes uncertainties need assessing	53
Research questions and hypotheses	54
Objectives	55
3.2 Methods.....	55
Resampling the existing calibration points.....	57
Uncertainties estimation on annual loads as hydrological indicator with resampling....	58
Uncertainties estimation on annual loads as hydrological indicator with infrequent sampling.....	59
Filling missing values in the spectral results	61
3.3 Results and Discussion	61
Redundant potential in the annual spectral data	61
The method used allows for very small uncertainties compared to standard monitoring methods.....	62
To obtain good results, the <i>in situ</i> WQ sensors need to be properly calibrated.....	64
3.4 Discussion and conclusions	69
3.5 References.....	72
CHAPTER 4: General Conclusion	80
Ability to create Water Quality Rating Curves for UV-Vis spectrophotometers	81
Reduce errors on measuring loads with using UV-Vis spectrophotometers and the most robust sampling method for affordable number of calibration points	82
Costs comparisons between continuous monitoring and stream restoration projects	84
4.2 References.....	87

LIST OF TABLES

Table 2.1 Summary of the laboratory analyses for discrete samples (Units: mg/L; concentrations with significant ‘degradation’ not included)	29
Table 2.2 Summary of linear regressions between Global Calibration and discrete samples results for NO ₃ -N and DOC among the three monitoring stations	32
Table 2.3 The summary of PLSR calibration for NO ₃ -N, NH ₄ -N, TDN, TKN, DOC TP, and PO ₄ -P at the three stations (UP, MD, DN).....	35
Table 2.4 Summary of NO ₃ -N cumulative loads for Manual method and Semi-Automatic method.....	42
Table 3.1 Methods applied to calculate the annual loads in this research (Adapted from Johnes, 2007; Birgand et al., 2010; Cassidy and Jordan, 2011)	60
Table 3.2: Uncertainty estimation expressed in percentage error compared to the reference annual load for the bootstrap method over a range of parameters	62
Table 3.3 Summary of calculated uncertainties for all parameters, stations and method leading to lowest uncertainty	68
Table 4.1 Costs analysis for the stream restoration projects (Unit: \$/ft.)	86

LIST OF FIGURES

Figure 2.1 Three monitoring stations along the Claridge Canal in Goldsboro, North Carolina (I: Contributed watershed area for UP station; I and II: Contributed watershed area for MD station; I, II, and III: Contributed watershed area for DN station) 19

Figure 2.2 Schematic of a trapezoidal flume cross-section, showing the Doppler velocity meter, the virtual vertical column and the location of the manual velocity measurement points (marked as +; Modified from Birgand, 2000)..... 21

Figure 2.3 Schematic of the elements of the water quality monitoring system (not in the scale; numbers detailed in text; Modified from Etheridge et al., (2013))..... 23

Figure 2.4 The linear regression relationship between Global Calibration and Discrete Samples for NO₃-N at UP. (Black solid line: 1 to 1 line; Black dash line: Regression line) . 31

Figure 2.5 The linear regression relationship between Global Calibration and Discrete Samples for DOC at UP. (Black solid line: 1 to 1 line; Black dash line: Regression line) 32

Figure 2.6: Large improvement in the prediction of Nitrate and DOC using the PLSR method..... 33

Figure 2.7: Improvement of the nitrate concentrations predicted by the PLSR method, compared to the ones predicted by the ‘global calibration’ directly from the instrument..... 34

Figure 2.8 An example of “stray” data points because of biofouling. Continuous NO₃-N concentrations from PLSR calibration (black line), NO₃-N concentrations from Discrete Samples (red dots), Flow rate (blue line), and Field visit date (gray vertical line) at UP on February 5-20 in 2014..... 39

Figure 2.9 Continuous NO₃-N concentrations from PLSR calibration with correction by using AQUARIUS software (black line), NO₃-N concentrations from Discrete Samples (red dots), Flow rate (blue line), and Field visit date (gray vertical line) at UP on February 5-20 in 2014..... 40

Figure 3.1: Distribution of uncertainties of annual loads for nitrate (top row; A and B) and DOC (bottom row ; C and D) corresponding to 24 stratified (left column; A and C) and random (B and D) samples used to create the calibrations 65

Chapter 1: Introduction

As stated by Bernhardt et al., (2005), there was 37,099 stream restoration projects implemented in the United States, and the top goals of restoration included improving water quality, developing or enhancing riparian zones, creating/improving instream habitat, facilitating fish passage, and stabilizing stream banks. However, the lack or lack of poor availability of data to make a robust assessment of the, and the reasons for, success or failure of stream restoration projects (e.g., Pander and Geist, 2013; Bennett et al., 2011). In most cases, the stream restoration effects are estimated, after restoration, by comparison between the restored and nearby ‘reference’ reaches (e.g., Colangelo, 2014; Daniluk et al., 2013; Howson et al., 2009), which can only add to the uncertainty of the estimation process. Comparing pre- and post-restoration, theoretically removes this type of uncertainty from the evaluation method, and we too hypothesize it should be the preferred approach. Using the mathematical analogy, discrete concentrations and load values correspond to the derivative values needed to calculate the integral or overall water quality benefits.

One would like to be able to claim that for this or that type of restoration, water quality has been improved by, e.g., ‘x%’. In the expression ‘x% water quality benefit’, the notion of integration over time is embedded. For very practical reasons, a common approach to estimate water quality benefits has been to compare the discrete sample concentrations and corresponding instantaneous loads over short and discontinuous periods of time. ² In summary, we believe that the concept of water quality benefits of stream restoration often refers to a concept of integral or cumulative effect. However, one has historically had access only to points in time, i.e., discrete flow and concentration values, corresponding to the derivative functions necessary to calculate the integral. Numerous authors have shown that

the uncertainty on material loads associated with infrequent sampling can be such that it is likely that the effects of stream restoration are confounded within the measurement uncertainty (e.g., Skarbøvik and Roseth, 2015; Skarbøvik et al., 2012; Cassidy and Jordan, 2011; Henjum et al., 2010; Birgand et al., 2010; Brauer et al., 2009).

New available continuous water quality sensors give access for the first time to the full dynamics of concentrations, or when combined with continuous flow measurements, the derivative functions needed to calculate the cumulative load, which information is necessary to have a chance to evaluate water quality benefits of stream restoration. The recent availability of continuous optical sensors, ultraviolet-visual spectrophotometers, for water quality has opened the possibility to obtain high frequency water quality data which can reveal concentration dynamics that were not available until now (e.g., Langergraber et al., 2003; Rieger et al., 2006, Etheridge et al., 2013).

However, these instruments are relatively new and a very limited return on experience is currently available for guidance. As a result, it is unclear whether these field spectrophotometers can be used for the conditions in the agricultural coastal plain stream, and what kind of performance one should expect. Etheridge et al. (2014, 2015) have shown that these instruments are subject to optical fouling, but these results were obtained in a tidal marsh, hence in very different conditions. The same authors have shown that it was possible to use the spectrophotometers to measure concentrations of parameters not known to absorb light. Etheridge et al. (2015) have essentially proposed to create water quality rating curves where the spectrophotometer absorbance data are used as index data from which concentrations of a suite of parameters can be calculated using a rating curve, itself generated using Partial Least Square Regression (PLSR) statistical methods (Etheridge et al., 2015).

To fill in the voids in the literature on the water quality benefits of stream restoration, North Carolina Department of Transportation (NC DOT), North Carolina Forest Service (NCFS) and North Carolina State University (NCSU) have partnered together to conduct a pilot study. The general approach to quantify the water quality benefits pursued consists in quantifying the changes in the stream solutes and material signatures of the “same” stream from before to after restoration with continuous water quality and hydrology monitoring.

1.2 Research Questions

This research attempted to answer the following questions. The first question associated with the continuous water quality data is: Can one use in situ ultraviolet-visual (UV-vis) spectrophotometers to measure water quality on a continuous basis in an agricultural coastal plain streams? The corollary question becomes: for what parameter and what does it take to obtain continuous data? The answers to these very important questions are necessary to enrich guidance in the use of these instruments and create Water Quality Rating Curves to interpret the absorbance data (Chapter 2).

The second question is: What are the uncertainties on concentrations and loads associated with data generated by the spectrophotometers? It is expected that the uncertainty on e.g., annual loads be much lower than when calculated from infrequent samples. The corollary questions become: How do the uncertainties associated with the instruments compare with those associated with infrequent sampling on the annual load indicator? The answers related to the second question is to quantify the uncertainties associated with the spectrophotometers itself and compare the uncertainties for estimating annual loads with infrequent sampling (Chapter 3).

The third research question is addressed as following: What does this continuous water quality information on the processes generating nutrient load in a coastal plain predominantly agricultural watershed, at the short term (e.g., storm events) and seasonal scales? Can the information obtained be used to infer about in-stream processes? The answers to the above questions are to characterize the nutrient dynamics from the continuous water quality information with seasonal effects and during the storm events (Chapter 4).

1.3 Objectives

To address our research questions, we monitored water quality and hydrology with high-frequency of the Claridge Canal due to be restored in Goldsboro, NC. In Chapter 2, we present the application of ultraviolet-visual spectrophotometers to monitor water quality on a continuous basis at three stations along the Claridge Canal for 16 months. The objectives of Chapter 2 are: (1) to evaluate the possibility of creating water quality rating curves using UV-Vis spectrophotometers to measure nitrate (NO₃-N), ammonium (NH₄-N), total dissolved nitrogen (TDN), Total Kjeldahl Nitrogen (TKN), dissolved organic carbon (DOC), phosphate (PO₄-P), total phosphorus (TP) and total suspended solids (TSS) in an agricultural coastal plain stream: (2) Report the potential challenges involved, and propose methods to obtain gap free time series and best results.

In Chapter 3, we assess the level of uncertainties associated with the use of such technology and how these compare to those induced by infrequent sampling on nutrient loads and concentrations. The objectives of Chapter 3 were addressed as follows: (1) Evaluate the uncertainties on concentration values and annual loads calculated by the ‘global calibration’ provided by the manufacturer, and calculated using PLSR depending upon the number and the distribution of the calibration points used. (2) Compare these uncertainties to those that

would be induced by infrequent sampling on estimating annual loads with best load algorithm methods; (3) Provide guidance to users to minimize uncertainties and maximize the potential of their UV-Vis spectrophotometers.

1.4 References

- Bennett, S. J., Simon, A., Castro, J. M., Atkinson, J. F., Bronner, C. E., Blersch, S. S., Rabideau, A. J. 201. The evolving science of stream restoration. *Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools* Geophysical Monograph Series 194, Published in 2011 by the American Geophysical Union.
- Bernhardt, E. S., Palmer, M. A., Allan, J. D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G. M., Lake, P. S., Lave, R., Meyer, J. L., O'Donnell, T. K., Pagano, L., Powell, B., Sudduth, E. 2005. Synthesizing U.S. river restoration efforts. *Science*, 308:636-637.
- Birgand F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., Bordenave, P. 2010.
- Brauer N., O'Geen, A. T., Dahlgren, R. A. 2009. Temporal variability in water quality of agricultural tailwaters: implications for water quality monitoring. *Agr Water Manage*, 96: 1001-1009.
- Cassidy, R. and Jordan, P. 2011. Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. *J Hydrol*, 405:18-193.
- Colangelo, D. J., 2014. Interim Response of Dissolved Oxygen to Reestablished Flow in the Kissimmee River, Florida, U.S.A. *Restoration Ecology*, 22: 376-387.

Chapter 2: Application of ultraviolet-visual spectrophotometers and water quality rating curves to monitor concentrations and loads on a continuous basis in the 'Claridge canal', an agricultural coastal plain stream

Background

One of the robust indicators which should be chosen to assess the water quality benefits of this stream restoration uses cumulative nutrient loads as a basis. We hypothesize that the bulk water quality effect can be assessed by the magnitude of the inflection or change of the correlation between cumulative loads of a treatment station vs that of a control station, between before and after restoration. A station just upstream of the restored section can serve as a control while the treatment stations can correspond to station within and also just downstream of the restored reach. But to obtain these cumulative curves, gap free flow and concentration data are necessary at all three stations at the same time. The continuous monitoring of hydrology and water quality reported in this report took place from December 2013 to March 2015.

Objectives of this chapter

The objectives of this research are: (1) to evaluate the possibility of creating water quality rating curves using UV-Vis spectrophotometers to measure nitrate ($\text{NO}_3\text{-N}$), ammonium ($\text{NH}_4\text{-N}$), total dissolved nitrogen (TDN), Total Kjeldahl Nitrogen (TKN), dissolved organic carbon (DOC), phosphate ($\text{PO}_4\text{-P}$), total phosphorus (TP) and total suspended solids (TSS) in an agricultural coastal plain stream: (2) Report the potential challenges involved, and reports methods to obtain gap free time series and best results.

2.2 Methods

Site description

The study site is the Claridge Canal, an agricultural coastal plain stream reach 2,200 m in length, in Goldsboro, North Carolina (35.42° N, 78.02° W). This reach was slated to be restored from 2015 to 2016 for 12 months starting in autumn to offset the construction of Highway US70 Bypass. Three monitoring stations were installed as Upstream Station (UP), which serves as the control station, Middle Stream Station (MD, located 1,390 m from UP), and Downstream Station (DN, located 807 m from MD), both corresponding to the treatment stations, for a total length of 2,197m of the study reach (Figure 2.1). The drainage areas at UP, MD, and DN are 236 ha, 414 ha, 573 ha, respectively. The average slope of the stream is 1.93%.

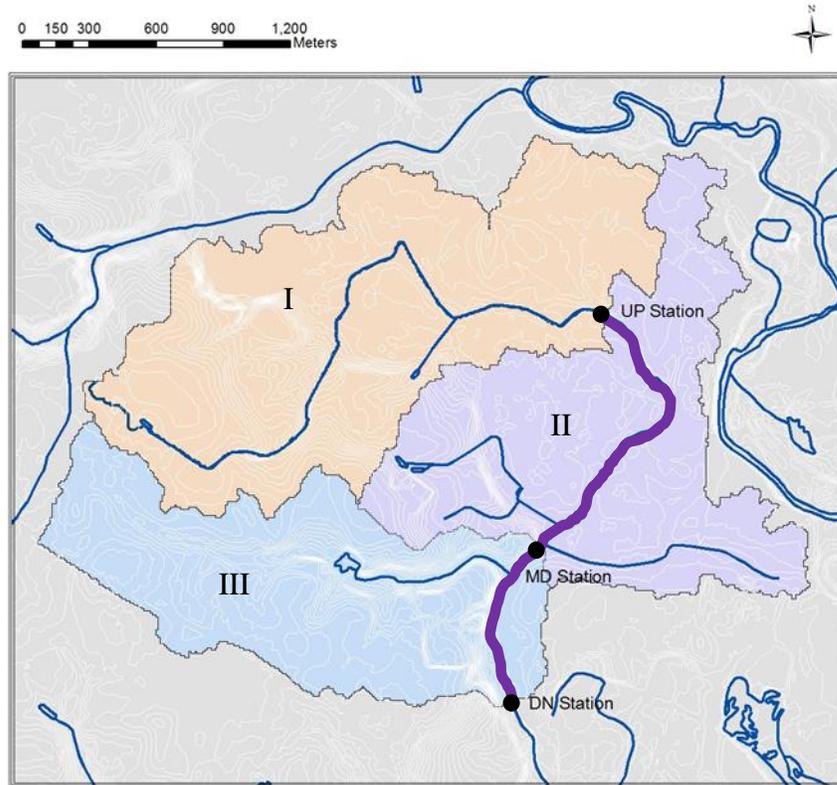


Figure 2.1

Three monitoring stations along the Claridge Canal in Goldsboro, North Carolina (I: Contributed watershed area for UP station; I and II: Contributed watershed area for MD station; I, II, and III: Contributed watershed area for DN station)

Flow measurements

Because of the low gradient of the stream, variable downstream control conditions can create large hysteresis in the stage-discharge relationships at the event and seasonal scales (e.g., Birgand et al., 2013). Stage measurements alone are thus insufficient to calculate discharge. To solve this problem, we measured stage and velocity using Doppler velocity meters (SonTek IQ Doppler meter, San Diego, CA, USA) installed in wooden trapezoidal

flumes. The trapezoidal flumes (details in Etheridge et al., 2013, and Birgand et al., 2005) funnel water unto a structure of known and stable geometries. The discharge (flow) rate (Q) is calculated as the product of cross-sectional average velocity (V) with the cross-sectional area (A) (Equation 1).

$$Q = V \times A \quad (\text{Equation 1})$$

The cross-sectional area is calculated using measured stages and the known dimensions of the flumes. The cross-sectional average velocity V is calculated using the index velocity rating (Morlock et al., 2002; Birgand et al., 2005; ISO 15769, 2010). An index velocity rating was established between the Sontek velocity data and the cross-section average velocities obtained during manual gauging.

