A Random Forest in the Great Lakes:
Exploring Nutrient Water Quality in the Laurentian Great Lakes Watersheds

By

John Dony

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Civil Engineering (Water)

Waterloo, Ontario, Canada, 2020

Author's declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

I would like to acknowledge my co-authors Drs. Kimberly Van Meter and Nandita Basu who contributed to the research described in this thesis.

Abstract

A data driven approach was used in this study to investigate the drivers of nutrient water quality across the Laurentian Great Lakes drainage basin. Monitored time series of nutrient water quality and discharge were modelled using a dynamic regression-based model. Random forest machine learning was used as a framework to assess drivers of nutrient water quality, using mean annual flow-weighted concentrations (FWCs) and ratios calculated from modelled water quality, combined with spatial factors from monitored watersheds. Analysis revealed that landscape variables of developed land use, tile drained land, and wetland area played important roles in controlling nitrate and nitrite (DIN) and soluble reactive phosphorus (SRP) FWCs, while soil type and wetland area was important for controlling particulate phosphorus (PP) FWCs. Fertilizer and manure practices were important controls in nutrient ratios of SRP:Total Phosphorus (TP), and DIN:TP, with developed land use, manure application, and tile drained land important for the former, and developed land use and manure application (vs synthetic fertilizer application) important for the latter. Plots of feature contribution were generated to isolate the effect that spatial variables had in machine learning models and revealed underlying behaviour of important controls in driving nutrient water quality across the basin. Random forest models were further developed to predict FWCs and ratios of nutrients across all watersheds within the Great Lakes drainage basin. Modelled results revealed hot spots of high DIN, SRP and PP in the watersheds along the southeastern shores of Lake Huron, on the eastern watersheds of the Huron-Erie corridor, and in the southwestern watersheds of Lake Erie. High SRP:TP ratio hot spots were seen in watersheds along the southeastern shores of Lake Huron and along the eastern side of the Huron-Erie corridor. Hot spots of low DIN:TP ratios with high nutrient export were seen in the southwestern watersheds of Lake Erie, which has implications for harmful algal growth. Nutrient ratios across the Great Lakes watersheds compared similarly to other heavily human impacted catchments of the Baltic Sea and western Europe. Annual basin loads of DIN, SRP, and TP were estimated from random forest models for each year from 2000-2016. Calculated annual nutrient loadings of SRP and TP were consistent with other published values of Great Lakes watershed estimates and revealed highest loadings during 2011 when the largest recorded algal bloom in Lake Erie occurred to date. Overall, this data-driven analysis of nutrient water quality reinforces and refines our process understanding of nutrient pollution dynamics across the Great Lakes drainage basin.

Acknowledgments

“I have found that it is the small everyday deed of ordinary folks that keep the darkness at bay. Small acts of kindness and love." - Gandalf (J. R. R. Tolkien ~ The Hobbit)

In my brief time at the University, I have found many small pockets of kindness, compassion and love in the never-ending hustle and bustle that is academic life. However, “ordinary” is the last word that I would use to describe my mentors, colleagues and friends that have surrounded me through this journey. In this quest for a Master’s degree, I feel like Bilbo accompanied by his band of proven, expert adventurers.

I first wanted to thank my supervisor and mentor, Dr. Nandita Basu, who has shown me warmth and wisdom along this meandering path we’ve taken together. I can’t express the gratitude I feel to work with someone who is not only a world-renowned expert and leader in her field, but whom also still manages to find time to encourage and counsel me despite her busy life. I also wanted to thank Dr. Kimberly Van Meter as she was vital in developing and supporting this project - you are both wicked-awesome.

Additionally, I wanted to thank Kirsten Grant for her tedious work in editing the many drafts of this thesis with a fine-tooth comb. Without her work, the quality of this paper would have suffered.

For the sake of conciseness, I wanted to briefly thank all my colleagues and friends in the BasuLab. I have leveraged the experience of so many, whether that’s asking about certain bits of code, or just casually distracting me from the ongoing burdens of graduate studies. Giving presentations is a lot easier when your audience is full of friendly faces and smiles. You have all helped foster a safe space for me to grow and learn, and I can’t thank you enough.

I also wanted to thank Global Water Futures, Lake Futures, and the Water Institute, for the many opportunities they have provided me in funding, travel, and research. From Saskatoon to San Francisco, I have had the resources to develop and widely share my research due to their generous support. I also wanted to thank the University of Waterloo and the provincial government for graciously awarding me the scholarships that put food on my table. They gave me financial stability during my program, which otherwise would have left quite a monstrous hole in my pocket.

Backstage to all this is my family and friends who lay as the foundation for my support network. These are the people who are directly responsible for my successes. Thank you Mom, Dad, Greg, and Lynn. I have struck the lottery to call you all as my family.

Last, but not least, I wanted to thank my partner, friend and treasure, Linden. My day shines brighter knowing I get to spend the next one with you.

Thank you.

Table of Contents

List of Figures

List of Tables

1.0 Introduction

1.1 Eutrophication and the Great Lakes

Anthropogenic nutrient inputs into water bodies pose serious threats to our water resources, particularly within the Laurentian Great Lakes basin. Excessive nitrogen and phosphorus loading into groundwater and surface water bodies can lead to significant water quality challenges downstream due to eutrophication (V. H. Smith, Joye, and Howarth 2006; Anderson, Glibert, and Burkholder 2002; Schindler 2006). Eutrophication describes the phenomenon of increased algal productivity associated with the nutrient enrichment of a water body. This enrichment promotes the rapid growth of harmful algal blooms (HABs) and nuisance algae blooms which are hazardous to both humans and ecosystem function. HABs produce toxins that are lethal to both humans and wildlife. Increased algal mass causes deep water (hypolimnetic) hypoxia in stratified water bodies, as oxygen is stripped out of the water column from increased respiration, creating ecologic “dead zones”. Consequences of eutrophication include ecosystem loss, fish kills, drinking water supply contamination, and diminished recreation, resulting in massive environmental, social and economic costs (Dodds et al. 2009; Pretty et al. 2003; Moss et al. 2011). Globally, billions of dollars are lost each year due to costs associated with eutrophication (Pretty et al. 2003; Dodds et al. 2009) .

The Great Lakes basin is particularly vulnerable to water quality threats due to its highly populated urban areas and substantial amount of agricultural land use (Environment and Climate Change Canada and Ontario Ministry of the Environment and Climate Change 2018a; Environment and Climate Change Canada and U.S. Environmental Protection Agency 2017). In the 1960’s and 70’s, Lake Erie’s water quality and biological diversity was severely degraded due to hypolimnetic hypoxia caused primarily by point sources of phosphorus from sewage discharge (Lee, Rast, and Jones 1978; Beeton 1965; Schelske 1979). In 1972, the Great Lakes Water Quality Agreement (GLWQA) was initiated as a binational agreement between Canada and the United States as an effort to control algal blooms and ecosystem losses, particularly within Lake Erie. The GLWQA initially focused on phosphorus reduction strategies of point source pollution , which were widely successful for improving water quality (Scavia et al. 2014; Colborne et al. 2019; Dove and Chapra 2015). Measures were implemented to reduce phosphorus in detergents and wastewater treatment plant discharge through legislated plant upgrades. These phosphorus reductions led to significant observed improvements in Lake Erie water quality during the 1980's and early 90's, although phosphorus sequestered from invasive zebra mussels and quagga mussels may have exaggerated improvements (Scavia et al. 2019). Despite these nutrient reduction efforts, eutrophication problems persist in the Great Lakes due to excessive nutrient loading in watersheds, primarily from non-point sources (Scavia et al. 2014; Dolan and Chapra 2012; Baker et al. 2014; Bootsma et al. 2015; Le Moal et al. 2019). In the last 25 years, Lake Erie has experienced a re-eutrophication, with increasing trends in algal blooms and hypolimnetic hypoxia despite reduced total phosphorus loads (Kane et al. 2014).

Water quality challenges from eutrophication in the Great Lakes are ongoing. Lake Erie routinely experiences algal blooms in the western and central basin during the summer months and in the summer of 2011, it experienced the largest recorded algal bloom in history as it stands (Michalak et al. 2013). In the summer of 2014, Toledo, Ohio's fourth largest city, issued a "do not drink" water advisory, due to HABs in the Lake (Fitzsimmons 2014). The city experienced elevated levels of microcystin in their drinking water, a lethal neurotoxin produced by the HABs' species of cyanobacteria (blue green algae) (Jetoo, Grover, and Krantzberg 2015). While much focus is on Lake Erie, HABs are also often recorded in shoreline areas of the other Great Lakes, including Muskegon Bay and Green Bay in Lake Michigan, Saginaw Bay and Georgian Bay in Lake Huron, and Hamilton Harbor, Oswego Harbor, and Bay of Quinte in Lake Ontario (Environment and Climate Change Canada and Ontario Ministry of the Environment and

Climate Change 2018a; Environment and Climate Change Canada and U.S. Environmental Protection Agency 2017). Cladophora, a nuisance algae, is problematic in the nearshore regions of Lake Michigan, Lake Ontario, and Lake Erie, and causes fouling of beaches and shorelines, and fouling of water intakes for drinking and cooling systems (Environment and Climate Change Canada and U.S. Environmental Protection Agency 2017; Bootsma et al. 2015).

Figure 1 – Aerial Image (brightened) of the Laurentian Great Lakes in 2015 showing algal blooms in Lake Erie ("NOAA Great Lakes Environmental Research Laboratory's Albums" 2015)

To combat the ongoing problem, binational agreements in 2016 between Canada and the U.S. have set the goal of reducing total phosphorus (TP) and soluble reactive phosphorus (SRP) loads into Lake Erie by 40 percent from 2008 levels by 2025 (Environment and Climate Change Canada and Ontario Ministry of the Environment and Climate Change 2018b; US EPA 2018) .

1.2 Challenges in Nutrient Management

Sources of nutrient pollution are generally categorized as point or non-point sources. Examples of the former include wastewater treatment outlets and septic systems, and typical examples of the latter include agricultural and urban runoff. While measures taken to control point sources in the Great Lakes basin have been generally considered successful, non-point sources are more difficult and costly to control (Schindler 2006; Lee 1973). Non-point sources of fertilizers applied on fields in excess of crop requirements accumulate in the landscape and enter riverine systems through surface and subsurface pathways. This contributes to current and future eutrophication problems in downstream water bodies (Han, Allan, and Bosch 2012). Non-point sources are deemed as the major source of nutrient pollution and driver of eutrophication in the Great Lakes, and strategies to improve water quality should focus on reducing these sources of nutrients, particularly phosphorus delivery (Lee 1973; Joosse and Baker 2011; Le Moal et al. 2019; Carpenter et al. 1998).

Phosphorus pollution in Lake Erie is widely accepted as the driver of eutrophication in the lake, and phosphorus is often the limiting nutrient in freshwater environments for algal growth (Correll 1998; Environment and Climate Change Canada and U.S. Environmental Protection Agency 2017; Schindler 1974; Sharpley et al. 1994). The form of phosphorus is key to consider for management, especially since dissolved forms (SRP) are more bioavailable and conducive to algal growth (Baker et al. 2014). While TP loadings to Lake Erie have remained stable in recent years, SRP loads have increased, correlating with increasing algal blooms (Douglas R. Smith, King, and Williams 2015; Joosse and Baker 2011; Daloğlu, Cho, and Scavia 2012). Management strategies for phosphorus should be tailored to the specific form they are targeting; TP management does not necessarily translate to SRP management. This is especially important as desired outcomes of management strategies may conflict. For example, while no tillage agricultural practices lead to TP reductions from less particulate forms entering waterways, this strategy increases SRP losses from fields (Lam et al. 2016; Douglas R. Smith et al. 2015). As such, targeted management of specific forms of nutrients, particularly phosphorus, is important to consider for implementing effective reduction strategies in the Great Lakes basin.

It is key to manage nitrogen pollution in conjunction with phosphorus since outcomes are often at odds with each other when considered separately. Broad strategies that only consider phosphorus reductions may be inadequate to address local nutrient processing and cycling conditions in some freshwater environments, especially where phosphorus is easily recycled (Conley et al. 2009). Furthermore, coastal and estuarine systems are often nitrogen limited, and strategies that focus solely on phosphorus upstream could exacerbate eutrophication challenges in these downstream environments (Conley et al. 2009; V. H. Smith and Schindler 2009). In the context of the Great Lakes drainage basin, this could have impacts downstream in the St. Lawrence River and the Gulf of St. Lawrence. Therefore, managing nitrogen and phosphorus together is crucial in aligning both broad and specific nutrient water quality goals.

Although increased nutrient export and concentrations are major drivers of eutrophication, nutrient ratios of phosphorus and nitrogen in aquatic systems are also important for management, especially when considering harmful algal growth and trophodynamics (Glibert and Burkholder 2011; Glibert et al. 2011; Saaltink et al. 2014). Freshwater systems that are enriched in phosphorus are more conducive to the growth of HAB organisms like cyanobacteria (Anderson, Glibert, and Burkholder 2002). Cyanobacteria favour low nitrogen to phosphorus ratios due to their ability to fix N_2 from the atmosphere (V. H. Smith and Schindler 2009). This is particularly concerning as cyanobacteria produce toxins that are lethal to humans, wildlife, and livestock, while also rendering drinking water resources unusable. Strategies that only focus on specific nutrients for management (i.e. only phosphorus or nitrogen) may overlook the eutrophication challenges associated with altered nutrient ratios. Therefore, focus on managing pollution should also consider the breadth of the nutrient water quality regime, rather than only single problematic pollutants.

Best management practices (BMPs) can be implemented to reduce nutrient loads from non-point sources. BMPs include measures such as buffer strips, cover crops, livestock fencing, tillage and fertilizer application practices. These practices intercept and treat nutrient rich runoff prior to entering waterways or prevent nutrient pollution at the source. Combinative approaches that use multiple different BMPs in conjunction are more effective in reducing nutrient loads from non-point sources than individual BMP use (Scavia et al. 2019; Bosch et al. 2013; Lam et al. 2016). Furthermore, BMP's performance has been shown to be most effective when targeting hot spots of nutrient pollution, rather than distributed or random placement on the landscape (Bosch et al. 2013; Park et al. 1994). Therefore, it is key to recognize and isolate hot spots of nutrient pollution to better implement reduction strategies.

It is challenging to consider all aspects of nutrient water quality for management. Targeted reduction strategies for specific forms of nutrients should be balanced with the objective of remediating the entire nutrient water quality regime so that desired outcomes do not conflict. This challenge in management is further compounded given our lack of knowledge in where adverse nutrient water quality problems exist in the Great Lakes basin. Identifying hot spots for the entire extent of poor nutrient water quality is essential in applying targeted and broad reduction strategies. In this study, our analysis investigated drivers and controls of nutrient pollution from a more comprehensive view of water quality. This was conducted by considering nitrogen water quality in conjunction with phosphorus, while also considering critical nutrient ratios. By doing so, we have identified hot spots and drivers of nutrient pollution through a more complete picture of the nutrient water quality challenges facing the Laurentian Great Lakes basin.

1.3 Gaps in Understanding and a New Approach

“We suggest that future research be focused on the cumulative effects of nutrient loading and other human-caused insults to lakes” (V. H. Smith and Schindler 2009)

Many factors add to the difficulty of parsing out the effects of non-point sources of nutrient inputs and their impacts to receiving waters. Nutrient export not only depends on the activities and sources within a watershed, but also on the local hydrology and biogeochemical processes, making source characterization and load estimation troublesome (Hamilton 2012). Much work in hydrology focuses on trying to capture the vast heterogeneity of landscapes and incorporate this into our process understanding (McDonnell et al. 2007; Wagener et al. 2010). However, when considering nutrient management, there are many gaps in our understanding from source to sink. Recent gaps include the long-term effectiveness of various agricultural practices for nutrient reduction at the field scale (H. Jarvie et al. 2017; Douglas R. Smith et al. 2015; H. P. Jarvie et al. 2013; Scavia et al. 2017). Uncertainty exists pertaining to the impacts of climate change on nutrient transport and the efficacy of BMPs (Kalcic et al. 2019; Bosch et al. 2014). Another particularly prominent gap is the lasting impacts and contribution of nutrient legacies in the landscape (Hamilton 2012; Van Meter and Basu 2017; Meals, Dressing, and Davenport 2010). Efforts have been made to quantify the lag time between the implementation of reduction strategies and when observed improvements can be expected, although uncertainty still exists in landscape responses. Sources, flow paths, and geochemical processes in groundwater are often neglected as a non-point source of nutrients, and processes at reactive interfaces, such as riparian and hyporheic zones, are not well understood, especially in the Great Lakes basin (Robinson 2015). Riverine fluxes of nutrients not only vary with discharge, but also with time and season, and sparse stream water quality measurements fail to adequately fully capture these dynamics (Hirsch, Moyer, and Archfield 2010). These gaps of knowledge are especially important when considering equifinality, as isolating the effect of multiple compounding factors in nutrient dynamics is difficult in poorly understood systems, especially when observations are limited. This contributes to the complexity of accurately capturing the heterogeneity of nutrient dynamics and responses within watersheds. Therefore, more holistic methods are needed to bridge our gaps in understanding to tackle the challenge of eutrophication in the Great Lakes.

Downstream water quality signatures can be assessed to explore the behaviour and drivers of nutrient export and to better understand the processes within upstream watersheds (Sivapalan 2006). For example, isotopic analysis can be used to determine the origins of water (e.g. runoff, groundwater) or dissolved constituents in downstream discharges (Gibson et al. 2002; Cole et al. 2004). Downstream water quality signatures gathered from these studies often offer one of the few metrics available for assessment of complex environmental systems upstream. However, using this approach remains challenging due to limited resources in monitoring, computation, and personnel.

Recently, machine learning methods have grown in popularity as tools to assess environmental signatures and systems, while avoiding challenges associated with other methods (Tyralis, Papacharalampous, and Langousis 2019). Machine learning offers an alternative to mechanistic modelling or experimental based work and can be used in conjunction to support findings and conclusions. Because of our gaps in understanding, many in the research community advocate for a holistic approach to evaluating environmental challenges in hydrology and nutrient pollution (McDonnell et al. 2007; Wagener et al. 2010; H. P. Jarvie et al. 2013; Moss 2008; Withers et al. 2014). This means an approach that evaluates the entirety of the challenge by looking at the cumulative effects of nutrient pollution, rather than focusing on individual processes or parts. This “big-picture” evaluation is crucial, especially given our incomplete understanding of the detailed physical, biological, and chemical interactions of nutrients in our environment. Machine learning can bridge our detailed process understanding of complex environmental systems to a more universal, encompassing point of view. These methods offer a simplistic, yet analytically robust framework to evaluate environmental challenges from holistic perspectives.

Random forest (RF) regression is one particular machine learning tool that has been used to evaluate nutrient water quality of surface waters and groundwater within North America (King, Cheruvelil, and Pollard 2019; Carlisle, Falcone, and Meador 2009; Read et al. 2015; Shen et al. 2020; Dugan et al. 2020). It has grown in popularity in recent years as a tool for hydrological modelling and assessment. These data driven models have been shown to have similar or increased predictive power when compared to other widely used and applied empirical and process-based models (Tyralis, Papacharalampous, and Langousis 2019; Solomatine and Ostfeld 2008). Advantages to RF models include their non-parametric nature and their ability to handle noisy, nonlinear, and intercorrelated data (Tyralis, Papacharalampous, and Langousis 2019; Breiman 2001; Meinshausen 2006). These models are also robust to overfitting, and can use both continuous and categorical variables (Breiman 2001; Meinshausen 2006; Liaw and Wiener 2002). To the authors' best knowledge, as it stands, this paper presents the first application of this tool for stream nutrient water quality of the entire drainage basin of the Laurentian Great Lakes.

1.4 Research Objectives

The objective of this analysis is to improve our understanding of the behaviour and responses of Great Lakes watersheds to nutrient pollution, by applying a data-driven approach to investigate trends in water quality. To move towards more targeted nutrient management, we must acknowledge and better characterize the heterogeneity of responses and behaviours of Great Lakes watersheds to nutrient inputs.

The following research questions have been posited to frame the objectives and context of this research:

1. What landscape variables drive nutrient pollution in the Great Lakes watersheds?
    a. What spatial factors are important in annual flow weighted concentrations?
    b. What spatial factors are important in annual flow weighted ratios?

2. What are the spatial patterns of nutrient pollution across the Great Lakes watersheds?
    a. What are the annual flow weighted concentrations of nutrients in the Great Lakes watersheds?
    b. What are the annual flow weighted ratios of nutrients in the Great Lakes watersheds?

3. What is the magnitude of annual nutrient loads into the Great Lakes?

2.0 Methodology

2.1 Source Datasets and Pre-processing

2.1.1 Study Area

The Laurentian Great Lakes drainage basin lies on the Canada-United States border, within the Canadian province of Ontario, and the American states of Minnesota, Wisconsin, Illinois, Indiana, Michigan, Ohio, Pennsylvania, and New York. The total basin area is more than 520,000 km^2, with 59% in the United States, and the remaining 41% in Ontario, Canada (Neff et al. 2005; MacDonagh-Dumler, Pebbles, and Gannon 2003). The Great Lakes span an area of 244,000 km^2, with more than 17,000 km of shoreline and contain approximately 23,000 km^3 of water, about one fifth of the world's surface freshwater supply. The basin is home to over 33 million people and the lakes provide drinking water to millions of Americans and Canadians who reside near its shores (Environment and Climate Change Canada and U.S. Environmental Protection Agency 2017; MacDonagh-Dumler, Pebbles, and Gannon 2003). The Great Lakes are also relied upon for transportation, fishing, industry, agriculture, and recreation.

The physical characteristics of the Great Lakes basin vary widely from north to south Basin (Canada, Shear, and Wittig 1995; Neff et al. 2005). In the northern parts of the basin, around Lake Superior and northern Lake Huron and Georgian Bay, a granite bedrock known as the Canadian Shield is the common feature of the terrain surface, with a thin cover of acidic soils and mostly conifer dominated forests. In the southern basin, soils are deeper and more fertile, with varying deposits of sand, silt, clay, and boulders from past glaciation, and underlaid by sedimentary rock of limestone, sandstone, shale, and gypsum. The south is also home to most of the agriculture and urban centres of the basin, with major cities including Milwaukee and

Chicago on Lake Michigan, Detroit, Buffalo, and Cleveland on Lake Erie, and Toronto and Hamilton on Lake Ontario.

Climate also varies widely in the Great Lakes basin (Canada, Shear, and Wittig 1995; Neff et al. 2005). Average July temperatures range from 25°C in Indiana, Ohio, Michigan, and Illinois to the south, to 17°C in the north and eastern regions from Lake Superior. Average January temperatures range from -2°C in Indiana and Ohio in the south, to -20°C in the northern basin of Lake Superior. Average annual precipitation ranges from greater than 1200 mm in regions east of Lake Ontario, to less than 690 mm in regions west of Lake Superior. Snowfall varies even more than precipitation, with the most hard-hit areas generally to the east of each Great Lake.

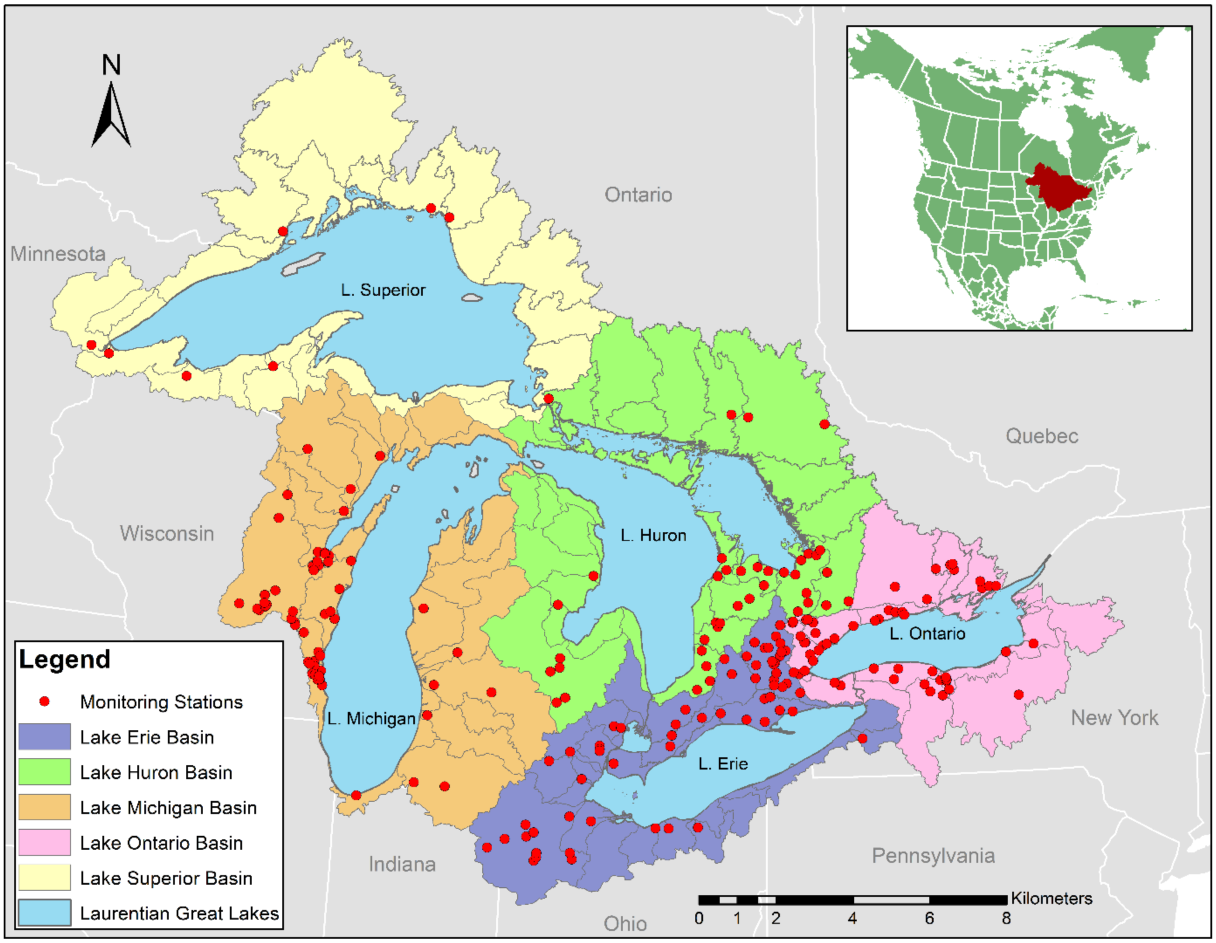

Figure 2 – Water quality monitoring stations and watersheds of the Laurentian Great Lakes drainage basin used in this study.

