Nutrient Load Modeling in the U.S.: Novel Considerations for Future Management

by

Lorrayne Miralha Marins da Silva

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2021 by the
Graduate Supervisory Committee:

Rebecca L. Muenich, Chair
Margaret Garcia
Tianfang Xu
Soe W. Myint

ARIZONA STATE UNIVERSITY

December 2021

ABSTRACT

Approximately 71% of the great lakes, lakes, reservoirs, and ponds, together with 51% of rivers and streams assessed in the US are impaired or threatened by pollution or do not meet the minimum water quality requirements. Pathogens, sediments, and nutrients are leading causes of impairment, with agriculture being a top source of pollution. Agricultural pollution has become a global concern overtaking urban contamination as the major factor of inland and coastal waters degradation in many parts of the world. High-yielding crop production has been achieved by the intensive use of inorganic fertilizers that are mainly composed of Nitrogen (N) and Phosphorus (P). N and P are essential nutrients for ecosystem structure, processes, and functions. However, N and P in excess can be problematic to the environment. One of the major impacts of the increasing amount of these nutrients in the environment is the global expansion of harmful algal blooms (HABs). Major agricultural nutrient pollution sources and climate change can exacerbate these risks. This dissertation aims to guide future policies to mitigate issues linked to excess nutrient loads in the U.S. by evaluating the impact of climate change on nutrient loads and assessing the environmental impact as well as the spatial patterns of one of the major agricultural sources of nutrient pollution - Concentrated Animal Feeding Operations (CAFOs). Specifically, I first investigated the impact of bias correction techniques when modeling mid-century nutrient loads in a watershed heavily impacted by CAFOs. Second, I evaluated the role of CAFOs in land use change and subsequent environmental degradation of the surrounding environment. Finally, I assessed the spatial organization of CAFOs and its links to water quality conditions. The findings revealed unique insights for future nutrient management strategies in the U.S.

DEDICATION

“To Jorge Antonio Marins da Silva (1957-2019), my father, the biggest fan of my science.”

ACKNOWLEDGMENTS

If you are reading this, you know I am grateful for you.

First, I would like to thank my advisor Dr. Rebecca L. Muenich for all the support and trust throughout my Ph.D. program. I am grateful for all her feedback, advice, and for always making sure she had time to answer my questions. She was fundamental to the success of this work and to my growth as an independent researcher. I also would like to thank my committee members Dr. Margaret Garcia, Dr. Tianfang Xu, and Dr. Soe W. Myint for all their assertive comments, support, and attention. I would like to thank my collaborators Dr. Danica Schaffer-Smith, Dr. Allison Steiner, Dr. Margaret Kalcic, Dr. Donald Scavia, Dr. Reagan Errera, and Dr. James Hood for contributing towards my growth as a scientist. They always brought up insightful questions that made me evolve as researcher and I will always remember every suggestion, questions, and comments each one of you provided. Thanks to all SSEBE professors, Dr. Vivoni, Dr. Wang, Dr. Mascaro, Dr. Boyer, and others, I learned significantly from each one of you. Thanks to all SSEBE staff, principally Melanie Duran, Jeff Ahlstrom, Stan Klonowski, and Kayla Vestal for making every step easier during my time at ASU. I will always be grateful for having met each one of you.

Thanks also to Arizona State University, GPSA, American Association of University Women, and AZ Water Association for providing the funding that made this degree possible.

I am thankful for the friends I made at ASU. Thanks to my lab mate Eleanor Rauh for becoming my Arizona guide, my ears when I needed someone to talk to, and for being by my side in both good and hard moments. My labmates Noah Rudko, Suraya Sidique, Tannis Breure, Maria Menchu, William Krukowski, and others for being enthusiastic about research with me and for always helping me look at the positive side of things. Thanks to Xin Guan, Peiyuan Li, Chenghao Wang, and Adil Mounir for being the best officemates someone could ask for. Thanks to my ASU friends Chinmay Nair, Adham Abozaeid, and Faizan Ahmad for making sure we would always have great board game nights and potlucks.

I would like to thank Hisham Jashami for being the arms and legs I needed in the difficult moments. Hisham was the one that saw the potential in me and pushed me towards graduate school. I partially owe this degree to him, and I am grateful to have had him by my side these past years.

Most importantly, I thank my family for providing me with the most they could to make me succeed, for making themselves present, and for celebrating with me, no matter the distance. Thanks to my mother, Deyse Miralha, for her unlimited strength, patience, love, support, and kindness. She is selfless, an extraordinary person, and I owe her every breath I take. I will never forget her comforting calls principally in moments when I was pushing myself to work harder but needed a break. A mother knows, they say, and she feels when her daughter needs some joy or break. She is the blessing I was able to have throughout my life. Thanks to my brother, Lawson Miralha, for accepting a younger sister with joy, for caring, for being the strength we needed when things got ugly. He has celebrated each step I take and has always assured me that I not only have a brother but also a friend for life. Thanks to my sister-in-law, Jessica Lee, for accepting us as her family, for celebrating with me like a true sister, for being kind, and for all her great suggestions.

Thanks to the one, who used to call me his little cute girl, who taught me how to ride a bike and drive a manual car, who taught me music and loved singing with me, who also fought for me countless times, who worked in several jobs to make sure my brother and I would have a proper childhood, who taught us nothing comes easy and that being hard-working, determined, and honest was all he wanted us to be. One of his last sentences was: “I am proud of both of you, for have been extraordinary children, and for achieving and becoming more than a father could have asked for.” He was a natural scientist, curious, spontaneous, ingenious, and the cheerful spirit I used to share some of my Ph.D. proposal ideas with. He even read a hydrology book for further discussion. Unfortunately, he won’t be able to print this dissertation as he used to print everything he could about his children. But wherever you are I am sure you are happy today. To Jorge Antonio Marins da Silva, my father, who was the best father someone could have ever had.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

APPENDIX Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

1.1. Background

According to the United States Environmental Protection Agency (USEPA), approximately 71% of the lakes, reservoirs and ponds, together with 51% of rivers and streams assessed in the US are impaired or threatened by pollution or do not meet the minimum water quality requirements (USEPA, 2020). Pathogens, sediments, and nutrients are leading causes of impairment, with agriculture being the top source of pollution. Agricultural pollution has become a global concern overtaking urban contamination as the major factor of inland and coastal waters degradation (FAO, 2017). High-yielding crop production has been achieved by the intensive use of inorganic fertilizers that are mainly composed by Nitrogen (N) and Phosphorus (P) (Gellings & Parmenter, 2016; Kondraju & Rajan, 2019; Stewart et al., 2019). N and P are essential nutrients for ecosystem processes and functions. Both N and P are necessary in processes linked to plant and animal survival, such as photosynthesis, cell growth, metabolism, and protein synthesis (Chapin et al., 2002; Guignard et al., 2017; Razaq et al., 2017). However, these nutrients in excess can be problematic to the environment.

The atmosphere is composed by 78% of N gas, but N in gaseous form cannot be directly used by most organisms (Stein & Klotz, 2016). To become biologically available, N fixation occurs through specialized bacteria or industrial processes, by ionizing phenomena such as lightning (Delwiche, 1970). Prior to large-scale synthetic fertilizers

production in the 1950s, the amount of N removed from the atmosphere via natural fixation was approximately balanced with the amount of N returning to the atmosphere via natural denitrifying processes. With industrial and agricultural advancements, specifically the development of the Haber-Bosch process, fixation exceeded denitrification leading to N excess and accumulation in the terrestrial and marine environment. N combines with molecules of oxygen and hydrogen available in the ecosystem, forming compounds such as ammonia and ammonium (NH_3 and NH_4) as well as nitrite and nitrate (NO_2 and NO_3) (Delwiche, 1970). When in excess, N may deplete the availability of oxygen, principally in aquatic ecosystems, resulting in algae blooms and intensified biological activity that kill oxygen-dependent organisms (Galloway et al., 2008). This biological productivity is also contingent upon the availability of P. Continental bedrocks are the primary natural source of P in the environment. P becomes naturally available for plant uptake via weathering processes and returns to the soil through decay or litterfall (Ruttenberg, 2003). The mining of phosphate rocks increased significantly to produce synthetic agricultural P fertilizer. In addition to the increase in the application of synthetic fertilizers, deforestation as well as urban and industrial waste disposal have enhanced phosphorus transport from terrestrial to aquatic ecosystems (Ruttenberg, 2003). The increasing amount of these nutrients, both N and P, in the environment has resulted in the impairment of drinking water sources and in the global expansion of harmful algal blooms (HABs), which produce toxins that lead to species mortality and threaten human health (Gill et al., 2018; Michalak, 2016; O'Neil et al., 2012).