Manual gauging was performed using the velocity area method (ISO 748, 1997) every field visit, except in non-wadable conditions. For that, the cross-section of the flume was divided into several vertical columns, marked by PVC strips were installed along the sides and on the bottom of the flume to mark the center of each column (Figure 2.2). The point velocity, measured using a portable velocity meter (Marsh-McBirney, Frederick, MD, USA; resolution 1 cm/s), in each column at 40 percent of the depth was taken as the column mean velocity (Figure 2.2). The manual mean velocity (V_m) was eventually calculated from dividing the overall flow rate, i.e., the sum of flow in each column, by the total cross-sectional area of the flume. The dimensions of the flumes at the three stations were measured using a total station after the flumes were installed. The index velocity from the Sontek used

to derive the velocity rating corresponded to the average of instantaneous measurements 1-3 hours before and after each manual gauging.

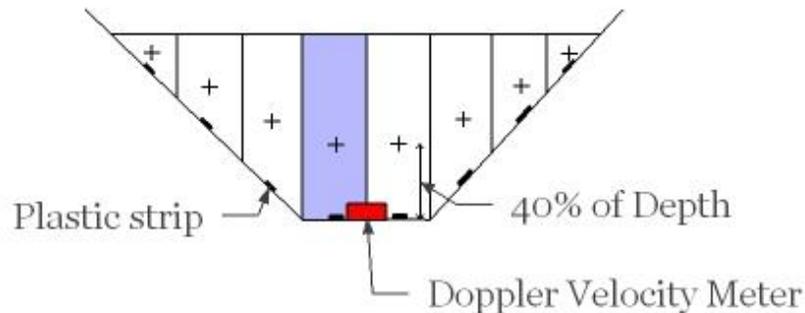


Figure 2.2

Schematic of a trapezoidal flume cross-section, showing the Doppler velocity meter, the virtual vertical column and the location of the manual velocity measurement points (marked as +; Modified from Birgand, 2000).

Water samples collections and measurements

Water quality at each station was monitored using a UV-Vis spectrophotometer (spectro::lyser model, s::can), a multi-parameter sonde (Eureka Manta 2), and an automatic discrete sampler (ISCO 6712). The UV-Vis spectrophotometer measured the absorbance of light between 220 to 750 nm for 5 and 35 mm path lengths. Thanks to the embedded Global Calibration from the manufacturer the instrument provided Turbidity (NTU), nitrate ($\text{NO}_3\text{-N}$; mg/L), DOC (mg/L), and total organic carbon (TOC; mg/L) measurements every 15 minutes. At the same time interval, the Manta sonde measured Temperature ($^{\circ}\text{C}$), pH, conductivity ($\mu\text{S/cm}$), turbidity (NTU), and colored dissolved organic matter (CDOM; $\mu\text{g/L}$). The

automatic discrete sampler took discrete water samples, up to 1 L every 12 hours (12 am and 12 pm). All field instruments were serviced on a biweekly basis.

Installation of the monitoring system

Although the water quality sensors are designed to be immersed, previous studies have shown that the reduced conditions surmised to exist in coastal plain streams may result in large chemical and biological fouling of the instruments (Etheridge et al., 2013). To limit the magnitude of fouling, we followed the method developed by Etheridge et al., (2013), i.e., to shorten the time of exposure of the instruments to water (Figure 2.3). In synchrony with the 15 minutes sensor measurement intervals, a micro controller (equipped with Arduino Uno Board; www.arduino.cc) (4) triggered a peristaltic pump (3) to draw stream water (1) to the multi-parameter sonde (5) and to the UV-Vis spectrophotometer (7), both of which fitted with flow through cells. The spectrophotometer was installed in a container (6) to avoid ambient light disturbance. After the sensors took their measurements, water was purged out of the measurement cells through a drain valve (8) under the spectrophotometer, and using the peristaltic pump (3) in a reverse mode back to the stream. The micro controller then sent a signal to a windshield washer pump (10) and valve (11) to rinse the optics the UV-Vis spectrophotometer (7) using tap water stored in a container (9) and resupplied during field visits. At 12 am and 12 pm every day, the discrete sampler (2) sampled up to ~900ml of water from the stream into the 24 bottles available.

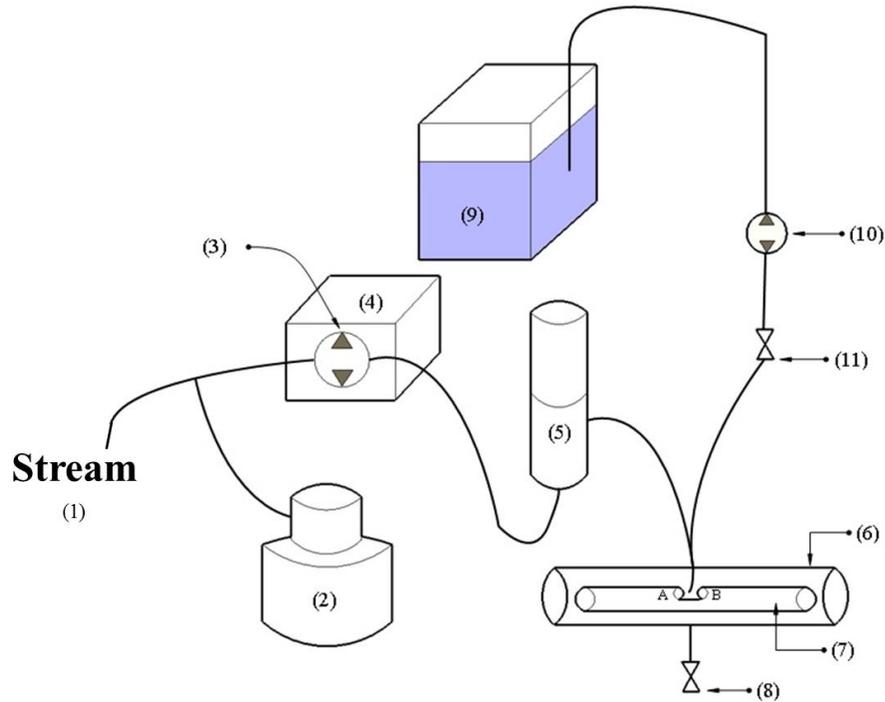


Figure 2.3

Schematic of the elements of the water quality monitoring system (not in the scale; numbers detailed in text; Modified from Etheridge et al., (2013))

Equipment maintenance

Biweekly field visits included flow and water quality data download, as well as other maintenance. The flumes and Doppler velocity meters were cleaned using brushes to remove sand, which was brought after storm events and clean algae that grew at certain times of the year. Because of bio- and chemical fouling on instruments and optics in particular (Flemming, 2011; Etheridge et al., 2014; Whelan and Regan, 2006), a cleaning procedure was applied onto the UV-Vis spectrophotometer and multi-parameter sonde based on that developed by Etheridge et al., (2014). For the UV-Vis spectrophotometers, we applied

cotton swabs soaked in 2% hydrochloric acid (HCl) to the optic lens for 2-3 minutes. Then, we rinsed the lenses with deionized water and dried with a cloth. This procedure was repeated until the absorbance spectrum values in air and deionized water measurements were ‘flat’ and within 10 m^{-1} , i.e., similar to the original instrument setting. For the multi-parameter sonde, we used a detergent and toothbrush to clean the sensors except for CDOM sensor. CDOM sensor was cleaned with 2 % HCl, a cotton swab and deionized water; the pH reference solution for pH sensor was changed every 2 months. The sensors in multi-parameter sonde were calibrated with standard solutions every 2-3 months.

Laboratory analyses

For each field visit, the discrete water samples were stored on ice and transported to the laboratory for analysis. The selections of the discrete samples for laboratory analysis depended on the initial water quality results from UV-Vis spectrophotometer and the flow events, i.e., the samples taken during flow events when concentrations change rapidly were preferentially selected.

The original 900 ml discrete samples were split into several aliquots for analyses. 500 mL were transferred for TSS analysis (Standard Method 2540 D; Rice et al., 1997a), and 150 mL were acidified with sulfuric acid for TKN (Standard Method 4500 N Org D; Rice et al., 2012) and TP analyses (Standard Method 4500 P F; Rice et al., 2012). The residual of the volume was filtered (pore size of $0.22 \mu\text{m}$; EMD Millipore Sterivex Sterile Pressure-Driven Devices, Darmstadt, Germany). Filtered solutions of 14 mL were transferred and acidified with sulfuric acid in 15 mL-centrifuge tubes for DOC (Standard Method 5310 B; Greenberg

et al., 2005) and TDN analysis (adapted by Standard Method 5310 B; Greenberg et al., 2005). Another 14 mL of filtered solutions without adding sulfuric acid were transferred in 15 mL-centrifuge tubes for NO₃-N (The Cadmium Reduction Method; Eaton et al., 1995), NH₄-N (Salicylate Method; Eaton et al., 1995) and PO₄-P analyses (Ascorbic Acid Method; Eaton et al., 1995).

We chose not to add acid or other preservation agents to unrefrigerated samples in the field. We ran a sample degradation study instead. For this, two grab samples were collected at the same time. One was labeled “Fresh”, immediately put on ice and brought back to the laboratory for analysis, and the other one was labeled “Degraded”, left in the sampler until the next field visit, when it was retrieved and analyzed following the procedure described above (TSS not analyzed).

Cumulative load calculations

The water quality indicator sought in this manuscript is cumulative load (L), which is the integral over a given time (t) of instantaneous loads calculated as the product of solute concentration (C) with flow rate (Q ; Equation 2). In reality, a good estimator of this integral is the summation of instantaneous 15-min loads calculated from the 15-min flow (Q_i) and concentration (C_i) data (Equation 3), where K is the unit adjustment factor and N is the number of 15 min intervals in t .

$$L = \int_0^t C(t)Q(t)dt \quad \text{(Equation 2)}$$

$$L \approx K \sum_{i=1}^{i=N} C_i Q_i \quad (\text{Equation 3})$$

Calibration methods for concentrations from spectral data

The new water quality sensor probes used in this project are among the best probes available out there, or at least that we are aware of. Contrary to chemistry based instruments, which have physical standard samples for calibrations, concentrations measured from spectrophotometers are calculated from a proprietary algorithm embedded in the probe. The algorithm is the equivalent of calibration of chemistry instruments, which calculates, using a pmay or may not fit well with local conditions. The ‘Global Calibration’ is the default algorithm embedded in the Spectro::lyser instrument to calculate concentrations *in situ* from the absorbance spectra. Nitrate, DOC, Total Organic Carbon (TOC) concentrations, and Turbidity values are thus calculated and saved in an output PARAmeter file, referred to as the PAR file. We compared the results from the PAR files to those of the discrete sample concentrations.

We also used PLSR as a regression model to correlate concentrations from the discrete samples to the spectral data (from UV-Vis spectrophotometers) measured at the same time. We then used the PLSR model to predict 15 min concentrations from the 15-min spectral data. PLSR essentially decreases the hundreds of wavelengths of spectral data to a smaller number of principal component vectors to obtain highest correlation with water quality concentrations (Etheridge et al., 2013). This statistical technique is well suited for data in

which the explanatory variables are highly auto correlated, as are absorbance values from sequential wavelengths (Mevik et al., 2011). PLSR, thus corresponds to our local calibration, with the great advantage that with one sensor, it is possible to calibrate for more than one nutrient or pollutant.

We followed the procedures for PLSR calibration described in Etheridge et al. (2013) and Langergraber et al. (2003), using the pls package (Mevik et al., 2011) in the R software (R Core Team, 2016). Briefly, this package takes 90% of the points to create a calibration model and cross validates on the remaining 10%, and calculates indicators such as Root Mean-Square Error of Prediction (RMSEP) and R^2 for a given number of components. The model was initially run for 20 components. The optimum number of PLSR components for each analysis was chosen as the lowest number of components for which the RMSEP was at or near its minimum value.

We hypothesize just as Etheridge et al. (2013) did, that this method essentially quantifies the co-variability of the ‘color matrix’ of the water with concentrations. We hypothesize that this co-variability has no theoretical reason to hold for all seasons and all hydrological conditions. We thus explored several models where in one all points were used, and in others where seasonal models were derived.

Methods for correcting erroneous data and filling missing data

Human errors and equipment failure invariably result in obvious errors and gaps in flow and/or concentration data in the field. Since both signals correspond to the derivative functions necessary to calculate the cumulative loads, it is essential to be able to find

solutions to correct for erroneous and fill in missing data. The details of this part would be lengthy in this report and can be found in Lin (2017).

2.3 Results and Discussion

Developing index velocity ratings to calculate flow

The Sontek velocimeter sensors tended to overestimate velocities by 2.5%, 7.0%, and 1.0% , the cross-section average velocities, for the UP, MD, and DN stations, respectively, justifying the need for correction and harmonization among stations. Details of the equations used can be found in Lin (2017).

Degradation sample study results

The degradation study showed that during Spring and Summer, ammonium and TKN could evolve over two weeks within the sampler. The consequence of this analysis is that when there was a significant difference ($\alpha=0.05$) between ‘fresh’ and ‘degraded’ samples for the water quality parameters in spring and summer (**Error! Reference source not found.**), only the discrete samples collected within 48 hours prior to the field date were used to improve the quality of PLSR calibration (suggested by Etheridge et al., 2013). All the discrete samples collected in fall and winter were used in the PLSR calibration, as no significant degradation was found at these times. All additional details in Lin (2017).

Discrete sample results

The laboratory analysis for discrete samples applied to PLSR calibrations are summarized in Table 2.1 which included the number of samples, mean concentrations, standard deviation, minimum concentrations, and maximum concentrations for NO₃-N, NH₄-N, TDN, TKN, DOC, TP, PO₄-P, and TSS among the three stations.

Table 2.1
Summary of the laboratory analyses for discrete samples (Units: mg/L; concentrations with significant ‘degradation’ not included)

	NO ₃ -N	NH ₄ -N	TDN	TKN	DOC	TP	PO ₄ -P	TSS
UP								
No. samples	230	67	228	218	213	216	138	212
Mean± Standard deviation	2.98±0.61	0.19±0.16	3.50±0.63	0.61±0.29	5.03±1.60	0.09±0.15	0.02±0.03	16.46±32.74
Minimum	1.0	0.10	1.8	0.28	2.6	0.01	0.01	0.33
Maximum	4.2	1.2	4.8	2.2	13.4	1.3	0.36	367
MD								
No. samples	102	79	101	98	102	58	54	81
Mean± Standard deviation	2.60±0.59	0.14±0.06	3.16±0.54	0.65±0.25	6.10±2.12	0.08±0.05	0.01±0.02	21.51±31.47
Minimum	1.1	0.10	1.7	0.3	3.0	0.02	0.01	2.0
Maximum	3.6	0.41	4.1	1.3	12.5	0.25	0.14	218
DN								
No. samples	176	72	176	161	193	159	130	166
Mean± Standard deviation	2.41±0.50	0.20±0.12	2.96±0.52	0.71±0.27	6.14±1.88	0.11±0.10	0.02±0.01	23.74±29.70
Minimum	0.8	0.10	1.5	0.3	3.7	0.02	0.01	1.0
Maximum	3.4	0.85	4.0	1.9	14.7	0.65	0.10	250

(Note: The detection limit is 0.1 mg/L for NO₃-N, NH₄-N, TDN, and DOC; 0.28 mg/L for TKN; 0.01 mg/L for TP and PO₄-P; no detection limit for TSS.)