2.1.2 Flow and Water Quality Data Sources

Discharge data for Canada was obtained from Environment and Climate Change Canada's (ECCC) discharge monitoring network ("Environment and Climate Change Canada Historical Hydrometric Data" 2016), while discharge data for the United States was obtained from the United States Geological Survey (USGS) ("National Water Information System" 2016). Water quality data for Canada was obtained from the Provincial Water Quality Monitoring Network (PWQMN) ("Provincial (Stream) Water Quality Monitoring Network" 2016), while that for the United States was obtained from the USGS ("National Water Information System" 2016).

Monitored water quality constituents used in this analysis were nitrogen (DIN) as combined nitrate and nitrite, Soluble Reactive Phosphorus (SRP) and Particulate Phosphorus (PP). PP concentrations were calculated by subtracting SRP from monitored Total Phosphorus (TP).

Monitoring stations in the Great Lakes drainage basin (Figure 2) were selected such that both discharge and water quality were measured in co-located areas on the same river body. Discharge stations were selected that were sufficiently close to the water quality stations, based on the two following decision criteria: 1) The water quality and streamflow station lay on the same river stem; 2) the percent difference in drainage area between the water quality and discharge station was less than 15%. Stations with more than 40 records and data between 2000 and 2016 were selected. Using these criteria, 159, 160, 159, and 195 stations were selected with water quality data for DIN, SRP, PP, and TP, respectively. In total, 202 co-located monitoring stations across the Great Lakes basin were used in this study.

Two nutrient ratios were calculated from monitored data and used in this study: SRP to TP (SRP:TP), and DIN to TP (DIN:TP). The SRP:TP ratio is a useful water quality metric that represents the fraction of phosphorus that is bioavailable. The DIN:TP ratio can yield insights into nutrient limitation, showing whether a water quality regime is potentially nitrogen or phosphorus limited. This is important for algal growth, as certain species and communities can favour different nutrient limiting regimes, such as cyanobacteria under low DIN:TP ratios.

2.1.3 Spatial Data Sources on Catchment Attributes

Spatial data was used for developing models of watershed nutrient water quality. Geoprocessing of spatial data was performed using ArcGIS, version 10. Spatial data sources were often specific to each country, as data coverings for the United States seldom overlapped with Ontario, and vice versa. A summary of the spatial data sources used can be found in Table A.1 of the Appendix. Catchments draining to United States monitoring stations were delineated using ARCGIS's hydrology toolbox. Catchments draining to Canadian monitoring stations were delineated using the Ontario Flow Assessment Tool (OFAT) from the Ontario Ministry of Natural Resources and Forestry.

Spatial data were aggregated into single spatial variable values for each watershed. These values were calculated using an area-weighted approach for geoprocessing spatial objects within a watershed, as given by the following equation:

$$X_{avg} = \frac{\sum_{i=1}^{n} X_i A_i}{A_{total}} \quad (1)$$

where X_{avg} is the calculated spatial variable value used for the watershed, X_i is the numerical value of the i^{th} spatial object within the bounds of the watershed, A_i is the area of with the i^{th} spatial object, and A_{total} is the total area of the watershed.

Land use data for the Great Lakes drainage basin was obtained from the 2015 Agriculture and Agri-Food Canada annual crop inventory for Ontario, and from the 2011 National Land Cover Database for the United States. Gridded land use data was aggregated into groups based on relevant land use and crop codes. Land use was grouped into spatial variables of percent forested land cover, percent wetland cover, and percent developed land cover (agricultural and urban land use). Tile drainage data for Ontario was obtained from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA), and for the United States from the United States Census of Agriculture, National Agricultural Statistics Services. Population density data for Ontario was obtained from 2011 Census data from Statistics Canada, and data for the United states was obtained from sub-county 2010 census data from the United States Census Bureau. Precipitation and temperature data were obtained from the WorldClim database at 1 km resolution (Fick and Hijmans 2017). Climate variables used were the average annual precipitation in millimetres per year in a watershed, and average annual temperature in degrees Celsius of a watershed. Slope data for the United States was obtained from the hydrologic landscape regions of the United States dataset (Wolock, Winter, and Mcmahon 2004). Ontario slope data was geoprocessed from a 30m digital elevation model obtained from OFAT. Slope variables were given as average percent slope in a watershed. Soil data for the United States was obtained from the Soil Survey Geographic Database by the United States Department of Agriculture, and soil data for Ontario was obtained from the National Soil Database by the Canadian Soil Information Service. There were gaps in soil data coverage for Northern Ontario and these were filled using data from the Harmonized World Soil Database (Fischer et al. 2008). Livestock density data for cattle, swine and chickens were obtained from the Food and Agriculture Organization of the United Nations (FAO) as part of their Gridded Livestock of the World v 2.01. Cattle, swine and chicken densities were then converted into a single "Livestock Equivalent Density" using animal unit coefficients for Ontario from Statistics Canada (Government of Canada 2001). These coefficients are based on the amount of manure each animal type would produce to fertilize a standardized acreage of cropland.

Watershed delineation, in addition to the geoprocessing of land use, tile drained land, population density, climate, slope, and soil type spatial data, used in this work was conducted by Chowdhury (2018). Processing of livestock density spatial data was performed by the author.

2.1.4 Weighted Regression on Time, Discharge and Season (WRTDS)

Water quality measurements are often sparse and therefore it is difficult to effectively represent the water quality regime in a stream. A Weighted Regression in Time Discharge and Season (WRTDS) model was applied for each tributary station using monitored water quality and discharge to generate continuous daily water quality for the co-located stations. WRTDS has grown in popularity in recent years as a water quality processing tool, and has been used for nutrient load estimation and analysis in the United States, including the Mississippi River (Sprague, Hirsch, and Aulenbach 2011), Lake Champlain tributaries (Medalie, Hirsch, and Archfield 2012), and the Susquehanna River Basin (Q. Zhang, Brady, and Ball 2013; Qian Zhang, Harman, and Ball 2016). WRTDS, developed by Hirsch et. al, is a dynamic model for water quality that not only accounts for concentration-discharge relationships, but also incorporates seasonal and time variability (Hirsch, Moyer, and Archfield 2010). It is developed using the following equation structure:

$$\ln(c) = \beta_0 + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \epsilon \tag{2}$$

where c is concentration, t is the time as a fraction of the year (e.g. Jan. 1st = 0, and Dec. 31st = 1), Q is the average daily streamflow, $\beta_{0\text{-}4}$ are regression coefficients, and ε is the residual error (Hirsch, Moyer, and Archfield 2010). This regression model is fitted to monitored discharge and water quality time series data. Regression coefficients, $\beta_{0\text{-}4}$, are not static, and are unique to each point from regression fitting. The equation structure has linear dependencies of concentration with discharge (natural log transformed) and time, in addition to sinusoidal dependencies with time of season. As such, this regression model calculates daily concentrations based on regressed relationships with discharge, time of year, and season. WRTDS was implemented for monitored water quality using the EGRET package in R from the USGS.

WRTDS fits the regression equation based on “distances” in time, discharge, and season from the monitored data. Time distance is the difference in the time at the estimation point to the time of monitored data, in units of fractions of a year. Discharge distance is the difference in the natural log-transformed discharge at the estimation point to the natural log-transformed discharge of monitored data, in units of natural log-transformed flow. Seasonal distance is the difference between seasons in the estimation point and monitored data, in units of fraction of a year (e.g. the seasonal distance between September 1, 2011 and September 1, 2012 equals zero, while the seasonal distance between September 1, 2011 and December 1, 2012 equals 0.25). Distances in time, discharge, and season for each monitored data point are weighted for fitting the regression equation based on Tukey's Tricube Weight Function given below:

$$\omega = \begin{cases} \left(1 - \left(\frac{d}{h}\right)^3\right)^3 & if\ |d| \leq h \\ 0 & if\ |d|\ > h \end{cases} \tag{3}$$

where ω is the unitless weight, d is the distance in time, discharge, or season, and h is the half window width in the same units as distance (Tukey 1977). The overall weight for each monitored data point is the product of the weights for time, discharge, and seasonal distances. This Tri-Cube weight function ensures that data points close to the estimation point are fitted to the regression better than those farther away. This weighting structure assigns greater weights to monitored data points that are closer to the estimation point in time discharge and season, while monitored data points that are farther away are cubically less weighted. Monitored data points with distances greater than the half window width are given weights of zero and are therefore not considered for fitting. Half window widths for discharge and season were set at the recommended values of 0.5 and 2 natural log units (Hirsch and De Cicco 2015). Half window widths for time were recommended as 7 years, although including the time varying component in WRTDS would induce non-stationarity in the regression (i.e. concentrations would be changing with time). To invoke stationarity into WRTDS, half-window widths were manipulated and set to 100 years. In this sense, all time distances in the monitored data would be weighted equal to 1 at any estimation point, and therefore the time component would be static in the regression (i.e. β_1 would not change at any estimation point). As such, the regression would only capture discharge and seasonal relationships with concentration from the monitored data.

WRTDS provides an approach that minimizes bias to estimating the concentration of a tributary given the date and discharge of estimation (Hirsch, Moyer, and Archfield 2010). While concentration data is log-transformed in the regression equation structure, re-transformation bias is corrected in actual concentrations through a weighted smearing estimator (Duan 1983). Regression coefficients in WRTDs are unique to each estimation point, and as such are unbiased to other estimations. Furthermore, weights for distances between monitored data points and estimation points are determined so that the nearest points have the greatest effect in estimated nutrient concentrations. WRTDS is a free-form regression method that is dynamic to a wide range of water quality regimes, while accounting for hydrologic factors specific to stream water chemistry.

Bias was assessed in WRTDS results by comparing monitored data with predicted data from the regression. Modelled data that exceeded our bias criteria was eliminated from further use. For each monitoring station, a Flux Bias value was assessed, as given by the following equations:

$$P = \sum_{i=1}^{n} C_{i,WRTDS}\, Q_i \tag{4}$$

$$O = \sum_{i=1}^{n} C_{i,Observed}\, Q_i \tag{5}$$

$$B = \frac{P - O}{O} \tag{6}$$

where B is the Flux Bias, P is the predicted flux, O is the observed flux, n is the number of observed data points, $C_{i,WRTDS}$ is the estimated concentration at day i from WRTDS, $C_{i,Observed}$

is the observed concentration at day i from the monitored station data, and Q_i is the observed discharge at day i from the monitored station data. Positive values of Flux Bias indicate overestimation of concentrations by WRTDS, while negative values indicate underestimation. Stations that had Flux Bias values of less than -0.15 and greater than +0.15 were deemed too biased in their WRTDS results, and as such were omitted from further use in this analysis.

2.1.5 Annual Flow Weighted Concentration (FWC)

Annual flow-weighted concentrations (FWC) of DIN, SRP and PP were used as a metric to assess the effect of various spatial factors on nutrient water quality.

$$FWC = \frac{\sum C_i Q_i}{\sum Q_i} \tag{7}$$

where, C_i is the concentration on day i and Q_i is the discharge on day i. Daily concentrations were estimated using WRTDS applied to monitored concentration data from the PWQMN ("Provincial (Stream) Water Quality Monitoring Network" 2016) and USGS ("Water Quality Portal" 2020), while daily discharge was obtained from ECCC (Environment and Climate Change Canada 2020) and the USGS ("Water Quality Portal" 2020). Nutrient ratios were calculated by dividing their respective FWCs. The FWCs and ratios for monitored watersheds were then used for developing nutrient water quality models.

FWCs are independent of discharge and can be used to capture a snapshot of water quality that is not affected by year to year variation in discharge and precipitation. Average annual FWCs are useful signatures for assessing the water quality within a watershed, provided land use and activities are stationary on an annual basis. FWCs can be used for calculating receiving loads, as discharge from a watershed can be used in conjunction with FWCs to calculate receiving loads.

2.1.6 Variable Selection and Collinearity

Spatial variables were screened prior to modelling to minimize collinearity and redundancy in the data set following methodologies adopted by Jolliffe (Jolliffe 1972; 1973) and Mansfield and Helms (Mansfield and Helms 1982). This was applied so that important spatial drivers of nutrient water quality could be better assessed; collinearity and redundancy pose challenges when trying to determine the effect a variable has on a model or system's response. A correlation matrix and variance inflation factors (VIF) were computed to assess collinearity between variables. A principal component analysis (PCA) was also conducted on select variables for further variable reduction to assess and limit redundant information in the independent variable set. The correlation matrix of all variables was generated and assessed for collinearity between specific variables. VIFs were calculated to quantify collinearity of a single variable with all independent variables. The correlation matrix and VIFs can be seen in Figure A.1 in the Appendix.

Forested land use was omitted from use in this study due to its significant correlation and intercorrelation with other independent variables. Forested land use showed high correlation with specific variables, as shown by the high correlation coefficients with developed land, wetlands, and silt and clay soils (Figure A.1). Forested land use also showed high intercorrelation amongst all variables as given by its large VIF. Forested land use was omitted because of this high correlation and intercorrelation, and intercorrelation was then reassessed among the remaining variables. As seen by the smaller VIF values without forested land use, the intercorrelation amongst the remaining variables was significantly lower (Figure A.1). VIF values were less than 10 without forested land use, which is considered an acceptable cut-off for intercorrelation (Stine 1995; Chatterjee and Yilmaz 2016). As such, forested land use was omitted from further use.

Average percent slope, average annual temperature, and average annual precipitation were considered to have redundant information with developed land use and were omitted from use in this analysis. Correlation coefficients between these variables were significant (p-values < 0.01), indicating that they likely shared information and were assessed for redundancy and removal. PCA offers insight into the information (i.e. variance) that variables share with one another, by projecting them into new variables, called "components", that do not share information with each other. Components are then ranked by the amount of information that they contain, with the first component (Component 1) having the most information, and the last component (Component N) having the least information. Each component is proportionally made up of projections of the original variables, with more significant variables having higher proportions in a component. Components were extracted from spatial variables of slope, temperature, precipitation and developed land use, and were assessed for variable reduction in accordance with traditional PCA methods (Jolliffe 1972; Al-Kandari and Jolliffe 2005). Variables with the highest proportions in components with the lowest information were omitted. These omitted variables were considered redundant, since information associated with these variables accounted for the least information in the entire data set. As such, average percent slope, average annual temperature, and average annual precipitation had the highest proportions in the three smallest components, as seen in Figure A.2 of the Appendix. This indicates these variables have little additional information when compared to developed land use. Therefore, these variables were deemed to have redundant information and were omitted from further use.

Watershed drainage area was omitted as a variable for use in this analysis. During preliminary stages of this research, watershed area was considered as a potential driver for nutrient water quality. However, it was consistently shown to be among the least important variables in random forest models for nutrient FWCs and ratios (variable importance is discussed in Section 2.2.2). Additionally, model performance remained unchanged with its omission. Because of its unimportance and lack of predictive power in nutrient water quality models, watershed area was omitted from further use in this analysis.

2.2 Modelling Framework

Random forest (RF) machine learning was used as a modelling framework to assess controls on nutrient FWCs and ratios in monitored watersheds of the Great Lakes. Variable importance and feature contribution metrics determined from this framework were used to assess drivers and behaviours of nutrient water quality across spatial variables within monitored watersheds. RF models underwent a training and validation process before then being applied to predict nutrient water quality across the entire ungauged basin of the Great Lakes, with metrics of uncertainty reflected in modelled predictions.

2.2.1 Random Forest Machine Learning Structure

RF models generate fitted binary decision trees based on random, independent sampling (with replacement) of independent variables and data (Figure 3) (Breiman 2001). Nodes where decisions are split are determined through minimizing mean square error in the sampled data. RF regression was applied in this analysis, meaning that an average value was taken from each numerical output of all the generated decision trees. Since many decision trees were generated, a distribution of predictions was also obtained, and this was used for estimating standard error and prediction uncertainty.

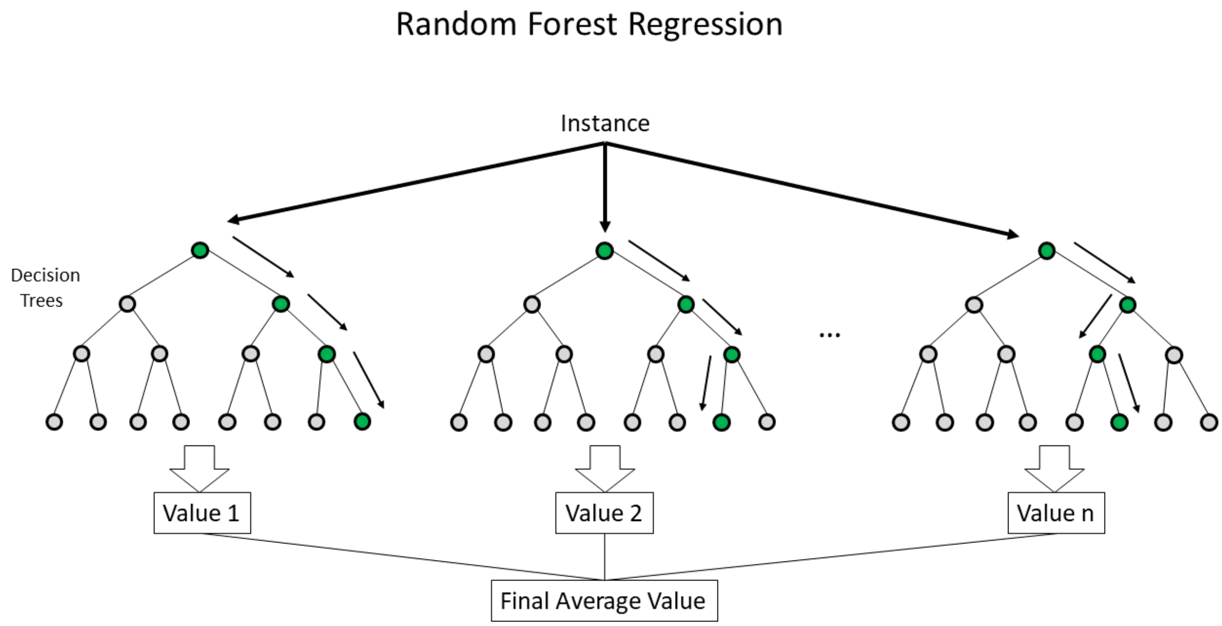

Figure 3 – Simplified diagram of random forest regression structure. Adapted from Jagannath (2017).

The independent variables used in this study were the spatial factors within watersheds remaining after variable selection: the percentage of developed land use, the percentage of wetlands, the percentage of land tile drained area, percent of silt and clay soils, and livestock densities. Variables were selected to minimize collinear dependencies and redundancy (see Section 2.1.5), so variable importance could be better assessed.

To perform the analysis, MATLAB's Statistics and Machine Learning Toolbox was used by applying their bagged regression tree ensemble functions. Models were specified to generate 1000 decision trees, with error converging to a minimum well below the number of trees generated in each model. Error convergence over the number of trees grown can be found in Figure A.3 of the Appendix. Standard values were used for the fraction of data sampling with replacement, and the fraction of variable sampling at decision splits, set as 1 and ⅓, respectively (Breiman 2001). A sample decision tree from an RF model used in this analysis can be seen in Figure A.5 of the Appendix.

2.2.2 Variable Importance

RF models are commonly used to assess which variables are most important for response. Variable importance is determined though perturbing input variables and determining which variable causes the greatest deviation in mean square error (Grömping 2009). Variable importance is given by the following equation structure:

$$V_{t,X_j} = \frac{1}{n_{OOB_t}} \sum_{i=1}^{n_{OOB_t}} \left(y_i - \hat{y}_i\left(X_{j\,permuted}\right)\right)^2 - OOBMSE_t \tag{8}$$

where V_{t,X_j} is the variable importance of the variable X_j for decision tree t in the RF model, n_{OOB_t} is the number of “Out of Bag” (OOB) observations for the decision tree, y_i is the prediction of the decision tree for the i^{th} OOB observation, and $\hat{y}_i\left(X_{j\,permuted}\right)$ is the prediction when the variable X_j is perturbed, and $OOBMSE_t$ is the mean square error for the decision tree across all OOB observations. “Out of Bag” refers to observations that were not sampled during training of the decision tree t, and therefore indicate an unbiased estimate of error for that tree. An overall variable importance for the variable X_j in entire RF model is then taken as the average value across all decision trees in the model. In this study, the variable importance was taken as an average value across all 1000 decision trees.

The perturbed variable that results in the greatest change in mean square error is considered the most important. This sensitivity analysis was performed to determine which independent spatial variables were most important in the RF framework. Variables that were expected to be the most dominant controls in nutrient water quality were also expected to have the greatest importance, since changes in these spatial variables would result in the greatest changes in concentrations and ratios. To assess the importance of spatial factors remaining after variable selection, a RF model for each of the three nutrient water quality parameters (DIN, SRP and PP) and two ratios (SRP:TP and DIN:TP) was generated using all spatial parameters of gauged watersheds as independent variables.

2.2.3 Feature Contribution (FC)

Feature Contribution (FC) was used to assess the influence and relationships spatial variables had on nutrient concentrations and ratios in the RF models. FC is a useful way to assess the effects of variables in “black box” modelling, like machine learning. FC reveals the behaviour of a model with respect to a model variable, while the effects of other variables are averaged out. FC was adapted from partial dependence interpretation (Friedman 2001)

For any given model, $f(X)$, where X is the entire variable set for that model, let X^S be a single variable of that set, and X^C be the complementary set (i.e. all the other variables), such that $f(X) = f(X^S, X^C)$. The FC for a variable of a model can be given by the following equations:

$$f^S(X^S) = E_C\,[f(X^S, X^C)] = \int f(X^S, X^C)\, p_c(X^C) dX^C \qquad (9)$$

$$p_c(X^C) \approx \int p(X) dX^s \qquad (10)$$

where $f^S(X^S)$ is the FC of the single variable in the model, and $p_c(X^C)$ is the marginal probability of the complimentary set X^C. Feature Contribution is the expected value of the model for a variable, X^S, when all the other variables, X^C, are not taken into consideration. This can be approximated by the following:

$$f^S(X^S) \approx \frac{1}{N}\sum_{i=1}^{N} f(X^S, X_i^C) \tag{11}$$

$$X_i = \left(X_i^S, X_i^C\right) \tag{12}$$

where X_i are observed data, X_i^S is the i^{th} observations for the variable, and X_i^C are the i^{th} observations for the complementary data sets. For interpretability, FC's were mean subtracted and normalized by the standard deviation, as given by the following:

$$f^S(X^S)' = \frac{f^S(X^S) - \mu_{observed}}{\sigma_{observed}} \tag{13}$$

Where $\mu_{observed}$ and $\sigma_{observed}$ are the mean and standard deviation of observed responses, respectively. This standardization of FCs was conducted so that they could be compared across different nutrient concentrations and ratios, where magnitudes and units differed for responses (e.g. µg/L vs mg/L vs mol:mol).

Plots of FC show the average effect a variable has on the response of a model, while isolated from the effects of other variables (Friedman 2001). This effect is shown as the variable's contribution to the response in standard deviations from the mean overall predicted value. For example, an FC value of 0 shows that a variable has no effect in nutrient concentrations or ratios predicted by the model at that point. However, a positive value shows an increasing effect on predicted nutrient concentrations and ratios compared to the mean overall prediction, and a value of less than zero shows a decreasing effect compared to the mean overall prediction. Since FC is associated with a specific variable, each observation in the data set has an associated FC value for each spatial variable used in the RF model. FCs for RF models were implemented using the plot partial dependence tool in MATLAB's Statistics and Machine Learning Toolbox.

2.2.4 Model Training and Validation

RF models do not require an independent data set as traditionally set aside during training and then used for cross-validation. Instead, "Out of Bag" (OOB) data, data that is not sampled and used in model development, can be used to provide estimates of unbiased model performance and validation (Breiman 2001; Tyralis, Papacharalampous, and Langousis 2019). Since the fraction of data sampling with replacement was set to 1, the OOB data used for model development is approximately ⅓ of the entire data set. The other approximately ⅔ of data was used for training of RF models. Since sampling was conducted randomly in the RF Model, random number generators were seeded with the same starting values for reproducibility of results during model training and validation.

Selecting and training RF models for use in prediction was an iterative process. A final RF model was developed for each of the three nutrient water quality parameters (DIN, SRP and PP) and two ratios (SRP:TP and DIN:TP) using the spatial variables deemed most important to water quality response from variable importance metrics. Models were then iteratively generated, starting with the two most important spatial variables, and then adding subsequent variables in decreasing order of variable importance. Final models were then selected using those with the spatial variables that yielded the minimum OOB mean square error, or with error changing by less than 1%. This methodology for RF model development is consistent with those outlined in other studies and publications (Ziegler and König 2014; Díaz-Uriarte and Alvarez de Andrés 2006). Results RF model training and development can be found in Table A.2 through Table A.6 of the Appendix, with final RF models for each nutrient water quality constituent selected indicated in bold. Final RF models were then applied to the entire unmonitored Great Lakes basin for spatial analysis of hot spots and loading calculations.

2.2.5 Model Prediction and Coefficient of Variation (CV)

Final RF models were used to predict FWCs and ratios across all ungauged watersheds of the Laurentian Great Lakes. This was done by using the spatial parameters of ungauged basins for input into the RF models to predict concentrations or ratios. Modelled nutrient water quality across the entire Great Lakes basin was then mapped and assessed for spatial patterns.

Additionally, coefficients of variation (CV) were calculated as a measure of uncertainty in predictions based on the variability of predictions from decision trees. Since RF models were specified to generate 1000 decision trees, a distribution of predictions was obtained and used for estimating standard error and prediction uncertainty. CVs were calculated by dividing the standard error of decision tree predictions by the mean predicted response for each watershed, as given by the following equation:

$$CV = \frac{\sigma_{prediction}}{\mu_{prediction}} \tag{14}$$

where $\sigma_{prediction}$ is the standard deviation of model predictions for the watershed from each decision tree in the RF model, and $\mu_{prediction}$ is the mean predicted nutrient water quality value for the watershed from the RF model. CVs allow us to estimate the uncertainty in predictions, with lower CV values indicating lower uncertainty. Higher CVs indicate greater variability in RF predictions and higher uncertainty.