Climate change can exacerbate these risks. Predicted increases in temperature and changes in rainfall can favor the emergence of HABs (Chapra et al., 2017; Moore et al., 2008). Extreme rainfall events can lead to excess runoff of N and P to rivers and streams, worsening the aquatic environment conditions. There is still a need, however, to understand the impact of climate change when quantifying future nutrient loads and HABs as studies show conflicting results as to whether nutrient loads will increase or decrease in a future climate (Kalcic et al., 2019).

The intensive livestock production industry is a major source of nutrients and other kinds of pollutants. This sector contributes to major environmental problems worldwide, including loss of biodiversity and water quality depletion (USEPA, 2004). The demand for meat products has increased with population growth, which has caused expansion and intensification of the livestock industry (Godfray et al., 2018). This industry requires water and feed concentrates, directly pressuring the environment through land clearing for feed production and water usage for animal needs (Austin et al., 2017; NRDC, 2019). These practices impact water resources as the water used returns to the environment in form of liquid manure, slurry, and wastewater (FAO, 2017). When the livestock production system becomes concentrated, known in the U.S. as concentrated animal feeding operations (CAFOs) that are regulated by the federal government since 1970s (USEPA, 2003), manure production tends to exceed the holding capacity of the surrounding ecosystem potentially degrading the environment and increasing water pollution (Oun et al., 2014). Moreover, manure is not only composed of nutrients such as N and P, but also contains bacteria, antibiotics, and heavy metals that can significantly impact the ecosystem and human health

(Burkholder et al., 2007; Harun & Ogneva-Himmelberger, 2013; Brands, 2014; Ogneva-Himmelberger et al., 2015; Guidry et al., 2018) . There remains a large gap in our understanding and tracking of this source of pollution so that improvements can be made to current policies and management practices.

CAFOs, primarily their manure lagoons and feedlots, are considered point source dischargers under the Clean Water Act (USEPA, 2018); however, these operations usually spread manures stored in lagoons or pits onto surrounding agricultural fields, making the manures susceptible to runoff as rainfall occurs (NRDC, 2019). Under this scenario, the land-applied manure is typically considered a non-point source pollutant exempted from most Clean Water Act regulations as agricultural stormwater. To monitor and regulate this pollution source to prevent water body contamination, the EPA requires CAFOs to obtain a National Pollution Discharge Elimination System (NPDES) permit to legally operate (USEPA, 2003; Centner, 2011). However, this regulation may only apply to CAFOs that have the potential to discharge manure directly in waterways excluding the manure spreading operations, depending on the state's interpretation of the federal rules (Duke Bass Connections, 2016). Due to high hauling prices and the fact these operations tend to cluster over space (Freeze & Sommerfeldt, 1985; Copeland, 2010; Yang et al., 2016), manure likely concentrates in nutrient-exhausted areas leading to environmental problems (Kellogg et al., 2000). There is still a need to observe spatially and temporally how CAFOs are impacting their surrounding environment based on the distance manure has been transported over the years. Although studies have reported that CAFO activities have the potential to deplete their surrounding natural resources and pollute water bodies (Martin et

al., 2018), no studies have yet reported the environmental trade-offs and conditions surrounding these operations.

One of the targets established by the Sustainable Development Goals (SDG Target 6.3) is a goal to "improve water quality by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials, halving the proportion of untreated wastewater and substantially increasing recycling and safe reuse globally" (United Nations, 2016). CAFOs produce significant amounts of manure that are often discharged in waterbodies or land applied without preliminary treatment (De Vries et al., 2010). Therefore, the spatial aggregation pattern of these operations may be an indication of the degree of pollution or degradation in their surroundings (Hribar, 2010; Yang et al., 2016; Martin et al., 2018; Baek & Smith, 2019). The identification of an optimal spatial organization for these operations together with best management practices could guide policy makers, improve the sustainability of the food production system, offer new opportunities for resource recovery, and ameliorate water quality conditions aligning with the SDG goal. Yet, no study has holistically reported if the clustering of these operations influences the degradation outcomes as well as if changes in the current policies could lead to better environmental conditions.

1.2. Climate, Bias Correction, and Nutrient Load Prediction

Water bodies in the U.S. and around the globe have increasingly exhibited symptoms of eutrophication, including hypoxia and harmful algal blooms (HABs) (Taranu et al., 2015; Wurtsbaugh et al., 2019). Eutrophication is driven by excess nutrients, N and P,

which are often attributed to intensified watershed activities such as agriculture and urban wastewater discharge (Stow et al., 2020; H. Xu et al., 2018). For example, research shows agricultural practices drive the increase of P loads in Lake Erie, Michigan USA (Michalak et al., 2013; Scavia et al., 2014) and of N loads in the Gulf of Mexico (Rabalais et al., 2009). Manure N and P is another factor leading to areas with nutrient surplus in Asia, South and North America, and Europe (Keplinger & Hauck, 2004; Li et al., 2007; Menzi et al., 2010; Sönmez et al., 2016). Moreover, habitat alteration, such as wetland removal and deforestation, has increased N and P loads in Klamath Lake, Oregon USA (Paerl et al., 2018). Climate change exacerbates these impacts. Warming temperatures have been linked to the earlier emergence of HABs in shallow lakes located in distinct regions of South America and Europe (Kosten et al., 2012), and precipitation changes are largely controlling the Gulf of Mexico load variability (Donner & Scavia, 2007). Changes in rainfall variability may also increase bloom emergence in Australia (O'Neil et al., 2012). While previous studies have suggested that impacts are likely to become more intense, there remains a gap in our understanding of how a changing climate may impact future nutrient loads.

Watershed models are often used to evaluate the effectiveness of different agricultural best management practices (BMPs) as well as to explain factors (i.e., land use practice, climate change) that reduce (or drive) excess nutrients (Daloğlu et al., 2012; Kalcic, Chaubey, et al., 2015; Kalcic, Frankenberger, et al., 2015; Muenich et al., 2016). The Soil and Watershed Assessment Tool (SWAT) is an example of a watershed-scale model that has been widely applied to investigate hydrological responses and nutrient load changes

based on land management, soil, climate, and topographic characteristics (Gassman et al., 2007; Douglas-Mankin et al., 2010; Dagnew et al., 2019). A recent study used multiple SWAT models to understand the impact of different management strategies that would decrease future P loads in Lake Erie, providing insights into how to achieve nutrient load reduction goals (Scavia et al., 2017). While management practices have been identified, the authors did not consider their performance under a changed climate. Little is known about how future climate may influence changes to lake nutrient loads. On the one hand, studies have suggested that future nutrient loads may decrease in western Lake Erie, associated with increased temperatures and evapotranspiration (Kalcic et al., 2019), as well as greater plant uptakes under higher CO2 concentrations (Culbertson et al., 2016). On the other hand, studies indicate that nutrient loads might increase under future climate regimes as changing annual temperature and precipitation increase flow rates and runoff in the Maumee watershed (Verma et al., 2015). Therefore, uncertainties remain surrounding the impacts of future land-atmosphere interactions on regional nutrient loads.