Global Calibration versus discrete sample results

The global calibration that came with the instrument was used directly without local calibration and compared with discrete concentrations. The results show significant linear correlations between the discrete and the raw instrument values, but show large departure from the one to one line (Figure 2.4 and Figure 2.5; Table 2.2). Additionally, the RMSE of each regression are equivalent to 10% to 25% of the absolute concentrations.

This entirely justify the need to find methods to create local calibration that can correct for slope, intercept and RMSE; hence the PLSR approach suggested by Torres and Bertrand-Krajewski (2008), Rieger et al., (2006), Langergraber et al. (2003), and Etheridge et al. (2014, 2015) to improve predictions for water quality concentrations.

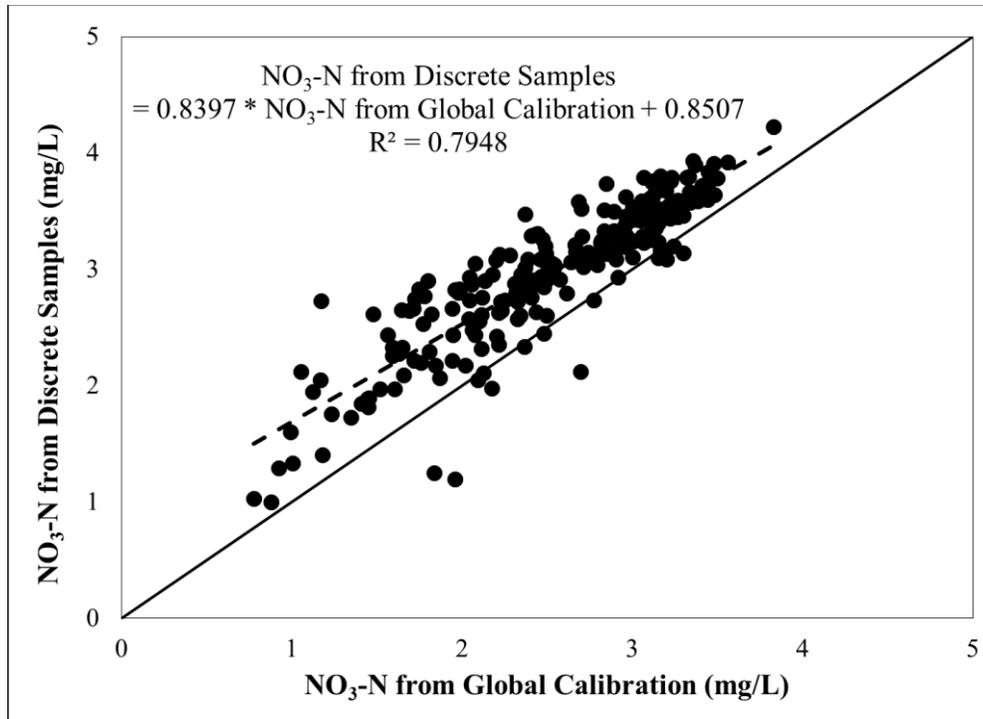


Figure 2.4

The linear regression relationship between Global Calibration and Discrete Samples for NO₃-N at UP. (Black solid line: 1 to 1 line; Black dash line: Regression line)

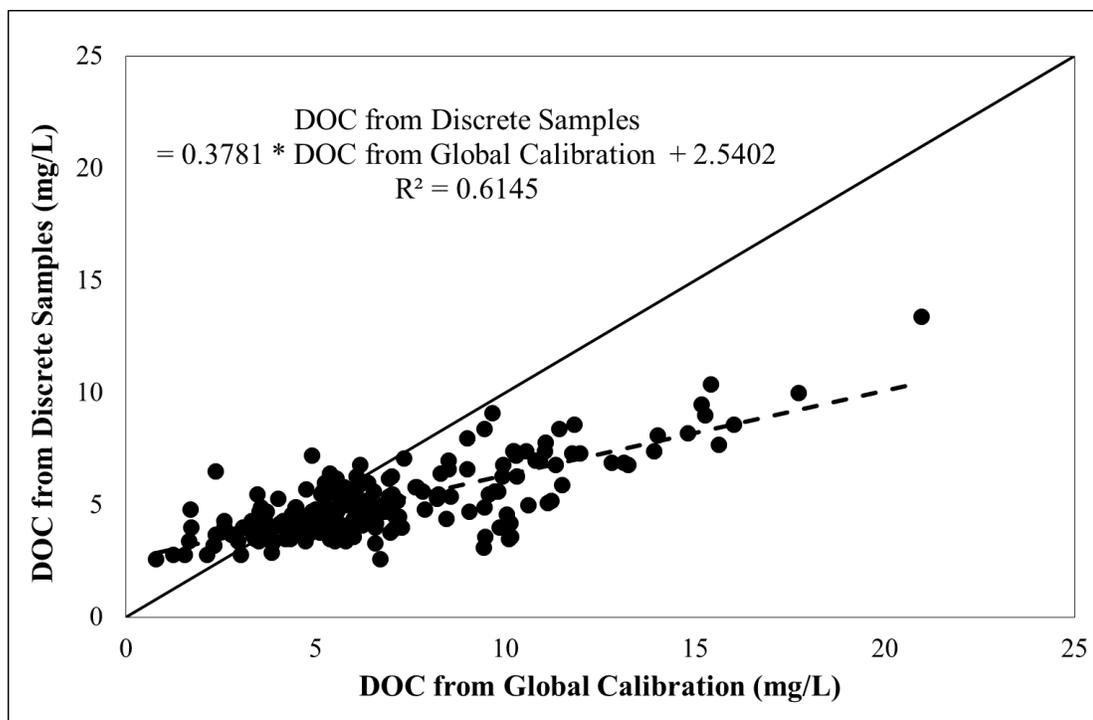


Figure 2.5

The linear regression relationship between Global Calibration and Discrete Samples for DOC at UP. (Black solid line: 1 to 1 line; Black dash line: Regression line)

Table 2.2

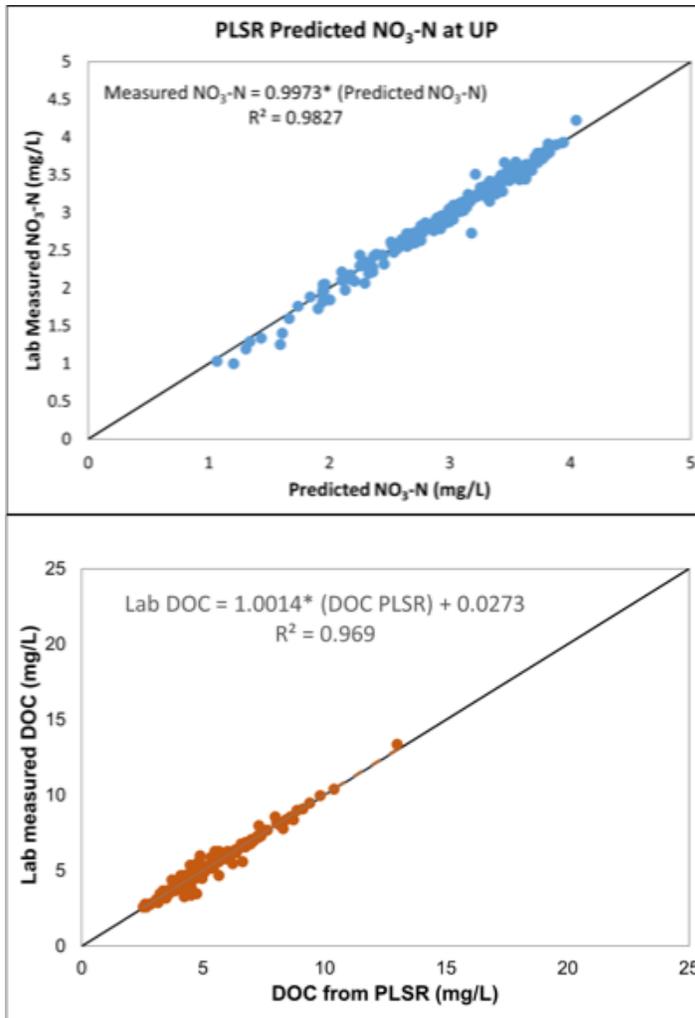
Summary of linear regressions between Global Calibration and discrete samples results for NO₃-N and DOC among the three monitoring stations

	NO ₃ -N				DOC			
Station	No. Observations	Regression Equation	R ²	RMSE	No. Observations	Regression Equation	R ²	RMSE
UP	230	Y = 0.8397 X + 0.8507	0.79	0.53	213	Y = 0.3781 X + 2.5402	0.61	2.78
MD	93	Y = 0.7438 X + 0.6002	0.87	0.27	91	Y = 0.6975 X + 1.7408	0.72	1.00
DN	154	Y = 0.5547 X + 0.9488	0.73	0.46	167	Y = 0.5577 X + 1.9064	0.65	1.56

(Note: 1. Y = Discrete Samples Results, and X = Global Calibration.)

Large improvement provided by the PLSR method on nitrate and DOC concentrations

Using the method developed by Etheridge et al. (2014, 2015), we have been able to create much better calibrations than the global calibration (Figure 2.6).



• NO₃-N

Station	Regression Equation	R ²
UP	Y = 1.00 X	0.98
MD	Y = 1.01 X	0.98
DN	Y = 1.00 X	0.98

• DOC

Station	Regression Equation	R ²
UP	Y = 1.01 X	0.97
MD	Y = 0.99 X	0.92
DN	Y = 1.00 X	0.94

Figure 2.6: Large improvement in the prediction of Nitrate and DOC using the PLSR method.

The improvement is quite obvious from the differences in chemographs between the instrument concentrations and from the PLSR generated concentrations (Figure 2.7).

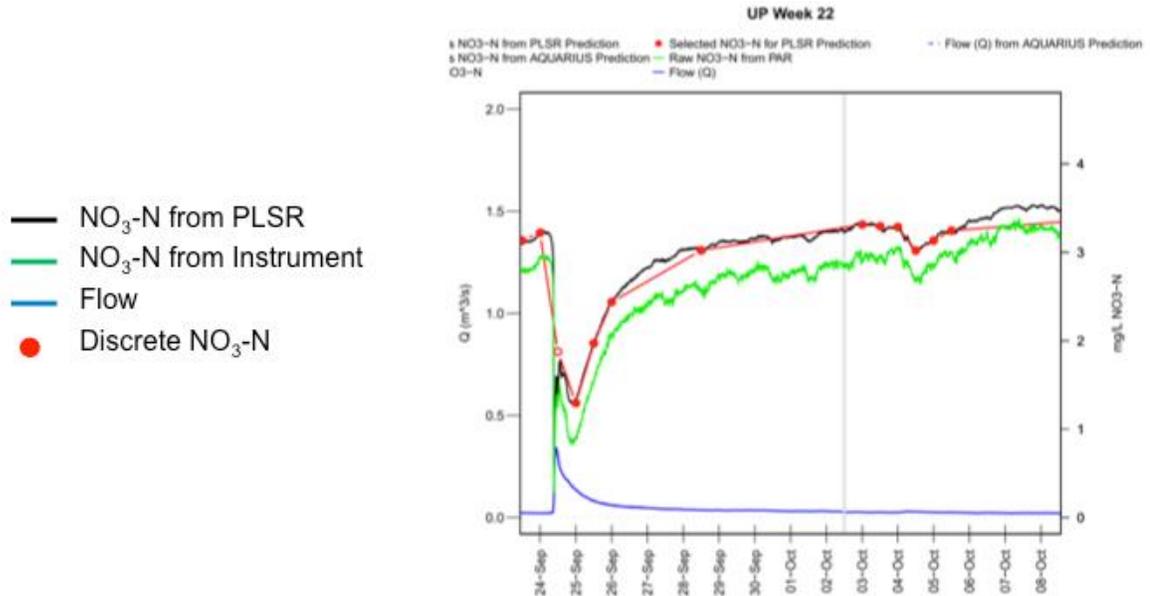


Figure 2.7: Improvement of the nitrate concentrations predicted by the PLSR method, compared to the ones predicted by the ‘global calibration’ directly from the instrument

It is possible to go beyond the instrument advertised capabilities and create Water Quality Rating Curves for most parameters tested

Our main hypothesis in this chapter was that the spectrophotometers could be used to predict not only nitrate, DOC and TSS, as advertised by the manufacturer, but go beyond that and use the PLSR method to create calibration for other parameters.

The details of the exact results are available in Lin (2017). We feel important to report here that the out of the 42 PLS regressions for the 7 parameters tested, 3 stations and 2 major climatic periods (fall-winter and spring-summer), 31 of them were found to predict

concentration on a robust basis. Eleven (11) were found not to predict concentration on a robust basis, these include mostly ammonium and phosphate. We used several criteria to evaluate the robustness of the predictions, the main one of which being the Nash Sutcliffe Efficiency (NSE) and its empirical distribution. In short, when the NSE values were lower than 0,65, more than 5% of the time, the regressions/predictions were deemed unsatisfied and rejected. The results from Lin (2017) are condensed in Table 2.3 below.

Table 2.3

The summary of PLSR calibration for NO₃-N, NH₄-N, TDN, TKN, DOC TP, and PO₄-P at the three stations (UP, MD, DN)

	NO ₃ -N	NH ₄ -N	TDN	TKN	DOC	TP	PO ₄ -P
UP in spring and summer							
NSE (% unsatif.)	0.93 (0%)	0.17 (100%)**	0.88 (0%)	0.87 (0%)	0.91 (0%)	0.95 (0%)	0.94 (7.4%)
UP in fall and winter							
NSE (% unsatif.)	0.94 (0%)	0.97 (15%)**	0.89 (0%)	0.87 (0.1%)	0.85 (0%)	0.97 (0%)	0.88 (33%)**
MD in spring and summer							
NSE (% unsatif.)	0.98 (0%)	0.80 (2%)	0.92 (0%)	0.87 (0%)	0.92 (0%)	0.88 (0.1%)	0.92 (9.4%)
MD in fall and winter							
NSE (% unsatif.)	0.95 (0%)	0.75 (35%)**	0.94 (0%)	0.71 (22%)**	0.93 (0%)	0.81 (8.8%)	0.74 (29%)**
DN in spring and summer							
NSE (% unsatif.)	0.93 (0%)	0.10 (100%)**	0.85 (0%)	0.80 (5.2%)	0.91 (0%)	0.95 (0%)	0.94 (0%)
DN in fall and winter							
NSE (% unsatif.)	0.93 (0%)	0.71 (39%)**	0.85 (0.1%)	0.84 (2.7%)	0.92 (0%)	0.97 (0%)	0.95 (1%)

TSS Predictions from PLSR and Turbidity

An intriguing result was the applications of two distinct and somewhat independent techniques, for the prediction of TSS. Indeed, turbidity is a parameter given directly by the instrument, and can be used as an index data, just like the whole absorbance spectrum to predict TSS.

We applied the PLSR method to TSS and found highly significant correlations and robust models for the UP and DN stations ($R^2 = 0.94$, RMSE between 8 and 10 mg/L) but not so for the MD station ($R^2=0.68$, RMSE = 29 mg/L), which model appears invalid using the 0.1 criterion for NSE. The better results for the UP and DN stations appear related, again, to the few high concentration values 5 to 10 times the median concentrations.

Using correlations reported before (e.g., Birgand et al., 2005; Skarbøvik and Roseth, 2015; Ramos et al., 2015; and Gippel, 1995), we also obtained highly significant correlations between turbidity values and TSS (R^2 greater than 0.96; details in Lin, 2017). The RMSE from turbidity predictions are 4, 3, 15 mg/L at the UP, MD and DN stations, respectively.