3.0 Results and Discussion

3.1 Monitored Nutrient Concentrations across the Great Lakes Basin

Monitored nutrient concentration data in the Great Lakes basin was processed using WRTDS in conjunction with monitored discharge data to generate mean annual flow-weighted concentrations (FWCs) across 17 years (2000-2016). The mean annual FWC values ranged widely across the Great Lakes basin, from 0.06 – 9.5 mg/L for DIN, 2 – 296 μg/L for SRP, and 6 – 808 μg/L for PP (Table 1). A subset of the 202 stations analyzed in this study (133 stations for DIN, 114 stations for SRP, and 130 stations for PP) had flux bias values between -/+ 0.15 and were used for further analysis in RF models.

Table 1 – Summary of nutrient FWCs (mg/L) from WRTDS processing for monitoring stations in the Great Lakes drainage basin

Nutrient	Number of Stations	Mean Observed FWC	Standard Deviation	Median FWC	Maximum Observed FWC	Minimum Observed FWC
DIN	133	2.33	2.01	1.71	9.51	0.06
SRP	114	0.041	0.053	0.018	0.296	0.002
PP	106	0.106	0.125	0.061	0.808	0.006

3.1.1 Dominant Controls on the Mean Annual FWC

Variable importance metrics, quantified using the RF modelling framework, were used to understand dominant controls on the mean annual FWC across the subset of monitored watersheds. Variable importance shows how sensitive modelled FWCs are to change when values of a variable are perturbed; variables that cause greater changes in FWCs are considered more important than those that do not (Grömping 2009; Tyralis, Papacharalampous, & Langousis 2019).

We found the percent developed land use, and the percent tile drained land to be the most important controls for dissolved nutrients, DIN and SRP (Figure 4). Strong positive correlations were observed between DIN and SRP and these spatial variables, and relationships clearly showed increased FWCs with increases in developed land use and tile drainage (Figure 5). Relationships with percent developed land use demonstrates the effect of non-point sources as a dominant driver of elevated nutrient concentrations in water bodies. We found percent wetland area and percent silt and clay soils to be the most important controls for PP (Figure 4). Wetlands were strongly negatively correlated with DIN, SRP, and PP FWCs, consistent with the understanding and observation of wetlands as nutrient sinks in anthropogenic landscapes. (Hansen et al. 2018; Dagnew, Scavia, Wang, Muenich, and Kalcic 2019). However, inverse cross correlation was also apparent between wetland cover and developed land cover, potentially hampering the trends observed in the data.

Tile drained land had significant positive correlations with nutrient FWCs. This shows the effect of engineered underground drainage systems as a driver in facilitating greater dissolved nutrient transport by circumventing the nutrient removal capacity of the soil column (Basu, Thompson, and Rao 2011). Variable importance for tile drained land was greatest for DIN, while lowest for PP. This is likely because DIN is transported through subsurface pathways in the landscape, and thus export would increase in the presence of tile drains. However, PP is transported through both surface and subsurface pathways as it exists in less mobile particulate form (vs. DIN is dissolved). The variable importance for tile drains is higher for SRP when compared to PP, further highlighting the greater transport of dissolved forms of nutrients through subsurface pathways enhanced by tile drains (Figure 4). Trends in FWCs from raw monitored PP data across developed land use, tile drained land, and wetlands appear less clear compared to DIN and SRP (Figure 5 and Figure 5). Scatter was more apparent, suggesting a weaker influence of landscape variables as drivers of PP when compared to DIN and SRP.

The percentage of silt and clay soil in a watershed was significantly positively correlated with all nutrient FWCs (Figure 4). This indicates elevated nutrient water quality from greater losses of nutrients in landscapes with finer soil types, likely due to the increased erosion potential from silt and clay soils when compared other coarser soil types (e.g. sand and gravel). This is further supported by its highest variable importance with PP, showing that finer soil types are major drivers of particulate concentrations due to the greater potential for erosion losses.

Population density was significantly positively correlated with SRP and PP, likely indicating the role of human sewage discharges in elevating phosphorus concentrations. It is of note that while positive correlation was observed with DIN, this correlation was not significant (p-value >0.05). Variable importance for population density was not among the greatest for SRP and PP, showing that it is a lesser driver of nutrient pollution compared to wetlands, tile drained land and developed land use. This could be attributed to success of point source controls, such as wastewater treatment upgrades and phosphorus reductions in detergents, implemented because of the GLWQA.

Livestock density was significantly positively correlated with nutrient FWCs, showing the role of manure in increasing nutrient concentrations. Livestock densities also had high variable importance in DIN and SRP FWCs. Fields applied with manure have been shown to have higher SRP losses than conventional fertilizers, especially when tile drains are present (Hodgkinson et al. 2002; Kinley et al. 2007; Kleinman et al. 2005). Dissolve inorganic nitrogen losses from fertilizer and manure applied in agricultural fields accounts for the majority of dissolved inorganic nitrogen exported globally from river mouths (Glibert 2020; Dumont et al. 2005). Correlations with livestock density and PP were less significant, and variable importance for PP was much less than DIN and SRP. This all corroborates the importance and significant positive correlations of livestock manure in DIN and SRP FWCs, indicating it plays a key role as a driver of dissolved nutrient water quality.

As previously mentioned, watershed area did not emerge as an important control in any of the nutrient FWCs. During preliminary stages of this research, variable importance for watershed area was consistently the least important spatial variable for DIN, SRP and PP, and trends in relationships were not evident. As such, watershed area was omitted from use in this analysis.

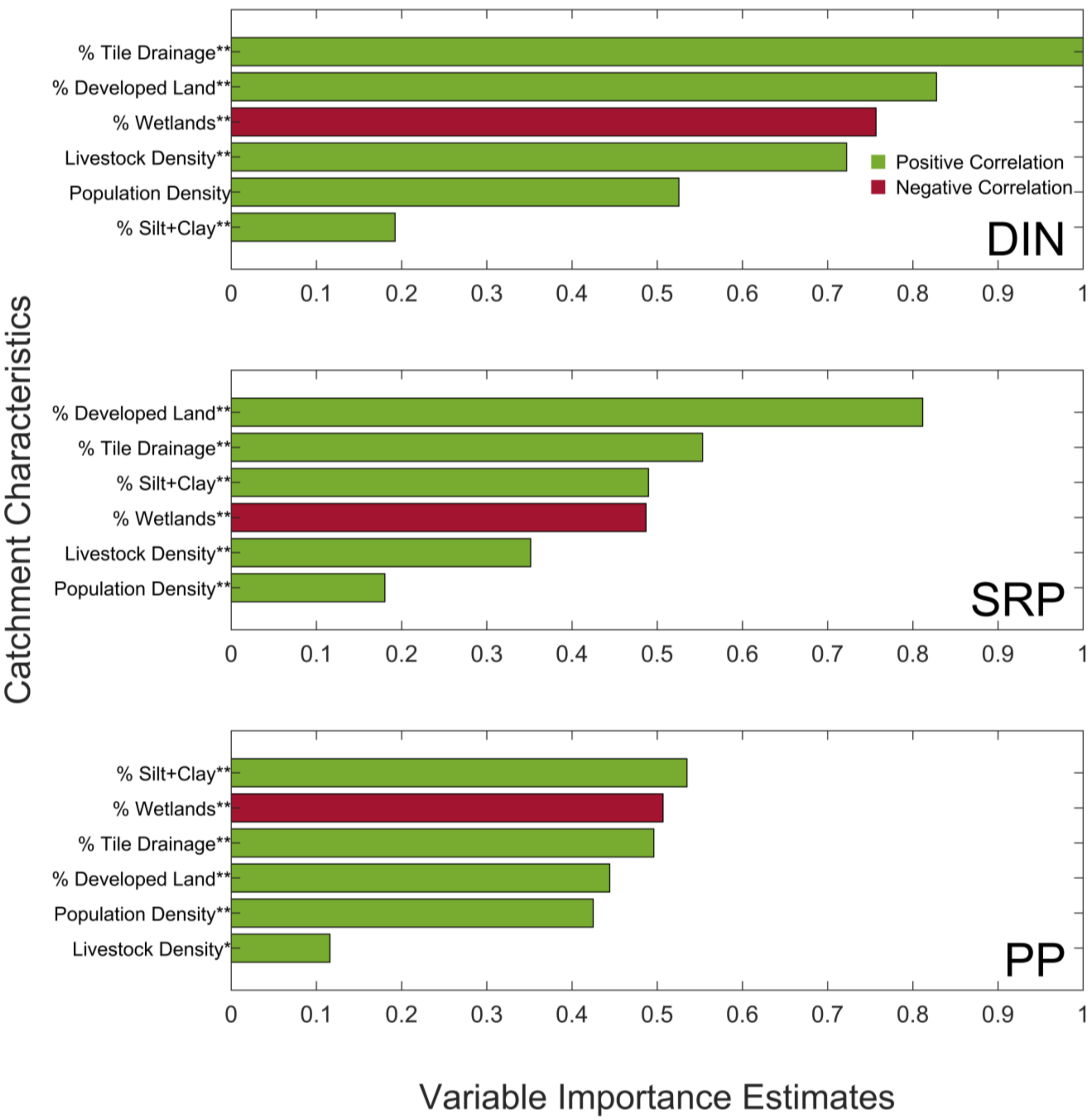

Figure 4 – Variable importance of spatial variables for modelled FWCs. FWCs are nitrogen (DIN) as combined nitrate and nitrite, soluble reactive phosphorus (SRP) and particulate phosphorus (PP) Correlations were determined using Mann-Kendall tests (denotes p<0.05, ** denotes p<0.01).*

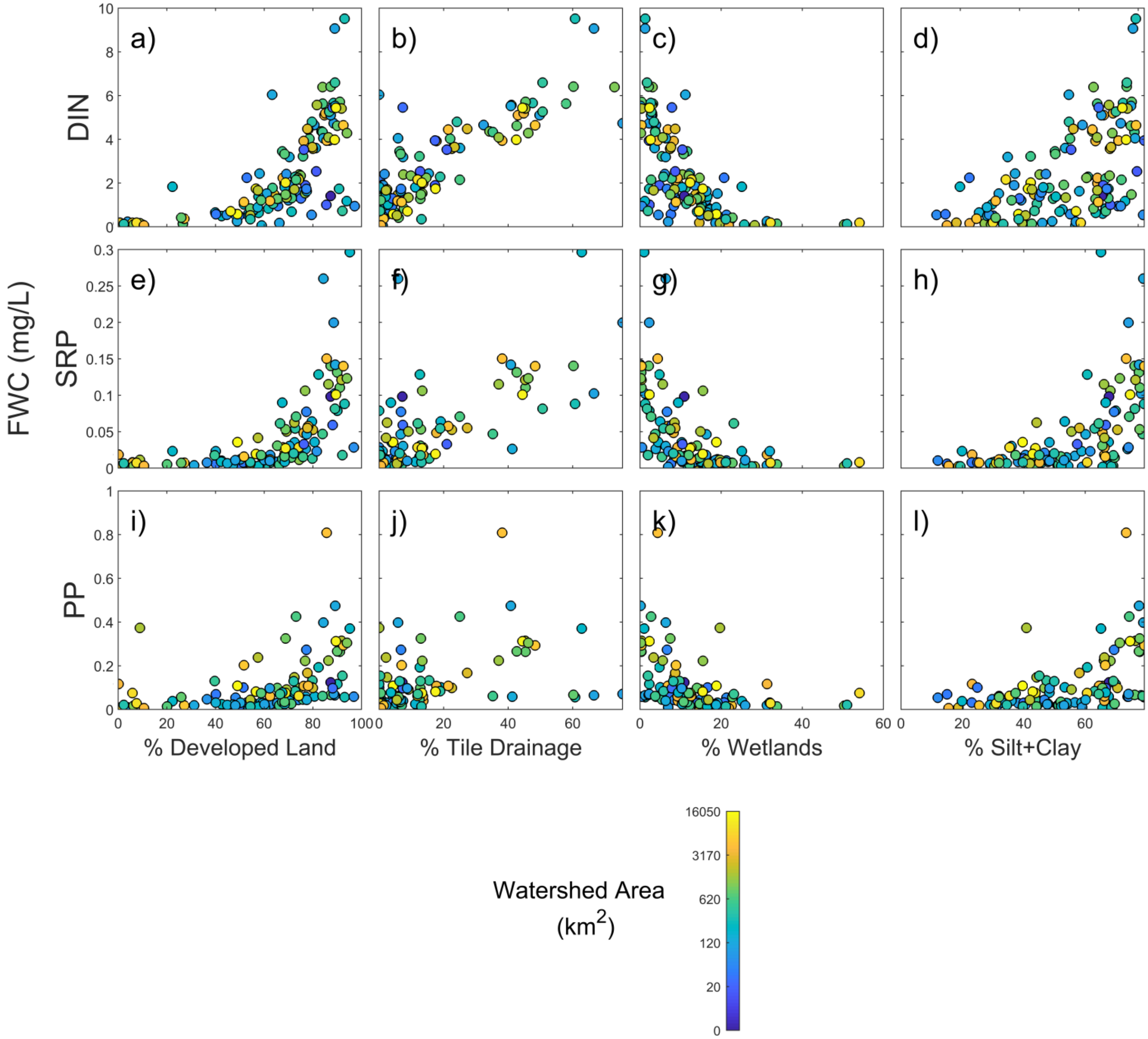

Figure 5 – Relationships between mean annual nutrient FWCs (mg/L) and select spatial variables across monitored watersheds in the Great Lakes basin (a-l). Marker colour indicates area of monitored watershed.

3.1.2 Feature Contribution Plots for FWCs

Feature contribution (FC) plots were generated to illustrate the effect of important spatial variables in the RF machine learning models (Figure 6). FC plots are a useful way to gain insights into the inner workings of machine learning models and other complex black box models. These plots allow the reader to see how predictor variables effect the model's prediction, while separated from the effects of other predictor variables. In this study, the predictor variables are the six spatial variables used for developing RF models, and the predictions are the average annual nutrient flow-weighted concentrations. Every monitored watershed has six FC values associated with it and the sum of these values show the difference between the predicted concentration for that watershed and the overall mean prediction of the model. This difference is given in units of standard deviations of the sample data. Therefore, FC values near zero show that a spatial variable has little to no effect on the modelled concentrations. FC values greater than zero indicate that the spatial variable has a positive effect, causing a greater predicted concentration than the overall mean prediction. FC values less than zero show the variable has a negative effect, causing a prediction less than the overall mean predicted concentration. Higher absolute values of FC show a larger effect of the variable in causing changes to predicted concentrations. For example, in Figure 6b and Figure 6c, 60% tile drained land (FC ≈ 1.5) plays a larger role in increasing predicted DIN concentrations than 20% wetland area does in decreasing predicted concentrations (FC ≈ -0.5). Furthermore, FCs are additive, such that a positive value from one spatial variable can cancel out a negative value from another spatial variable.

General trends were apparent in how spatial variables influenced nutrient concentrations when looking at the FC plots. Increasing developed land use at high percent land cover caused higher increases in nutrient concentrations, as seen by the higher positive FC values with increasing percent developed land use at higher proportions (Figure 6a, Figure

6e, & Figure 6i). This supports our understanding that non-point sources are the major driver of nutrient pollution in the Great Lakes. Similarly, greater presence of tile drained land resulted in higher increases in nutrient concentrations, as seen by FCs increasing from zero with increasing percent tile drainage (Figure 6b, Figure 6f, & Figure 6j). This illustrates the impact of increased nutrient export from subsurface drainage networks. Wetlands showed a decreasing effect on nutrient concentrations, where more wetland area caused FCs to decrease from zero, until a certain threshold, whereby FCs then plateaued (Figure 6c, Figure 6g, & Figure 6k). This shows the effects of wetlands as nutrient sinks, and their ability to reduce nutrient pollution with more wetland cover until a certain threshold. These overall trends show that hotspots with the highest nutrient concentrations for all species occur in watersheds where there is a high proportion of developed land use, high proportion of tile drained land, and low proportion of wetland cover.

The spatial variables also exhibited threshold effects in certain cases. FCs for wetlands decreased from zero as percent area increased to a threshold value of approximately 15%, after which values appeared to stabilize (Figure 6c, Figure 6g, & Figure 6k). This showed a reduction in modelled nutrient concentrations due to increased wetland area until this threshold, highlighting their role as sinks in nutrient removal. However, after wetland area exceeded 15%, FCs plateaued at negative values. This demonstrated that there was a negligible effect in modelled concentrations when wetland area exceeded 15%. These results perhaps suggest there may be little difference in nutrient load reduction for strategies that restore wetland area beyond 15% of a watershed. This threshold value of 15% wetland area may be valuable to consider for nutrient managers looking at watershed scale strategies in maximizing nutrient reductions. Threshold behaviour was also apparent when looking at developed land use, as increases in FCs from zero were not seen until after approximately 50%, 65% and 75% land cover, for DIN, SRP, and PP, respectively (Figure 6a, Figure 6e, & Figure 6i). This perhaps suggests that at the watershed scale, the effect of developed land use of less than these threshold values may be negligible with respect to increasing nutrient loads, and subsequently, that developed land use greater than these thresholds would have adverse effects to downstream nutrient water quality for each nutrient species. Again, this threshold behaviour may be valuable for nutrient managers looking at watershed scale strategies for load reductions. Threshold behaviour was also apparent with soil type, as increases in FCs from zero were not seen until approximately 75%, 55% and 65% silt and clay soil cover for DIN, SRP and PP, respectively (Figure 6d, Figure 6h, & Figure 6l). This perhaps shows that high percent cover of finer soil types causes elevated concentrations for these nutrient species, while the effects from lower percent cover may be negligible.

Threshold behaviour in nutrient water quality is valuable to consider from a management perspective, especially for strategies like wetland restoration and land use changes for nutrient load reduction. Management strategies that apply thresholds could better allocate resources, while achieving similar performance in overall nutrient load reduction, compared to more blanketed, wide sweeping approaches. Using the threshold behaviour seen by FC, an ideal point of 15% wetland area and 50-75% developed land use emerges (dependant on nutrient species), that perhaps reveals a tipping point in nutrient water quality from when land use begins to have adverse impacts. This is especially important for balancing interests among stakeholders in nutrient pollution; strategies should be taken to not overburden agriculture and urban development while still maximizing the potential for nutrient load reductions. However, due to the limited number of monitored watersheds in this study encompassing this threshold, it is recommended that more investigation be taken prior to leveraging these thresholds for nutrient management.

Individual FC plots revealed that the influence of spatial variables varies as a function of nutrient species. This can be seen by comparing the magnitude of FC between different nutrient species for certain spatial variables. The decreasing effects of wetlands for nutrient concentrations appear greatest for PP (Figure 6k), as seen by the greater absolute FC values compared to DIN and SRP (Figure 6c & Figure 6g). This highlights the major function of wetlands in settling out suspended sediments from watershed runoff, thereby removing particulate forms of nutrients from entering waterways. Tile drainage density had the strongest effects for DIN, then SRP, while its effect was minimal for PP (Figure 6b, Figure 6f, & Figure 6j). This shows the ability of tile drains in favouring transport of dissolved forms of nutrients over particulate forms. High percentages of developed land use had a greater effect in increasing DIN and SRP concentrations when compared to PP (Figure 6b, Figure 6f, & Figure 6j). This shows the more dominant effect of dissolved nutrient losses from agricultural activities and urban land use in increasing dissolved nutrient concentrations downstream. Soil type appeared to have the greatest effect in PP (Figure 6l), with greater proportions of silt and clay soils showing large FC values when compared to dissolved nutrients of DIN and SRP (Figure 6d & Figure 6h). This highlights the greater potential of erosion and sediment transport for finer soil types in elevating downstream particulate forms of nutrient water quality when compared to dissolved forms. Overall, this demonstrates that watersheds with high developed land use, high proportions of tile drained land, and low wetland cover result in hot spots of high DIN and SRP concentrations. Additionally, watersheds with high proportions of fine soil cover and low wetland cover result in hot spots of high PP concentrations.

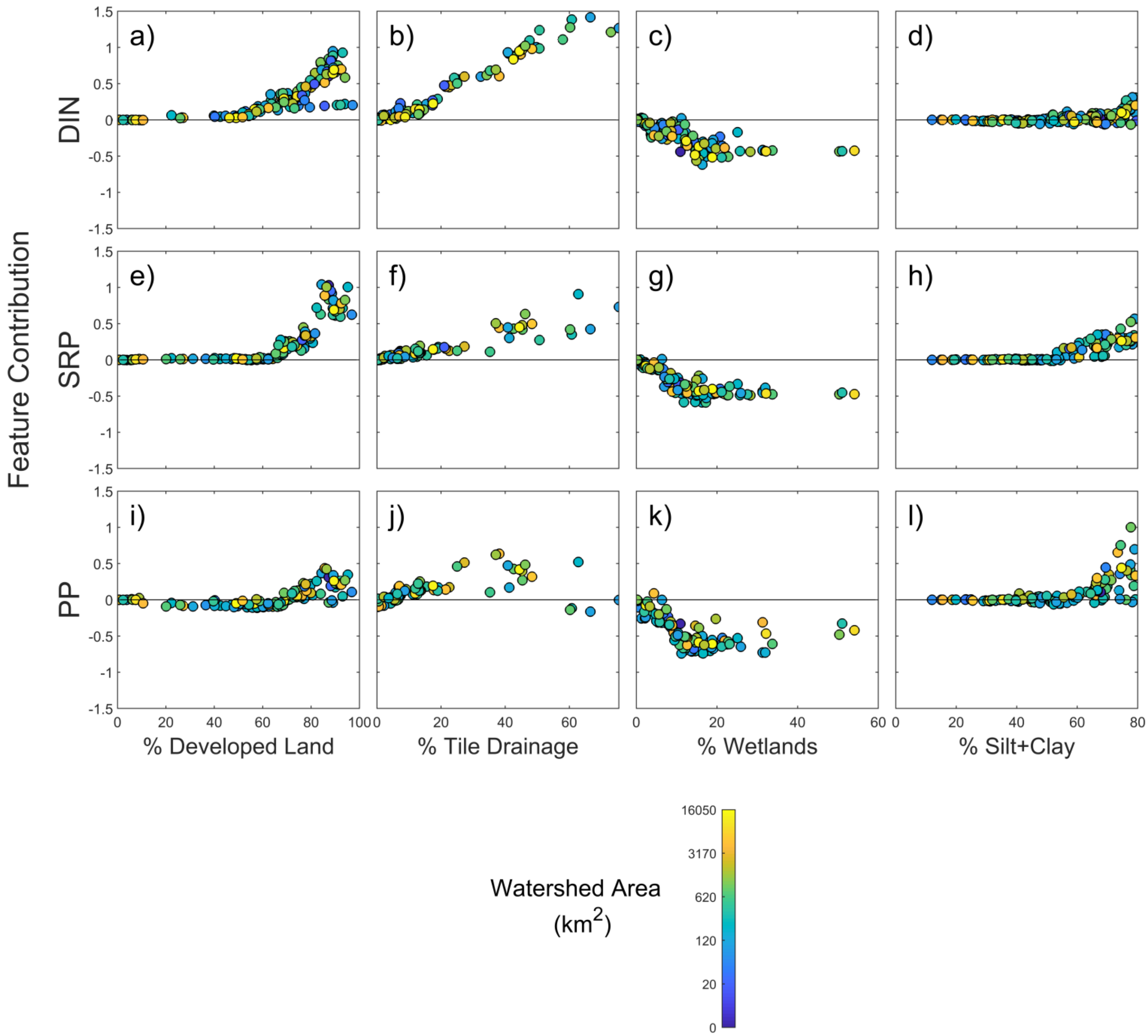

Figure 6 – Feature contribution plots for the top four important predictor spatial variables for nutrient concentrations (a-l). Feature contribution shows the effect a spatial variable has in model predictions, separate from the effects of other spatial variables in the model. Feature contribution values greater than zero indicate that the spatial variable causes a predicted nutrient concentration greater than the mean prediction. Feature contribution values less than zero indicate that the spatial variable causes a predicted nutrient concentration less than the mean prediction. Solid line indicates a feature contribution value of zero. Data points were coloured based on watershed area. Vertical axis units are given as standard deviations from the overall mean predicted nutrient concentration.

3.1.3 Random Forest Model Prediction for FWCs

The RF models performed adequately for all three constituents during validation; models for dissolved nutrients, especially DIN, performed significantly better when compared to PP (Figure 7). Models for DIN, SRP and PP accounted for 81%, 54% and 31% of the variation in unbiased FWCs, respectively. For SRP and PP, the model underestimated the small number of highest FWCs. This underprediction could be attributed to the difficulty of RF models in extrapolating outside the normal range of training data, as is the case with these few observed high FWCs (Tyralis, Papacharalampous, and Langousis 2019). Additionally, phosphorus dynamics in the landscape are more complex than nitrogen, with sorption, desorption, and redox chemistry integral to its translation and transformation through the landscape. This process complexity for phosphorus may explain the lower model performance in capturing downstream water quality from generic spatial variables, especially for less mobile particulate forms.

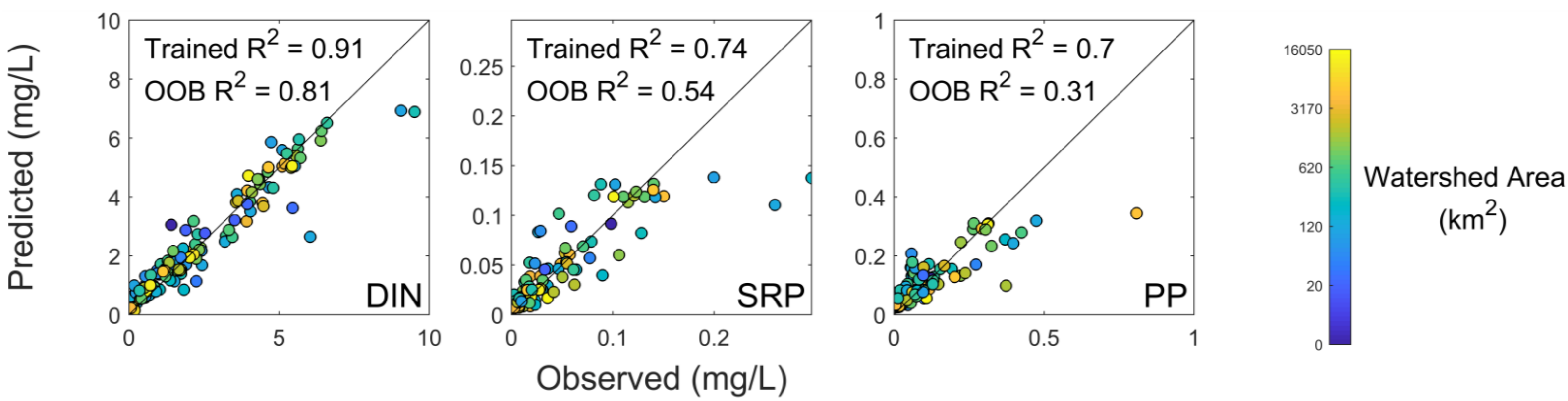

Figure 7 – Predictive (1:1) measure of RF flow-weighted concentrations against WRTDS-estimated flow-weighted concentrations for DIN, SRP, and PP. Data points were coloured based on watershed area. R^2 values are given for Trained and OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation.