When considering the implementation of future climate model simulations in watershed modeling, a common approach includes bias correction (Hakala et al., 2019; Cannon et al., 2020) which is applied to correct systematic errors in climate model outputs, primarily due to the difference in scale (Hakala et al., 2019), that serve as inputs to watershed models. For example, most climate models are developed at either a global (GCMs; 100 – 500 km grid resolution) or regional (RCMs; 10-50 km grid resolution) scale (Leung et al., 2003), whereas watershed models operate with a full domain at typically an 8-digit hydrologic unit code scale (total model area ~ 100 –1500 km2) or smaller (NRCS,

2020). Studies have investigated the impacts of bias correcting climate data for input to hydrological models in watershed processes, such as changes in flow, evapotranspiration, and rainfall (Teutschbein & Seibert, 2012; Cannon et al., 2015; Meyer et al., 2019). For instance, comparing two bias correction techniques, Quantile Mapping (QM) and Quantile Delta Mapping (QDM), Cannon et al (2015) found that QM can inflate the magnitude of rainfall extremes in Canada. Another study compared 6 bias correction approaches (e.g., linear scaling (LS), distribution mapping (DM), QM) when correcting RCM precipitation and temperature outputs to simulate streamflow in an arid catchment in China. The authors found that while all methods improved raw RCM-simulated precipitation, there were variations in their corrected statistics, which resulted in differences when simulating flow via SWAT. Among all methods, LS overestimated flow by 100% in the simulation period and had the greatest bias (Fang et al., 2015).

However, there is a lack of understanding on how bias correction influences the prediction of nutrient loads. Kalcic et al. (2019) explored the influence of climate change on nutrient load predictions in the Maumee River watershed using mid-century climate projections from one global and four regional climate models as input to SWAT. While their results suggested nutrient load would decrease under a future climate, they did not bias correct the climate model outputs. Other studies have applied bias correction to predict future nutrient loads but did not investigate how different bias correction techniques would impact modelled nutrient outputs (Culbertson et al., 2016; Mehan et al., 2019). For Example, Teutschbein et al. (2017) modeled changes in total inorganic N across 19 sites in Sweden. The authors combined GCM and RCM outputs and used only one bias correction

method (i.e., distribution scaling) to correct for systematic bias. Results indicated significant increases in total inorganic N loads in the future (Teutschbein et al., 2017). A recent study in Germany on the effects of bias correcting climate model outputs to predict future changes in discharge suggested that predicted future flow differs by bias correction method (Wörner et al., 2019). This investigation has yet to be done for nutrient load models. There remains a need to understand how bias correcting climate model outputs (i.e., precipitation and temperature) influence nutrient load model outputs to advance this field of study and guide decision-making processes.

1.3. CAFOs and the Environment

The animal protein and products sector is growing and intensifying faster than crop production in almost all countries (FAO, 2017). It was projected that the global demand for meat and milk in 2050 will grow by 73 and 58 percent, respectively (FAO, 2013). Consequently, large animal feeding operations have arisen as a cost-effective, efficient alternative to provide food for a growing population that demands more animal products (McDonald et al., 2020). When these facilities, which vary by animal type (Ribaudo et al., 2003), confine animals for more than 45 days per year, don't have an actively growing crop on the facility, and meet a threshold number of animals, they are regulated as CAFOs in the U.S. Other parts of the world may describe these operations differently, with terms such as feedlots or intensive livestock operations. The growth of these operations worldwide is likely to exacerbate its linked environmental impacts. Gas, aerosol, and particulate emissions from CAFO activities can affect air-quality conditions (Wilson & Serre, 2007;

Ogneva-Himmelberger et al., 2015). Other pollutants from CAFOs include nutrients, sediments, pathogens, heavy metals, hormones, antibiotics, and ammonia in manures (Mole, 2013; Scanes, 2018; Kronberg & Ryschawy, 2019; Randad et al., 2019). While the significant amounts of manure CAFOs generate can be a valuable source of nutrients such as P, N, and organic carbon for agriculture (Yan et al., 2017), the excess accumulation of manure can impact the environment, for example (Burkholder et al., 2007; Garcia et al., 2019; Muenich et al., 2016; Tullo et al., 2019; Miralha, Muenich, Scavia, et al., 2021; Raff & Meyer, 2021a), and lead to long-term, large-scale ecological degradation (Qi et al., 2017).

A previous study of CAFOs in Michigan found that operators were primarily applying manure within regulatory limits, yet were often applying amounts above crop nutrient needs, indicating that manure is treated as a waste product they need to dispose of, rather than as a valuable fertilizer for crops (Long et al., 2018). Most states regulate CAFO manure applications based on environmental risk or nutrient limits based on crop needs, but these inefficient applications likely occur due to high costs associated with manure hauling (Sims et al., 2005; Centner, 2012), lack of markets due to nutrient ratio variability, as well as presence of pathogens, metals, antibiotics and other undesirable qualities (Ribaudo et al., 2003; Keplinger & Hauck, 2006; Liu et al., 2018; Sonne et al., 2019; Pepper et al., 2019). To meet state and federal waste management regulations, CAFO operators must either acquire or rent sufficient land nearby for applications, or transfer or sell manure to nearby farms. The need for more land for manure applications has the potential to change land use patterns and environmental conditions in the vicinity of CAFOs.

To mitigate potential CAFO impacts, it is necessary to better understand how these operations affect the environment over time and space. Whether changes in local land use patterns and environmental quality are intensified due to presence of CAFOs has not previously been explored at large spatial scales. Land conversions due to human influence (e.g., loss of forest or wetlands to cropland and shrubland) often result in degraded quality relative to the prior reference condition. These physical landscape changes impact ecological function, which may in turn influence environmental quality. Satellite remote sensing products have been used extensively to quantify spatiotemporal changes in environmental conditions (Ishtiaque et al., 2016a; Estoque et al., 2018a). Several products offer advantages for environmental change analysis (e.g., evapotranspiration (ET), land surface temperature (LST)), including repeated monitoring over large areas, relatively long historical records, and free data access (LP DAAC, 2019). A variety of indices derived from satellite remote sensing data have been used to measure degradation in previous studies (Eckert et al., 2015; Alatorre et al., 2016a; C. Wang & Myint, 2016; Estoque et al., 2017), but these have not yet been specifically applied to assess the localized impacts of CAFOs.

1.4. CAFOs Spatial Organization and Water Quality

Standard operational activities, accidental discharges such as CAFO lagoon ruptures (Mallin & Cahoon, 2003), as well as leaching and runoff of contaminants from manure-applied fields into nearby water bodies (Sousan et al., 2021) can lead to severe ecological

degradation. As CAFOs tend to cluster in space for production and logistical purpose, their associated environmental impacts may also intensify.

Clustering has become the cornerstone of efficiency and economic success in industrialized production and development (Porter, 1998; Ayres & Ayres, 2002; Deutz & Gibbs, 2008). Although it is argued that industrial clusters may provide environmental benefits (Lifset & Graedel, 2002), few studies in industrial ecology have investigated the relationship between clustering of point sources and the impact in the surrounding environment (Kennedy, 1999; Lall & Mengistae, 2005; Anh et al., 2011; Yoon & Nadvi, 2018). Regardless of the specific industry, spatially concentrated production can potentially generate negative environmental outcomes. Like many other industries, CAFOs tend to cluster in space, which may lead to similar drawbacks. The clustering of animal agriculture may lead to cumulative adverse environmental effects because individual production facilities add animals over time, whether or not they own sufficient cropland to handle the additional manure produced (Thurow & Thompson, 1998). However, studies have not explored the spatial clustering of CAFOs and its relationship with the conditions of the surrounding environment. Daniels (1997) determined that for cluster zoning to be an effective method of land protection, planners must delineate reasonable densities for development such that the carrying capacity of the local environment is not exceeded, especially in a watershed. Excessive density, inappropriate locations, a combination of these two, or unreasonable expectations about what cluster development can do, are the major four potential clustering abuses of a given operation (Daniels, 1997; Porter, 1998). These abuses may apply to the spread of CAFOs both in the

U.S. and worldwide, principally when it comes to socio-environmental impacts as well as to the argument of economic and production benefits from the clustering of these operations. While watersheds with high concentration of CAFOs are potentially at higher risk of degradation than others (Martin et al., 2018; C. Brown et al., 2020), studies have yet to investigate if the spatial organization of CAFOs is a characteristic of higher negative environmental outcomes. Specifically, this spatial clustering pattern has yet to be investigated by water quality studies that account for agricultural source of pollution.