It becomes interesting to compare their performance side by side, to predict the total TSS annual loads. The TSS cumulative loads calculated from PLSR and turbidity calibrations show very similar results at the UP station, i.e 224 and 223 kg/ha, respectively. However, loads computed with the PLSR calibrations at the MD and DN stations overestimate by about 30% those predicted using turbidity (MD: 340 vs 262 kg/ha, 29.6% difference; DN: 544.4 vs 405.2 kg/ha; 34% difference). These results are discussed below.

Dicsussion: can we reliably measure nutrient concentrations and loads with Water Quality Rating Curves?

The use of PLSR with UV-vis spectro::photometers had been reported in waste water (Torres, and Bertrand-Krajewski,), in a tidal marsh (Etheridge et al., 2013; 2014), and in a lake (Birgand et al., 2016). This study shows an application in the more standard setting of monitoring stations in streams and rivers. We have coined the concept of Water Quality Rating Curves by analogy to other rating curves often used in hydrology, such as stage-discharge relationships.

We have found that in an agricultural stream of the Coastal Plain, it was possible to establish significant and robust correlations between light absorbance between the UV and the visible and many parameters of interest for the N, P, and C cycles, although we were unable to establish valid correlations between absorbance and $\text{NH}_4\text{-N}$ and $\text{PO}_4\text{-P}$ for most of the time and stations, however.

The results for TSS predictions using PLSR as a rating method or turbidity as index data, point out another extremely important point and potential risk when using water quality rating curves. After review of the concentration time series, the large overestimations of the TSS loads at the MD and DN stations using PLSR compared to the turbidity based predictions, are associated with concentration peaks at or near flow peaks, which *predictions are outside the calibration ranges*.

The first consequence of this observation is that it is very important, for all parameters, to try to capture the largest concentration range during calibration, and identify the

concentrations predicted outside the calibration range. Ideally, this should correspond to a minimum percentage of the time. For parameters that exhibit a concentration effect, the concentrations tend to peak during flow peaks, which makes their ‘capture’ or sampling difficult, and also their incorporation in the calibration pool, because these events are rare by definition. The errors associated with extrapolating outside the calibration range are particularly amplified because they are associated with the flow peaks. The consequences of predictions outside the calibration range are inherently less problematic for parameters that exhibit ‘dilution effects’ during events, such as nitrate for the studied watersheds (see chapter 2 and 3).

The second consequence for TSS specifically, is that some of the concentration peaks predicted with the PLSR appear unrealistically high (>1000 mg/L). For Turbidity values less than 500 NTU, it is generally accepted that the TSS and turbidity increase linearly. For higher values (>500 NTU), the turbidity do not increase linearly with TSS anymore as some of the particles are in the shadows of others that have already absorbed light. It is thus likely that the predictions of TSS concentrations outside the calibration range using turbidity as index data are underestimated, yet bounded. In the end, there seems to be too much risk with the PLSR predicted TSS loads and Turbidity predicted loads are the preferred choice.

Beyond water quality rating curves: additional local concentration corrections

Most of the time, PLSR could reliably predict continuous water quality concentrations and fit well with the discrete samples results, despite the fact that some fouling was observed after two weeks most of the time. This suggests that the PLSR did correct for most of the

fouling. Nonetheless, the time series revealed that just before and/or just after field servicing and optics cleaning, the predicted concentrations appeared offset compared to the lab values. This coincides with PLSR ability to correct for fouling, as on average it corrects well but might ‘not correct enough’ just before field servicing, and ‘too much’ just after servicing. The concentration predictions from PLSR thus appear as an artifact of the method, which we decided to correct manually using the AQUARIUS software, such as in the figures below.

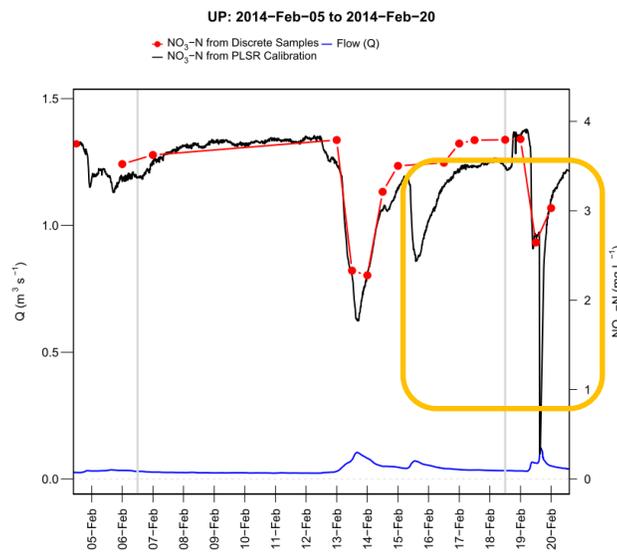


Figure 2.8

An example of “stray” data points because of biofouling. Continuous NO₃-N concentrations from PLSR calibration (black line), NO₃-N concentrations from Discrete Samples (red dots), Flow rate (blue line), and Field visit date (gray vertical line) at UP on February 5-20 in 2014.

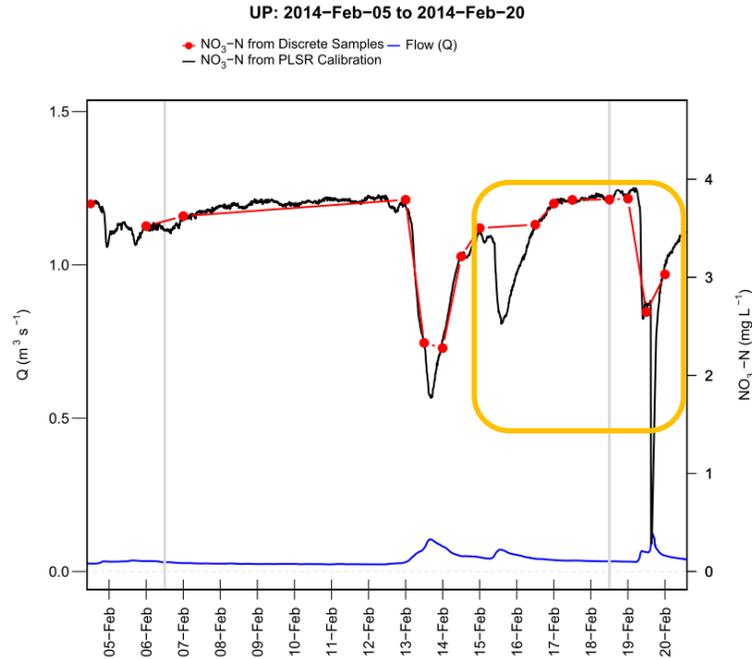


Figure 2.9

Continuous NO₃-N concentrations from PLSR calibration with correction by using AQUARIUS software (black line), NO₃-N concentrations from Discrete Samples (red dots), Flow rate (blue line), and Field visit date (gray vertical line) at UP on February 5-20 in 2014.

Results for filling missing flow and concentration data

Missing data is a true problem with these instruments. We have learned the hard way that one must find ways to fill in the data, with some robust methods. All the details are provided in Lin (2017). We feel important to summarize the methods used for this purpose.

In order to compensate the loss of flow data, we developed regression relationships among the monitoring stations to fill the gap. This was a simple but tedious task, but shows the importance of the experimental design with 3 stations.

As discussed previously, the monitoring system did not always function properly during several periods. The percentages of properly functioning period for the monitoring systems were 60.7, 54.4, and 64.0 % at UP, MD, and DN respectively and most of periods with gap were not synchronous for the three stations.

A total of three methods were compared to fill in the data. The first ‘do nothing’ method consisted in linearly interpolating between continuous discrete samples, regardless of the changes of flow. The second, the ‘Manual method’ consisted, during events, in taking into account the likely changes of concentrations associated with flow changes, by adding concentration points just before an event and at the flow peaks, using ‘visual expertise’ using adjacent events as guides. Although fast, the manual method is highly subjective. The third ‘Semi-Automatic method’ took the subjectivity out of the manual method by creating regression relationships between flow rates and water quality concentrations from adjacent/similar storm events (details in Lin, 2017).

Overall, the results show that all three methods are rather comparable and do not generate large relative uncertainties (Table 2.4). The ‘do nothing’ method is not particularly attractive as it is obvious that the chemographs generated do not represent well, at times, concentration variations during storms

Table 2.4

Summary of NO₃-N cumulative loads for Manual method and Semi-Automatic method

Station	Manual Method		% Absolute difference between A and B	NO ₃ -N cumulative loads for Semi-Automatic Method (C)	% Absolute difference between A and C
	NO ₃ -N cumulative loads with Manual Method (A)	NO ₃ -N cumulative loads the 'do nothing' method (B)			
UP	19.6	19.7	+0.7%	19.3	+1.2%
MD	15.6	15.8	+1.7%	14.9	-4.2%
DN	16.5	16.9	+2.4%	16.6	+0.5%

Challenges and maintenance for the application of UV-Vis spectrophotometers *in situ*

There are several main challenges associated with obtaining high-frequency water quality and hydrology data continuously, and keeping the monitoring system (UV-Vis spectrophotometers, multi-parameter sonde and flow meter) working properly *in situ* including: (1) biofouling on optics surface, (2) sufficient power supply, (3) heat supply in adverse weather, (4) animal effects and equipment maintenance. The details were addressed as following: (1) Biofouling on optics surface, would affect the measurements quality of UV-Vis spectrophotometers, therefore, with 2 % HCl was applied. (2) Sufficient power supply: adequate power was required for the several electronic instruments (ex: UV-Vis spectrophotometers, flow meter, multi-parameter sonde, and discrete sampler) installed at monitoring stations. To meet this need, rechargeable batteries and solar panel boards were set up at each monitoring station. (3) Heat supply in adverse weather: in winter, the water would freeze in the tubes if the temperature was below 0°C and affect the transportation of water in tubes, thus preventing water quality measurements by UV-Vis spectrophotometers. It has

been suggested to use tube insulation to cover tubes. Furthermore, Skarbøvik and Roseth (2015) suggest applying a heating cable or heating lamp in-situ. (4) Animal effects and equipment maintenances: There were several animals (e.g., mouse and turtles) that bit the cables, wires, and tubes, hence, the applications of hard covers to protect the cables, wires and tubes were necessary. Additionally, changing the pump tubing was necessary since the tubing would wear out within a month. Overall, the maintenance of the monitoring stations *in situ* was essential to ensure that the measuring equipment would collect high quality data.

2.4 Conclusion

We observe that the effects of stream restoration on e.g., bank erosion, habitat, or macroinvertebrate communities seems to be relatively well established. We observe that in these examples, the indicators used are ‘integrative indicators’ as they integrate over time the effect of the pulsation nature of streams and the effect of restoration. There has been no equivalent consensus on the effects of stream restoration on water quality for at least two main reasons: either because they are very small and/or because the methods used until now were not robust enough and did not integrated over long continuous periods of time the nutrient fluxes.

We hypothesize that one way to address the water quality impact of restoration is to provide methods to obtain reliable integrative water quality indicators, e.g., nutrient and material loads. The arrival of continuous water quality sensors to obtain concentration values on a continuous basis provides a solution to what had remained an obstacle until now. We have proposed to use UV-Vis spectrophotometers as index data source and the creation of

Water Quality Rating Curves to obtain concentration data on a high frequency basis to calculate robust N, P, C, and material fluxes.

We have shown that in an agricultural stream of the coastal plain of North Carolina, it was possible to construct robust Water Quality Rating Curves to measure nitrate, TDN, TKN, DOC, and TP, using the absorbance data as index data and PLSR as a rating method. We have also shown that this method did not work well for $\text{NH}_4\text{-N}$ and $\text{PO}_4\text{-P}$, although it is possible that this was the result of not optimal calibration points pool. Our results suggest that to obtain robust Water Quality Rating Curves, it is necessary to have as large as possible a calibration concentration range, where calibration concentrations are as stratified as possible. Extrapolating beyond the calibration range is inherently risky and may result in errors, and in the case of parameters that exhibit concentration effect during events, large overestimation as we showed for TSS predicted using PLSR. For TSS, we suggest that it is preferable to use the more conservative turbidity based method.

We have also shown that the concentrations calculated by the algorithms embedded on the instruments are not reliable without local calibration, but that after calibration, can yield more acceptable results for nitrate and DOC. However, the PLSR rating seems to be more robust, is not technically over complicated and can be applied to other parameters. The uncertainties induced by the combination of absorbance data and Water Quality Rating Curves is explored in Chapter 2. To obtain fully continuous nutrient flux data, it is necessary to provide methods to fill the gaps. We have proposed methods to fill flow and concentration data, the latter of which seem robust.

2.5 References

- Antón, A., Elozegi, A., García-Arberas, L., Díez, J., Rallo, A. 2010. Restoration of dead wood in Basque stream channels: effects on brown trout population. *Ecology of Freshwater Fish*, 20: 461-471.
- Aquatic Informatics Inc. 2009. Aquarius Hydrologic Workstation Software. 200–322. Water St., Vancouver, B.C. V6B 1B6, Canada. info@aquaticinformatics.com.
- Bennett, S. J., Simon, A., Castro, J. M., Atkinson, J. F., Bronner, C. E., Blersch, S. S., Rabideau, A. J. 201. The evolving science of stream restoration. *Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools* Geophysical Monograph Series 194, Published in 2011 by the American Geophysical Union.
- Bernhardt, E. S., Palmer, M. A., Allan, J. D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G. M., Lake, P. S., Lave, R., Meyer, J. L., O'Donnell, T. K., Pagano, L., Powell, B., Sudduth, E. 2005. Synthesizing U.S. river restoration efforts. *Science*, 308:636-637.
- Birgand F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., Bordenave, P. 2010. Uncertainties in assessing annual nitrate load and concentration indicators: Part 1. Impact of sampling frequency and load estimation algorithms. *Trans. ASABE*, 53(2):437-446.
- Birgand, F. 2000. Quantification and modeling of in-stream processes in agricultural canals of the lower coastal plain, North Carolina State University Theses.
- Birgand, F., Appelboom, T. W., Chescheir, G. M., Skaggs, R. W. 2011. Estimating nitrogen, phosphorus, and carbon fluxes in forested and mixed-use watersheds of the lower

- coastal plain of North Carolina: uncertainties associated with infrequent sampling. *Trans. ASABE.*, 54(6): 2099-2110.
- Birgand, F., Aveni-Deforge, K., Smith, B., Maxwell, B., Horstman, M., Gerling, A. B., Carey, C. C. 2016. First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency in situ water chemistry measurements at multiple sites. *Limnol. Oceanogr. Methods.*
doi:10.1002/lom3.10122.
- Birgand, F., Benoist, J.-C., Novince, E., Gilliet, N., Saint-Cast, P., Le Saos, E. 2005. Guide d'application de la technique du Doppler continu pour une section calibrée en bois, *Ingénierie-EAT*, 41 : 77-82.
- Birgand, F., Lellouche, G., Appelboom, T.W. 2013. Measuring flow in non-ideal conditions for short-term projects: Uncertainties associated with the use of stage-discharge rating curves. *J. Hydrol.*, 50(3): 186-195.
- Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E., Newman, J. R., Palmer-Felgate, E. J. 2015. Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration–flow relationships. *Science of the Total Environment*, 511: 608-620.
- Brauer N., O'Geen, A. T., Dahlgren, R. A. 2009. Temporal variability in water quality of agricultural tailwaters: implications for water quality monitoring. *Agr Water Manage*, 96: 1001-1009.
- Cassidy, R. and Jordan, P. 2011. Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. *J Hydrol*, 405:18-193.
- Colangelo, D. J., 2014. Interim Response of Dissolved Oxygen to Reestablished Flow in the Kissimmee River, Florida, U.S.A. *Restoration Ecology*, 22: 376-387.

Crumpton, W. G., Isenhardt, T. M., Mitchell, P. D. 1992. Nitrate and organic N analyses with second-derivative spectroscopy. *Limnol. Oceanogr.*, 37:907-913.