3.2 Predicted Nutrient Concentrations across the Great Lakes Basin

Final RF models for each nutrient water quality parameter were applied to estimate FWCs for all ungauged watersheds of the Great Lakes drainage basin. RF models were generated using 1000 decision trees, with each tree having a prediction of average annual FWC. A mean value from all 1000 decision trees was used to predict average annual FWCs for each of the Great Lakes watersheds. Additionally, coefficients of variation (CV) were calculated as a measure of uncertainty in predictions based on the variability of predictions from decision trees. The CVs allow us to estimate the uncertainty in predictions, with lower CV values indicating lower uncertainty. CVs were calculated by dividing the standard error of decision tree predictions by the mean predicted response for each watershed. The spatial distribution of predicted average annual FWCs and their associated CVs for each nutrient parameter can be seen in Figure 8. Modelled FWCs across the entire Great Lakes basin, as well as those for each of the five Great Lakes, are summarized in Table 2.

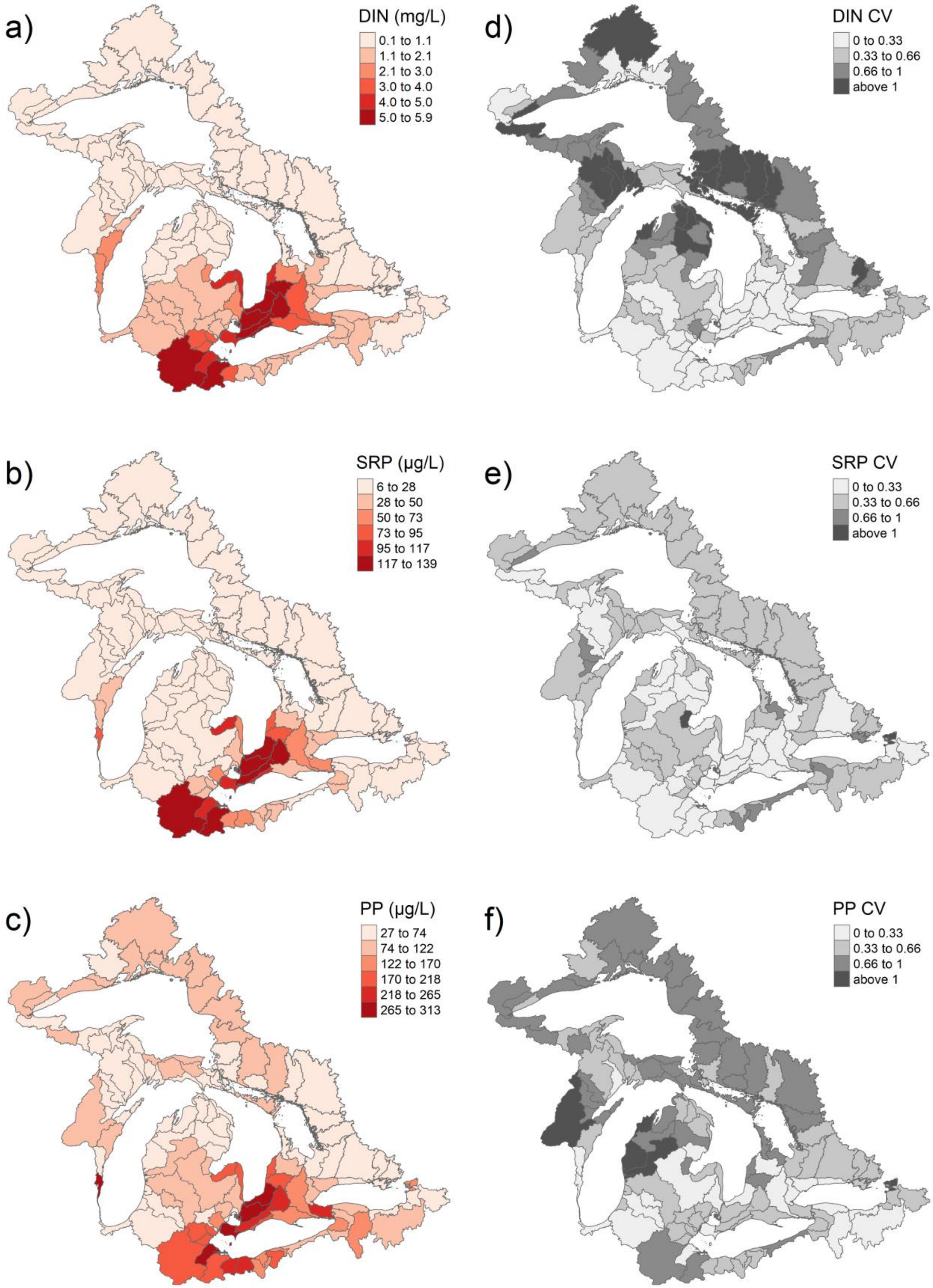

Figure 8 – Spatial distribution of RF modelled average annual FWCs (a-c) with CVs of predictions (d-f) for the Laurentian Great Lakes watersheds. CVs show uncertainty in predictions, with higher (darker) values indicating higher relative variation in modelled results.

Predicted FWCs were generally highest in the southwestern drainage basin of Lake Erie, particularly in the Maumee, Sandusky, and Cedar-Portage watersheds, and Maitland and Ausable watersheds on the southeastern shores of Lake Huron, and in the Thames, Sydenham, Cedar, and Rondeau watersheds along the eastern side of the Huron-Erie corridor. These hot spots of nitrogen and phosphorus generally coincide with areas of high developed land use (>80%), high tile drained land area (>40%), and low wetland area (<5%). Again, this highlights the dominance of non-point source nutrient pollution in the Great Lakes drainage Basin. Low CVs for predictions generally coincide with watersheds where FWCs were highest, thus lowering the uncertainty in these high predictions. This was likely attributed to the large number of monitoring stations in these areas that were used for training models (Figure 2). In contrast, fewer monitoring stations were present in low concentration watersheds, such as around Lake Superior and Georgian Bay in Lake Huron (Figure 2). This has practical merit for nutrient managers, as there is lower uncertainty in predictions where nutrient concentrations are highest, making these areas low risk for targeted nutrient reduction strategies.

High phosphorus FWCs in the Lake Erie drainage basin support the consensus that excessive phosphorus loadings from non-point sources cause eutrophication and summer algal blooms in the Lake. The highest PP concentrations were predicted in the southwestern watersheds of Lake Erie, in the Maumee, Cedar-Portage, Huron-Vermilion, and Black-Rocky watersheds, along the eastern (Canadian) side of the Huron-Erie corridor, in the Ausable, Cedar, Sydenham, and Thames watersheds, and in the Niagara watershed of Lake Ontario. (Figure 8c). The highest SRP concentrations were predicted in the southwestern watersheds of Lake Erie, specifically in the Maumee, Sandusky, and Cedar-Portage watersheds, and along the eastern (Canadian) side of the Huron-Erie corridor in the Ausable, Sydenham, Thames and Rondeau watersheds (Figure 8b). SRP hot spots with lower PP concentrations may get overlooked if focus on phosphorus reduction only considers TP targets. This is especially significant considering the greater bioavailability of soluble forms of phosphorus for the promotion of algae growth, when compared to particulate forms (Baker et al. 2014). Furthermore, management practices in these hot spots of SRP and PP can be targeted towards their respective elevated forms of phosphorus, rather than general TP reduction strategies.

The data driven results were also consistent with findings from mechanistic, processed based models. Higher SRP FWCs from RF models on the Canadian side of the Huron-Erie corridor were consistent with recent SWAT models that also show high dissolved phosphorus losses in these areas (Figure A.6) (Scavia et al. 2019; Dagnew, Scavia, Wang, Muenich, and Kalcic 2019; Dagnew, Scavia, Wang, Muenich, Long, et al. 2019). Consistency with mechanistic findings adds confidence in the data driven results of nutrient water quality across the Great Lakes basin.

Table 2 – Summary of area-weighted mean FWC values modelled (mg/L) for drainage basins across the Great Lakes watersheds.

Nutrient	All Basins Mean	Erie Mean	Huron Mean	Michigan Mean	Ontario Mean	Superior Mean
DIN	1.33	3.91	1.01	1.09	1.02	0.26
SRP	0.026	0.085	0.017	0.017	0.019	0.008
PP	0.088	0.191	0.075	0.059	0.066	0.074

3.3 Monitored Nutrient Ratios across the Great Lakes Basin

Mean annual flow weighted ratios for monitored watersheds were generated using nutrient FWCs from monitored data processed using WRTDS across 17 years (2000-2016). The mean annual ratios across the Great Lakes watersheds ranged widely in the Great Lakes basin, from 0.03 – 0.74 for SRP:TP, and 1.1–124 for DIN:TP in mol:mol (Table 3). A subset of the 202 stations analyzed in the study (106 stations for SRP:TP, and 109 stations for DIN:TP) had flux bias values between -/+ 0.15 and were used for further analysis in RF models.

Table 3 – Summary of nutrient ratios (mol:mol) from WRTDS processing for monitoring stations in the Great Lakes drainage basin

Ratio	Number of Stations	Mean Ratio	Standard Deviation	Median Ratio	Maximum Observed Ratio	Minimum Observed Ratio
SRP:TP	106	0.28	0.13	0.27	0.74	0.03
DIN:TP	109	54.9	48.3	42.5	273.5	0.9

3.3.1 Dominant Controls on the Mean Annual Ratios

The importance of spatial variables for nutrient ratios was assessed using the machine learning framework to understand dominant controls and drivers of nutrient water quality. Figure 9 shows the results of the variable importance analysis and shows correlations of spatial variables with the raw ratios. Figure 10 shows raw scatter plots of ratios processed from monitored watersheds in the Great Lakes basin across important spatial variables.

The proportion of developed land use, livestock density, tile drained land, silt and clay soil and population density were positively correlated with the proportion of bioavailable P (SRP:TP), while negative correlations were present with the proportion of wetland area. These correlations were the same when looking at the proportion of DIN relative to TP (DIN:TP). When looking at the importance of spatial variables to nutrient ratios, the impacts of agricultural manure and fertilizer appeared evident. Overall, more scatter was apparent when looking at relationships between nutrient ratios and spatial variables (Figure 10), compared to those for nutrient FWCs (Figure 5).

Livestock density, developed land use, and tile drained land were found to be the most important spatial variables for SRP:TP and they were significantly positively correlated to SRP:TP ratios (Figure 9). Increases in tile drain density increases the flow through the subsurface pathways, and thus likely increases SRP:TP ratios from greater dissolved transport. Fields applied with manure have been shown to have higher SRP losses, when compared to synthetic fertilizers, especially in the presence of tile drains, and thus would lead to increased SRP:TP ratios downstream (Hodgkinson et al. 2002; Kinley et al. 2007; Kleinman et al. 2005). As the amount of developed (agricultural) land in a watershed increases, so too would the SRP losses from tile drains and manure application, explaining its importance and positive correlation with SRP:TP ratios. It is of note that application methods, rates and timing of fertilizer and manure, in addition to tillage practices, would also have an effect on the phosphorus losses from agricultural fields, but such information was not available at this scale (Bundy, Andraski, and Powell 2001; D. R. Smith et al. 2007; Tabbara 2003; Withers, Clay, and Breeze 2001).

Positive correlations were observed between DIN:TP ratios and livestock densities and they also exhibited high variable importance (Figure 9). Fertilizer and manure from agriculture accounts for the majority of dissolved inorganic nitrogen export (Glibert 2020; Dumont et al. 2005) and manure often has higher nitrogen to phosphorus ratios than synthetic fertilizer blends (Allan, Murray, and Child 2018; Munroe et al. 2018). As such, nutrient losses from fields applied with manure would generally shift the DIN:TP ratio higher compared to fields applied with synthetic fertilizers. This demonstrates that manure is a major driver in nitrogen to phosphorus ratios downstream. Increased local atmospheric deposition from ammonia volatilization of animal waste would also support the high importance and positive correlation of these livestock densities with DIN:TP. However, it is expected that atmospheric deposition would be small compared to the overall contribution of nitrogen losses from agricultural fields (Whitall, Hendrickson, and Paerl 2003; Paerl 1997).

The variable importance of wetlands was low for SRP:TP ratios suggesting that it is a weaker driver in SRP:TP ratios when compared to developed land use, livestock density and tile drained land. However, wetland area showed significant negative correlation with SRP:TP ratios. While wetlands were negatively correlated with DIN:TP ratios, this correlation was not significant (p-value >0.05), and variable importance was also low. This suggests that wetlands are not a strong driver of nitrogen to phosphorus ratios as well.

Silt and clay soils were significantly positively correlated with SRP:TP ratios yet showed low variable importance. While variable importance was higher for DIN:TP ratios, the positive correlations observed were not significant (p-value >0.05). This all demonstrates that soil type is a weak driver of nutrient ratios downstream.

While population density had positive correlations with SRP:TP and DIN:TP, these correlations were not significant (p-value >0.05). Additionally, population density had low variable importance, further suggesting that the effects of human sewage discharges do not drive nutrient ratios in the Great Lakes basin. This could be due to success of the point source controls implemented from the GLWQA.

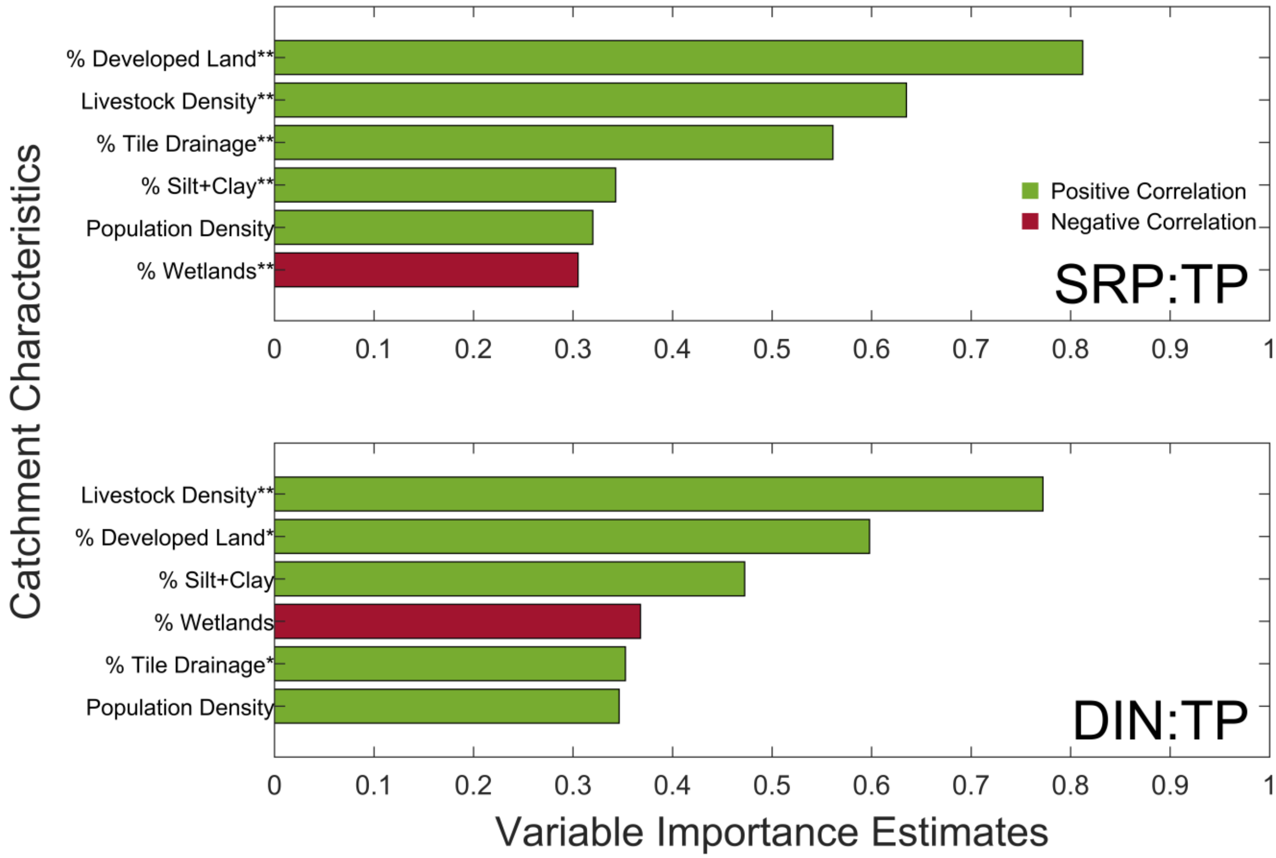

Figure 9 – Variable importance of spatial variables for modelled nutrient ratios. Ratios are soluble reactive phosphorus over total phosphorus (top) and nitrogen as nitrate and nitrite over total phosphorus (bottom). Correlations were determined using Mann-Kendall tests (denotes p<0.05, ** denotes p<0.01). Bold variable names show those selected for final models.*

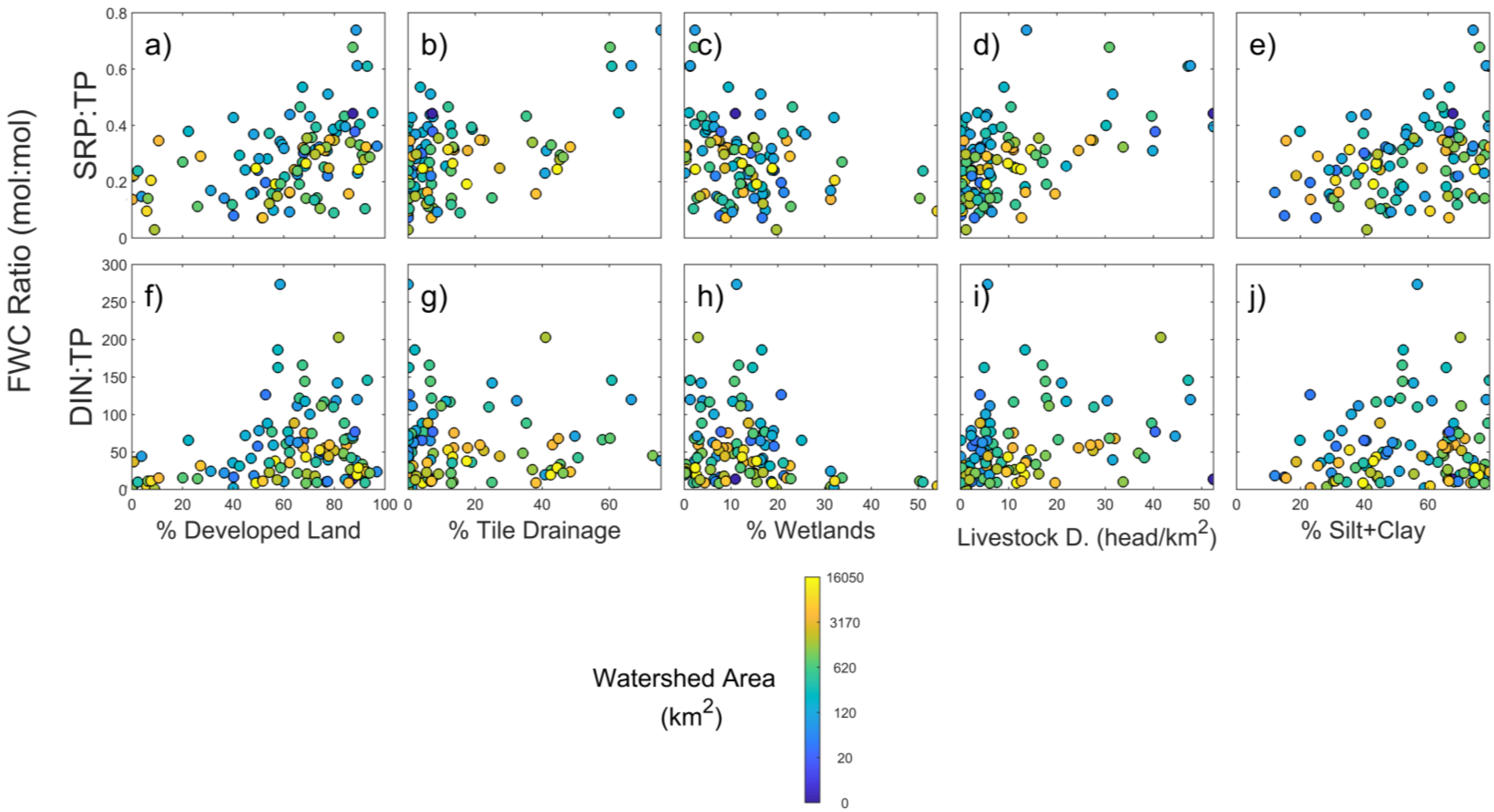

Figure 10 – Relationships between mean annual flow-weighted nutrient ratios (mol:mol) and select important spatial variables across monitored watersheds in the Great Lakes basin (a-j). Marker colour indicates area of monitored watershed.

3.3.2 Feature Contribution Relationships for Nutrient Ratios

FC plots revealed general trends in how spatial variables influenced nutrient ratios (Figure 11). Overall, trends appeared less clear compared to the those observed for nutrient FWCs (Figure 6). High percentages of developed land use and high densities of livestock had high positive FC values for both SRP:TP and DIN:TP ratios (Figure 11a, Figure 11d, Figure 11f, & Figure 11i). High proportions of tile drained land and high proportions of silt and clay soils had positive FC values for SRP:TP ratios (Figure 11b & Figure 11e). High proportions of silt and clay soil cover had negative FC values for DIN:TP ratios (Figure 11j). This demonstrates that the highest SRP:TP ratios occurred in watersheds with high developed land use, high livestock densities, high tile drained land, and high percentages of silt and clay soil cover, while the highest DIN:TP ratios were seen in watersheds with high developed land use, high livestock densities, and low proportions of silt and clay soils.

We found percent developed land use to exhibit a threshold behaviour for both SRP:TP ratios, and DIN:TP ratios, although its effect on SRP:TP ratio was greater. Percent developed land use had a negligible effect of SRP:TP ratios till a threshold value of 60%, beyond which it increased approximately linearly to an FC value of 1 at 80% developed land use (Figure 11a). Developed land use also exhibited threshold behaviour in its effect on increasing DIN:TP ratios. Positive FC values were observed for developed land use between 50% and 90%, while FC values were equal to zero at both low (< 50%) and high (> 90%) proportions of developed land use (Figure 11f).

In contrast to this threshold behaviour for percent developed land use, FC values for livestock density appeared to increase monotonically for SRP:TP (Figure 11d). This highlights the impacts of manure application in increasing SRP losses from agricultural fields. This threshold behaviour for SRP:TP ratios also coincided well with previously discussed threshold behaviour for nutrient FWCs in Section 3.1.2 and may provide additional value for nutrient management. Livestock density also appears to have a significant effect in increasing DIN:TP ratios. At densities of greater than 5 equivalent head per square kilometer, FC values are consistently high, although there is scatter in the high values (Figure 11i). This can be attributed to the higher nitrogen to phosphorus content in manure compared to synthetic fertilizers.

Percent tile drained land showed a threshold effect in SRP:TP ratios – FC values are near zero until about 40% tile drainage, after which values increase monotonically and plateau at 1 at 60% tile drains (Figure 11b). Tile drains facilitate increased dissolved phosphorus export through subsurface pathways, and thus increases in tile drain density increases SRP:TP ratios. Tile drained land, however, did not appear to have a major effect in DIN:TP ratios, as FC values fluctuated near zero across all proportions of drained land (Figure 11g).

Wetland area within a watershed showed little effect in influencing SRP:TP and DIN:TP ratios. For SRP:TP, FC values fluctuated near zero across the entire range of wetland areas (Figure 11c). This suggests wetlands were not a major driver for SRP:TP ratios, despite their role in reducing SRP and PP concentrations. For DIN:TP, FC values showed a small negative trend for wetland area greater than 20% (Figure 11h). Higher proportions of wetland area decrease both DIN and TP, but these suggests that it decreases DIN at a proportionally lower rate compared to P.

Soil type had a small and threshold driven effect on nutrient ratios. For SRP:TP ratios, beyond a threshold of 50% silt and clay soil, increasing proportions of fine soils increased the predicted ratio (Figure 11e). For DIN:TP, beyond a threshold of 60% silt and clay soil, increasing proportion of fine soils decreased DIN:TP ratios (Figure 11j). Increasing proportions of fine soils would increase erosion and thus increase TP transport leading to decrease in DIN:TP ratios.