Studies have previously explored the impact of CAFOs on water quality. For instance, concentrations of nitrate, ammonium, total nitrogen (TN), and other ions were higher in CAFO-impacted streams than in control streams where CAFOs are not present (Harden, 2015). Additionally, overaccumulation of total phosphorus (TP) by manure application has allowed for export into surface and subsurface water bodies impacting the aquatic ecosystem (Withers & Jarvie, 2008). For instance, the Neuse River watershed in North Carolina received approximately 41,000 metric tons of N and 16,000 metric tons of P from CAFO waste in the 2000s (Glasgow & Burkholder, 2000). These additions of large amounts of nutrients often introduce eutrophic/hypoxic conditions to surface water bodies (Muenich et al., 2016; Tullo et al., 2019; Miralha, Muenich, Scavia, et al., 2021; Raff & Meyer, 2021b), an impact that must be prevented and monitored to avoid permanent ecological damages. Assessing the spatial organization of CAFOs and its links to water quality conditions could improve our understanding on how to better manage and regulate these facilities to prevent future permanent environmental damages.

1.5. Research Objectives and Dissertation Structure

Overall, understanding the factors influencing future N and P concentration both in terms of climate change modeling approaches and sources of pollution – such as CAFOs - may help optimize future nutrient reduction strategies and policies. **This dissertation aims to guide future policies to mitigate issues linked to excess nutrient loads in the U.S. by evaluating the impact of climate change on nutrient loads and assessing one of the major agricultural sources of nutrients -CAFOs (Figure 1)**. Chapter 1 has described the problems addressed in this dissertation in addition to motivations, and provided the literature review per each question evaluated in the following chapters. In Chapter 2, the goal was to identify the impact of bias correction (BC) methods for climate-modeled data in the prediction of nutrient loads using an existing Soil and Water Assessment Tool (SWAT) model for the Maumee Watershed in the Western Lake Erie Basin, an agricultural watershed with current CAFO activities. *Chapter 1 - section1.2. and Chapter 2 were published in Science of the Total Environment as Miralha et al 2021.* In the third chapter, I investigated whether regulated CAFOs drive land use change over time and space as well as if their presence degrades local environmental conditions; specifically, I used MODIS land products from 2000 to 2018 to examine these questions in Michigan and North Carolina. *Chapter 1 -section 1.3 and Chapter 3 were published in Science of the Total Environment as Miralha et al 2021*. In the fourth chapter, I aimed to investigate if CAFO-clustered watersheds are likely to present higher concentrations of TP and TN than CAFO-dispersed basins. For CAFOs with animal number data available, I additionally evaluated the relationship between the spatial autocorrelation of the number of animals per CAFO

and water quality conditions to determine if both spatial clustering and size of CAFOs differently impact water quality. *Chapter 4 is currently in preparation for a peer-reviewed journal submission.* Chapter 5 summarizes the takeaways of this work and the future research direction focused on developing a national database of AFOs.

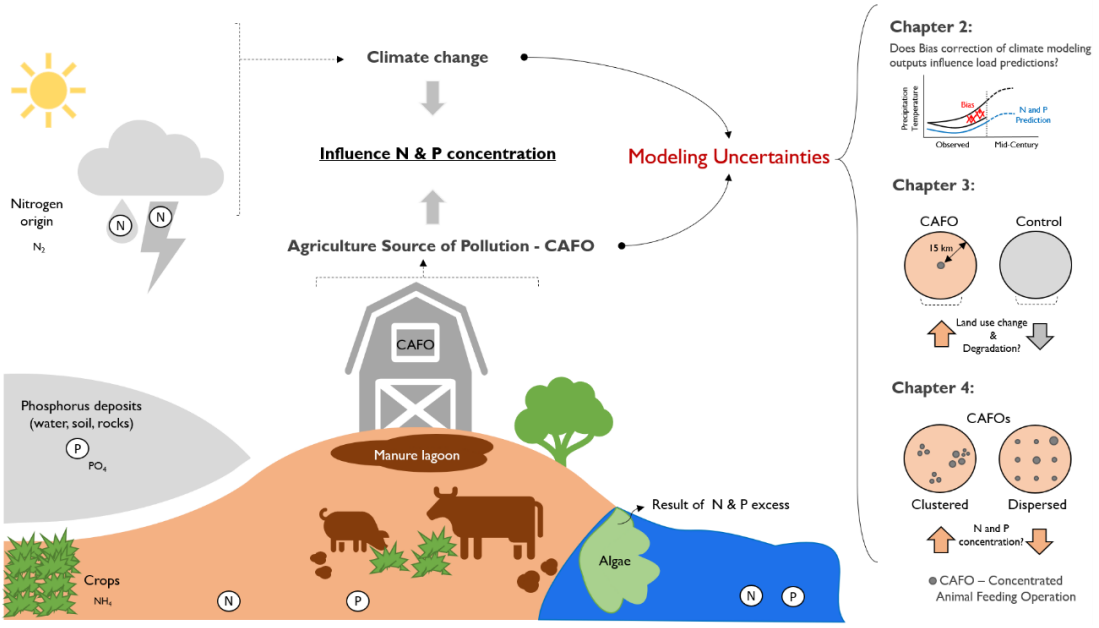


Figure 1. Dissertation sketch on how climate and the agriculture source of pollution (CAFO) ideas interconnect. This research aims to bring insights for future nutrient modeling approaches and nutrient management policies.

CHAPTER 2

BIAS CORRECTION OF CLIMATE MODEL OUTPUTS INFLUENCES WATERSHED MODEL NUTRIENT LOAD PREDICTIONS

As discussed in chapter one (section 1.2), bias correction is a common technique used by modelers. However, with the variety of methods existent in the literature it is still unclear how the choice of a bias correction method impacts nutrient load prediction. This study evaluates the impacts of bias correcting precipitation and temperature from 4 GCMs from the Climate Model Intercomparison Project version 5 (CMIP5) model database (Taylor et al., 2012) on SWAT outputs of nutrient loads to the Western Lake Erie Basin (WLEB) from the Maumee River, Michigan USA. I modeled observed dissolved reactive P (DRP), total P (TP), and total N (TN) loads from 1985 to 1999, which were compared to SWAT loads driven by not bias-corrected and bias-corrected GCM outputs. The climate model scenario and bias correction method that most closely matched the observations were selected and used for the evaluation of hydrological processes and absolute nutrient load changes in the mid-century (2051-2065). Results from this study advance knowledge on the impacts of bias-corrected climate model outputs driving watershed models. It may also serve as a guide for watershed modelers worldwide on how bias correction techniques and climate model choices may influence modeled nutrient outputs as well as serve as a cautionary tale for the decision-making process of nutrient load management strategies.

2.1. METHODS

2.1.1. Study area

The study site is the Maumee River basin located in parts of Michigan, Indiana, and Ohio, US (Figure 2). This agriculturally dominated basin has flat topography characteristics as well as poorly drained soils, which requires installation of subsurface tile drains to ensure viability of agriculture. The basin is predominantly corn, soybean, and wheat crops, with a mix of dairy, swine, and poultry livestock operations. The watershed receives about 984 mm of rainfall on average, with an annual average temperature of 10°C (Williams & King, 2020). Rainfall patterns have changed over the past 30 to 40 years influencing tributary discharge and P delivery in Lake Erie. Across the basin, annual rainfall increased by 102 mm from 1975 to 2017, with increases in intense rainfall occurring primarily during spring and summer seasons (Williams & King, 2020).