Daniluk, T. L., Lautz, L. K., Gordon, R. P., Endreny, T. A., 2013. Surface water-groundwater interaction at restored streams and associated reference reaches. *Hydrology Process*, 27: 3730-3746.

Davis, R. T., Tank, J. L., Mahl, U. H., Winikoff, S.G., Roley, S. S. 2015. The influence of two-stage ditches with constructed floodplains on water column nutrients and sediments in agricultural streams. *Journal of the American Water Resources Association (JAWRA)*, 51(4): 941-955.

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

Duncan, D., Harvey, F., Walker, M. and Australian Water Quality. 2007. EPA Guidelines: Regulatory monitoring and testing- Water and wastewater sampling. Environment Protection Authority in Australia.

Eaton, A. D., Clesceri, L. S., Greensberg, A. E. 1995. Standard methods for the examination of water and wastewater, 19th ed. American Public Health Association.

ESRI. ArcGIS, <http://www.esri.com/>

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

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

- Etheridge, J. R., Birgand, F., Osborne, J. A., Osburn, C. L., Burchell, M. R., Irving, J. 2014. Using in situ ultraviolet–visual spectroscopy to measure nitrogen, carbon, phosphorus, and suspended solids concentrations at a high frequency in a brackish tidal marsh. *Limnol. Oceanogr. Methods*, 12:10-22.
- Etheridge, J. R., 2013. Quantifying the water quality benefits of a constructed brackish marsh and tidal stream system using continuous water quality and flow monitoring. In: Ph.D. Dissertation. North Carolina State University, Raleigh, NC.
- Fichot, C. G. and Benner, R. 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. *Geophys. Res. Lett.* 38: L03610 doi:10.1029/2010GL046152.
- Fischenich, J. C. 2011. Stream restoration benefits. *Stream Restoration in Dynamic Fluvial Systems: Scientific Approaches, Analyses, and Tools Geophysical Monograph Series 194*, Published in 2011 by the American Geophysical Union.
- Flemming, H. C. 2011. Microbial biofouling: Unsolved problems, insufficient approaches, and possible solutions. In: H.C. Flemming, J. Wingender, and U. Szewzyk, editors, *Biofilm highlights*. Springer, New York. 81–109.
- Gippel, G. J. 1995. Potential of turbidity monitoring for measuring the transport of suspended solids in streams. *Hydrol Process.*, 9:83-97.
- Greenberg, A. E., Clesceri, L. S., Eaton, A. D. 2005. Standard methods for the examination of water and wastewater, 25th ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.
- Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., Loewenthal, M., Palmer-Felgate, E. J., Jarvie, J. P. 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary

- production and implications for the Water Framework Directive. *Hydrological Process*, 29: 3388-3407.
- Henjum, M. B., Hozalski, R. M., Wennen, C. R., Arnold, W.A., Novak, P. J. 2010. Correlations between in situ sensor measurements and trace organic pollutants in urban streams. *J Environ Monitor*, 12: 225-233.
- Howson, T. J., Robson, B. J., Mitchell, B. D., 2009. Fish assemblage response to rehabilitation of a sand-slugged lowland river. *River. Research Applications*, 25: 1251-1267.
- Hughes, A. O. 2016. Riparian management and stream bank erosion in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 50(2), 277-290.
- ISO 15769. 2010. Hydrometry - Guidelines for the application of acoustic velocity meters using the Doppler and echo correlation methods.
- ISO 748. 1997. Measurement of liquid flow in open channel - velocity-area methods.
- Langergraber, G., Fleischmann, N., Hofstädter, F. 2003. A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. *Water Sci. Technol.*, 47(2): 63-71.
- Mevik, B., Wehrens, R., and Liland, K. H. 2015. pls: Partial least squares and principal component regression. R package 2.4-3. <http://CRAN.R-project.org/package=pls>.
- Morlock, S. E., Nguyen, H. T., Ross, J. H. 2002. Feasibility of Acoustic Doppler Velocity Meters for the Production of Discharge Records from US Geological Survey Streamflow-Gaging Stations. US Geological Survey. Water-Resources Investigations Report 01-4157.

- Obolewski, K., Glińska-Lewczuk, K., Strzelczak, A., Burandt, P. 2014. Effects of a Floodplain Lake Restoration on Macroinvertebrate Assemblages- A Case Study of the Lowland River (The Słupia River, N Poland). *Polish Journal of Ecology* 62(3): 557-573.
- Osburn, C. L. and Morris, D. P. 2003. Photochemistry of chromophoric dissolved organic matter in natural waters. In: E.W. Helbling and H. Zagarese, editors, *UV effects in aquatic organisms and ecosystems*. The Royal Society of Chemistry, Cambridge, UK. p. 185–217.
- Palmer, M. A., E. S. Bernhardt, J. D. Allan, P. S. Lake, G. Alexander, S. Brooks, J. Carr, S. Clayton, C. N. Dahm, J. Follstad Shah, D. L. Galat, S. G. Loss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, G. M. Kondolf, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano and E. Sudduth. 2005. Standards for Ecologically Successful River Restoration. *Journal of Applied Ecology*, 42: 208-217.
- Pander, J. and Geist, J. 2013. Ecological indicators for stream restoration success. *Ecological Indicators*, 30: 106-118.
- R Core Team. 2016. R: A language and environment for Statistical Computing. <http://www.R-project.org>.
- Ramos, T. B., Gonçalves, M. C., Branco, M. A., Brito, D., Rodrigues, S., Sánchez-Pérez, J.-M., Sauvage, S., Prazeres, Â., Martins, J. C., Fernandes, M. L., Pires, F. P. 2015, Sediment and nutrient dynamics during storm events in the Enxoé temporary river, southern Portugal. *Catena*, 127:177-190.
- Rice, E. W., Baird, R. B., Eaton, A. D., Clesceri, L. S. 1997a. Standard methods for the examination of water and wastewater: 2540 Solids, 22nd ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.

- Rice, E. W., Baird, R. B., Eaton, A. D., Clesceri, L. S. 2012. Standard methods for the examination of water and wastewater, 22nd ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.
- Rieger, L., Langergraber, G., Siegrist, H. 2006. Uncertainties of spectral in situ measurements in wastewater using different calibration approaches. *Water Sci. Technol.*, 53:187-198.
- Rochelle-Newall, E. and Fisher, T. 2002. Chromophoric dissolved organic matter and dissolved organic carbon in Chesapeake Bay. *Mar. Chem.*, 77:23-41.
- Saraceno, J. F., Pellerin, B. A., Downing, B. D., Boss, E., Bachand, P. A., Bergamaschi, B. A. 2009. High-frequency in situ optical measurements during a storm event: Assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes. *J. Geophys. Res.* 114: G00F09, doi:10.1029/2009JG000989.
- Skarbøvik, E. and Roseth, R. 2015. Use of sensor data for turbidity, pH and conductivity as an alternative to conventional water quality monitoring in four Norwegian case studies. *Acta Agriculturae Scandinavica, Section B- Soil & Plant Science*, 65(1): 63-73.
- Skarbøvik, E., Stålnacke, P., Bogen, J., Bønsnes, T. E. 2012. Impact of sampling frequency on mean concentrations and estimated loads of suspended sediment in a Norwegian river: implications for water management. *Sci Total Environ*, 433:462-471.
- StreamStats from USGS. http://streamstatsags.cr.usgs.gov/v3_beta/viewer.htm?stabbr=NC
- Suzuki, N. and Kuroda, R. 1987. Direct simultaneous determination of nitrate and nitrite by ultraviolet second-derivative spectrophotometry. *Analyst* 112:1077-1079.

Torres, A. and Bertrand-Krajewski, J.-L. 2008. Partial Least Squares local calibration of a UV-visible spectrometer used for in situ measurements of COD and TSS concentrations in urban drainage systems. *Water Sci. Technol*, 57:581-588.

Wayne County, NC. <http://www.waynegov.com/page/214>

Whelan, A., and Regan, F. 2006. Antifouling strategies for marine and riverine sensors. *J. Environ. Monit*, 8:880-886.

Chapter 3: Uncertainties assessment for continuous and infrequent sampling in measuring nutrient loads with the application of UV-Visual spectrophotometers in an agricultural coastal plain stream

3.1 Introduction

Water quality monitoring *in situ* probes uncertainties need assessing

There is a fast increasing number of articles that have reported the use of water quality sensors in tidal marshes since 2010 (Etheridge et al., 2013, 2014, 2015), rural (Bowes et al., 2015; Skarbøvik and Roseth, 2015; Cassidy and Jordan, 2011), and urban catchments (Halliday et al., 2015). However, and to our knowledge, there are few reports on the uncertainties associated with these sensors, although this is crucial to warrant the investment into these very expensive pieces of equipment.

Among the very interesting applications of PLSR is that it seems possible to obtain good calibrations with parameters not known to absorb light such as Organic Nitrogen, (ON), bromide (Br^-), phosphate ($\text{PO}_4\text{-P}$), total phosphorus (TP), Iron and Silica (Etheridge et al., 2014; 2015; Birgand et al., 2016; Chapter 2). We have thus proposed the concept of ‘water quality rating curves’, where absorbance spectra is used as ‘index data’ (*sensu* Morlock et al., 2002) and PLSR corresponds to the ‘rating method’. We have shown that water quality rating might be used for parameters not known to absorb light, because we hypothesized that in many watersheds, there might be covariability of the ‘color matrix’ of the water with concentrations (Birgand et al., 2016; Chapter 2).

Research questions and hypotheses

Measured, or rather predicted concentrations are thus fraught with uncertainty directly linked to the residuals associated with the regressions found. Like in all calibrations, the values and robustness of the regressions depend on the number and the distribution of calibration points, which may vary from parameter to parameter, and from station to station. Additionally, despite cleaning systems and techniques, the optics of these instruments, tend to foul due to metal oxide precipitation and/or the growth of biological and microbial agents over time (e.g., Whelan and Regan, 2006; Flemming, 2011; Etheridge et al., 2013; Tait et al., 2014; this report).

By estimating residuals, the PLSR approach can be directly used to calculate the uncertainty associated with these continuous water quality instruments, which does not appear to have been done before. Rieger et al. (2006) and Langergraber et al. (2003) suggested the calibration points cover the entire prediction range and be distributed equally. More calibration points may suggest more robust rating curves, but the number of points must remain small enough not to add to the already high investment and maintenance costs. The research questions become: 1) what kind of uncertainty should one expect on concentration and load values from continuous water quality instruments? 2) Can one develop guidance on how to minimize uncertainty with the use of these instruments using an affordable number of calibration points? 3) How do uncertainties associated with the UV-Vis sensors compare with those associated with infrequent sampling on the annual load indicator?

Objectives

The objective of this chapter is to report the uncertainties that one might expect using UV-Vis instruments in streams and rivers on concentration and load values for nitrate [(NO₃-N)], total dissolved nitrogen [TDN], Total Kjeldahl Nitrogen [TKN], dissolved organic carbon [DOC], total phosphorus [TP], and total suspended solids [TSS]. We used *in situ* UV-Vis spectrophotometers and continuous monitoring of hydrology at two stations in the upper coastal plain of North Carolina, USA.

As a result, the objectives of the research were addressed as follows: 1) Evaluate the uncertainties on concentration values and of annual loads calculated by the ‘global calibration’ provided by the manufacturer, and calculated using PLSR depending upon the number and the distribution of the calibration points used. 2) Compare these uncertainties to those that would be induced by infrequent sampling on estimating annual loads with best load algorithm methods; 3) Provide guidance to users to minimize uncertainties and maximize the potential of their UV-Vis spectrophotometers.

3.2 Methods

We followed the procedures for PLSR calibration described in chapter 2 using the *pls* package (Mevik et al., 2011) in the R software (R Core Team, 2016). Briefly, this package takes 90% of the points to create a calibration model and cross validates on the remaining 10%, and calculates indicators such as Root Mean-Square Error of Prediction (RMSEP), percentage of variance of wavelengths being explained and R² for a given number of

components. The model was initially run for 20 components.” The optimum number of PLSR components for each analysis was chosen based on the two conditions. First, RMSEP values decreased dramatically as the increasing of number of components; then, RMSEP values became stable after the first turning point of the number of the component. Thus, the number of component related to the first transition point for RMSEP values was considered in the first condition. Second, the first number of component where the percentage of variance of wavelengths was higher than 99.95 % was considered. At last, the minimum of the number of component obtained from the two condition was determined for PLSR calibration. With PLSR calibrations, the 15-minutes continuous concentrations for NO₃-N, TDN, TKN, DOC, TP, and TSS were predicted from the spectral data and discrete sample results.

Reference load

The cumulative nutrient and material loads (L) is the integral over a given time (t) of instantaneous loads calculated as the product of solute concentration ($C(t)$) with flow rate ($Q(t)$; Equation 2). In reality, this integral can be approached by the summation of instantaneous 15-min loads calculated from the 15-min flow (Q_i) and concentration (C_i) data (Equation 3), where K is the unit adjustment factor and N is the number of 15 min intervals in t .

$$L = \int_0^t C(t)Q(t)dt \quad (\text{Equation 2})$$

$$L \approx K \sum_{i=1}^{i=N} C_i Q_i \quad (\text{Equation 3})$$

All loads calculated from calculated 15-min concentration data were calculated using Equation 3.

Resampling the existing calibration points

In order to evaluate the uncertainty associated with the use of UV-vis spectrophotometers, we numerically resampled the existing calibration points by using bootstrap, random and stratified sample subsets from the original dataset, and applied PLSR on these subsets, and predicted 15-min concentration data from the 15-min absorbance spectra. We then compared the computed loads and compared them to the reference ones and expressed the differences in percentages.

Bootstrap sampling

Bootstrap sampling refers to random sampling with replacement from all the original existing calibration points. The sample size of the new datasets is identical with the number of original calibration points, and the bootstrap sampling in this research is referred to case bootstrap resampling. Bootstrap sampling was applied to $\text{NO}_3\text{-N}$, TDN, TKN, DOC, TP and TSS at both the UP and DN stations.

Random subset sampling

Random subset sampling refers to generating a new dataset by selecting a subset randomly without replacement from the original dataset. In this research, we chose the samples sizes with subset sampling to be 52, 24 and 12 data points, so that the uncertainties would be comparable to those induced by weekly, biweekly, and monthly discrete sampling,

respectively. The sample sizes with 52, 24 and 12 points were applied to NO₃-N, TDN, TKN, DOC, TP and TSS at UP and DN stations. The sample sizes with 52 were not applied to TSS at DN because the maximum number of original calibration points was only 45.

Stratified sampling

In order to obtain higher quality of PLSR calibrations, Rieger et al. (2003) and Langergraber et al. (2006) suggested that the discrete samples be distributed equally over the given concentration range. Stratified sampling refers to separating the original dataset into mutually exclusive subgroups and using random sampling without replacement within each bin (or stratum). The details of how this was done are available in Lin (2017).

Uncertainties estimation on annual loads as hydrological indicator with resampling

We obtained 1000 different calibration point datasets for each sampling method. We then ran PLSR to create a calibration for each of the dataset. The optimum number of components was selected based on the description in the Methods section. From each of these calibrations, we predicted the continuous concentrations and thus obtained 1000 load estimates for each of the parameters tested. Additionally to the PLSR calibrations, we created 1000 TSS vs turbidity rating curves. For TSS, we thus obtained 1000 estimates of TSS loads using PLSR and 1000 estimates using turbidity as index data.

We estimated the uncertainties for annual loads as a percentage error to reference loads (Equation 5):

$$\text{Error percentage (\%)} = \frac{L_e - L_r}{L_r} \times 100\% \quad (\text{Equation 5})$$

Where L_r is the reference load using the continuous concentrations predicted from all existing original calibration points (for TSS, we obtained two reference load values for PLSR and turbidity based methods) and where L_e is one of the 1000 estimated loads calculated as described above. We thus obtained a distribution of 1000 error percentages from which we extracted the average error (e_{avg}), the median value (e_{50}) (both estimates of accuracy or bias), as well as the 5th (e_5) and 95th (e_{95}) percentiles between which 90% of the total uncertainty is included, as a way to express the range of uncertainties.