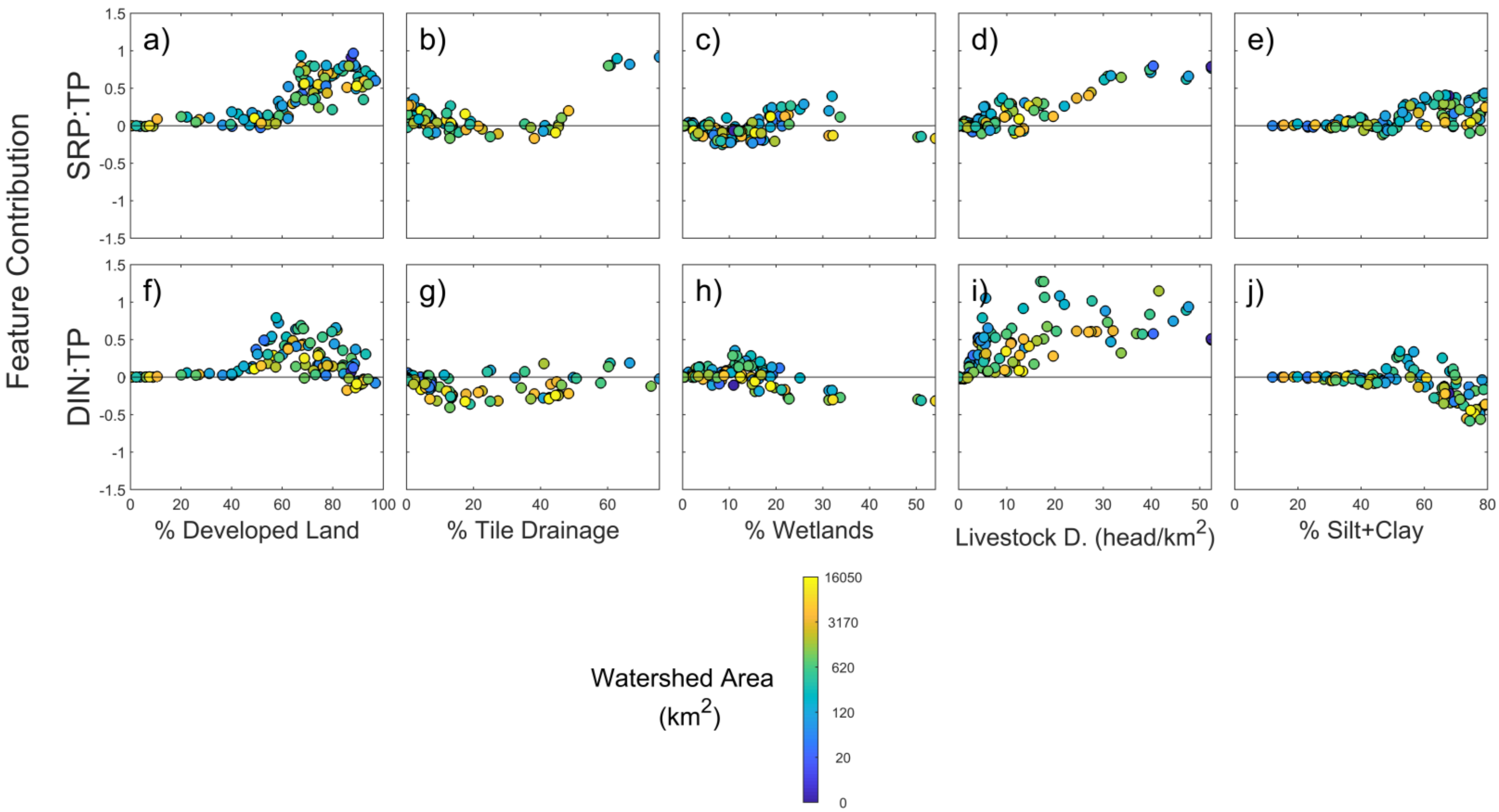

Figure 11 – Feature contribution plots for the top five important predictor spatial variables for nutrient ratios (a-j). Feature contribution shows the effect a spatial variable has in model predictions, separate from the effects of other spatial variables in the model. Feature contribution values greater than zero indicate that the spatial variable causes a predicted nutrient ratio greater than the mean prediction. Feature contribution values less than zero indicate that the spatial variable causes a predicted nutrient ratio less than the mean prediction. Solid line indicates a feature contribution value of zero. Data points were coloured based on watershed area. Vertical axis units are given as standard deviations from the overall mean predicted nutrient ratio.

3.3.3 Random Forest Model Prediction for Nutrient Ratios

Final RF models for SRP:TP and DIN:TP ratios performed similarly with lower R^2 values compared to predictions of the individual nutrient concentrations. Models for SRP:TP and DIN:TP accounted for 34%, and 22% of the variation in water quality ratios for out-of-bag observations. The model appeared to overestimate the lowest SRP:TP ratios (<0.2) and underestimate the highest ones. The spatial distribution of the watersheds with the worst predictions were investigated and no spatial trends were seen to explain their low SRP:TP ratios (e.g. low developed land use, high wetland area, low livestock densities etc.). This suggests they were outliers in the relationships drawn from spatial data by the machine learning models, possibly explaining their poor prediction performance in these low values. Site specific conditions for these watersheds may explain their low SRP:TP ratios, and these local effects would not be captured in the generic spatial variables used in this study for the RF model. This could include the nuanced effects of small-scale precipitation, local nutrient cycling conditions, and/or seasonal dynamics that are specific to these monitored watersheds.

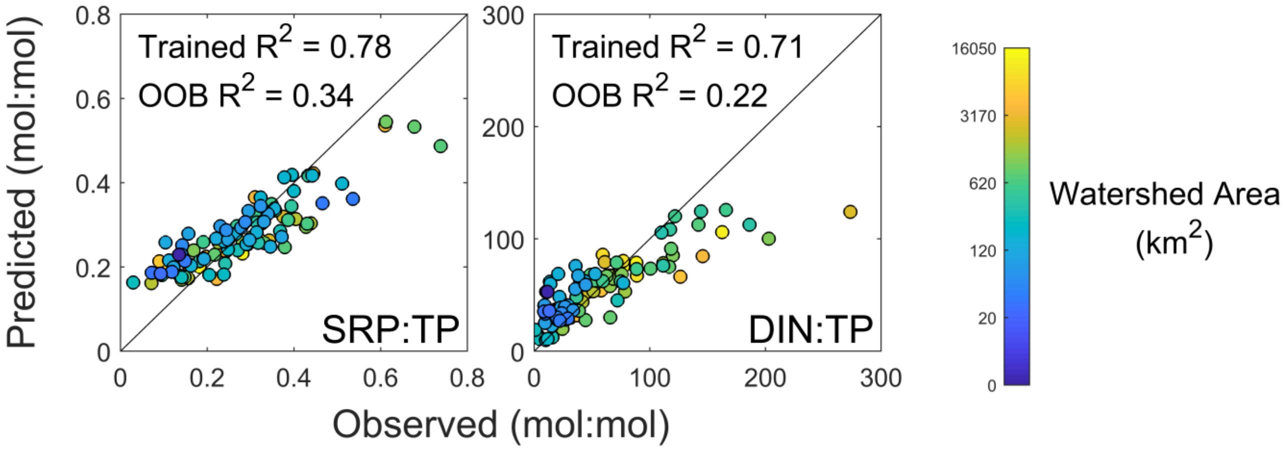

Figure 12 – Predictive (1:1) measure of RF flow-weighted ratios against WRTDS-estimated flow-weighted ratios for SRP:TP and DIN:TP. Data points were coloured based on watershed area. R^2 values are given for Trained and OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation.

3.4 Predicted Nutrient Ratios across the Great Lakes Basin

Using the final developed RF models, nutrient ratios were predicted for the entire Great Lakes drainage basin (Figure 13). Predicted ratios across the Great Lakes basin were spatially consistent with predicted nutrient FWCs (Figure 8). Modelled ratios across the entire Great Lakes basin, as well as those for each of the five Great Lakes, are summarized in Table 4.

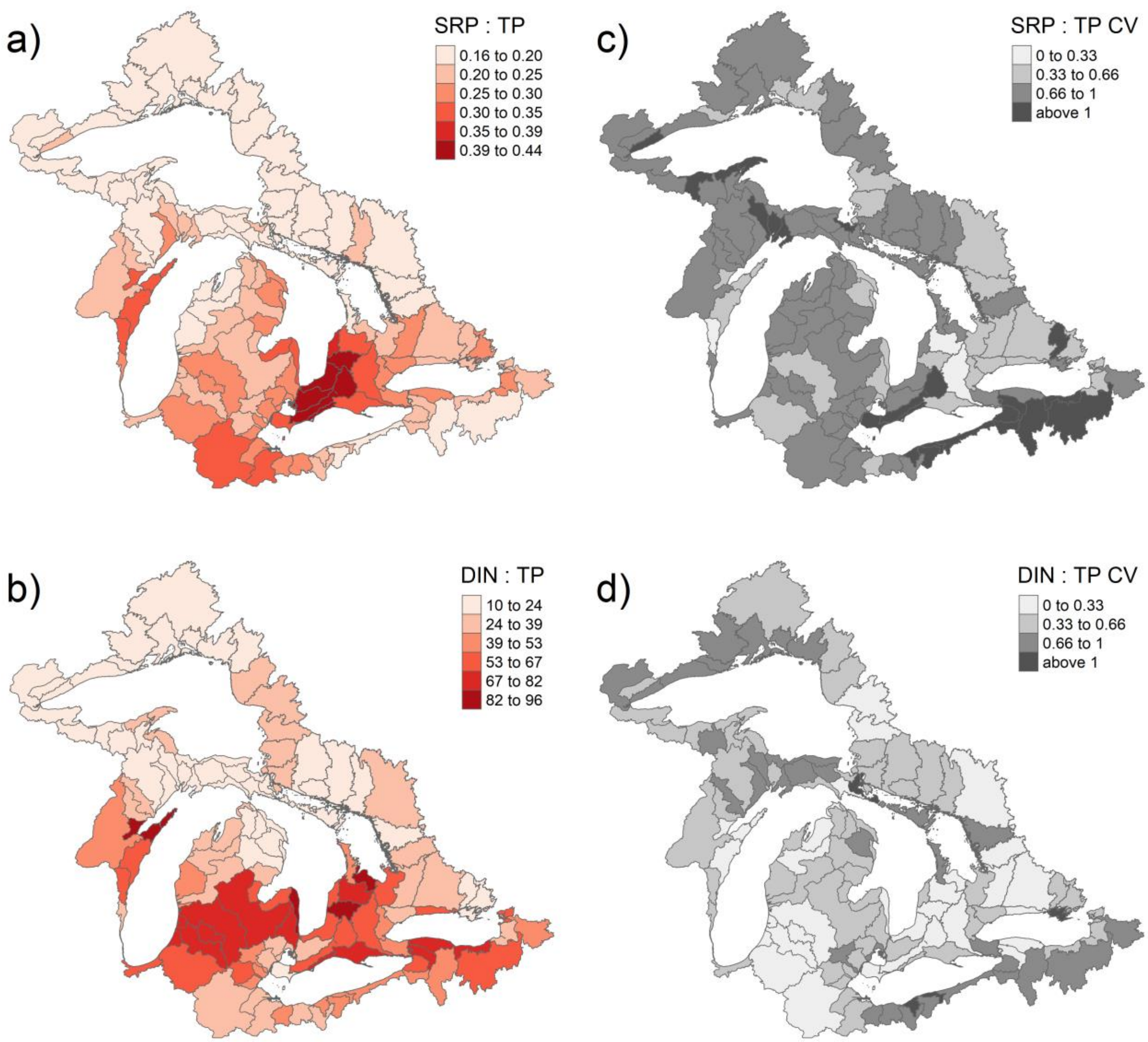

Figure 13 – Spatial distribution of RF modelled average annual ratios (a-b) with CVs of prediction (c-d) for the Laurentian Great Lakes watersheds.

The area with the highest SRP:TP ratios was seen in the Pentangore, Maitland, Sydenham, and Thames watersheds on the southeastern shore of Lake Huron and along the eastern side of the Huron-Erie corridor (Figure 13a). These watersheds are dominated by agricultural land use (>75%) and have high livestock densities. This hot spot may be of interest to nutrient managers because of its high relative input of bioavailable phosphorus into Lake Huron and the Huron-Erie corridor. This bioavailable phosphorus could be a significant contributor to algal growth in the lake. High DIN:TP was also seen in the Birch-Willow, Southwest Georgian Bay and Maitland watersheds of Lake Huron, and the Duck-Pensaukee and Door-Kewaunee watersheds of Lake Michigan, near Green Bay (Figure 13). Again, hot spots of SRP:TP and DIN:TP generally had low CVs (<1), reducing uncertainty in the modelled results.

Watersheds in southwestern Lake Erie, such as the Maumee, Sandusky, Cedar-Portage, Detroit, and the Ottawa-Stony watersheds, and in the eastern watersheds of the Huron-Erie corridor, such as the Cedar and Sydenham watersheds, had low DIN:TP ratios, highlighting potentially phosphorus enriched regimes. These watersheds also had high nutrient FWCs (Figure 8), showing high nutrient export with low DIN:TP ratios. Lower DIN:TP inputs from these watersheds could shift DIN:TP ratios lower downstream in Lake Erie, promoting the growth of harmful cyanobacteria blooms that favour these low ratios (V. H. Smith and Schindler 2009).

Watersheds along northern Lake Superior and Georgian Bay had low SRP:TP and DIN:TP ratios. These areas also coincided with low nutrient FWCs (Figure 8) and illustrates the nutrient water quality regime of less human impacted catchments.

Table 4 – Summary of area-weighted mean ratio values modelled (mol:mol) for drainage basins across the Great Lakes watersheds.

Ratio	All Basins	Erie Basin	Huron Basin	Michigan Basin	Ontario Basin	Superior Basin
SRP:TP	0.23	0.31	0.22	0.24	0.22	0.18
DIN:TP	40	45	40	48	50	21

Modelled ratios across the Great Lakes basin (Table 4) were compared with ratios from other large drainage systems in published studies. Ratios in the Great Lakes basin were similar to values recorded in the Narragansett Bay watershed in Rhode Island and Massachusetts, from July-October 2012, where mean DIN:TP ratios were 40, and mean SRP:TP ratios were 0.40 (Smucker et al. 2016). DIN:TP ratios for the Great Lakes watersheds were similar to the average N:P ratio of 49.1 observed in the Baltic sea drainage basin from 1970-2000 (Saaltink et al. 2014). DIN:TP ratios of the Great Lakes watersheds compared well with the human impacted western European drainage basins of the Seine, Somme and Scheldt Rivers, where ratios for the wet year of 1996 and the dry year of 2001 were 35 and 51, 45.4 and 66.5, and 30.1 and 38.7, respectively (Thieu, Billen, and Garnier 2009). While average SRP:TP values for the Great Lakes watersheds were much lower compared to the average SRP:TP ratio of 0.60 observed between 2000-2005 for the San Francisco Estuary (Sacramento – San Joaquin River Bay Delta), DIN:TP ratios of the Great Lakes watersheds were much higher than their average value of 18.0 (Glibert et al. 2011).

Modelled ratios across the Great Lakes basin were also compared to ratios for impacted and unimpacted monitored streams and rivers across the U.S. by Maranger et. al (2018). Modelled average DIN:TP ratios for the Lake Erie, Huron, Michigan, and Ontario basins were consistent with the 75th percentile TN:TP ratio of 44.6 for monitored streams across the United States. In contrast, average DIN:TP ratios for streams in the Lake Superior basin were more aligned with the median TN:TP ratio of 24.7 for United States' rivers and streams (Maranger, Jones, and Cotner 2018).

3.5 Nutrient Loads from the Great Lakes Basin

3.5.1 Modelled Annual Loadings

Annual nutrient loads were estimated to determine the export to the Great Lakes from tributary sources for years 2000 to 2016. Loads were calculated using RF modelled FWCs across the entire drainage basin, and annual area-discharge regression relationships for years 2000 to 2016 from monitored watersheds (Figure A.4). As such, estimated loads were driven by differences in annual discharge, consistent with precipitation as the main driver of interannual variability in nutrient loads (Sinha and Michalak 2016). Annual average FWC's are independent of discharge and can be scaled with annual discharge to calculate annual loads, assuming stationarity in the landscape (i.e. no changes in land use between years). Average annual basin loads for each of the Great Lakes is presented in Table 5 below. Estimated annual nutrient loads for each year from 2000-2016 for DIN, SRP, and TP are summarized in Table A.7 through Table A.9 in the Appendix.

Table 5 – Mean annual modelled basin loads to the Laurentian Great Lakes in tonnes/year for period of 2000-2016. SE refers to the standard error of predictions.

	Lake Erie		**Lake Huron**		**Lake Michigan**		**Lake Ontario**		**Lake Superior**	
	Annual Load	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**
DIN	120,586	2,832	53,624	10,517	48,775	7,814	27,931	3,483	10,241	12,568
SRP	2,584	65	928	300	776	306	565	66	321	353
TP	8,541	92	4,895	425	3,673	433	2,358	94	3,506	500

Lake Erie had the largest modelled annual basin loads for DIN, SRP, and TP of all the Great Lakes by a wide margin (Table 5). These massive loads explain the underlying causes of the eutrophication problems that persist in Lake Erie. Lake Erie had the lowest relative and absolute standard error in model predictions for DIN, SRP, and TP for each of the Great Lakes, indicating the lowest uncertainty in these estimated loads. This is likely attributed to the large number of monitored watersheds in the Lake Erie basin used to train the RF models compared to the other Great Lakes basins (Figure 2). The second largest loadings occurred in Lake Huron for all nutrient constituents (Table 5). These high loadings in Lake Erie and Huron reflect the hot spots of nutrient pollution seen in watersheds of these drainage basins (Figure 8) and shows the effect of anthropogenic landscapes in driving nutrient pollution. Lake Superior had the lowest nutrient loads for dissolved nutrients, DIN and SRP. The Lake Superior basin is the least human impacted of all the Great Lakes (Table 5), and these low DIN and SRP loads highlight that low dissolved nutrient export coincides with low anthropogenic impacts within a watershed. Lake Ontario had the lowest TP loadings of any of the Great Lake (Table 5). This is likely due to Lake Ontario having the smallest drainage basin of the Great Lakes (MacDonagh-Dumler, Pebbles, and Gannon 2003), and thus less TP export from less drainage discharge.

3.5.2 Comparison with Literature Estimates

Great Lakes basin loading estimates were compared to loads published in literature to assess modelled results. Figure 14 and Figure 15 show annual modelled loads and annual published basin load values, for SRP and TP, respectively. Overall, estimated loading compared similarly with other published estimates.

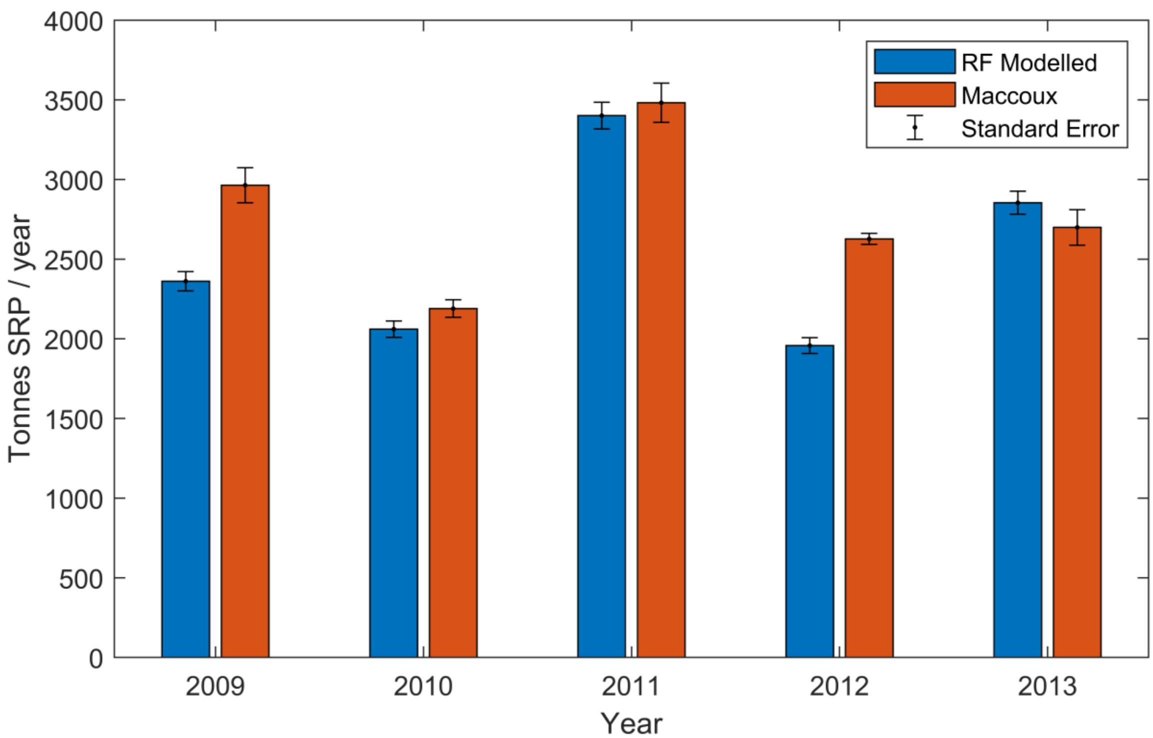

Figure 14 – Annual SRP loading to Lake Erie from RF model (blue) compared to annual loads from Maccoux et al. (2016) (orange).

Average annual SRP export to Lake Erie was estimated to be approximately 2,600 tonnes per year, which was consistent with loads published by Maccoux et al. (2016), where the annual load estimates ranged from 2,627 to 3,482 tonnes per year between 2009 and 2013 (Figure 14). Trends in SRP loading year to year were also similar with Maccoux et al. (2016) estimates. Both modelled and published values showed peak SRP loadings in 2011, which coincided with the largest algal bloom recorded in Lake Erie to date (Michalak et al. 2013).

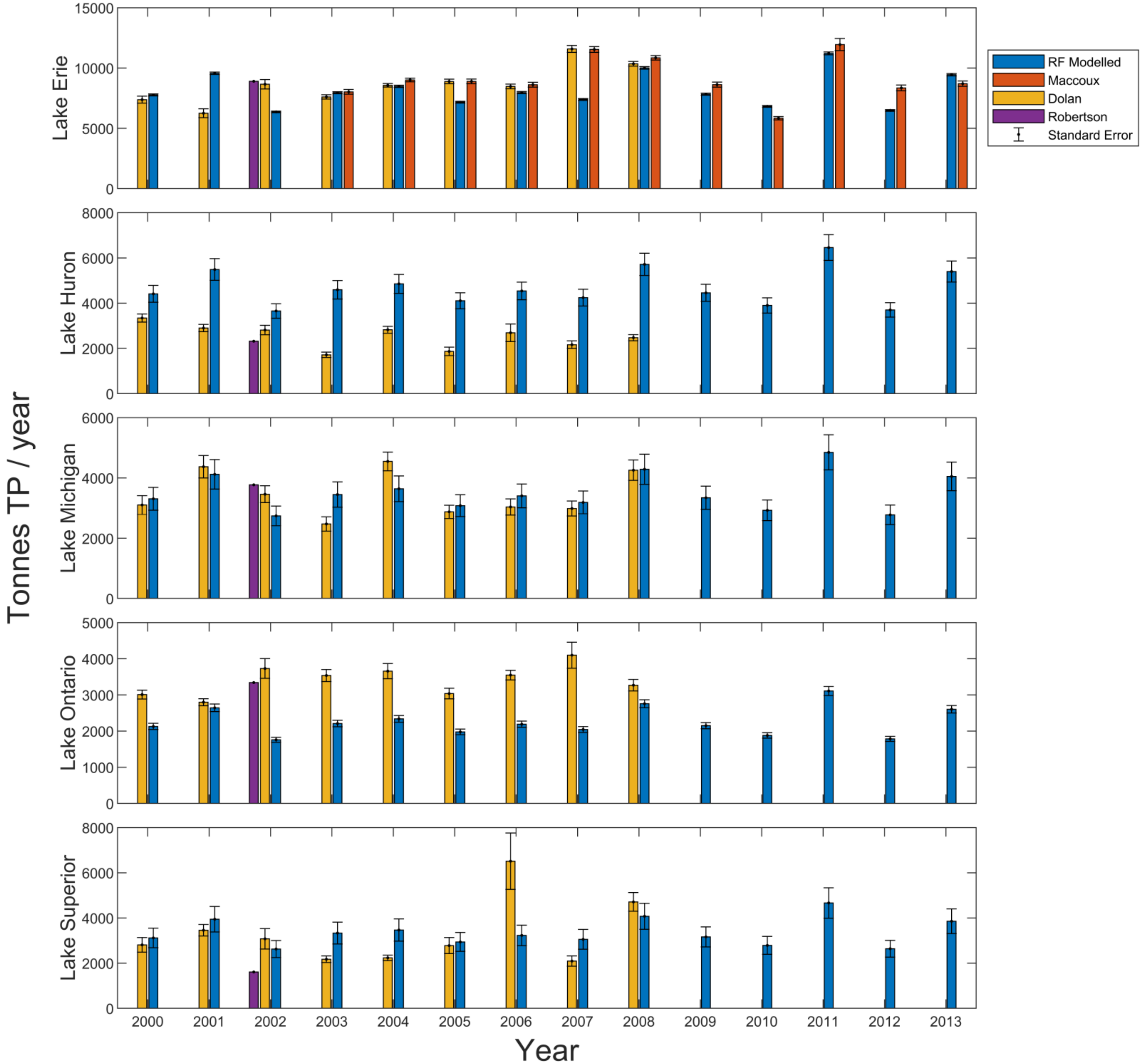

Figure 15 – Annual TP loading to Great Lakes from RF models (blue) compared to annual loads from Maccoux et al. (2016) (orange), Dolan and Chapra (2012) (yellow), and Robertson et al. (2019) (purple).

Average annual TP estimates to Lake Erie were approximately 8,500 tonnes per year. This is similar to estimates reported by Maccoux et al. (2016) and Dolan and Chapra (2012) where Lake Erie TP loads ranged from 8,024 to 11,946 tonnes per year between 2003 and 2013, and 6,252 to 11,584 tonnes per year between 2000 and 2008, respectively. Calculated loads were also like Robertson et al. (2019), who estimated Lake Erie watershed loadings of 8,900 tonnes for the year 2002. Again, modelled TP loadings were greatest for the year 2011, which coincided with the largest recorded algal bloom in Lake Erie to date (Figure 15).

TP loadings for other Great Lakes were also consistent with published values, however larger uncertainty exists in both modelled results in this study and other published results, as seen by the greater standard error in annual loads to Lakes Michigan, Ontario, Superior, and Huron, when compared to Lake Erie (Figure 15). This is likely attributed to the fewer number of monitoring stations for the drainage basins of these lakes as opposed to the greater number in the highly impacted Lake Erie basin (Figure 2). Average annual TP load estimates for Lake Michigan were approximately 3,700 tonnes per year, compared to similar load estimates reported by Dolan and Chapra (2012) of 2,472 to 4,548 tonnes for 2000-2008, and Robertson et al. (2019) of 3,770 tonnes for 2002. TP estimates for Lake Superior were approximately 3,500 tonnes per year, consistent with estimates from Dolan and Chapra (2012) where TP loads estimates from 2000-2008 were approximately 2,091 to 6,512 tonnes per year. Average annual TP estimates for Lake Ontario were approximately 2,400 tonnes per year and were generally less than estimates by Dolan and Chapra (2012) of 2,798 to 4,098 tonnes per year from 2000-2008. Modelled annual TP loadings for Lake Huron were much higher than annual loadings calculated by Dolan and Chapra (2012), and Robertson et al. (2019). The discharge driven approach to load estimation likely accounts for the observed differences between published values and load estimates for Lake Ontario and Lake Huron. Lake Huron has the largest drainage basin of the Great Lakes, while Lake Ontario has the smallest. As such, due to the area-discharge relationships used for load estimation, the basin area of these lakes could explain why estimated Lake Ontario loads were lower than published values, while Lake Huron loads were higher.