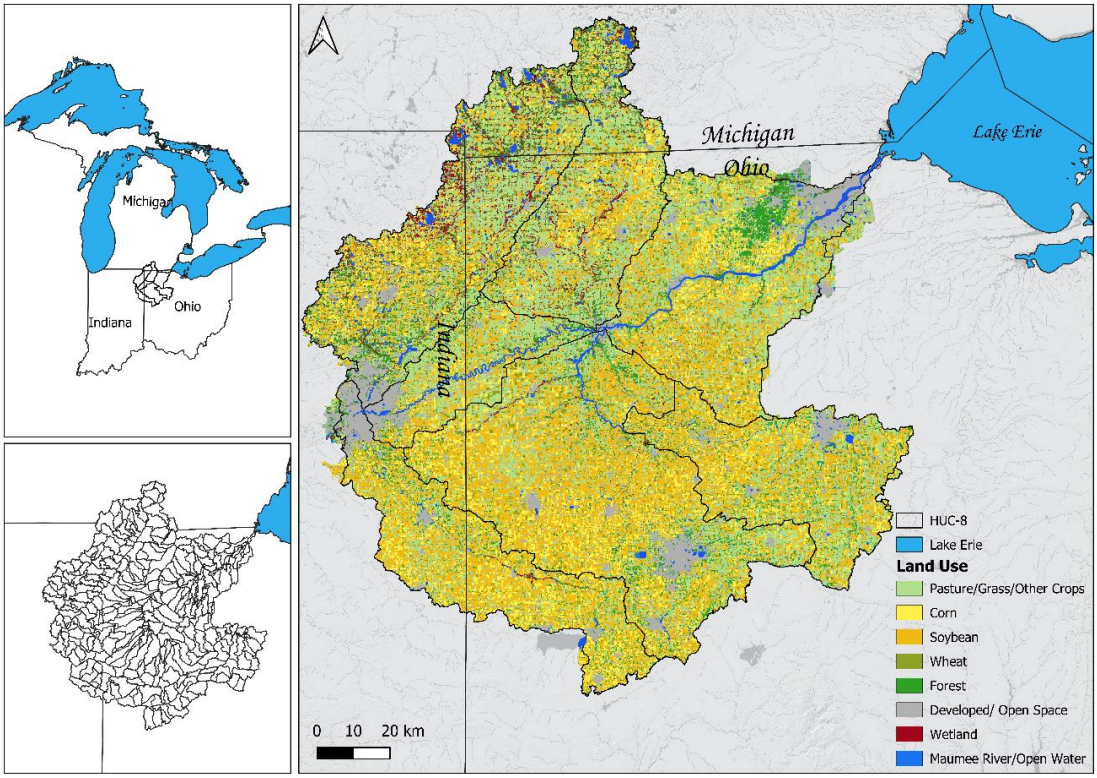


Figure 2. Maumee River basin: location (top left panel), SWAT sub-basins (bottom-left panel), land use characteristics, and main HUC-8 sub-basins (main panel).

2.1.2. SWAT model characteristics

SWAT is a semi-distributed hydrologic and water quality model driven by daily time-scale inputs including precipitation, temperature, solar radiation, relative humidity and wind speed. This is a widely used model in the investigation of water resource issues associated with land use management, especially in agricultural lands (Tan et al., 2020). The model used in this study was calibrated and validated for the Maumee River watershed considering field-scale best management practices (BMPs), and fertilizer applications (Apostel et al., 2021). This SWAT model sub-divided the 358 sub-basins into 24,256 hydrologic response units (HRUs), each with ~70 acres area on average. The model

accounted for unique combinations of crop rotations simplified to meet watershed crop characteristics; both inorganic and organic (including manure) fertilizer application; tile drainage, tillage management, and other management practices such as fertilizer incorporation and buffer strips; the increasing prevalence of soil stratification in the region; labile P calculation for each HRU; and new snow parameters based on data from NOAA's Global Historical Climatology Network. The main objective of the calibration was ensuring the model simulates reasonable streamflow, nutrient, and sediment load values. The calibration was performed for 2005-2015 period, with validation for 2000-2004. This model was run using the SWAT 2012 revision 635, which includes modification on the movement of soluble P through subsurface tile drains. The daily model calibration and validation metric results for TP, DRP and, TN are in the supplementary information (Appendix A- SI 1.1. Table 1).

2.1.3. Climate variables and Climate Models Selection

Four global climate models from the Coupled Model Intercomparison Project Phase 5 (CMIP5) were used in this study: (1) CCSM4, (2) MPI-ESM-MR (i.e. MPI), (3) CNRM-CM5 (i.e. CNRM), and (4) IPSL-CM5A-MR (i.e. IPSL). Though the models have distinct atmospheric grid resolutions, CCSM4 has the highest number of grid points within the study area compared to the other three models (Table 1). In this study, I used daily temporal resolution outputs with the Representative Concentration Pathway 8.5 (RCP8.5) as the future emissions scenario (Pachauri et al., 2014).

Table 1. Description of the climate models used in this study.

Model	ID in this study	Description	Atmospheric Grid (lat x lon)	Grid points (lat x lon)
CCSM4	CCSM4	National Center for Atmospheric Research (NCAR) - Community Earth System Model 4	0.94º x 1.25º	4 x 3 = 12
MPI-ESM-MR	MPI	Max Planck Institute for Meteorology - Earth System model running on mixed resolution grid	1.86º x 1.87º	2 x 2 = 4
CNRM-CM5	CNRM	Centre National de Recherches Météorologiques and Centre Européen de Recherche et de Formation Avancée	1.40º x 1.40º	2 x 2 = 4
IPSL-CM5A-MR	IPSL	Institut Pierre-Simon Laplace - Earth System Model for the 5th IPCC report - Mid resolution	1.27º x 2.5º	2 x 2 = 4

These models were selected from a set of 18 CMIP5 models, because they minimized error in historical seasonal-average temperature and precipitation over the WLEB (Appendix A – SI 1.2. Table 2). Overall, the model that provided the closest patterns to gridded observations (Appendix A – SI 1.3. Figure 3) across the temperature and precipitation metrics (the CCSM4 model) was also the one that had a realistic spatial representation of the Great Lakes. From a subset of modeled historical (1980-1999) and future (2046 -2065, RCP 8.5 emission scenario) simulations, we used model precipitation (PCP), maximum temperature (TMAX), and minimum temperature (TMIN) as input to SWAT. I distinguish the SWAT run with climate model outputs from 1980-1999 as the "model-historical simulation", and from 2046-2065 as the "future simulation". A separate simulation using daily observations of PCP, TMAX and TMIN from 71 weather stations for 1980-1999 was also simulated for the historical time period, and this is referred to as

the "model-observed simulation". The "model-observed simulation" is treated as the baseline in this study.