Uncertainties estimation on annual loads as hydrological indicator with infrequent sampling

For comparison purposes, we calculated the uncertainties expected on the annual loads that would have been induced by infrequent sampling. Many algorithms have been developed to estimate loads from infrequent concentration and continuous flow data (e.g., Phillips et al., 1999; Johnes, 2007; Birgand et al., 2010; Cassidy and Jordan, 2011). In this research, we used two algorithms: the flow-weighted concentration average estimator method and the linear interpolation method as they have been shown to perform best in coastal plain streams (Birgand et al., 2010).

For the first method, the annual flow weighted concentration average was estimated for a given sampling frequency from the summation of instantaneous loads obtained at the times of sampling divided by the summation of flow at the same times. This concentration was

then multiplied by the flow volume calculated from continuous flow records. The first method is often referred to as the M5 method (Table 3.1; Phillips et al., 1999; Birgand et al., 2010). For the second method, 15 min concentrations were calculated from linear interpolation from discrete concentrations obtained at regular intervals. The second method is often referred to as the M6 method (Table 1; Phillips et al., 1999; Birgand et al., 2010).

Table 3.1

Methods applied to calculate the annual loads in this research (Adapted from Johnes, 2007; Birgand et al., 2010; Cassidy and Jordan, 2011)

Method Number	Equation	Description
M5	$Load = KV \frac{\sum_{i=1}^n C_i Q_i}{\sum_{i=1}^n C_i}$	Product of the annual flow volume and the flow weighted average of the concentration (Littlewood, 1992).
M6	$Load = K \sum_{j=1}^{365} C_j^{int} Q_j$	Product of linear interpolation of concentrations and continuous flow rates (Q_j) (Moatar and Meybeck, 2005).
n	Number of samples	
K	Conversion factor accounts for units and intervals of sampling	
C_i	Instantaneous concentration measured at the time and day of the i th samples ($mg L^{-1}$)	
Q_i	Instantaneous flow rate measured at the time and day of the i th samples ($m^3 s^{-1}$)	
Q_j	Continuous flow rate ($m^3 s^{-1}$)	
V	Annual cumulative flow volume ($m^3 s^{-1}$)	
C_{int}	Linearly interpolation concentration between the two consecutive samples	

To estimate uncertainties induced by infrequent sampling, the reference continuous concentrations were numerically resampled at the tested fixed intervals (weekly, biweekly and monthly). We thus generated 1000 sets of fixed intervals discrete concentrations, for each of which the M5 and M6 loads were calculated. The error percentage between the estimated loads and the reference loads were calculated using Equation 5. The uncertainties induced by infrequent sampling were evaluated for NO_3-N , TDN, TKN, DOC, TP, and TSS at UP and DN.

Filling missing values in the spectral results

During the 16-month monitoring period (from December 2013 to March 2015), human errors and equipment failure (e.g. power outage) caused distinct gaps in spectral data measured by the UV-Vis spectrophotometers. For the purpose of the uncertainty calculations, we created a synthetic dataset over a full hydrological year (as defined by USGS, the hydrological year in the U.S. is from October 1 for any given year to September 30 in the following year.) for each station. Fifteen-minute flow data were almost always available, so gap filling was applied mostly to concentration data. First, we identified the missing spectral data periods. We then searched in the previous/next year if there were spectral data available that roughly corresponded to the same periods with similar storm events. Great care was taken to make sure that the flow and concentration peaks/troughs were fully synchronous. For example, data were missing from October 2013 to November 2013. As a result, the spectral data measured from October 2014 to November 2014 was used to patch the missing 2013 period. When no fitting concentration data were available, the concentration data were duplicated to best fit the flow events. In the end, there were very few duplication of instantaneous loads because most of the flow data were original.

3.3 Results and Discussion

Redundant potential in the annual spectral data

The details of the value of the synthetic dataset used for the uncertainty analysis are detailed in Lin (2017) and the conclusions are that the synthetic data created can be used as a

realistic representation of flow and concentration dynamics that could have been recorded in these watersheds and therefore suitable datasets for uncertainty calculations. From the synthetic datasets created at the UP and DN stations, it was possible to evaluate the uncertainties. The details of the reference/synthetic datasets statistics are in Lin (2017).

The method used allows for very small uncertainties compared to standard monitoring methods

The indicators for all simulations for parameters tested are detailed in Lin (2017). For the bootstrap method, which assesses rather closely the uncertainty expected on the annual loads to be used to quantify the effects of stream restoration, the uncertainties appear small, which gives high hopes that the system we have put together will be good enough to detect restoration effects, even if they are small (**Table 3.2**). Besides for TP, most of the uncertainties are within $\pm 3.5\%$ or better.

Table 3.2: Uncertainty estimation expressed in percentage error compared to the reference annual load for the bootstrap method over a range of parameters

Bootstrap	UP	DN
Nitrate	-1.3 to +1.0	-1.2 to +3.8
TDN	-0.7 to +0.9	-1.4 to +3.1
TKN	-3.5 to +3.6	-3.5 to +7.6

DOC	-3.0 to +2.5	-2.2 to +2.1
TP	-6.5 to +13.8	-4.5 to +8.0
TSS (PLSR)	-4.9 to +2.3	-4.3 to +3.1
TSS (Turbidity)	-2.4 to +2.4	-4.4 to +4.3

Across all simulations for both stations and all parameters, the bootstrap resampling gave the smallest uncertainties, most of which within -3.5% to +3.5% (**Table 3.2**). This was expected as the number of calibration points are 5 to 30 times more numerous. We obtained the smallest uncertainties for TDN (-0.7% to +0.9%) and the largest for TP (-6.5% to +13.8%).

The uncertainty distributions with bootstrap sampling were symmetrical for all parameters at UP and DN stations and generally centered around 0, except for TP at UP station, suggesting that the bias (e_{avg}) and median-bias (e_{50}) have the same trend.

The relatively large uncertainty for TP probably comes from the skewness of calibration points, as 90% of the points are within 15% of the total calibration range corresponding to the lowest concentrations at UP station. Even with bootstrap sampling, there were high probabilities for calibration points to be selected from the lower range (0.01 to 0.30 mg/L at UP). As a result, some of the bootstrap subsampling do not capture the high concentrations, which can generate higher uncertainties, particularly since the high concentrations occurred around flow peaks (Chapter 2). Consequently, the uncertainty range for estimating loads is larger for TP than TDN.

Our results also show that standard monthly sampling would have produced, on annual loads, uncertainties of the order of $\pm 10\%$ for nitrate, $\pm 12\%$ for TDN, $\pm 30\%$ for TKN, \pm

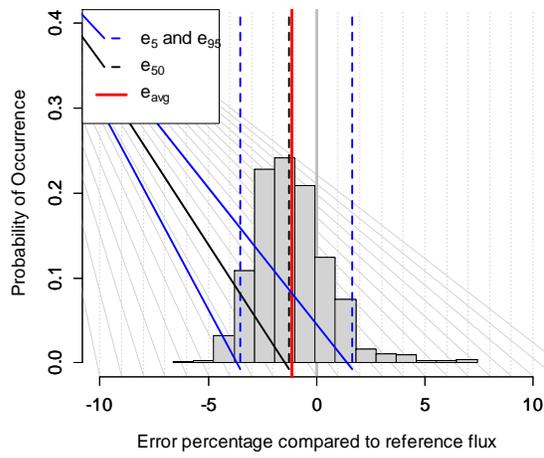
17% for DOC, -60 to +100% for TP and -80 to +150% for TSS (details in Lin, 2017). This amply justify the needs to change from common sampling strategies and the benefits of using our proposed system.

To obtain good results, the *in situ* WQ sensors need to be properly calibrated

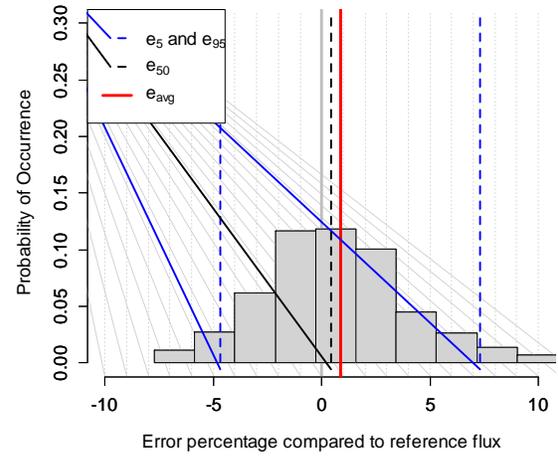
While the previous results are highly encouraging, they presuppose that a lot of samples be collected for calibration. Many agencies such as DOT cannot afford but the equivalent of monthly or maybe bimonthly samples to calibrate their probes. For this reason, we explored the uncertainties that would be associated with a limited number of calibration samples and how the distribution (randomly or stratified) of these samples would matter on the uncertainties.

As stated in the methods part, the uncertainties can be characterized by their range (90% of the uncertainty included between the 5th – e₅ – and 95th – e₉₅ – percentile of the distribution, and their median – e₅₀). The uncertainty indicators and the uncertainty distributions are illustrated in Figure 3.1 below for nitrate and DOC for random and stratified sampling for 24 samples. All distribution histograms and all the numbers corresponding to each histogram are available in Lin (2017).

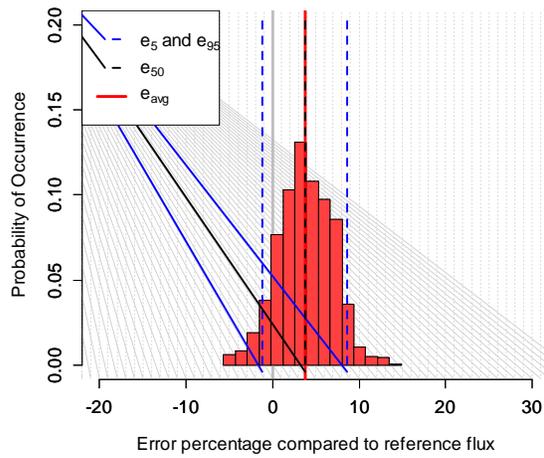
A



B



C



D

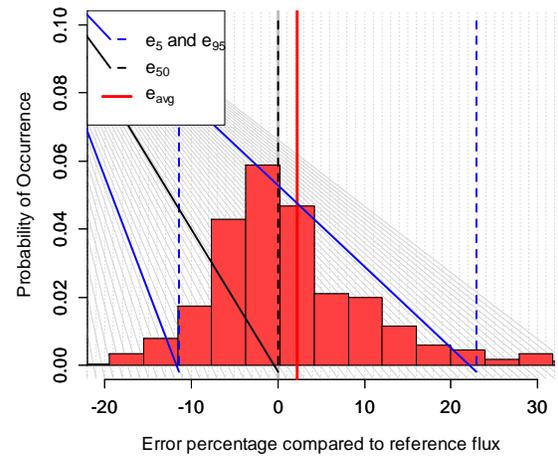


Figure 3.1: Distribution of uncertainties of annual loads for nitrate (top row; A and B) and DOC (bottom row ; C and D) corresponding to 24 stratified (left column; A and C) and random (B and D) samples used to create the calibrations

Much lower uncertainties from stratified compared to random sampling and infrequent sampling

Similarly to what Rieger et al. (2003) and Langergraber et al. (2006) reported, we found that uncertainties associated with stratified sampling are always better than those associated with random sampling, regardless of the number of calibration samples, the parameter, and the station (details in Lin, 2017).

For TKN and DOC, the uncertainty associated with 12 or 24 random sampling for PLSR calibration are large and of the same order as those that would be induced with regular interval discrete sampling calculated using M5 and M6. For TP and TSS, discrete infrequent sampling generates much higher uncertainties, than the PLSR results. However, the uncertainties even with 52 samples can still be high using random sampling for TP at the upstream station (-11.2% to +27.6%).

With increasing number of samples (from 12 to 52), the bias associated with random sampling tends to converge toward zero. This is expected as this tends toward the reference calibration set. Two notable exception are TP at UP and TKN at DN, although the bias absolute values diminish, they are at -5.6% and +7%, respectively for 52 samples.

Interestingly, the bias associated with stratified sampling tends to converge towards negative numbers for nitrate and TDN and positive numbers for DOC and TKN, which reference sample distributions have more high and low concentrations, respectively. This suggests that the annual nitrate and TDN loads, and DOC and TKN loads, tend to be underestimated and overestimated, respectively, with a smaller stratified subset of sample for calibration. This could be due to an overrepresentation of the lower frequency concentrations

in the stratified calibration subsets, and/or an overrepresentation of the higher frequency concentrations in the reference calibration set.

Because the biases, tend to stay relatively stable with stratified sampling, and because the skewness of regression residuals of the respectively lower and higher ranges (data not shown) of concentrations for nitrate and TDN, and DOC and TKN, we conclude that the reference loads are likely systematically overestimating and underestimating, respectively, the actual loads. We also conclude that the annual loads computed from stratified calibration sets are likely less biased than the reference loads.

More stratification is better

For nitrate and TDN, it was possible to test whether the number of bins or the stratification of the calibration samples improves prediction. Although this corresponds to a small improvement, it does seem that stratifying from 2 to 3 to 5 bins, does lower the overall uncertainties for these two parameters at the two stations, using 30 calibration samples. For example, the uncertainties for TDN at UP improved from -2.7% to +3.5% using two bins, to -2.5% to +2.2% using three bins, and to -2.0% to +1.5% using 5 bins (**Error! Reference source not found.**). The relatively small improvement might be due to the small uncertainties to start with for TDN and nitrate parameters.

Uncertainties and calibration robustness depend on parameters

Among all parameters, stations, and best estimation methods, uncertainties calculated on annual loads are lowest to highest in that order: nitrate, TDN (uncertainties very close to

those of nitrate), DOC, TSS (both methods performing about the same), TKN and TP
(**Error! Reference source not found.**).

Table 3.3
Summary of calculated uncertainties for all parameters, stations and method leading to lowest uncertainty

	Best method	UP	Best method	DN
Nitrate	Strat 52	-2.8 to +1.0	Strat 52	-2.3 to +3.0
TDN	Strat 52	-2.0 to +1.3	Strat 52	-3.6 to +5.1
TKN	Rand 52	-11.7 to +16.1	Rand 52	-7.0 to +33.5
DOC	Strat 24	-1.3 to +8.6	Strat 24	-0.1 to +5.2
TP	Rand 52	-11.2 to +27.6	Rand 52	-7.8 to +17.0
TSS (PLSR)	Rand 52	-7.0 to +2.9	NA	NA
TSS (turbidity)	Rand 52	-3.8 to +4.9	NA	NA

Not surprisingly, uncertainties on nitrate and TDN are very similar since TDN was predominantly composed by NO₃-N (50% to 90% for both stations) in the Claridge Canal based on the continuous concentrations. The improvement of the uncertainty from 12 to 24, which suggests that with few samples, PLSR is able to better reproduce the relationship between absorbance and concentrations with nitrate, than it is with TDN or DOC. Nitrate is known to strongly absorb light between 190 nm to 250 nm (Crumpton et al. 1992; Suzuki and Kuroda 1987; Finch et al. 1998; Olsen, 2008), while TDN and DOC represent an ensemble of molecules of many different origins. There theoretically is a lesser correlation

with identified wavelengths. Nonetheless, the uncertainties for DOC at DN are of the same order of the ones for nitrate. For this station, PLSR was able to provide a more robust predictions of the DOC annual loads. This might be due to the intermittent addition of DOC from the Little River during very high flow events.