Methodologies for published loads were different than those used by this study, likely explaining the differences in estimated loads. Dolan and Chapra (2012), and Maccoux et al. (2016) calculated loads from monitored tributaries using the Stratified Beale's Ratio Estimator (SBRE), while loads from unmonitored tributaries were calculated using the unit area load (UAL) method. SBRE estimates loads by sorting and separating monitored water quality data into groups based on the monitored daily discharge for each tributary. Annual loads for each tributary are then estimated from relationships drawn between the average monitored load and the average monitored discharge in each of these groups (Beale 1962; Dolan, Yui, and Geist 1981). The WRTDS method, which explicitly captures the relationships between concentration and discharge that vary with time and season, is a more robust method for estimating annual loads from sparse concentration data (Hirsch, Moyer, and Archfield 2010). The UAL method estimates loads for unmonitored watersheds based on loads from nearby monitored watersheds. However, unlike the RF approach, this method fails to capture the land use types that are unique to unmonitored basins. Robertson et al. (2019) also calculated loads from monitored watersheds using Stratified Beale's Ratio Estimator, however, they then used these loads for training their SPARROW model. SPARROW is a multiple linear regression model that uses watershed attributes to calculate nutrient loads. These attributes are associated with sources and sinks of nutrients and include wastewater treatment plants, fertilizer application and forested, wetland and shrubland area. Differences between loads calculated from Robertson et al. (2019) and in this paper could be attributed to the linear nature of SPARROW versus the non-parametric capability of RF models, in addition to the fewer variables, and larger watersheds used in this study.

The estimated loads presented here are key for understanding the current conditions of the Great Lakes and potential paths forward in nutrient management. They offer insight into the magnitude of the underlying driver of modern-day eutrophication challenges. While TP loads, primarily from point sources, have been reduced due to the GLWQA in 1978, water quality challenges from nutrient pollution still persist in the Great Lakes (Robertson and Saad 2011). By quantifying nutrient loads, managers can be better informed on the performance of past reduction strategies, while also establishing a baseline to develop and adapt new strategies to achieve future targets. This is especially important for measuring progress on the recent 40% reduction targets for SRP and TP, set binationally to improve Lake Erie's water quality.

4.0 Conclusions

The consequences of eutrophication from anthropogenic nutrient pollution pose serious threats to our water resources in the Great Lakes basin. A data driven approach in this study was undertaken to assess the drivers and behaviour of nutrient water quality within the Great Lakes basin. Monitored water quality and discharge across the basin were used in conjunction with statistical methods and machine learning models to evaluate relationships between nutrient water quality and spatial variables (e.g. soil type, land use, population density etc.). This analysis highlighted the impacts of anthropogenic landscapes and non-point sources in altering the nutrient water quality regime downstream.

The results from our analysis showed land use (i.e. percent developed land, tile drained land, percent wetlands, etc.) as the strongest driver of flow-weighted concentrations for DIN, SRP, and PP. Variable importance metrics revealed that developed land use and percent tile drained land, were the important drivers of DIN and SRP, while soil type and wetland area were the most important drivers of PP concentrations. Feature contribution plots showed that the highest DIN and SRP concentration hot spots in watersheds were due to high developed land use, high proportions of tile drained land, and low wetland cover. The highest PP concentration hot spots were due to high proportions of fine soil cover and low wetland cover. Threshold behaviour in drivers of nutrient concentrations was observed, particularly when looking at percent wetland area and developed land use in a catchment. This behaviour may be valuable for nutrient managers to use when considering wetland restoration strategies for nutrient reduction while balancing interests of stakeholders. Variable importance metrics highlighted the role of manure and fertilizer practices in driving nutrient ratios of SRP:TP and DIN:TP. Livestock density, developed land use and tile drained land, were shown to be important drivers in SRP:TP ratios, while livestock density, developed land use, and soil type were shown to be important drivers in DIN:TP. Feature contribution plots for ratios revealed that the highest

SRP:TP and DIN:TP ratios in the basin were due to high developed land use and high livestock densities, while high proportions of tile drainage also played a significant role in elevating DIN:TP ratios. Overall, these anthropogenic drivers of nutrient water quality highlight the underlying impacts of human sources and practices of nutrient pollution in the Great Lakes drainage basin.

RF models were used to estimate nutrient concentrations and ratios across all watersheds in the Great Lakes basin to identify nutrient hot spots. Hot spots for DIN, SRP, and PP were found in the southwestern watersheds of Lake Erie, the southeastern shores of Lake Huron, and in the watersheds along the eastern side of the Huron-Erie corridor. Lowest nutrient FWCs were observed in northern Lake Huron and Lake Superior watersheds, where there are fewer human impacted catchments. Hot spots for high SRP:TP were also found in the watersheds along the eastern side of the Huron-Erie corridor, and in the southeastern shores of Lake Huron. Despite high nutrient export, low DIN:TP ratios were seen in the watersheds of southwestern Lake Erie, which could have implications in promoting the growth of harmful algal blooms in the lake. Overall, nutrient hot spots coincided with heavily human impacted catchments, further highlighting the anthropogenic impact of nutrient pollution in the Great Lakes. Additionally, nutrient ratios across the Great Lakes basin were consistent with ratios found in heavily human impacted catchments like the Baltic Sea and French river systems in Western Europe.

Annual DIN, SRP, and TP loads were estimated using RF modelled concentrations and area-discharge relationships from monitored data to determine the total annual export for each lake. Lake Erie was observed to have the highest load for all nutrients. Lake Superior showed the lowest average annual export for DIN and SRP, and Lake Ontario had the lowest TP export. Highest calculated export of SRP and TP occurred in 2011, which also coincided with the largest recorded algal bloom in Lake Erie to date. Calculated annual loadings of SRP and TP from the developed models were found to be consistent with other published loadings, adding confidence to the modelled results.

Limitations of this study included the uncertainty in the data itself and the applied assumptions. Error and uncertainty inherent in measured nutrient water quality and discharge data, in addition to the error inherent in spatial data, were not considered. Monitoring stations were constrained to have both water quality and discharge in collocated areas, thus restricting the size of the data sets used for WRTDS and machine learning. A greater number of monitoring stations would have been ideal in this analysis, especially given the tendency of machine learning to use large data sets. Monitoring stations were also biased to the Lake Erie drainage basin, and a more uniform distribution of stations across the Great Lakes basin would have been ideal. Stationarity with time was assumed in landscape variables and in water quality from the years 2000 to 2016; changes in the landscape and water quality regime during the time period were ignored. Loadings were calculated using annual area-discharge relationships from monitored stations, and the uncertainty of these relationships was not incorporated into estimates. While simplicity of spatial data was key to the approach of this study, other spatial variables that represent important factors to nutrient water quality would be valuable to include, such as spatial data showing groundwater contribution, or the abundance of BMPs. Furthermore, while efforts were made to minimize collinearity in the data set, variable intercorrelation still existed and potentially hampered assessments of drivers for nutrient water quality and statistical assumptions inherent in models.

Potential paths forward include the assessment of drivers of nutrient water quality under different conditions. Do drivers compare in the spring, summer and fall months? Are drivers different in low flow vs high flow conditions? Differences in drivers under these conditions may yield valuable insights for nutrient managers in temporally targeted reduction strategies. Another avenue for future work includes seeing how drivers change with evolving land use and water quality regimes. By revoking stationarity assumptions, drivers may differ over time, which would highlight the performance of past reduction strategies while highlighting current nutrient pollution challenges in the Great Lakes. It may also be interesting to take a deeper dive into the generic spatial variables used in this study and look at more detailed spatial information. For example, how would the nutrient reduction capabilities of wetlands change when looking at the connectivity and size of wetlands within watersheds? How do different agricultural practices and urban land uses play a role in increasing nutrient concentrations from developed land use? Loading estimates could be further improved, as the annual discharge used to scale FWCs was only a function of watershed area. Greater efforts to better capture discharge could be taken, such as incorporating monitored annual precipitation, or monitored discharge from adjacent watersheds. Additionally, error could be better aggregated and represented in the models used for analysis. While flux bias was assessed in WRTDS, the error in the regression relationships could be incorporated and compounded into the error estimates from RF machine learning models. Furthermore, more efforts should be made to investigate the potential for nutrient management in leveraging the threshold behaviour of wetland and developed land use in their effects of reducing nutrient concentrations and loads.

Recognizing the drivers of nutrient pollution and hot spot locations is critical for nutrient managers in implementing targeted management. The results of our analysis support our current understanding of nutrient pollution dynamics and are specific to the responses and behaviours of the Laurentian Great Lakes drainage basin. This reinforces and refines our knowledge of the underlying drivers of nutrient pollution in the Great Lakes and demonstrates that stochastic perspectives and machine learning tools can be used to reveal the “big picture” behaviour of complex environmental systems.

References

Al-Kandari, Noriah M., and Ian T. Jolliffe. 2005. “Variable Selection and Interpretation in Correlation Principal Components.” *Environmetrics* 16 (6): 659–72. https://doi.org/10.1002/env.728.

Allan, J. David, Michael W. Murray, and Matthew Child. 2018. “Fertilizer Application Patterns and Trends and Their Implications for Water Quality in the Western Lake Erie Basinrk Group Members and Reviewers.” E95-2/30-2018E-PDF. International Joint Commision.

Anderson, Donald M., Patricia M. Glibert, and Joann M. Burkholder. 2002. “Harmful Algal Blooms and Eutrophication: Nutrient Sources, Composition, and Consequences.” *Estuaries* 25 (4): 704–26. https://doi.org/10.1007/BF02804901.

Baker, D. B., R. Confesor, D. E. Ewing, L. T. Johnson, J. W. Kramer, and B. J. Merryfield. 2014. “Phosphorus Loading to Lake Erie from the Maumee, Sandusky and Cuyahoga Rivers: The Importance of Bioavailability.” *Journal of Great Lakes Research* 40 (3): 502–17. https://doi.org/10.1016/j.jglr.2014.05.001.

Basu, Nandita B., Sally E. Thompson, and P. Suresh C. Rao. 2011. “Hydrologic and Biogeochemical Functioning of Intensively Managed Catchments: A Synthesis of Top-down Analyses: MANAGED CATCHMENTS.” *Water Resources Research* 47 (10). https://doi.org/10.1029/2011WR010800.

Beale, E.M.L. 1962. “Some Uses of Computers in Operational Research.” *Journal of Industrial Organization*, no. 31: 27–28.

Beeton, Alfred M. 1965. “Eutrophication of the St. Lawrence Great Lakes1.” *Limnology and Oceanography* 10 (2): 240–54. https://doi.org/10.4319/lo.1965.10.2.0240.

Bootsma, Harvey A., Mark D. Rowe, Colin N. Brooks, and Henry A. Vanderploeg. 2015. “Commentary: The Need for Model Development Related to Cladophora and Nutrient Management in Lake Michigan.” *Journal of Great Lakes Research*, Complex interactions in Lake Michigan’s rapidly changing ecosystem, 41 (January): 7–15. https://doi.org/10.1016/j.jglr.2015.03.023.

Bosch, Nathan S., J. David Allan, James P. Selegean, and Donald Scavia. 2013. “Scenario-Testing of Agricultural Best Management Practices in Lake Erie Watersheds.” *Journal of Great Lakes Research* 39 (3): 429–36. https://doi.org/10.1016/j.jglr.2013.06.004.

Bosch, Nathan S., Mary Anne Evans, Donald Scavia, and J. David Allan. 2014. “Interacting Effects of Climate Change and Agricultural BMPs on Nutrient Runoff Entering Lake Erie.” *Journal of Great Lakes Research* 40 (3): 581–89. https://doi.org/10.1016/j.jglr.2014.04.011.

Breiman, Leo. 2001. “Random Forests.” *Machine Learning* 45 (1): 5–32.

Bundy, L. G., T. W. Andraski, and J. M. Powell. 2001. “Management Practice Effects on Phosphorus Losses in Runoff in Corn Production Systems.” *Journal of Environmental Quality* 30 (5): 1822–28. https://doi.org/10.2134/jeq2001.3051822x.

Canada, Canada Environment, Harvey Shear, and Jennifer Wittig. 1995. *The Great Lakes: An Environmental Atlas and Resource Book*. Great Lakes National Program Office, U.S. Environmental Protection Agency.

Carlisle, Daren M., James Falcone, and Michael R. Meador. 2009. “Predicting the Biological Condition of Streams: Use of Geospatial Indicators of Natural and Anthropogenic Characteristics of Watersheds.” *Environmental Monitoring and Assessment* 151 (1): 143–60. https://doi.org/10.1007/s10661-008-0256-z.

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. “NONPOINT POLLUTION OF SURFACE WATERS WITH PHOSPHORUS AND NITROGEN.” *Ecological Applications* 8 (3): 559–68. https://doi.org/10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2.

Chatterjee, Sangit, and Mustafa Yilmaz. 2016. “A Review of Regression Diagnostics for Behavioral Research:” *Applied Psychological Measurement*, July. https://doi.org/10.1177/014662169201600301.

Chowdhury, Qualbe Shadman T. 2018. “Biogeochemical Signatures of the Great Lakes Watersheds.” University of Waterloo. http://hdl.handle.net/10012/13624.

Colborne, S.F., T.J. Maguire, B. Mayer, M. Nightingale, G.E. Enns, A.T. Fisk, K.G. Drouillard, et al. 2019. “Water and Sediment as Sources of Phosphate in Aquatic Ecosystems: The Detroit River and Its Role in the Laurentian Great Lakes.” *Science of The Total Environment* 647 (January): 1594–1603. https://doi.org/10.1016/j.scitotenv.2018.08.029.

Cole, Marci L., Ivan Valiela, Kevin D. Kroeger, Gabrielle L. Tomasky, Just Cebrian, Cathleen Wigand, Richard A. McKinney, Sara P. Grady, and Maria Helena Carvalho da Silva. 2004. “Assessment of a δ 15 N Isotopic Method to Indicate Anthropogenic Eutrophication in Aquatic Ecosystems.” *Journal of Environmental Quality* 33 (1): 124–32. https://doi.org/10.2134/jeq2004.1240.

Conley, Daniel J., Hans W. Paerl, Robert W. Howarth, Donald F. Boesch, Sybil P. Seitzinger, Karl E. Havens, Christiane Lancelot, and Gene E. Likens. 2009. “Controlling Eutrophication: Nitrogen and Phosphorus.” *Science* 323 (5917): 1014–15. https://doi.org/10.1126/science.1167755.

Correll, David L. 1998. “The Role of Phosphorus in the Eutrophication of Receiving Waters: A Review.” *Journal of Environmental Quality* 27 (2): 261–66. https://doi.org/10.2134/jeq1998.00472425002700020004x.

Dagnew, Awoke, Donald Scavia, Yu-Chen Wang, Rebecca Muenich, and Margaret Kalcic. 2019. “Modeling Phosphorus Reduction Strategies from the International St. Clair-Detroit River System Watershed.” *Journal of Great Lakes Research* 45 (4): 742–51. https://doi.org/10.1016/j.jglr.2019.04.005.

Dagnew, Awoke, Donald Scavia, Yu-Chen Wang, Rebecca Muenich, Colleen Long, and Margaret Kalcic. 2019. “Modeling Flow, Nutrient, and Sediment Delivery from a Large International Watershed Using a Field-Scale SWAT Model.” *JAWRA Journal of the American Water Resources Association* 55 (5): 1288–1305. https://doi.org/10.1111/1752-1688.12779.

Daloğlu, Irem, Kyung Hwa Cho, and Donald Scavia. 2012. “Evaluating Causes of Trends in Long-Term Dissolved Reactive Phosphorus Loads to Lake Erie.” *Environmental Science & Technology* 46 (19): 10660–66. https://doi.org/10.1021/es302315d.

Díaz-Uriarte, Ramón, and Sara Alvarez de Andrés. 2006. “Gene Selection and Classification of Microarray Data Using Random Forest.” *BMC Bioinformatics* 7 (1): 3. https://doi.org/10.1186/1471-2105-7-3.

Dodds, Walter K., Wes W. Bouska, Jeffrey L. Eitzmann, Tyler J. Pilger, Kristen L. Pitts, Alyssa J. Riley, Joshua T. Schloesser, and Darren J. Thornbrugh. 2009. “Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages.” *Environmental Science & Technology* 43 (1): 12–19. https://doi.org/10.1021/es801217q.

Dolan, David M., and Steven C. Chapra. 2012. “Great Lakes Total Phosphorus Revisited: 1. Loading Analysis and Update (1994–2008).” *Journal of Great Lakes Research* 38 (4): 730–40. https://doi.org/10.1016/j.jglr.2012.10.001.

Dolan, David M., Alexander K. Yui, and Raymond D. Geist. 1981. “Evaluation of River Load Estimation Methods for Total Phosphorus.” *Journal of Great Lakes Research* 7 (3): 207–14. https://doi.org/10.1016/S0380-1330(81)72047-1.

Dove, Alice, and Steven C. Chapra. 2015. “Long-Term Trends of Nutrients and Trophic Response Variables for the Great Lakes: Great Lakes Nutrient Trends.” *Limnology and Oceanography* 60 (2): 696–721. https://doi.org/10.1002/lno.10055.

Duan, Naihua. 1983. “Smearing Estimate: A Nonparametric Retransformation Method.” *Journal of the American Statistical Association* 78 (383): 605–10. https://doi.org/10.2307/2288126.

Dugan, Hilary A., Nicholas K. Skaff, Jonathan P. Doubek, Sarah L. Bartlett, Samantha M. Burke, Flora E. Krivak-Tetley, Jamie C. Summers, Paul C. Hanson, and Kathleen C. Weathers. 2020. “Lakes at Risk of Chloride Contamination.” *Environmental Science & Technology* 54 (11): 6639–50. https://doi.org/10.1021/acs.est.9b07718.

Dumont, E., J. A. Harrison, C. Kroeze, E. J. Bakker, and S. P. Seitzinger. 2005. “Global Distribution and Sources of Dissolved Inorganic Nitrogen Export to the Coastal Zone: Results from a Spatially Explicit, Global Model.” *Global Biogeochemical Cycles* 19 (4). https://doi.org/10.1029/2005GB002488.

Environment and Climate Change Canada. 2020. “Historical Hydrometric Data.” Environment and Climate Change Canada. https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html.

“Environment and Climate Change Canada Historical Hydrometric Data.” 2016. Environment and Climate Change Canada. https://wateroffice.ec.gc.ca/mainmenu/historical_data_index_e.html.

Environment and Climate Change Canada, and Ontario Ministry of the Environment and Climate Change. 2018a. “Canada-Ontario Lake Erie Action Plan: Partnering on Achieving Phosphorus Loading Reductions to Lake Erie from Canadian Sources.” https://www.watercanada.net/lake-erie-action-plan-to-reduce-phosphorus-loads-by-40/.

———. 2018b. “Canada-Ontario Lake Erie Action Plan.” https://www.canada.ca/en/environment-climate-change/services/great-lakes-protection/action-plan-reduce-phosphorus-lake-erie.html.

Environment and Climate Change Canada, and U.S. Environmental Protection Agency. 2017. “State of the Great Lakes 2017 Technical Report.” https://binational.net/2017/06/19/sogl-edgl-2017/.

Fick, S.E., and R.J. Hijmans. 2017. “WorldClim 2: New 1km Spatial Resolution Climate Surfaces for Global Land Areas.” *International Journal of Climatology* 37 (12): 4302–15.

Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, and D. Wiberg. 2008. “Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008).” IIASA, Laxenburg, Austria and FAO.

Fitzsimmons, Emma G. 2014. “TAP WATER BAN FOR TOLEDO RESIDENTS.” *New York Times*, August 3, 2014.

Friedman, Jerome H. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” *The Annals of Statistics* 29 (5): 1189–1232.

Gibson, John J., Pradeep Aggarwal, James Hogan, Carol Kendall, Luiz A. Martinelli, Willi Stichler, Dieter Rank, et al. 2002. “Isotope Studies in Large River Basins: A New Global Research Focus.” *Eos, Transactions American Geophysical Union* 83 (52): 613–17. https://doi.org/10.1029/2002EO000415.

Glibert, Patricia M. 2020. “Harmful Algae at the Complex Nexus of Eutrophication and Climate Change.” *Harmful Algae*, Climate change and harmful algal blooms, 91 (January): 101583. https://doi.org/10.1016/j.hal.2019.03.001.

Glibert, Patricia M., and JoAnn M. Burkholder. 2011. “Harmful Algal Blooms and Eutrophication: ‘Strategies’ for Nutrient Uptake and Growth Outside the Redfield Comfort Zone.” *Chinese Journal of Oceanology and Limnology* 29 (4): 724–38. https://doi.org/10.1007/s00343-011-0502-z.

Glibert, Patricia M., David Fullerton, Joann M. Burkholder, Jeffrey C. Cornwell, and Todd M. Kana. 2011. “Ecological Stoichiometry, Biogeochemical Cycling, Invasive Species, and Aquatic Food Webs: San Francisco Estuary and Comparative Systems.” *Reviews in Fisheries Science* 19 (4): 358–417. https://doi.org/10.1080/10641262.2011.611916.

Government of Canada, Statistics Canada. 2001. “A Geographic Profile of Manure Production in Canada - ARCHIVED.” February 22, 2001. https://www150.statcan.gc.ca/n1/en/catalogue/16F0025X.

Grömping, Ulrike. 2009. “Variable Importance Assessment in Regression: Linear Regression versus Random Forest.” *The American Statistician* 63 (4): 308–19. https://doi.org/10.1198/tast.2009.08199.

Hamilton, Stephen K. 2012. “Biogeochemical Time Lags May Delay Responses of Streams to Ecological Restoration: *Time Lags in Stream Restoration*.” *Freshwater Biology* 57 (July): 43–57. https://doi.org/10.1111/j.1365-2427.2011.02685.x.

Han, Haejin, J. David Allan, and Nathan S. Bosch. 2012. “Historical Pattern of Phosphorus Loading to Lake Erie Watersheds.” *Journal of Great Lakes Research* 38 (2): 289–98. https://doi.org/10.1016/j.jglr.2012.03.004.

Hansen, Amy T., Christine L. Dolph, Efi Foufoula-Georgiou, and Jacques C. Finlay. 2018. “Contribution of Wetlands to Nitrate Removal at the Watershed Scale.” *Nature Geoscience* 11 (2): 127–32. https://doi.org/10.1038/s41561-017-0056-6.

Hirsch, Robert M., and Laura A. De Cicco. 2015. “User Guide to Exploration and Graphics for RivEr Trends (EGRET) and DataRetrieval: R Packages for Hydrologic Data.” USGS Numbered Series 4-A10. *User Guide to Exploration and Graphics for RivEr Trends (EGRET) and DataRetrieval: R Packages for Hydrologic Data*. Vol. 4-A10. Techniques and Methods. Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/tm4A10.

Hirsch, Robert M., Douglas L. Moyer, and Stacey A. Archfield. 2010. “Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs1: Weighted Regressions on Time, Discharge, and Season (WRTDS), With an Application to Chesapeake Bay River Inputs.” *JAWRA Journal of the American Water Resources Association* 46 (5): 857–80. https://doi.org/10.1111/j.1752-1688.2010.00482.x.

Hodgkinson, R. A, B. J Chambers, P. J. A Withers, and R Cross. 2002. “Phosphorus Losses to Surface Waters Following Organic Manure Applications to a Drained Clay Soil.” *Agricultural Water Management* 57 (2): 155–73. https://doi.org/10.1016/S0378-3774(02)00057-4.

Jagannath, Venekata. 2017. “Random Forest Template for TIBCO Spotfire® - Wiki Page | TIBCO Community.” 2017. https://community.tibco.com/wiki/random-forest-template-tibco-spotfirer-wiki-page.

Jarvie, Helen, Laura Johnson, Andrew Sharpley, Douglas Smith, David Baker, Tom Bruulsema, and Remegio Confesor. 2017. “Increased Soluble Phosphorus Loads to Lake Erie: Unintended Consequences of Conservation Practices?” *Journal of Environmental Quality* 46 (January): 123–132. https://doi.org/10.2134/jeq2016.07.0248.

Jarvie, Helen P., Andrew N. Sharpley, Paul J. A. Withers, J. Thad Scott, Brian E. Haggard, and Colin Neal. 2013. “Phosphorus Mitigation to Control River Eutrophication: Murky Waters, Inconvenient Truths, and ‘Postnormal’ Science.” *Journal of Environmental Quality* 42 (2): 295–304. https://doi.org/10.2134/jeq2012.0085.

Jetoo, Savitri, Velma I. Grover, and Gail Krantzberg. 2015. “The Toledo Drinking Water Advisory: Suggested Application of the Water Safety Planning Approach.” *Sustainability* 7 (8): 9787–9808. https://doi.org/10.3390/su7089787.

Jolliffe, I. T. 1972. “Discarding Variables in a Principal Component Analysis. I: Artificial Data.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 21 (2): 160–73. https://doi.org/10.2307/2346488.

———. 1973. “Discarding Variables in a Principal Component Analysis. Ii: Real Data.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 22 (1): 21–31. https://doi.org/10.2307/2346300.

Joosse, P. J., and D. B. Baker. 2011. “Context for Re-Evaluating Agricultural Source Phosphorus Loadings to the Great Lakes.” *Canadian Journal of Soil Science* 91 (3): 317–27. https://doi.org/10.4141/cjss10005.

Kalcic, Margaret M., Rebecca Logsdon Muenich, Samantha Basile, Allison L. Steiner, Christine Kirchhoff, and Donald Scavia. 2019. “Climate Change and Nutrient Loading in the Western Lake Erie Basin: Warming Can Counteract a Wetter Future.” *Environmental Science & Technology* 53 (13): 7543–50. https://doi.org/10.1021/acs.est.9b01274.