2.1.4. Bias Correction Techniques

Bias correction (BC) was performed with daily climate model output (i.e., PCP, TMAX, TMIN) and equivalent daily observed data for the historical period (1980 to 1999). To bias correct, the observed values from the 71 weather stations were averaged to the nearest grid point of the climate model to ensure the spatial resolution of both the observed climate data and climate model output would be the same. To enable the comparison among BC techniques, we selected four well-established univariate BC methods and three recently developed multivariate BC methods. The four univariate methods are: (1) Delta; (2) Scaling; (3) Empirical Quantile Mapping (EQM); (4) Quantile Delta Mapping (QDM). The *delta* method adds the mean change signal between the simulated climate in the training period (*p)* and the test period (*s)* to the observations (*o)* (Equation 2.1), where the training period considers the data from 1980-1999 and the test period includes 2046- 2065. The delta method is based on a difference for variables such as temperature, while for precipitation we apply a quotient approach.

$$Delta_{TMAX/TMIN} = o + (mean\,(s) - mean\,(p))$$

$$Delta_{PCP} = o * (mean\,(s)/mean\,(p)) \quad (2.1)$$

The *scaling* method scales the simulated variables based on the difference or quotient between the observed and simulated means (Equation 2.2). The difference is based on an additive equation that is usually applied to unbounded variables such as temperature,

while the quotient or multiplicative is applied to variables such as precipitation, so frequency can be preserved.

$$Scaling_{\,TMAX,TMIN;\,additive} = s - mean\,(p) + mean\,(o)$$

$$Scaling_{\,PCP;\,multiplicative} = \frac{s}{mean\,(p)} \ * mean\,(o) \qquad (2.2)$$

EQM consists of calibrating the simulated cumulative distribution function (CDF) by adding the observed quantiles (i.e., mean delta change and the individual delta changes) in the corresponding simulated quantiles. This method adjusts 99 percentiles and linearly interpolates inside this range every two consecutive percentiles (Gutiérrez et al., 2019). EQM also has the option of extrapolation, which enables keeping the extreme values of the distribution; however, this option was not used in this study because of the potential risk of bias in the extremes of the distribution. Details of this method can be found in the documentation of the R package *hyfo-biasCorrect* (Y. Xu, 2020). *QDM* is a quantile mapping method developed to avoid trend deterioration as in the traditional quantile mapping technique (Meyer et al., 2019). First, QDM extracts the climate change trend from the projected future quantiles. Then, the quantile mapping technique is applied to the detrended series. The quantile mapping is based on a transfer function that converts the CDF of the modeled data *m*, to match the CDF of the observed series *o* in a historical period *h* (Equation 2.3).The CDF function is denoted by *F*, and the bias correction of modeled PCP or TMP at time *t* within some projected period *p* is denoted by $\widehat{PCP\ or\ TMP}_{m,p}\ (t)$. This transfer function is only based on historical period information, excluding future model projection relationships (Cannon et al., 2015). Thus, QDM was developed as a solution to preserve model-projected relative changes in quantiles while correcting

systematic biases in the quantiles of modeled series considering the observed values. The QDM transfer function is indicated by Equation 2.4, where the additive sign changes to multiplicative if the variable to be corrected is precipitation. More details on QDM can be found in Cannon et al. (2015).

$$\widehat{PCP\ or\ TMP}_{m,p}\ (t) = F_{o,h}^{-1}\{F_{m,h}\left[PCP\ or\ TMP_{m,p}\ (t)\right]\} \quad (2.3)$$

$$\widehat{PCP\ or\ TMP}_{m,p}\ (t) = F_{o,h}^{-1}\left\{F_{m,p}^{(t)}\left[PCP\ or\ TMP_{m,p}\ (t)\right]\right\} + PCP\ or\ TMP_{m,p}\ (t) - F_{m,h}^{-1}\left\{F_{m,p}^{(t)}\left[PCP\ or\ TMP_{m,p}\ (t)\right]\right\} \quad (2.4)$$

The multivariate bias correction (MBC) methods chosen were: (1) *MBCp* - Multivariate bias correction (Pearson correlation) that matches marginal distributions using QDM and the Pearson correlation dependence structure; (2) *MBCr* - Multivariate bias correction (Spearman rank correlation) that matches marginal distributions using QDM and the Spearman rank correlation dependence structure; and (3) *MBC-N* - Multivariate bias correction (N-pdft) that matches the multivariate distribution using QDM and the N-dimensional probability density function transform (Cannon et al., 2015; Cannon, 2018; Meyer et al., 2019). This method applies an orthogonal rotation to the data before QDM is executed. This rotation exposes QDM to a linear combination of the original variables, so QDM eventually corrects the probability distribution of the rotated data. For this method, the number of iterations used were the default (i.e., 30 iterations).

While the univariate methods are based on mean deviations as well as quantile calculations, the multivariate techniques are based on the correlation or relationship among several variables to correct the variable set as target. MBC analyses require several inputs, but here we aimed to correct PCP based only on TMAX and TMIN inputs. The same was

done to correct TMAX, using only TMIN and PCP inputs, while TMIN was corrected using PCP, and TMAX. I chose this simplified approach as these are typically the variables used in future climate studies in watershed modeling, and often the most readily available across climate models. Two main packages in R studio were used for these analyses: 1) *hyfo*- used to bias correct daily climate data, and 2) MBC. We did not use *delta* from the hyfo package, because it does not consider the quotient approach for precipitation. Plots of monthly precipitation and temperature variables from 1980 to 1999 per scenario (observed, no-bias corrected (historical), and bias corrected) can be found in the supplementary information (Appendix A – SI 1.4. Figure 2).

2.1.5. SWAT Simulations

I ran several SWAT simulations to estimate TN, TP, and DRP loads. One simulation used the observed (i.e., 71 weather stations data) PCP, TMAX, and TMIN from 1980 to 1999 as inputs. I also ran SWAT simulations using PCP, TMAX, and TMIN from the climate models as inputs without bias corrections for both the historical (1980-1999) and future (2051-2065) periods, and then with the 7 methods for bias correct on for both the historical and future periods.

This resulted in 69 SWAT simulations, 28 runs using the 4 climate models and 7 bias-correction methods for each time period, 1 run with observed climate data, 4 runs with observed climate data at each climate model resolution, and 8 model historical and future runs without bias correction. Because the first 5 years from the climate inputs were used as a spin-up period in SWAT, the historical and future comparisons were based on the

period from 1985 to 1999 and from 2051 to 2065. Each historical run was plotted using Quantile-Quantile plots (Q-Q plots) of average monthly values of TN, TP and DRP per year from 1985-1999. The future Q-Q plots were based on the best historical scenarios and served to evaluate absolute nutrient change in the mid-century period from 2051-2065.

2.1.6. Evaluation of Model Fit, Bias Correction Method, and Future Load Impacts

A systematic approach was used to evaluating model fit, bias correction methods, and load impacts (Figure 3). Prior to the modeling analysis, I compared the climate model outputs (i.e., PCP, TMAX, and TMIN) with the mean of the observed climate variables in the basin. This comparison was done prior to bias correction and after bias correcting the climate model product. After that, I ran the SWAT model scenarios. I first evaluated which climate model outputs without bias correction (i.e., model-historical) resulted in the best SWAT-modeled loads when compared to the model-observed loads (SWAT driven by the observed climate in 71 stations) from 1985 to 1999. Second, I evaluated which bias correction technique provided the best SWAT-generated loads using Q-Q plots and three metrics: (1) Percent Bias (P_{BIAS}), (2) Nash-Sutcliffe Efficiency (NSE), and (3) R^2. P_{BIAS} is defined as the deviation between the simulated and observed values. Negative P_{BIAS} indicates underestimation, and positive value indicates overestimation. NSE ranges from -∞ to 1, with 1 reflecting a perfect match to observations. R^2 is the square of correlation coefficient *r*, which is a measure of the linear relationship between observed and simulated values. Average monthly values of TN, TP, and DRP were the inputs for P_{BIAS}, NSE, and R^2 calculations. I calculate these metrics based on SWAT simulations with the observed

climate matched to each climate model resolution (Appendix A – 1.3. Figure 1) for optimal comparison with the climate-BC scenarios. These calculations were based on the Monthly-averaged loads modeled between 1985 and 1999 (i.e., historical period). To evaluate if direction and magnitude of change vary with bias correction, I calculated absolute change between the nutrient loads in the historical period (1985-1999) and loads for the mid-century (2051-2060) based on the SWAT outputs with and without bias correction. The replication of this methodological approach in other regions will depend on the availability of observed water quality data and a calibrated watershed model.