The much higher uncertainties for TKN and TP are due in large part to the fact that they report calibration with random sampling. The bootstrap results do show however, that for these two parameters, PLSR calibrations were the least robust

3.4 Discussion and conclusions

To our knowledge, this report is the first to report uncertainties on concentrations and annual loads using *in situ* spectrophotometers. The PLS rating approach allows the systematic exploration of the level of uncertainty expected on concentrations and loads. The strength of the PLSR ratings are very similar to those found in chapter 2, which suggests that the synthetic dataset built was realistic and the gap filling method using data duplication in a small proportion of the times did not have obvious apparent consequences. We therefore believe that the results of this study are representative of those expected on the data collected between 2013 and 2015.

The first striking result is that the uncertainties are much lower using stratified sampling rather than random sampling. In other words, the first recommendation to obtain robust water quality rating curves is to obtain stratified concentrations in the calibration sample pool. The second striking result is that it is likely that even the robust, i.e., precise, calibrations obtained with 100+ samples, as the bootstrapping results suggest, might be

inaccurate. Calibration might be biased towards overestimation for parameters exhibiting dilution effects during storms (such as nitrate and TDN in our case), and biased towards underestimation for parameters that exhibit concentration effects during storms, such as DOC, TP, and TSS. The third striking result is that a potential major source of uncertainty is the extrapolation of concentration predictions beyond the calibration range, as the results of the PLSR-based vs turbidity-based results suggest.

Now, within the guidelines described above, it is possible to make other complementary observations. Our results suggest that for nitrate (or TDN mostly made of nitrate) that is known to absorb light in a narrow UV range, very small uncertainties can be expected, i.e., within $\pm 2\%$ for 52 stratified samples. However, in these reactive watersheds 12 stratified samples to calibrate the instrument would yield uncertainties on annual loads within $\pm 7\%$, which is not small, although three times less than uncertainty using standard monthly sampling.

For DOC, which is also known to generally absorb light but which correspond to a large array of molecules, the uncertainty increase to -1% to $+9\%$ or a 10 percentage point spread, compared to a 6 percentage point spread for nitrate for 24 samples for calibration. We believe that the $+4\%$ bias corresponds to the systematic underestimation of the reference load, while we believe that stratified sampling induces less bias.

For other parameters such as TKN and TP, which are not known to absorb light and for which we hypothesize that there is a covariability of light absorbing constituents with concentrations to derive Water Quality Rating Curves, the uncertainties can be much higher

than for nitrate and DOC. For these two, we were unable to test our recommended stratified sampling from the sample pool we were able to obtain. Nonetheless, the uncertainties for 52 calibration points (randomly sampled) could yield uncertainties within a 30 to 40 percentage points, which is not too different from the equivalent 40 and 50 percentage points uncertainties induced from weekly regular interval discrete sampling. It thus seems that although we were able to establish robust water quality rating curves for all parameter tested, one should expect large uncertainties on the annual loads for TKN and TP unless the number of calibration points be much more than 52, and that their concentrations be stratified. As a result, water quality rating curves might not be used as a routine practice for standard monitoring stations, and might remain an efficient research tool.

In chapter 1, we showed that it might be possible to obtain water quality rating curves for ammonium and phosphate. The small calibration pool did not allow exploring uncertainties with these parameters, but it is fair to say that the conclusions for TKN and TP also apply for these parameters too.

We should note that the level of uncertainties computed with the M5 and M6 methods correspond to highly reactive watersheds, where uncertainties tend to be higher (e.g., Birgand et al., 2010). It is thus possible that the level of uncertainties on annual loads for all parameters and all subsampling techniques are on the higher range and that smaller uncertainties might be expected for less reactive watersheds.

Finally, we believe that the bootstrap and other subsampling techniques should be routinely used to estimate the uncertainty on computed annual loads when using in situ field spectrophotometers.

3.5 References

Bernhardt, E. S., Palmer, M. A., Allan, J. D., Alexander, G., Barnas, K., Brooks, S., Carr, J., Clayton, S., Dahm, C., Follstad-Shah, J., Galat, D., Gloss, S., Goodwin, P., Hart, D., Hassett, B., Jenkinson, R., Katz, S., Kondolf, G. M., Lake, P. S., Lave, R., Meyer, J. L., O'Donnell, T. K., Pagano, L., Powell, B., Sudduth, E. 2005. Synthesizing U.S. river restoration efforts. *Science*, 308:636-637.

Birgand F., Appelboom, T. W., Chescheir, G. M., Skaggs, R. W. 2011 b. Estimating nitrogen, phosphorus, and carbon fluxes in forested and mixed-use watersheds of the lower coastal plain of North Carolina: uncertainties associated with infrequent sampling. *Trans. ASABE*, 54(6): 2099-2110.

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

Birgand F., Faucheux, C., Gruau, G., Moatar, F., Meybeck, M. 2011 a. Uncertainties in assessing annual nitrate loads and concentration indicators: Part 2. Deriving sampling frequency charts in Brittany, France. *Trans. ASABE*, 54(1): 93-104.

Birgand, F. 2000. Quantification and modeling of in-stream processes in agricultural canals of the lower coastal plain, North Carolina State University Theses.

Birgand, F., Aveni-Deforge, K., Smith, B., Maxwell, B., Horstman, M., Gerling, A. B., Carey, C. C. 2016. First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency *in situ* water chemistry

- measurements at multiple sites. *Limnol. Oceanogr. Methods*.
doi:10.1002/lom3.10122.
- Birgand, F., Lellouche, G., Appelboom, T.W. 2013. Measuring flow in non-ideal conditions for short-term projects: Uncertainties associated with the use of stage-discharge rating curves. *J. Hydrol.*, 50(3): 186-195.
- Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., Gozzard, E., Newman, J. R., Palmer-Felgate, E. J. 2015. Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration–flow relationships. *Science of the Total Environment*, 511: 608-620.
- Brito, R. S., Pinheiro, H. M., Ferreira F., Saldanha Matos, J., Lourenço, N. D. 2014, *In situ* UV-Vis spectroscopy to estimate COD and TSS in wastewater drainage systems. *Urban Water Journal*, 11: 261-273.
- Cassidy, R. and Jordan, P. 2011. Limitations of instantaneous water quality sampling in surface-water catchments: comparison with near-continuous phosphorus time-series data. *J Hydrol*, 405:18-193.
- Crumpton, W. G., Isenhardt, T. M., Mitchell, P. D. 1992. Nitrate and organic N analyses with second-derivative spectroscopy. *Limnol. Oceanogr.*, 37:907-913.
- Coynel, A., Schäfer, J., Hurtrez, J.-E., Dumas, J., Etchebera, H., Blanc, G. 2004. Sampling frequency and accuracy of SPM flux estimates in two contrasted drainage basins. 2004. *Science of the Total Environment*, 330: 233-247.
- Eaton, A. D., Clesceri, L. S., Greensberg, A. E. 1995. *Standard methods for the examination of water and wastewater*, 19th ed. American Public Health Association.
- Etheridge, J. R., Birgand, F., Burchell, M. R. 2015. Quantifying nutrient and suspended solids fluxes in a constructed tidal marsh following rainfall: The value of capturing the rapid changes in flow and concentrations. *Ecological Engineering*, 78: 41-52.

- Etheridge, J. R., Birgand, F., Osborne, J. A., Osburn, C. L., Burchell, M. R., Irving, J. 2014. Using *in situ* ultraviolet–visual spectroscopy to measure nitrogen, carbon, phosphorus, and suspended solids concentrations at a high frequency in a brackish tidal marsh. *Limnol. Oceanogr. Methods*, 12:10-22.
- Etheridge, J. R., 2013. Quantifying the water quality benefits of a constructed brackish marsh and tidal stream system using continuous water quality and flow monitoring. In: Ph.D. Dissertation. North Carolina State University, Raleigh, NC.
- EU. 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Brussels, Belgium: European Parliament.
- Fichot, C. G. and Benner, R. 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. *Geophys. Res. Lett.* 38: L03610 doi:10.1029/2010GL046152.
- Flemming, H. C. 2011. Microbial biofouling: Unsolved problems, insufficient approaches, and possible solutions. In: H.C. Flemming, J. Wingender, and U. Szewzyk, editors, *Biofilm highlights*. Springer, New York. 81–109.
- Gippel, G. J. 1995. Potential of turbidity monitoring for measuring the transport of suspended solids in streams. *Hydrol Process.*, 9:83-97.
- Glasgow, H. B., Burkholder, J. M., Reed, R. E., Lewitus, A. J., Kleinman, J. E. 2004. Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. *J Exp Mar Biol Ecol.*, 200:409-448.
- Grayson, R. P. and Holder, J. 2016. Improved automation of dissolved organic carbon sampling for organic-rich surface waters. *Science of the Total Environment*, 543: 44-51.

- Greenberg, A. E., Clesceri, L. S., Eaton, A. D. 2005. Standard methods for the examination of water and wastewater, 25th ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.
- Gruber, G., Bertrand-Krajewski, J.-L., de Be´ne´dittis, J., Hochedlinger, M., Lettl, W. 2006 Practical aspects, experiences and strategies by using UV/VIS sensors for long-term sewer monitoring. *Water Practice and Technology*, 1(1), doi10.2166/wpt.2006.020.
- Halliday, S. J., Skeffington, R. A., Wade, A. J., Bowes, M. J., Gozzard, E., Newman, J. R., Loewenthal, M., Palmer-Felgate, E. J., Jarvie, J. P. 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. *Hydrological Process*, 29: 3388-3407.
- Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., Arnold, J. G. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE*, 49(3):689-701.
- Henjum, M. B., Hozalski, R. M., Wennen, C. R., Arnold, W.A., Novak, P. J. 2010. Correlations between *in situ* sensor measurements and trace organic pollutants in urban streams. *J Environ Monitor*, 12: 225-233.
- ISO 15769. 2010. Hydrometry - Guidelines for the application of acoustic velocity meters using the Doppler and echo correlation methods.
- Jiang Y., Frankenberger, J. R., Bowling, L. C., Sun, Z. 2014. Quantification of uncertainty in estimated nitrate-N loads in agricultural watersheds. *Journal of Hydrology*, 519: 106-116.
- Johnes, P. J. 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. *Journal of Hydrology*, 332: 241-258.

- Langergraber, G., Fleischmann, N., Hofstädter, F. 2003. A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. *Water Sci. Technol.*, 47(2): 63-71.
- Lepot, M., Torres, A., Hofer, T., Caradot, N., Gruber, G., Aubin, J.-B., Bertrand-Krajewski, J.-L. 2016. Calibration of UV/Vis spectrophotometers: A review and comparison of different methods to estimate TSS and total and dissolved COD concentrations in sewers, WWTPs and rivers. *Water Research*, 101:519-534.
- Littlewood, I. G. 1992. Estimating constituent loads in rivers: A review. Report No. 117. Wallingford, U.K.: Institute of Hydrology.
- Mevik, B., Wehrens, R., and Liland, K. H. 2015. pls: Partial least squares and principal component regression. R package 2.4-3. <http://CRAN.R-project.org/package=pls>.
- Moatar, F., and Meybeck, M. 2005. Compared performances of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. *Hydrol. Proc.*, 19: 429-444.
- Moatar, F., Meybeck, M. 2007. Riverine fluxes of pollutants: Towards predictions of uncertainties by flux duration indicators. *Comptes Rendu Geosci.*, 339(6):367-382.
- Morlock, S. E., Nguyen, H. T., Ross, J. H. 2002. Feasibility of Acoustic Doppler Velocity Meters for the Production of Discharge Records from US Geological Survey Streamflow-Gaging Stations. US Geological Survey. Water-Resources Investigations Report 01-4157.
- 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. Proc.*, 13: 1035-1050.
- R Core Team. 2016. R: A language and environment for Statistical Computing. <http://www.R-project.org>.

- Ramos, T. B., Gonçalves, M. C., Branco, M. A., Brito, D., Rodrigues, S., Sánchez-Pérez, J.-M., Sauvage, S., Prazeres, Â., Martins, J. C., Fernandes, M. L., Pires, F. P. 2015, Sediment and nutrient dynamics during storm events in the Enxoé temporary river, southern Portugal. *Catena*, 127:177-190.
- Reinelt, L. E. and Grimvall, A. 1991. Estimation of nonpoint source loadings with data obtained from limited sampling programs. *Environmental Monitoring and Assessment*, 21: 173-192.
- Rekolainen, S., Posch, M., Kämäri, J., and Ekholm, P. 1991. Evaluation of the accuracy and precision of annual phosphorus load estimates from two agricultural basins in Finland. *Journal of Hydrology*, 128: 237-255.
- Rice, E. W., Baird, R. B., Eaton, A. D., Clesceri, L. S. 1997a. Standard methods for the examination of water and wastewater: 2540 Solids, 22nd ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.
- Rice, E. W., Baird, R. B., Eaton, A. D., Clesceri, L. S. 2012. Standard methods for the examination of water and wastewater, 22nd ed. American Public Health Association; American Water Works Association; Water Environment Federation. American Public Health Association, Washington, D.C.
- Rieger, L., Langergraber, G., Siegrist, H. 2006. Uncertainties of spectral *in situ* measurements in wastewater using different calibration approaches. *Water Sci. Technol.*, 53:187-198.
- Rochelle-Newall, E. and Fisher, T. 2002. Chromophoric dissolved organic matter and dissolved organic carbon in Chesapeake Bay. *Mar. Chem.*, 77:23-41.
- Rode, M., and Suhr, U. 2007. Uncertainties in selected river water quality data. *Hydrol. and Earth Systems Sci.*, 11(2):863-874.

- Saraceno, J. F., Pellerin, B. A., Downing, B. D., Boss, E., Bachand, P. A., Bergamaschi, B. A. 2009. High-frequency *in situ* optical measurements during a storm event: Assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes. *J. Geophys. Res.* 114: G00F09, doi:10.1029/2009JG000989.
- Skarbøvik, E. and Roseth, R. 2015. Use of sensor data for turbidity, pH and conductivity as an alternative to conventional water quality monitoring in four Norwegian case studies. *Acta Agriculturae Scandinavica, Section B- Soil & Plant Science*, 65(1): 63-73.
- Skarbøvik, E., Stålnacke, P., Bogen, J., Bønsnes, T. E. 2012. Impact of sampling frequency on mean concentrations and estimated loads of suspended sediment in a Norwegian river: implications for water management. *Sci Total Environ*, 433:462-471.
- Suzuki, N. and Kuroda, R. 1987. Direct simultaneous determination of nitrate and nitrite by ultraviolet second-derivative spectrophotometry. *Analyst* 112:1077-1079.
- Torres, A. and Bertrand-Krajewski, J.-L. 2008. Partial Least Squares local calibration of a UV-visible spectrometer used for *in situ* measurements of COD and TSS concentrations in urban drainage systems. *Water Sci. Technol*, 57:581-588.
- United States Geological Survey, "Explanations for the National Water Conditions", http://water.usgs.gov/nwc/explain_data.html
- Webb, B. W., Phillips, J. M., Walling, D. E., Littlewood, I. G., Watts, C. D., Leeks, G. J. L. 2007. Load estimation methodologies for British rivers and their relevance to the LOIS RACS(R) programme. *Science of The Total Environment*, 194-195: 379-389.
- Whelan, A., and Regan, F. 2006. Antifouling strategies for marine and riverine sensors. *J. Environ. Monit*, 8:880-886.

Worrall, F., Howden, N. J. K., Burt, T. P. 2013. Assessment of sample frequency bias and precision in fluvial flux calculations- An improved low bias estimation method. *Journal of Hydrology*, 503: 101–110.

Chapter 4: General Conclusion

The context of this research is the quantification of water quality benefits of stream restoration. This information is needed to associate nutrient credits to this practice. To fill this void in the literature, North Carolina Department of Transportation (NC DOT), North Carolina Forest Service (NCFS) and North Carolina State University (NCSU) have collaborated to conduct a pilot study.