Kane, Douglas D., Joseph D. Conroy, R. Peter Richards, David B. Baker, and David A. Culver. 2014. “Re-Eutrophication of Lake Erie: Correlations between Tributary Nutrient Loads and Phytoplankton Biomass.” *Journal of Great Lakes Research* 40 (3): 496–501. https://doi.org/10.1016/j.jglr.2014.04.004.

King, Katelyn, Kendra Spence Cheruvelil, and Amina Pollard. 2019. “Drivers and Spatial Structure of Abiotic and Biotic Properties of Lakes, Wetlands, and Streams at the National Scale.” *Ecological Applications* 29 (7): e01957. https://doi.org/10.1002/eap.1957.

Kinley, Robert D., Robert J. Gordon, Glenn W. Stratton, Gary T. Patterson, and Jeff Hoyle. 2007. “Phosphorus Losses through Agricultural Tile Drainage in Nova Scotia, Canada.” *Journal of Environmental Quality* 36 (2): 469–77. https://doi.org/10.2134/jeq2006.0138.

Kleinman, Peter J. A., Ann M. Wolf, Andrew N. Sharpley, Douglas B. Beegle, and Lou S. Saporito. 2005. “Survey of Water-Extractable Phosphorus in Livestock Manures.” *Soil Science Society of America Journal* 69 (3): 701–8. https://doi.org/10.2136/sssaj2004.0099.

Lam, W. V., M. L. Macrae, M. C. English, I. P. O’Halloran, and Y. T. Wang. 2016. “Effects of Tillage Practices on Phosphorus Transport in Tile Drain Effluent under Sandy Loam Agricultural Soils in Ontario, Canada.” *Journal of Great Lakes Research* 42 (6): 1260–70. https://doi.org/10.1016/j.jglr.2015.12.015.

Le Moal, Morgane, Chantal Gascuel-Odoux, Alain Ménesguen, Yves Souchon, Claire Étrillard, Alix Levain, Florentina Moatar, et al. 2019. “Eutrophication: A New Wine in an Old Bottle?” *Science of The Total Environment* 651 (February): 1–11. https://doi.org/10.1016/j.scitotenv.2018.09.139.

Lee, G. Fred. 1973. “Role of Phosphorus in Eutrophication and Diffuse Source Control.” *Water Research* 7 (1): 111–28. https://doi.org/10.1016/0043-1354(73)90156-5.

Lee, G. Fred, Walter Rast, and R. Anne Jones. 1978. “Water Report: Eutrophication of Water Bodies: Insights for an Age Old Problem.” *Environmental Science & Technology* 12 (8): 900–908. https://doi.org/10.1021/es60144a606.

Liaw, Andy, and Matthew Wiener. 2002. “Classification and Regression by RandomForest” 2: 6.

Maccoux, Matthew J., Alice Dove, Sean M. Backus, and David M. Dolan. 2016. “Total and Soluble Reactive Phosphorus Loadings to Lake Erie.” *Journal of Great Lakes Research* 42 (6): 1151–65. https://doi.org/10.1016/j.jglr.2016.08.005.

MacDonagh-Dumler, Jon, Victoria Pebbles, and John Gannon. 2003. “Lake Basin Management Initiative Experience and Lessons Learned Brief North American Great Lakes.” In .

Mansfield, Edward R., and Billy P. Helms. 1982. “Detecting Multicollinearity.” *The American Statistician* 36 (3a): 158–60. https://doi.org/10.1080/00031305.1982.10482818.

Maranger, Roxane, Stuart E. Jones, and James B. Cotner. 2018. “Stoichiometry of Carbon, Nitrogen, and Phosphorus through the Freshwater Pipe.” *Limnology and Oceanography Letters* 3 (3): 89–101. https://doi.org/10.1002/lol2.10080.

McDonnell, J. J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, et al. 2007. “Moving beyond Heterogeneity and Process Complexity: A New Vision for Watershed Hydrology.” *Water Resources Research* 43 (7). https://doi.org/10.1029/2006WR005467.

Meals, Donald W., Steven A. Dressing, and Thomas E. Davenport. 2010. “Lag Time in Water Quality Response to Best Management Practices: A Review.” *Journal of Environment Quality* 39 (1): 85. https://doi.org/10.2134/jeq2009.0108.

Medalie, Laura, Robert M. Hirsch, and Stacey A. Archfield. 2012. “Use of Flow-Normalization to Evaluate Nutrient Concentration and Flux Changes in Lake Champlain Tributaries, 1990–2009.” *Journal of Great Lakes Research*, Lake Champlain in 2010, 38 (January): 58–67. https://doi.org/10.1016/j.jglr.2011.10.002.

Meinshausen, Nicolai. 2006. “Quantile Regression Forests,” 17.

Michalak, Anna, Eric Anderson, Dmitry Beletsky, Steven Boland, Nathan Bosch, Thomas Bridgeman, Justin Chaffin, et al. 2013. “Record-Setting Algal Bloom in Lake Erie Caused by Agricultural and Meteorological Trends Consistent with Expected Future Conditions.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (April). https://doi.org/10.1073/pnas.1216006110.

Moss, Brian. 2008. “Water Pollution by Agriculture.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1491): 659–66. https://doi.org/10.1098/rstb.2007.2176.

Moss, Brian, Sarian Kosten, Mariana Meerhoff, Richard W. Battarbee, Erik Jeppesen, Néstor Mazzeo, Karl Havens, et al. 2011. “Allied Attack: Climate Change and Eutrophication.” *Inland Waters* 1 (2): 101–5. https://doi.org/10.5268/IW-1.2.359.

Munroe, Jake, Christine Brown, John Lauzon, Tom Bruulsema, Dale Cowan, and Ivan O’Halloran. 2018. “Soil Fertility Handbook Publication 611, 3rd Edition.” Ontario Ministry of Agriculture, Food and Rural Affairs. http://www.omafra.gov.on.ca/english/crops/pub611/pub611.pdf.

“National Water Information System.” 2016. U.S. Geological Survey. http://waterdata.usgs.gov/nwis/.

Neff, Brian, S.M. Day, A.R. Piggott, and L.M. Fuller. 2005. *Base Flow in the Great Lakes Basin*. *U.S. Geol. Surv. Sci. Invest. Rep*. https://doi.org/10.3133/sir20055217.

"NOAA Great Lakes Environmental Research Laboratory's Albums." 2015. Flickr. 2015. /photos/noaa_glerl/albums/.

Paerl, Hans W. 1997. "Coastal Eutrophication and Harmful Algal Blooms: Importance of Atmospheric Deposition and Groundwater as 'New' Nitrogen and Other Nutrient Sources." *Limnology and Oceanography* 42 (5part2): 1154–65. https://doi.org/10.4319/lo.1997.42.5_part_2.1154.

Park, S. W., S. Mostaghimi, R. A. Cooke, and P. W. McClellan. 1994. "Bmp Impacts on Watershed Runoff, Sediment, and Nutrient Yields1." *JAWRA Journal of the American Water Resources Association* 30 (6): 1011–23. https://doi.org/10.1111/j.1752-1688.1994.tb03349.x.

Pretty, Jules N., Christopher F. Mason, David B. Nedwell, Rachel E. Hine, Simon Leaf, and Rachael Dils. 2003. "Environmental Costs of Freshwater Eutrophication in England and Wales." *Environmental Science & Technology* 37 (2): 201–8. https://doi.org/10.1021/es020793k.

"Provincial (Stream) Water Quality Monitoring Network." 2016. Government of Ontario. https://data.ontario.ca/dataset/provincial-stream-water-quality-monitoring-network.

Read, Emily K., Vijay P. Patil, Samantha K. Oliver, Amy L. Hetherington, Jennifer A. Brentrup, Jacob A. Zwart, Kirsten M. Winters, et al. 2015. "The Importance of Lake-Specific Characteristics for Water Quality across the Continental United States." *Ecological Applications* 25 (4): 943–55. https://doi.org/10.1890/14-0935.1.

Robertson, Dale M., and David A. Saad. 2011. "Nutrient Inputs to the Laurentian Great Lakes by Source and Watershed Estimated Using SPARROW Watershed Models1." *JAWRA Journal of the American Water Resources Association* 47 (5): 1011–33. https://doi.org/10.1111/j.1752-1688.2011.00574.x.

Robertson, Dale M., David A. Saad, Glenn A. Benoy, Ivana Vouk, Gregory E. Schwarz, and Michael T. Laitta. 2019. "Phosphorus and Nitrogen Transport in the Binational Great Lakes Basin Estimated Using SPARROW Watershed Models." *JAWRA Journal of the American Water Resources Association* 0 (0). https://doi.org/10.1111/1752-1688.12792.

Robinson, Clare. 2015. "Review on Groundwater as a Source of Nutrients to the Great Lakes and Their Tributaries." *Journal of Great Lakes Research* 41 (4): 941–50. https://doi.org/10.1016/j.jglr.2015.08.001.

Saaltink, Rémon, Ype van der Velde, Stefan C. Dekker, Steve W. Lyon, and Helen E. Dahlke. 2014. "Societal, Land Cover and Climatic Controls on River Nutrient Flows into the Baltic Sea." *Journal of Hydrology: Regional Studies* 1 (July): 44–56. https://doi.org/10.1016/j.ejrh.2014.06.001.

Scavia, Donald, Serghei A. Bocaniov, Awoke Dagnew, Yao Hu, Branko Kerkez, Colleen M. Long, Rebecca L. Muenich, Jennifer Read, Lynn Vaccaro, and Yu-Chen Wang. 2019. "Detroit River Phosphorus Loads: Anatomy of a Binational Watershed." *Journal of Great Lakes Research*, November. https://doi.org/10.1016/j.jglr.2019.09.008.

Scavia, Donald, J. David Allan, Kristin K. Arend, Steven Bartell, Dmitry Beletsky, Nate S. Bosch, Stephen B. Brandt, et al. 2014. “Assessing and Addressing the Re-Eutrophication of Lake Erie: Central Basin Hypoxia.” *Journal of Great Lakes Research* 40 (2): 226–46. https://doi.org/10.1016/j.jglr.2014.02.004.

Scavia, Donald, Margaret Kalcic, Rebecca Logsdon Muenich, Jennifer Read, Noel Aloysius, Isabella Bertani, Chelsie Boles, et al. 2017. “Multiple Models Guide Strategies for Agricultural Nutrient Reductions.” *Frontiers in Ecology and the Environment* 15 (3): 126–32. https://doi.org/10.1002/fee.1472.

Schelske, Claire L. 1979. “Role of Phosphorus in Great Lakes Eutrophication: Is There a Controversy?” *Journal of the Fisheries Research Board of Canada* 36 (3): 286–88. https://doi.org/10.1139/f79-045.

Schindler, D. W. 1974. “Eutrophication and Recovery in Experimental Lakes: Implications for Lake Management.” *Science* 184 (4139): 897–99. https://doi.org/10.1126/science.184.4139.897.

———. 2006. “Recent Advances in the Understanding and Management of Eutrophication.” *Limnology and Oceanography* 51 (1part2): 356–63. https://doi.org/10.4319/lo.2006.51.1_part_2.0356.

Sharpley, Andrew N., S. C. Chapra, R. Wedepohl, J. T. Sims, T. C. Daniel, and K. R. Reddy. 1994. “Managing Agricultural Phosphorus for Protection of Surface Waters: Issues and Options.” *Journal of Environment Quality* 23 (3): 437. https://doi.org/10.2134/jeq1994.00472425002300030006x.

Shen, Longzhu, Giuseppe Amatulli, Tushar Sethi, Peter Raymond, and Sami Domisch. 2020. “Estimating Nitrogen and Phosphorus Concentrations in Streams and Rivers, within a Machine Learning Framework.” *Scientific Data* 7 (May). https://doi.org/10.1038/s41597-020-0478-7.

Sinha, Eva, and Anna M. Michalak. 2016. “Precipitation Dominates Interannual Variability of Riverine Nitrogen Loading across the Continental United States.” *Environmental Science & Technology* 50 (23): 12874–84. https://doi.org/10.1021/acs.est.6b04455.

Sivapalan, Murugesu. 2006. “Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale.” In *Encyclopedia of Hydrological Sciences*. American Cancer Society. https://doi.org/10.1002/0470848944.hsa012.

Smith, D. R., P. R. Owens, A. B. Leytem, and E. A. Warnemuende. 2007. “Nutrient Losses from Manure and Fertilizer Applications as Impacted by Time to First Runoff Event.” *Environmental Pollution* 147 (1): 131–37. https://doi.org/10.1016/j.envpol.2006.08.021.

Smith, Douglas R., Wendy Francesconi, Stan J. Livingston, and Chi-hua Huang. 2015. “Phosphorus Losses from Monitored Fields with Conservation Practices in the Lake Erie Basin, USA.” *AMBIO* 44 (2): 319–31. https://doi.org/10.1007/s13280-014-0624-6.

Smith, Douglas R., Kevin W. King, and Mark R. Williams. 2015. “What Is Causing the Harmful Algal Blooms in Lake Erie?” *Journal of Soil and Water Conservation* 70 (2): 27A-29A. https://doi.org/10.2489/jswc.70.2.27A.

Smith, Val H., Samantha B. Joye, and Robert W. Howarth. 2006. “Eutrophication of Freshwater and Marine Ecosystems.” *Limnology and Oceanography* 51 (1part2): 351–55. https://doi.org/10.4319/lo.2006.51.1_part_2.0351.

Smith, Val H., and David W. Schindler. 2009. “Eutrophication Science: Where Do We Go from Here?” *Trends in Ecology & Evolution* 24 (4): 201–7. https://doi.org/10.1016/j.tree.2008.11.009.

Smucker, Nathan J., Anne Kuhn, Michael A. Charpentier, Carlos J. Cruz-Quinones, Colleen M. Elonen, Sarah B. Whorley, Terri M. Jicha, Jonathan R. Serbst, Brian H. Hill, and John D. Wehr. 2016. “Quantifying Urban Watershed Stressor Gradients and Evaluating How Different Land Cover Datasets Affect Stream Management.” *Environmental Management* 57 (3): 683–95. https://doi.org/10.1007/s00267-015-0629-3.

Solomatine, Dimitri P., and Avi Ostfeld. 2008. “Data-Driven Modelling: Some Past Experiences and New Approaches.” *Journal of Hydroinformatics* 10 (1): 3–22. https://doi.org/10.2166/hydro.2008.015.

Sprague, Lori, Robert Hirsch, and Brent Aulenbach. 2011. “Nitrate in the Mississippi River and Its Tributaries, 1980 to 2008: Are We Making Progress?” *Environmental Science & Technology* 45 (August): 7209–16. https://doi.org/10.1021/es201221s.

Stine, Robert A. 1995. “Graphical Interpretation of Variance Inflation Factors.” *The American Statistician* 49 (1): 53–56. https://doi.org/10.1080/00031305.1995.10476113.

Tabbara, Hadi. 2003. “Phosphorus Loss to Runoff Water Twenty-Four Hours after Application of Liquid Swine Manure or Fertilizer.” *Journal of Environmental Quality* 32 (3): 1044–52. https://doi.org/10.2134/jeq2003.1044.

Thieu, Vincent, Gilles Billen, and Josette Garnier. 2009. “Nutrient Transfer in Three Contrasting NW European Watersheds: The Seine, Somme, and Scheldt Rivers. A Comparative Application of the Seneque/Riverstrahler Model.” *Water Research* 43 (6): 1740–54. https://doi.org/10.1016/j.watres.2009.01.014.

Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading, Massachusetts: Addison-Wesley.

Tyralis, Hristos, Georgia Papacharalampous, and Andreas Langousis. 2019. “A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources.” *Water* 11 (5): 910. https://doi.org/10.3390/w11050910.

US EPA. 2018. “U.S. Action Plan for Lake Erie.” Overviews and Factsheets. Great Lakes. https://www.epa.gov/glwqa/us-action-plan-lake-erie.

Van Meter, K J, and N B Basu. 2017. “Time Lags in Watershed-Scale Nutrient Transport: An Exploration of Dominant Controls.” *Environmental Research Letters* 12 (8): 084017. https://doi.org/10.1088/1748-9326/aa7bf4.

Wagener, Thorsten, Murugesu Sivapalan, Peter A. Troch, Brian L. McGlynn, Ciaran J. Harman, Hoshin V. Gupta, Praveen Kumar, P. Suresh C. Rao, Nandita B. Basu, and Jennifer S. Wilson. 2010. “The Future of Hydrology: An Evolving Science for a Changing World.” *Water Resources Research* 46 (5). https://doi.org/10.1029/2009WR008906.

"Water Quality Portal." 2020. National Water Quality Monitoring Council. https://www.waterqualitydata.us/portal/.

Whitall, David, Brad Hendrickson, and Hans Paerl. 2003. "Importance of Atmospherically Deposited Nitrogen to the Annual Nitrogen Budget of the Neuse River Estuary, North Carolina." *Environment International*, Future Directions in Air Quality Research : Ecological,Atmospheric,Regulatory/Policy/Economic, and Educational Issues, 29 (2): 393–99. https://doi.org/10.1016/S0160-4120(02)00175-7.

Withers, Paul J. A., Stephen D. Clay, and Victor G. Breeze. 2001. "Phosphorus Transfer in Runoff Following Application of Fertilizer, Manure, and Sewage Sludge." *Journal of Environmental Quality* 30 (1): 180–88. https://doi.org/10.2134/jeq2001.301180x.

Withers, Paul J. A., Colin Neal, Helen P. Jarvie, and Donnacha G. Doody. 2014. "Agriculture and Eutrophication: Where Do We Go from Here?" *Sustainability* 6 (9): 5853–75. https://doi.org/10.3390/su6095853.

Wolock, David, Thomas Winter, and Gerard Mcmahon. 2004. "Delineation and Evaluation of Hydrologic-Landscape Regions in the United States Using Geographic Information System Tools and Multivariate Statistical Analyses." *Environmental Management* 34 Suppl 1 (February): S71-88. https://doi.org/10.1007/s00267-003-5077-9.

Zhang, Q., D. C. Brady, and W. P. Ball. 2013. "Long-Term Seasonal Trends of Nitrogen, Phosphorus, and Suspended Sediment Load from the Non-Tidal Susquehanna River Basin to Chesapeake Bay." *Science of The Total Environment* 452–453 (May): 208–21. https://doi.org/10.1016/j.scitotenv.2013.02.012.

Zhang, Qian, Ciaran J. Harman, and William P. Ball. 2016. "An Improved Method for Interpretation of Riverine Concentration-Discharge Relationships Indicates Long-Term Shifts in Reservoir Sediment Trapping." *Geophysical Research Letters* 43 (19): 10,215-10,224. https://doi.org/10.1002/2016GL069945.

Ziegler, Andreas, and Inke R. König. 2014. "Mining Data with Random Forests: Current Options for Real-World Applications." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4 (1): 55–63. https://doi.org/10.1002/widm.1114.

Appendix

Table A.1 – Sources of spatial data used to determine drivers of nutrient water quality within the Great Lakes drainage basin, adapted from Chowdury (2018).

Data Variable(s)	Title	Spatial Coverage	Author(s)/Organization
Land-use (Agriculture, Forested, Urban, Wetlands, etc.)	Annual Crop Inventory (2015)	Ontario	Agriculture and Agri-Food Canada
Land-use (Agriculture, Forested, Urban, Wetlands, etc.)	National Land Cover Database (2011)	U.S.	Multi-Resolution Land Characteristics (MRLC) Consortium
Soil Texture (Percent Sand, Silt, and Clay)	Harmonized World Soil Database	Global	Food and Agriculture Organization of the United Nations (FAO), International Institute for Applied Systems Analysis (IIASA), ISRIC-World Soil Information, Institute of Soil Science – Chinese Academy of Sciences (ISSCAS), Joint Research Centre of the European Commission (JRC)
Soil Texture (Percent Sand, Silt and Clay)	Detailed Soil Survey (DSS)	Ontario	National Soil Database (NSDB)
Soil Texture (Percent Sand, Silt and Clay)	Area- and Depth-Weighted Averages of Selected SSURGO Variables for the Conterminous United States and District of Columbia	U.S.	Michael E. Wieczorek, USGS-WRD MDWSC, Geographer
Tile Drainage Percentages	Tile Drainage Area shapefile	Ontario	Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA)
Tile Drainage Percentages	Tile Drainage Area shapefile	U.S.	USDA, NASS, 2012 Census of Agriculture
Climate (Precipitation and Temperature)	Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas (1970-2000)	Global	Fick, S.E. and R.J. Hijmans, 2017
Slope	Ontario Flow Assessment Tool (OFAT)	Ontario	Ministry of Natural Resources and Forestry
Slope	Hydrologic Landscape Regions of the US	U.S.	Wolock, D.M., Thomas, C.W., Gerard, M.

Population Density	Ontario Population 2011 Census data	Ontario	UWaterloo Geospatial Center Library
Population Density	Sub-County 2010 Census data	U.S.	United States Census Bureau
Cattle, Chicken and Pig Density	Gridded Livestock of the World v 2.01	Global	FAO Robinson, T.P., Wint G.R.W., Conchedda G., Van Boeckel T.P., Ercoli V., Palamara E., Cinardi G., D'Aietti L., Hay S.I., and Gilbert M. (2014)

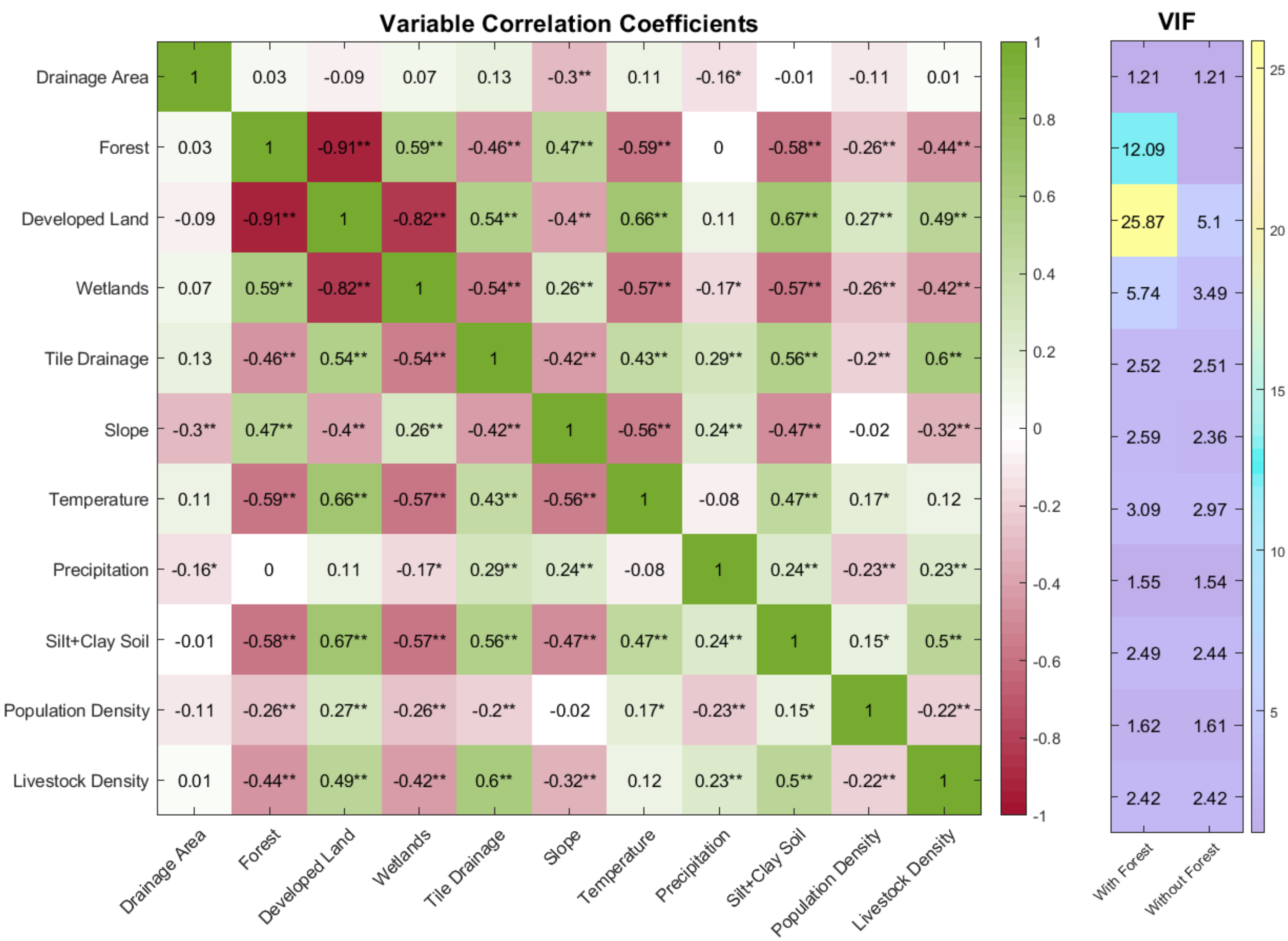

Figure A.1 – Correlation coefficients (left) and Variance Inflation Factors (right) for all preliminary variables assessed. VIFs are shown for all variables (first column), and all variables except forested land use (second column). Large change in VIFs after removal of forested land use shows the significant multicollinearity of forested land use amongst all variables

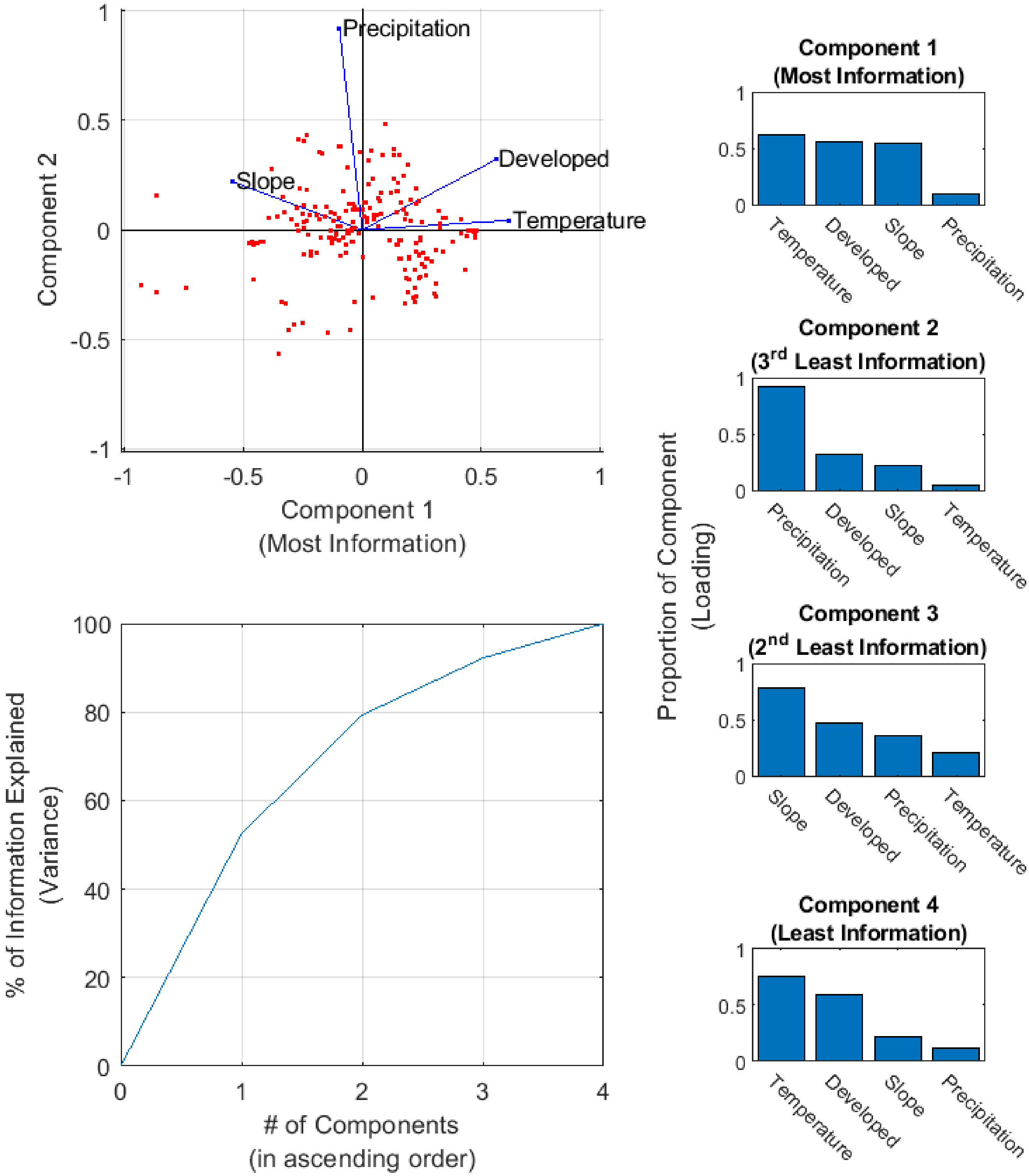

Figure A.2 – PCA of select variables assessed for redundant information. Figures show biplots for variables of first two components (top left), cumulative information explained by components in ascending order (bottom left), and variable proportions (loadings) for each of the four components (right).