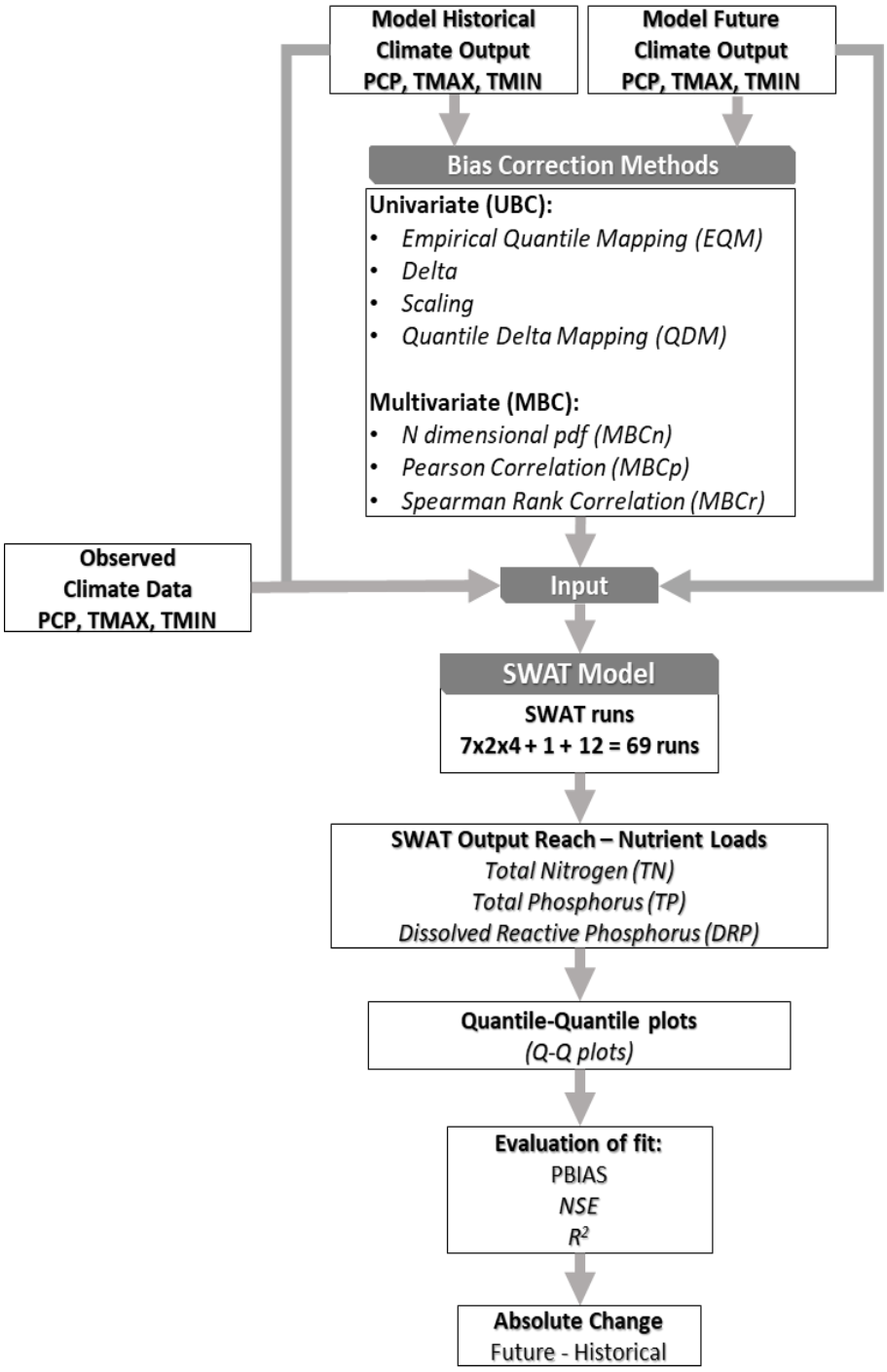


Figure 3. Flow diagram illustrating all the analysis as well as the climate model outputs used in this study.

2.2. RESULTS

2.2.1. Assessment of Climate data without Bias-Correction

Comparing averages of observed PCP, TMAX, and TMIN values with the climate model output (i.e., historical with no BC) revealed that bias differs among climate models (Figure 4). PCP bias generally increased with the magnitude of the observed values, where MPI and IPSL overestimated observed PCP by 38% and 20%, respectively. Due to better alignment with observed values, bias in CNRM PCP was the lowest among the model ensemble (7%, deviation from the observed climate (σ) = 5.03) while its standard deviation was higher than CCSM4 (P_{BIAS} =11%; σ = 3.98), which performed better for larger values. The MPI model had the highest degree of deviation from the observed values (σ = 6.2), followed by IPSL (σ = 5.38). For temperature, on average, the climate models appeared to predict the historical TMAX well, with deviation from the observed values occurring most in the extremes of the distribution. Overall, TMAX and TMIN were underestimated in all models except for CCSM4 TMAX (Figure 4). CCSM4 TMIN also show strong underestimation with PBIAS = -61.1%. IPSL ranked first (P_{BIAS} = -1.5%) when predicting historical TMAX, followed by MPI (-4%), CNRM (-5%), and CCSM4 (5%). Prior to bias correction, CCSM4 and CNRM historical PCP outputs fit the observed PCP better, while MPI and IPSL were best in fitting observed temperature. Underestimation in CNRM was the most severe, with the model ranking last in performance (-74% bias for TMIN). MPI and IPSL had the lowest bias (both ~ -57%), but this was still high compared to PCP and TMAX. CNRM had the overall lowest bias for PCP, while MPI and IPSL were best for TMIN and TMAX, respectively.

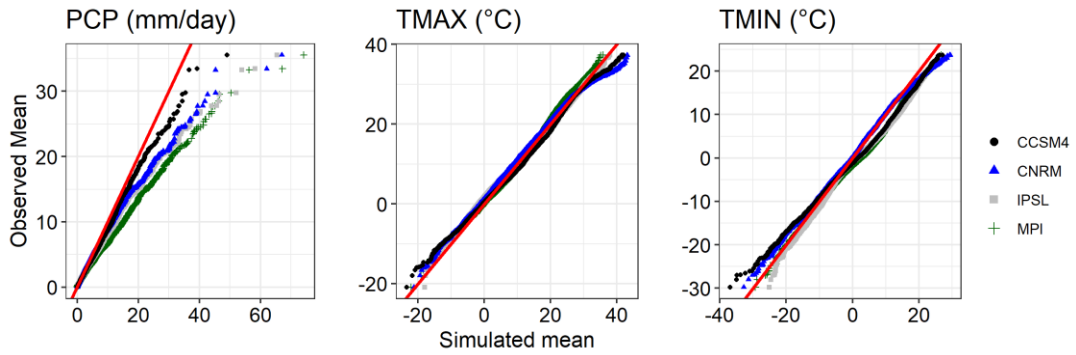


Figure 4. Comparison of the four climate models to the observed (red 1:1 line) precipitation (PCP), maximum and minimum temperature (TMAX and TMIN) values. Climate model performance varies by model and by variable in the historical period (1980-1999) for the Maumee watershed.

2.2.2. Bias Correction of Climate output

The application of BC generally moved the historical no-BC values closer to observed values (Figure 5). The Delta method was able to correct most of the bias between the historical and the observed PCP. In general, overestimates increased with increasing PCP values as well as with extremes for TMAX and TMIN. For TMAX and TMIN, the mismatch mostly occurred when using the scaling approaches. Results varied for each BC technique by model and climate output. For example, while EQM performed well when correcting PCP outputs from CNRM, IPSL, and MPI, the technique's fit worsened when correcting CCSM4 PCP.