Until now, there has not been any clear evidence that stream restoration can provide water quality benefits (e.g., Bernhardt et al., 2005; Craig et al., 2008; Alexander and Allan, 2006; Castillo et al., 2016; Daigneault et al., 2017; Eshleman and Sabo, 2016), much less accepted numbers on the water quality benefits. The reasons for this might be that the benefits are too small to be measurable, and/or that the uncertainties on the quantification methods were just too large to detect benefits. We have shown that the uncertainties on robust indicators, which we believe annual cumulative loads to be, can easily reach 20 to 40% using standard weekly or biweekly sampling methods (Chapter 3), leaving little chance to quantify with any confidence water quality improvement of the same magnitude. This reason alone could explain the reasons for the lack of recognized water quality benefits. One chance to dramatically reduce the uncertainties on loads, and therefore on the bulk water quality effect, is to obtain concentration values at a pace that can capture the dynamics of the sometime rapid concentration changes with flow.

Because water quality benefits largely depend on residence time (e.g. Craig et al., 2008), it is likely that the water quality improvements in restored streams are inherently small. We

estimated the residence in our 2,200 m stream to be around 5 hours during April base flow conditions (see details in Lin, 2017). In constructed wetland with long water residence time, the removal efficiencies have been measured to reach up to 97%, 32%, 25%, and 53% for nitrate, TDN, ammonium, and soluble active phosphorous, respectively (Ardón et al., 2010). The advent of new water quality sensors able to measure concentration at frequency in par with those of flow or stage sensors, opens the possibility to reduce uncertainty to much lower levels, and should the water quality effects be much higher than the measurement uncertainties, to quantify these benefits. Continuous hydrology and water quality monitoring were applied in the Claridge Canal which was restored in Goldsboro, North Carolina.

This report is thus a methodological piece of work at its core. It first required to test whether new optical based sensors could even be used in the low gradient streams of the coastal plain of North Carolina. In particular, we were very interested to know whether we could (1) Apply *in situ* ultraviolet-visual (UV-Vis) spectrophotometers and create water quality rating curves with using Partial Least Squares Regression (PLSR) to measure water quality parameters ($\text{NO}_3\text{-N}$, $\text{NH}_4\text{-N}$, TDN, TKN, DOC, $\text{PO}_4\text{-P}$, TP and TSS) on a continuous basis in the agricultural coastal plain stream; whether we could (2) Quantify the uncertainties on measuring the cumulative loads associated with the data generated by UV-Vis spectrophotometers and provide guidelines for robust sampling strategies.

Ability to create Water Quality Rating Curves for UV-Vis spectrophotometers

UV-Vis spectrophotometers were used as index data source and to create Water Quality Rating Curves to obtain concentration data on a high frequency basis to calculate robust N, P,

C, and material fluxes. The PLSR rating seems to be more robust than the algorithms embedded on the instruments for our studied stream. The results demonstrated that it was possible to construct robust Water Quality Rating Curves to measure nitrate, TDN, TKN, DOC, and TP, using the absorbance data as index data and Partial Least Squares Regression (PLSR) as a rating method in an agricultural stream of the coastal plain of North Carolina. The results also indicated that this method did not work well for $\text{NH}_4\text{-N}$ and $\text{PO}_4\text{-P}$, although it is possible that optimal calibration point pools were not obtained. Extrapolating beyond the calibration range is inherently risky and may result in errors, and in the case of parameters that exhibit concentration effect during events, large overestimation for TSS predicted using PLSR. For predicting TSS, it is suggested to use the more conservative turbidity based method. To obtain fully continuous nutrient flux data, it is necessary to provide methods to fill the gaps. This research demonstrated that acceptable methods were developed to fill missing flow and concentration data.

Reduce errors on measuring loads with using UV-Vis spectrophotometers and the most robust sampling method for affordable number of calibration points

This research is the first to report uncertainties on concentrations and annual loads using *in situ* spectrophotometers, induced by the errors on sensor concentrations. The existing calibration points were numerically resampled by using bootstrap, random and stratified sample subsets from the original dataset, and applied PLSR on these subsets to estimate the uncertainties on annual loads. From the bootstrap resampling results, continuous monitoring with the application of UV-Vis spectrophotometers can reduce errors on measuring loads,

e.g., within $\pm 5\%$ measuring error for nitrate, TDN, TKN, DOC and TSS; and approximately $\pm 10\%$ for TP.

The uncertainties are much lower using stratified sampling rather than random sampling. It is likely that even the robust, i.e., precise, calibrations obtained with over 100 samples, as the bootstrapping results suggest, might be inaccurate. Calibration might be biased towards overestimation for parameters exhibiting dilution effects during storms (such as nitrate and TDN), and biased towards underestimation for parameters that exhibit concentration effects during storms, such as DOC, TP, and TSS in this research.

The results suggested that for nitrate (or TDN mostly made of nitrate) that is known to absorb light in a narrow UV range, very small uncertainties can be expected, i.e., within $\pm 2\%$ for 52 stratified samples. For DOC, which is also known to generally absorb light but which correspond to a large array of molecules, the uncertainty increase from -1% to $+9\%$ or a 10-percentage point spread, compared to a 6-percentage point spread for nitrate for 24 samples for calibration. For other parameters, such as TKN and TP, which are not known to absorb light, we hypothesized that there is a co-variability of light absorbing constituents with concentrations to derive Water Quality Rating Curves, the uncertainties can be much higher than for nitrate and DOC.

The level of uncertainties computed with the flow-weighted concentration average estimator method (M5) and the linear interpolation method (M6) correspond to highly reactive watersheds, where uncertainties tend to be higher (e.g., Birgand et al., 2010). It is thus possible that the level of uncertainties on annual loads for all parameters and all

subsampling techniques are on the higher range and that smaller uncertainties might be expected for less reactive watersheds. This research suggested the bootstrap and other subsampling techniques should be routinely used to estimate the uncertainty on computed annual loads when using *in situ* field spectrophotometers.

Our results thus show that the level of uncertainty on the annual load cumulative indicator of is indeed largely reduced compared to the infrequent sampling methods that have been applied in past stream restoration monitoring. Now, our monitoring design and methods involved a lot of time and resources. We thought interesting to evaluate the relative effort that this type of monitoring would represent in the whole stream restoration costs.

Costs comparisons between continuous monitoring and stream restoration projects

From 1990, United States government has invested at least \$1 billion every year in stream restoration projects (Bernhardt et al., 2005). In addition, there were 1,345 stream restoration projects implemented from 1970 to 2004 with 444 million investment in the Upper Midwest, USA. (Michigan, Ohio, Wisconsin; Alexander and Allan, 2006). Water quality management as one of top restoration goals have been invested in \$60 million of total costs for 108 projects (median cost is \$234,500). There were only 33% of water quality management projects being monitored in the Upper Midwest. More expensive projects had more opportunities to be monitored. Furthermore, the lack of monitoring information became inhibitions to evaluate the improvements for stream restoration projects; thus, establishing standardization of monitoring is essentially needed (Alexander and Allan, 2006). The same authors (Alexander and Allan, 2007) interviewed the staff in 39 projects among 1,345 stream

restoration projects in Upper Midwest, USA. The results indicated that the highest constraint was “lack of people power or staff time” for the projects without monitoring.

The costs for stream restoration projects were mostly affected by the type and size of the watersheds (Bonham and Stephenson, 2004). Table 4.1 summarizes the costs analysis with linear foot for several stream restoration projects. The construction costs (\$30-\$70/ft.) for the different projects are comparable except for the project in Long Creek, Bessemer City, North Carolina. In addition, the total costs for stream restoration projects during 2016-2017 reported by NC Division of Mitigation Services (NCDMS) is \$391/ft.

Based on the estimated budget from the research proposal (Birgand, 2013), the total fees for continuous monitoring in pre-restoration period in three years (2013-2016) were about \$353,000 for the Claridge Canal with 2,200 m in length. Thus, the continuous monitoring fees in the Claridge Canal were \$16.23/ft., including \$5.53/ft. (34%) for personnel, and \$10.7/ft. (66%) for field supplies/material, monitoring equipment, and transportation. When comparing the continuous monitoring costs for the Claridge Canal and the total costs for stream restoration estimated by NCDMS, the continuous monitoring costs is 4.15% of the total costs for stream restoration. Thus, it is fair to say that the costs for pre-restoration continuous monitoring represented a small proportion of the overall project. The post-restoration costs increase this percentage to about 10% (Birgand, personal communication).

Table 4.1 Costs analysis for the stream restoration projects (Unit: \$/ft.)

Description	Pre-construction	Site Acquisition	Construction	Post-Construction	Total Costs	Restoration Goal	References
Small Stream (<3,001 ft.)	26.14	5.65	68.35	18.81	118.96	Reduce the impacts by surface coal mining in the Southern Appalachian Region	Bonham and Stephenson (2004)
Medium Stream (3,001 – 10,000 ft.)	21.25	4.21	57.28	10.01	92.74		
Large stream (>10,000 ft.)	13.04	-	45.82	6.37	65.22		
Long Creek, Bessemer City, NC	-	-	13.33	-	-	Improve stability and habitat	Jennings et al. (2005)
Kentucky Creek, Newland, NC	-	-	33.33	-	-		
East Prong Roaring River, Stone Mountain State Park, NC	-	-	60	-	-		
Little Garvin Creek, Clemson, NC	-	-	50	-	-		
Stream restoration fees in North Carolina (NC) during 2016-2017	-	-	-	-	391		

-: data not applicable

4.2 References

- Alexander, G. G. and Allan, J. D. (2007). Ecological Success in Stream Restoration: Case Studies from the Midwestern United States. *Environmental Management*, 40(2), 245–255. <https://doi.org/10.1007/s00267-006-0064-6>
- Alexander, G. G., and Allan, J. D. 2006. Stream restoration in the Upper Midwest, USA. *Restoration Ecology*, 14(4): 595–604.
- Alexander, G. G., and Allan, J. D. 2006. Stream restoration in the Upper Midwest, USA. *Restoration Ecology*, 14(4): 595–604.
- Ardón, M., Morse, J. L., Doyle, M. W., Bernhardt, E. S. 2010. The Water Quality Consequences of Restoring Wetland Hydrology to a Large Agricultural Watershed in the Southeastern Coastal Plain. *Ecosystems*, 13(7): 1060–1078.
- Aubert, A. H., Gascuel-Oudou, C., Gruau, G., Akkal, N., Faucheux, M., Fauvel, Y., and others. (2013). Solute transport dynamics in small, shallow groundwater-dominated agricultural catchments: insights from a high-frequency, multisolute 10 yr-long monitoring study. *Hydrology and Earth System Sciences*, 17(4), 1379–1391. <https://doi.org/10.5194/hess-17-1379-2013>
- Bernhardt, E. S., Palmer, M. A., Allan, J. D., Alexander, G., Barnas, K., Brooks, S., and others. (2005). Synthesizing US river restoration efforts. *Science*, 308(5722), 636–637.
- Birgand, F. (2013). Intensive Monitoring of Nutrient and Material Load in Claridge Nursery Stream “The Canal” pre-, during and post construction of Highway 70 Bypass in Wayne County, NC. NCDOT Research Proposal.

- Birgand, F., Faucheux, C., Gruau, G., Augeard, B., Moatar, F., Bordenave, P., and others. (2010). Uncertainties in assessing annual nitrate loads and concentration indicators: Part 1. Impact of sampling frequency and load estimation algorithms. *Trans. Asabe*, 53(2), 437–446.
- Blanco, A. C., Nadaoka, K., Yamamoto, T., Kinjo, K. (2010). Dynamic evolution of nutrient discharge under stormflow and baseflow conditions in a coastal agricultural watershed in Ishigaki Island, Okinawa, Japan. *Hydrological Processes*, 24(18), 2601–2616. <https://doi.org/10.1002/hyp.7685>
- Bonham, J. and Stephenson, K. (2004). A Cost Analysis of Stream Compensatory Mitigation Projects in the Southern Appalachian Region. The American Society of Agricultural Engineers, Self-Sustaining Solutions for Streams, Watersheds, and Wetlands Conference, St. Paul Minnesota, September 12-15, 2004.
- Bowes, M. J., Jarvie, H. P., Halliday, S. J., Skeffington, R. A., Wade, A. J., Loewenthal, M., and others. (2015). Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration–flow relationships. *Science of The Total Environment*, 511, 608–620. <https://doi.org/10.1016/j.scitotenv.2014.12.086>
- Castillo, D., Kaplan, D., Mossa, J. 2016. A Synthesis of Stream Restoration Efforts in Florida (USA): Florida Stream Restoration. *River Research and Applications*, 32(7): 1555–1565.
- Cerro, I., Sanchez-Perez, J. M., Ruiz-Romera, E., Antigüedad, I. (2014). Variability of particulate (SS, POC) and dissolved (DOC, NO₃) matter during storm events in the Alegria agricultural watershed: The importance of storm events in agricultural watersheds. *Hydrological Processes*, 28(5), 2855–2867. <https://doi.org/10.1002/hyp.9850>
- Craig, L. S., Palmer, M. A., Richardson, D. C., Filoso, S., Bernhardt, E. S., Bledsoe, B. P., Doyle, M. W., Groffman, P. M., Hassett, B. A., Kaushal, S. S., Mayer, P. M., Smith, S. M. 2008. Stream restoration strategies for reducing river nitrogen loads. *Frontiers in Ecology and the Environment*, 6(10): 529–538.

Craig, L. S., Palmer, M. A., Richardson, D. C., Filoso, S., Bernhardt, E. S., Bledsoe, B. P., Doyle, M. W., Groffman, P. M., Hassett, B. A., Kaushal, S. S., Mayer, P. M., Smith, S. M. 2008. Stream restoration strategies for reducing river nitrogen loads. *Frontiers in Ecology and the Environment*, 6(10): 529–538.

Daigneault, A. J., Eppink, F. V., Lee, W. G. 2017. A national riparian restoration programme in New Zealand: Is it value for money? *Journal of Environmental Management*, 187: 166–177.

Eshleman, K. N., and Sabo, R. D. (2016). Declining nitrate-N yields in the Upper Potomac River Basin: What is really driving progress under the Chesapeake Bay restoration? *Atmospheric Environment*, 146: 280–289.

<https://deq.nc.gov/about/divisions/mitigation-services/dms-customers/fee-schedules>

Jennings, G., Clinton, D., Bidelspach, D., Penrose, D., Shaffer, M., Rozzell, L., Hall, K., Bass, K., Doll, B., Zink, J., Calabria, J., Foster, S. (2005). Stream Restoration Lessons Learned in North Carolina. <http://srwqis.tamu.edu/media/1731/greg%20jennings%20ky%20stream%20oct05.pdf>.

Kadlec, R. H. and Hey, D. L. (1994). Constructed wetlands for river water quality improvement. *Water Science and Technology*, 29(4), 159–168.

Poor, C. J. and McDonnell, J. J. (2007). The effects of land use on stream nitrate dynamics. *Journal of Hydrology*, 332(1–2), 54–68. <https://doi.org/10.1016/j.jhydrol.2006.06.022>

Ramos, T. B., Rodrigues, S., Branco, M. A., Prazeres, A., Brito, D., Gonçalves, M. C., and others. (2015a). Temporal variability of soil organic carbon transport in the Enxóe agricultural watershed. *Environmental Earth Sciences*, 73(10), 6663–6676. <https://doi.org/10.1007/s12665-014-3888-z>

Ramos, Tiago B., Gonçalves, M. C., Branco, M. A., Brito, D., Rodrigues, S., Sánchez-Pérez, J.-M., and others. (2015b). Sediment and nutrient dynamics during storm events in

the Enxoé temporary river, southern Portugal. *CATENA*, 127, 177–190.
<https://doi.org/10.1016/j.catena.2015.01.001>

Stedmon, C. A., Markager, S., Sondergaard, M., Vang, T., Laubel, A., Borch, N. H., Windelin, A. (2006). Dissolved organic matter (DOM) export to a temperate estuary: Seasonal variations and implications of land use. *Estuaries and Coasts*, 29(3), 388–400.

Webb, B. W. and Walling, D. E. (1985). Nitrate behaviour in streamflow from a grassland catchment in Devon, UK. *Water Research*, 19(8), 1005–1016.