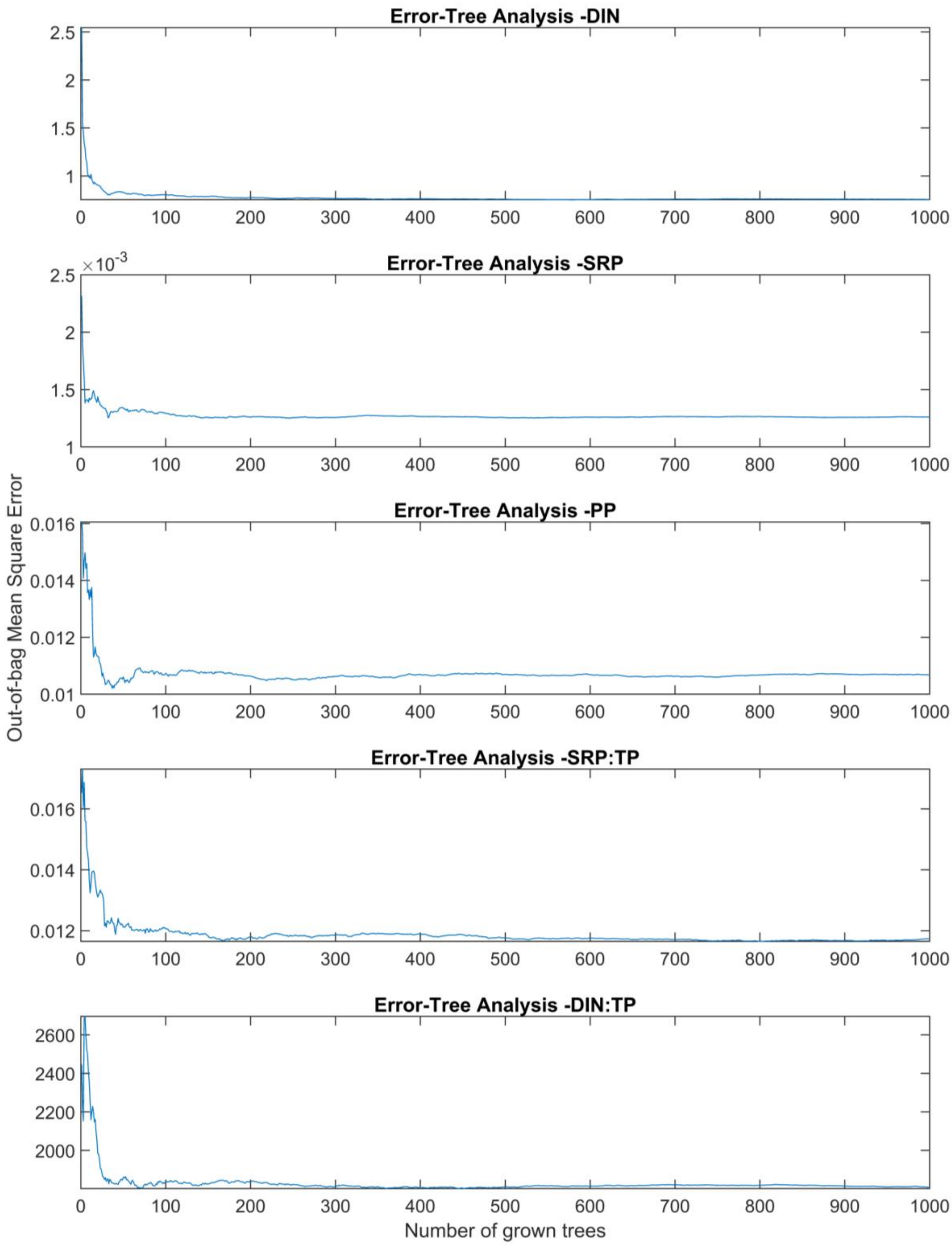

Figure A.3 – Error convergence for RF models grown over 1000 decision trees.

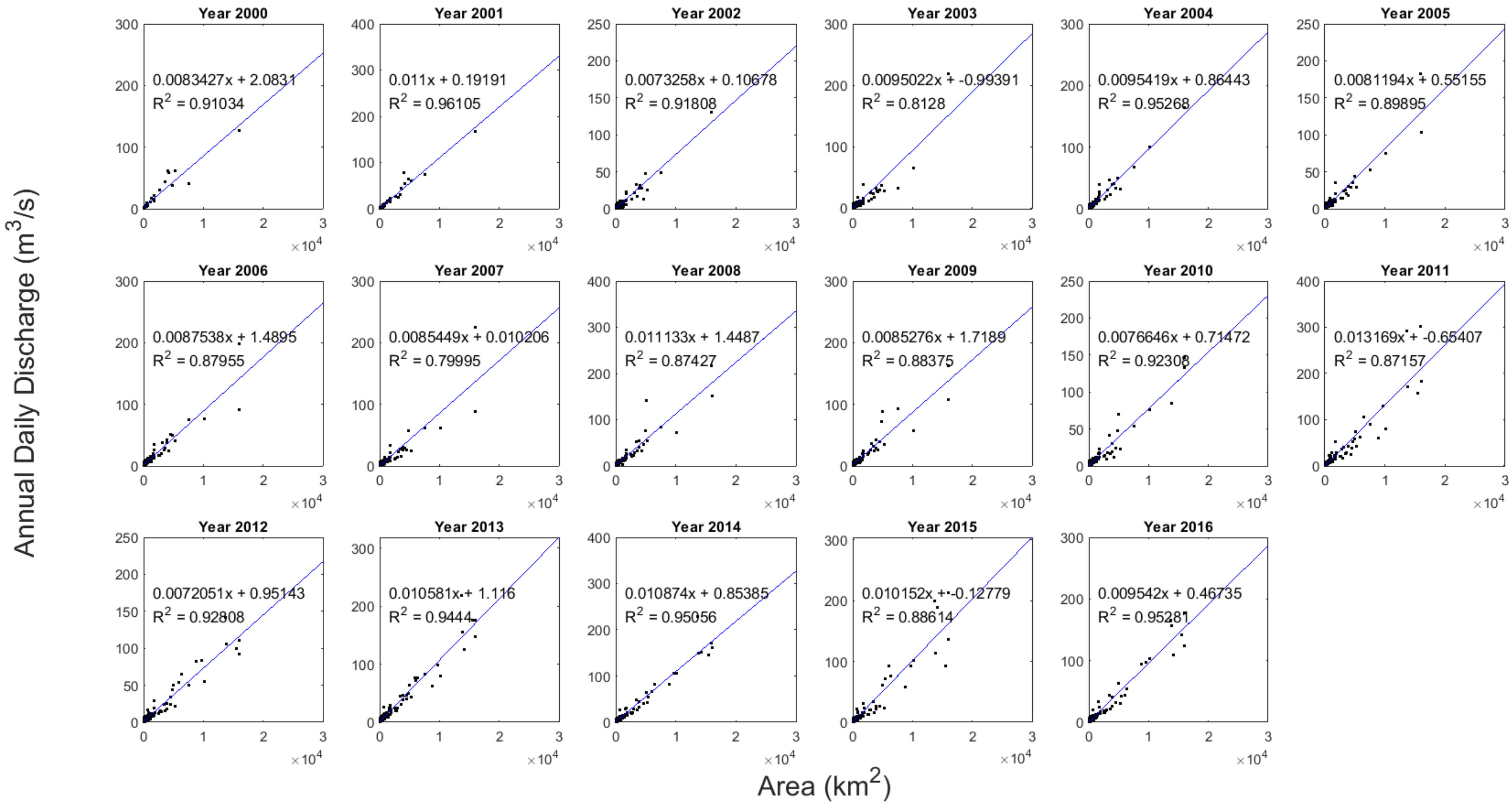

Figure A.4 – Linear area-discharge relationships for years 2000 to 2016 used to calculatate loadings in the Laurentian Great Lakes drainage basin.

Table A.2 – Model training and development results for DIN concentrations. R^2 values are given for Trained and OOB datasets. Percent Bias and Mean Square Error (MSE) are given for OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation. Models in bold indicate variables used for final RF model application.

Variables	Trained R^2	OOB R^2	OOB Percent Bias	OOB MSE
Tile Drainage,	0.82	0.72	96	1.132
Tile Drainage, Developed Land	0.86	0.76	59	0.985
Tile Drainage, Developed Land, Wetlands	0.87	0.75	55	1.022
Tile Drainage, Developed Land, Wetlands, Livestock Density	0.91	0.80	45	0.809
Tile Drainage, Developed Land, Wetlands, Livestock Density, Population Density	**0.91**	**0.81**	**42**	**0.755**
Tile Drainage, Developed Land, Wetlands, Livestock Density, Population Density, Silt+Clay	0.91	0.81	45	0.775

Table A.3 – Model training and development results for SRP concentrations. R^2 values are given for Trained and OOB datasets. Percent Bias and Mean Square Error (MSE) are given for OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation. Models in bold indicate variables used for final RF model application.

Variables	Trained R^2	OOB R^2	OOB Percent Bias	OOB MSE
Developed Land,	0.68	0.46	69	0.00153
Developed Land, Tile Drainage,	0.73	0.53	73	0.00130
Developed Land, Tile Drainage, Silt+Clay	**0.74**	**0.54**	**85**	**0.00126**
Developed Land, Tile Drainage, Silt+Clay, Wetlands	0.76	0.51	77	0.00137
Developed Land, Tile Drainage, Silt+Clay, Wetlands, Livestock Density	0.77	0.53	80	0.00131
Developed Land, Tile Drainage, Silt+Clay, Wetlands, Livestock Density, Population Density	0.78	0.53	71	0.00128

Table A.4 – Model training and development results for PP concentrations. R^2 values are given for Trained and OOB datasets. Percent Bias and Mean Square Error (MSE) are given for OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation. Models in bold indicate variables used for final RF model application.

Variables	Trained R^2	OOB R^2	OOB Percent Bias	OOB MSE
Silt+Clay,	0.46	0.12	132	0.0145
Silt+Clay, Wetlands	0.57	0.21	111	0.0125
Silt+Clay, Wetlands, Tile Drainage	0.61	0.27	100	0.0114
Silt+Clay, Wetlands, Tile Drainage, Developed Land	0.64	0.21	100	0.0125
Silt+Clay, Wetlands, Tile Drainage, Developed Land, Population Density	0.69	0.31	94	0.0108
Silt+Clay, Wetlands, Tile Drainage, Developed Land, Population Density, Livestock Density	**0.70**	**0.31**	**94**	**0.0107**

Table A.5 – Model training and development results for SRP:TP ratios. R^2 values are given for Trained and OOB datasets. Percent Bias and Mean Square Error (MSE) are given for OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation. Models in bold indicate variables used for final RF model application.

Variables	Trained R^2	OOB R^2	OOB Percent Bias	OOB MSE
Developed Land	0.52	0.14	28	0.0159
Developed Land, Livestock Density,	0.61	0.22	29	0.0139
Developed Land, Livestock Density, Tile Drainage	0.70	0.32	26	0.0120
Developed Land, Livestock Density, Tile Drainage, Silt+Clay	0.76	0.32	26	0.0119
Developed Land, Livestock Density, Tile Drainage, Silt+Clay, Population Density	0.76	0.32	26	0.0119
Developed Land, Livestock Density, Tile Drainage, Silt+Clay, Population Density, Wetlands	**0.78**	**0.34**	**27**	**0.0117**

Table A.6 – Model training and development results for DIN:TP ratios. R^2 values are given for Trained and OOB datasets. Percent Bias and Mean Square Error (MSE) are given for OOB datasets. Trained data refers to all data used in model development and provides a metric of biased performance. OOB refers to "Out of Bag" data and indicates data not used in model development and provides a metric of unbiased performance and validation. Models in bold indicate variables used for final RF model application

Variables	Trained R^2	OOB R^2	OOB Percent Bias	OOB MSE
Livestock Density,	0.42	0.07	113	2328
Livestock Density, Developed Land	0.62	0.19	110	1880
Livestock Density, Developed Land, Silt+Clay	0.65	0.22	114	1814
Livestock Density, Developed Land, Silt+Clay, Wetlands	**0.71**	**0.22**	**116**	**1808**
Livestock Density, Developed Land, Silt+Clay, Wetlands, Tile Drainage	0.72	0.22	120	1815
Livestock Density, Developed Land, Silt+Clay, Wetlands, Tile Drainage, Population Density	0.74	0.21	123	1817

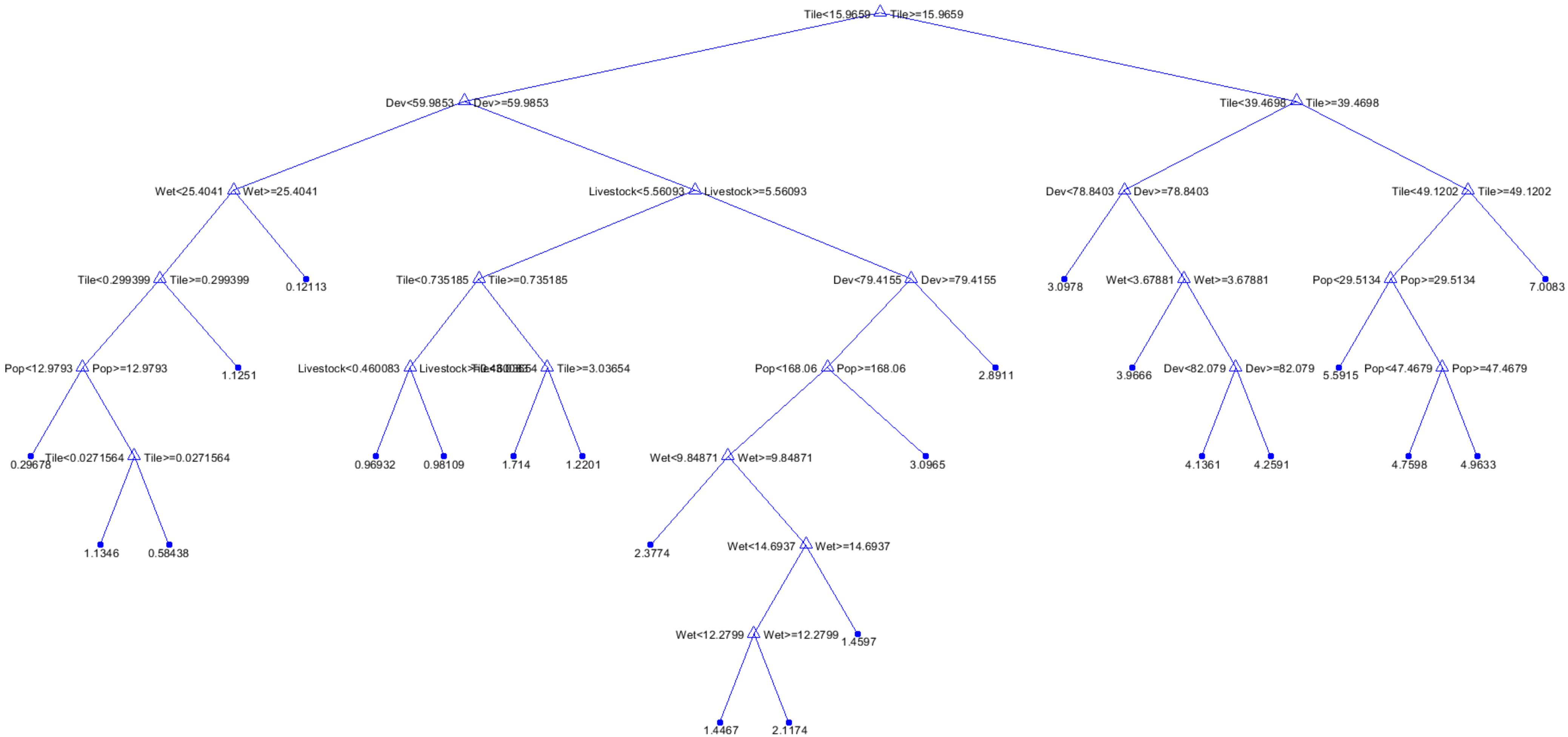

Figure A.5 – Example of a decision tree used for RF models of DIN FWC.

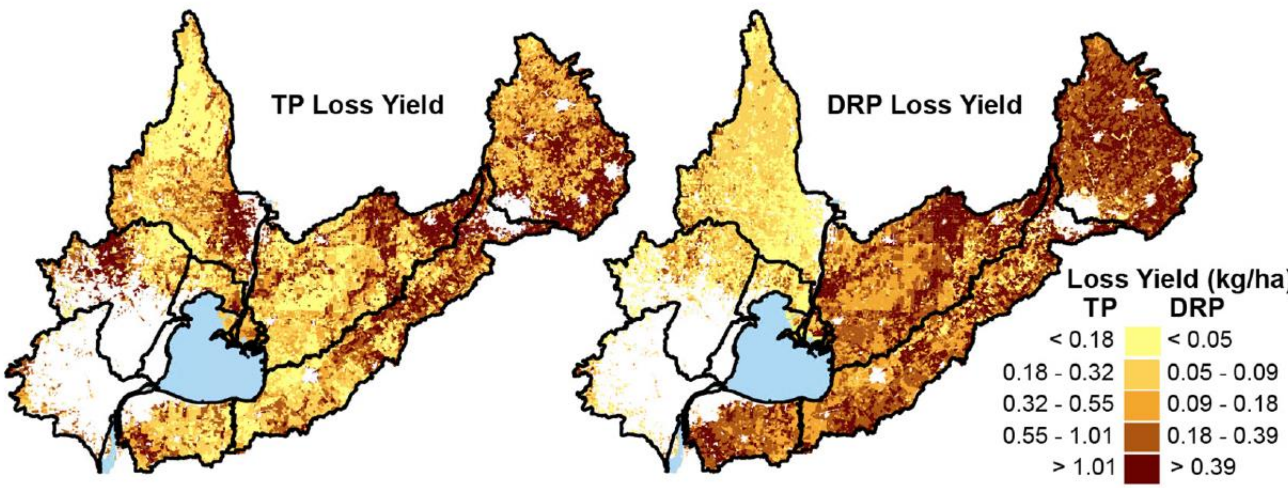

Figure A.6 – Modelled TP and SRP (DRP) loss yields from SWAT modelling performed by Scavia et al. (Scavia et al. 2019). Results show stark contrast of lower to higher losses when comparing the eastern watersheds U.S. to the western watersheds of Canada, respectively.

Table A.7 – Modelled annual DIN basin loads to the Laurentian Great Lakes in tonnes/year. SE refers to the standard error of predictions.

	Lake Erie		**Lake Huron**		**Lake Michigan**		**Lake Ontario**		**Lake Superior**	
Year	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**
2000	72,531	8,711	182,199	17,996	137,569	17,074	77,993	11,566	172,052	17,841
2001	88,970	10,888	228,502	23,074	172,053	21,880	96,793	14,713	217,195	22,986
2002	59,197	7,246	152,082	15,361	114,507	14,566	64,413	9,794	144,568	15,304
2003	73,830	9,141	192,064	19,640	144,388	18,618	80,871	12,470	183,237	19,616
2004	78,997	9,606	201,418	20,193	151,797	19,151	85,614	12,908	191,043	20,085
2005	66,740	8,131	170,546	17,136	128,495	16,251	72,415	10,945	161,867	17,052
2006	74,289	8,975	187,980	18,703	141,804	17,741	80,190	11,988	177,898	18,572
2007	68,750	8,426	176,865	17,889	133,145	16,962	74,862	11,400	168,194	17,827
2008	93,317	11,311	237,025	23,672	178,717	22,453	100,931	15,153	224,564	23,526
2009	73,068	8,807	184,353	18,289	139,119	17,350	78,751	11,735	174,313	18,149
2010	63,508	7,721	161,884	16,225	122,006	15,388	68,818	10,373	153,534	16,138
2011	104,207	12,832	269,501	27,400	202,749	25,978	113,789	17,431	256,680	27,335
2012	60,430	7,323	153,463	15,324	115,713	14,535	65,353	9,810	145,386	15,229
2013	88,010	10,689	224,075	22,432	168,902	21,275	95,309	14,347	212,443	22,305
2014	89,683	10,917	228,935	22,978	172,509	21,792	97,255	14,683	217,218	22,861
2015	81,316	9,979	209,487	21,218	157,675	20,118	88,610	13,515	199,298	21,150
2016	77,962	9,514	199,597	20,092	150,347	19,053	84,675	12,825	189,544	20,002

Table A.8 – Modelled annual SRP basin loads to the Laurentian Great Lakes in tonnes/year. SE refers to the standard error of predictions.

	Lake Erie		Lake Huron		Lake Michigan		Lake Ontario		Lake Superior	
Year	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**
2000	2,402	319	3,944	475	3,186	421	2,054	320	3,311	447
2001	2,944	404	4,942	608	3,983	538	2,555	407	4,177	575
2002	1,959	269	3,289	405	2,651	358	1,700	271	2,780	383
2003	2,442	342	4,152	517	3,341	458	2,137	346	3,522	491
2004	2,615	355	4,358	532	3,514	471	2,258	357	3,675	503
2005	2,209	301	3,689	452	2,975	400	1,910	303	3,113	427
2006	2,460	330	4,068	493	3,284	437	2,113	331	3,423	465
2007	2,275	313	3,825	471	3,082	417	1,976	316	3,234	446
2008	3,089	417	5,129	624	4,138	553	2,661	419	4,320	589
2009	2,420	323	3,990	483	3,222	427	2,075	324	3,354	455
2010	2,102	285	3,502	428	2,825	379	1,815	287	2,953	404
2011	3,447	478	5,827	722	4,692	639	3,006	483	4,935	684
2012	2,001	270	3,321	404	2,679	358	1,723	271	2,797	381
2013	2,913	394	4,848	592	3,911	524	2,513	397	4,086	559
2014	2,968	403	4,953	606	3,994	536	2,565	406	4,178	572
2015	2,691	371	4,530	559	3,649	495	2,340	374	3,832	529
2016	2,580	352	4,317	530	3,480	469	2,234	355	3,645	501

Table A.9 – Modelled annual TP basin loads to the Laurentian Great Lakes in tonnes/year. SE refers to the standard error of predictions.

	Lake Erie		Lake Huron		Lake Michigan		Lake Ontario		Lake Superior	
Year	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**	**Annual Load**	**SE**
2000	6,779	451	11,431	672	8,871	595	5,792	452	9,517	632
2001	8,306	571	14,327	860	11,088	761	7,205	576	12,016	814
2002	5,527	380	9,535	572	7,379	507	4,795	383	7,998	542
2003	6,888	483	12,038	731	9,302	648	6,027	489	10,138	694
2004	7,378	502	12,631	753	9,785	667	6,368	505	10,568	711
2005	6,233	425	10,695	639	8,282	566	5,387	428	8,954	604
2006	6,941	467	11,791	698	9,142	618	5,960	469	9,841	658
2007	6,418	442	11,089	667	8,580	590	5,573	446	9,305	631
2008	8,717	589	14,866	883	11,521	782	7,504	593	12,422	833
2009	6,828	457	11,565	683	8,970	604	5,851	459	9,642	643
2010	5,932	403	10,152	605	7,864	536	5,118	406	8,493	571
2011	9,726	676	16,894	1,021	13,064	904	8,476	683	14,201	967
2012	5,645	382	9,625	572	7,459	506	4,859	384	8,042	539
2013	8,221	558	14,053	837	10,888	741	7,088	561	11,752	790
2014	8,376	571	14,357	857	11,119	759	7,234	575	12,016	810
2015	7,591	524	13,133	791	10,161	700	6,598	529	11,026	749
2016	7,280	498	12,516	749	9,690	663	6,301	502	10,486	708

[this page has been intentionally left blank]