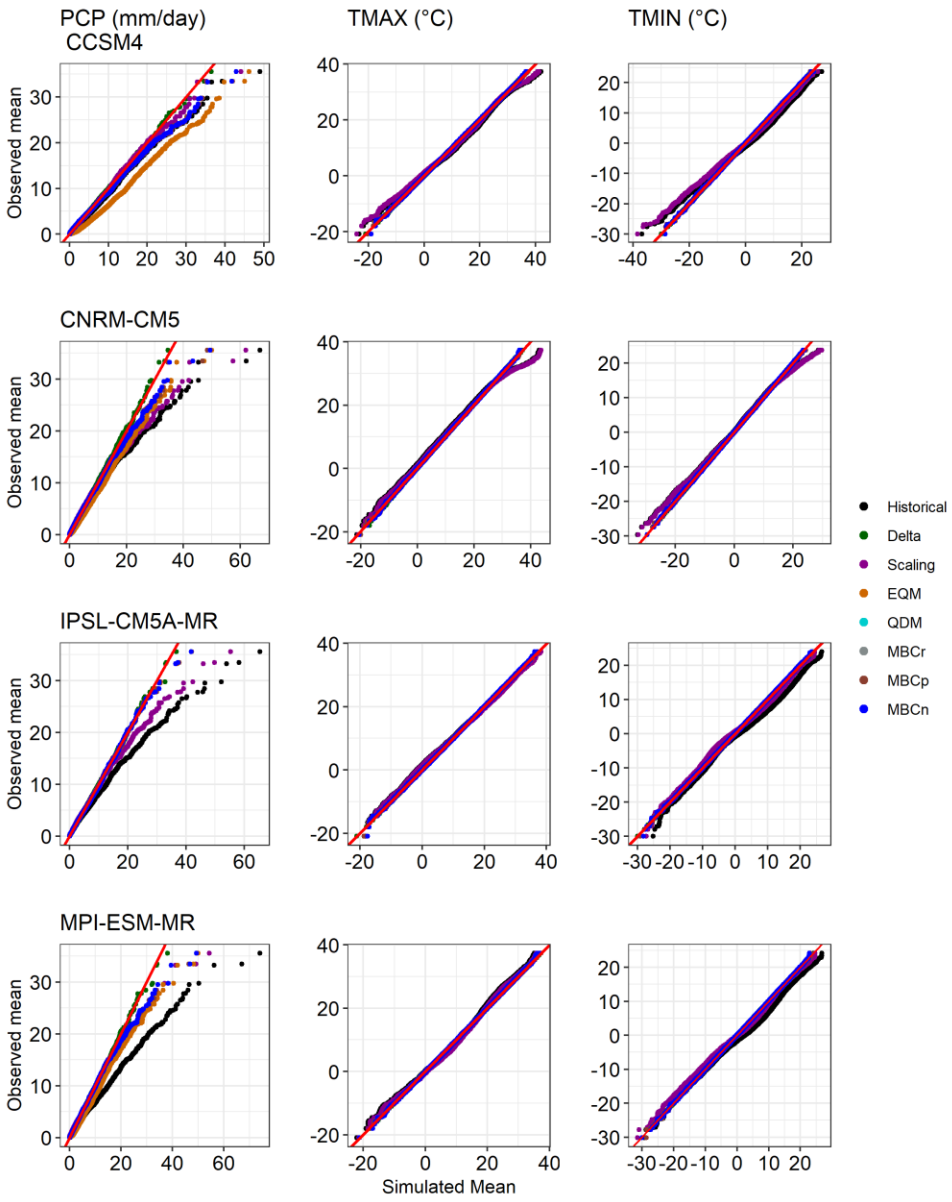


Figure 5. Quantile-Quantile (Q-Q) plots between observed climate data (y-axis – solid black line) and each bias corrected method and climate model output (x-axis). PCP, TMAX, and TMIN are the overall average among the stations (for the observed) and the climate model grid (for the not bias-corrected and bias-corrected scenarios). The *y*-axis and solid red lines indicate the observed climate data from the 71 weather stations. For

comparison, the model-historical (black dots) indicates the non-bias-corrected climate data.

2.2.3. SWAT- Scenarios Without Climate Bias Correction

Monthly SWAT TP and DRP loads, driven by historical daily precipitation and temperature from the climate models without bias correction, were over-predicted compared to observed loads (Figure 6). While the overestimates increased with increasing loads in all scenarios, performance also varied based on climate model and nutrient. DRP and TP loads based on CCSM4 inputs were closer to observed compared to other models (Figure 6). Modeled TN showed better fits despite the excess for other loading variables within the ensemble, and IPSL-CM5A-MR inputs resulted in the best overall TN fits (Figure 6).

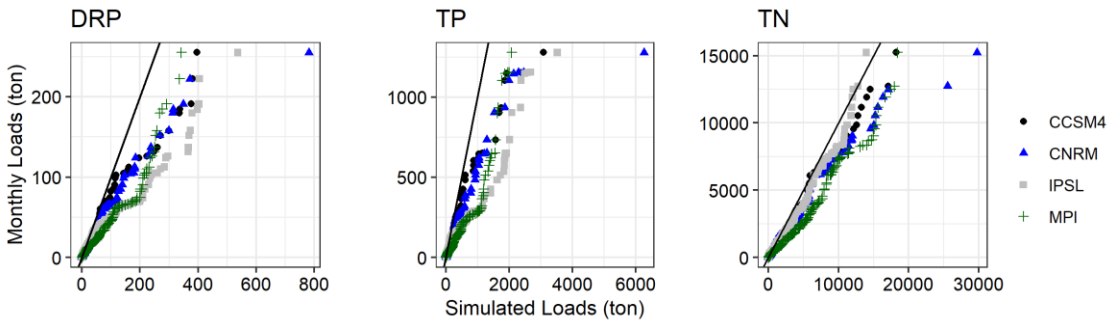


Figure 6. Quantile-Quantile plots (Q-Q plots) of monthly loads from SWAT driven by the observed climate data from the 71 stations (solid black line) and the SWAT loads driven by climate model outputs without bias correction (i.e., model-historical scenario) from 1985 to 1999.

For seasonal loads averaged over the time period of the analysis, all models overestimated DRP and TP loads from November to March (Figure 7). IPSL-CM5A-MR underestimated TN loads most of the spring months, although IPSL-CM5A-MR TN results

were a closer fit to model-observed values. Interestingly, CCSM4 simulation results appeared to follow closely March and April model-observed TN average load values.

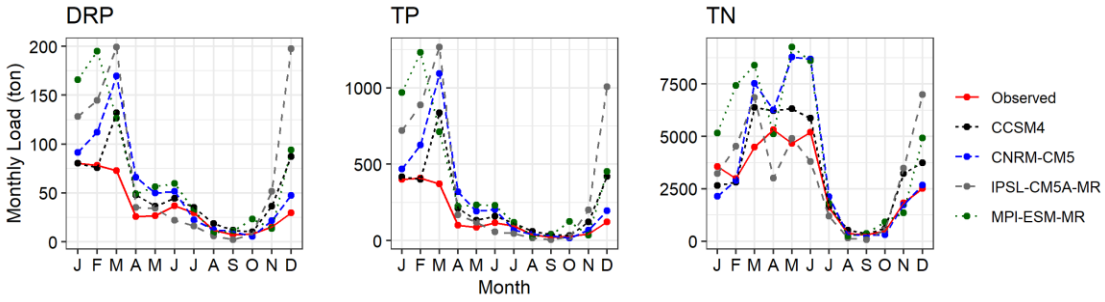


Figure 7. Comparison of monthly averaged observed loads for each nutrient among the SWAT driven by the non-bias corrected climate model outputs (model-historical scenario) and SWAT driven by the 71 climate stations.

2.2.4. SWAT Output with Climate Bias Correction

After bias correcting PCP, TMAX, and TMIN and re-running the SWAT model, there were clear differences in load estimates among bias correction methods (Figure 8). Among the bias correction methods, Delta resulted in loads closest to observations, followed by MBCn and QDM, respectively. MBCn and the other multivariate BC techniques (i.e. MBCp and MBCr), however, performed similarly to QDM, with MBCn outperforming QDM in most cases. In general, while the majority of the BC scenarios showed improvements over the simulations without bias correction, EQM resulted in larger deviation from observed loads except for TN.

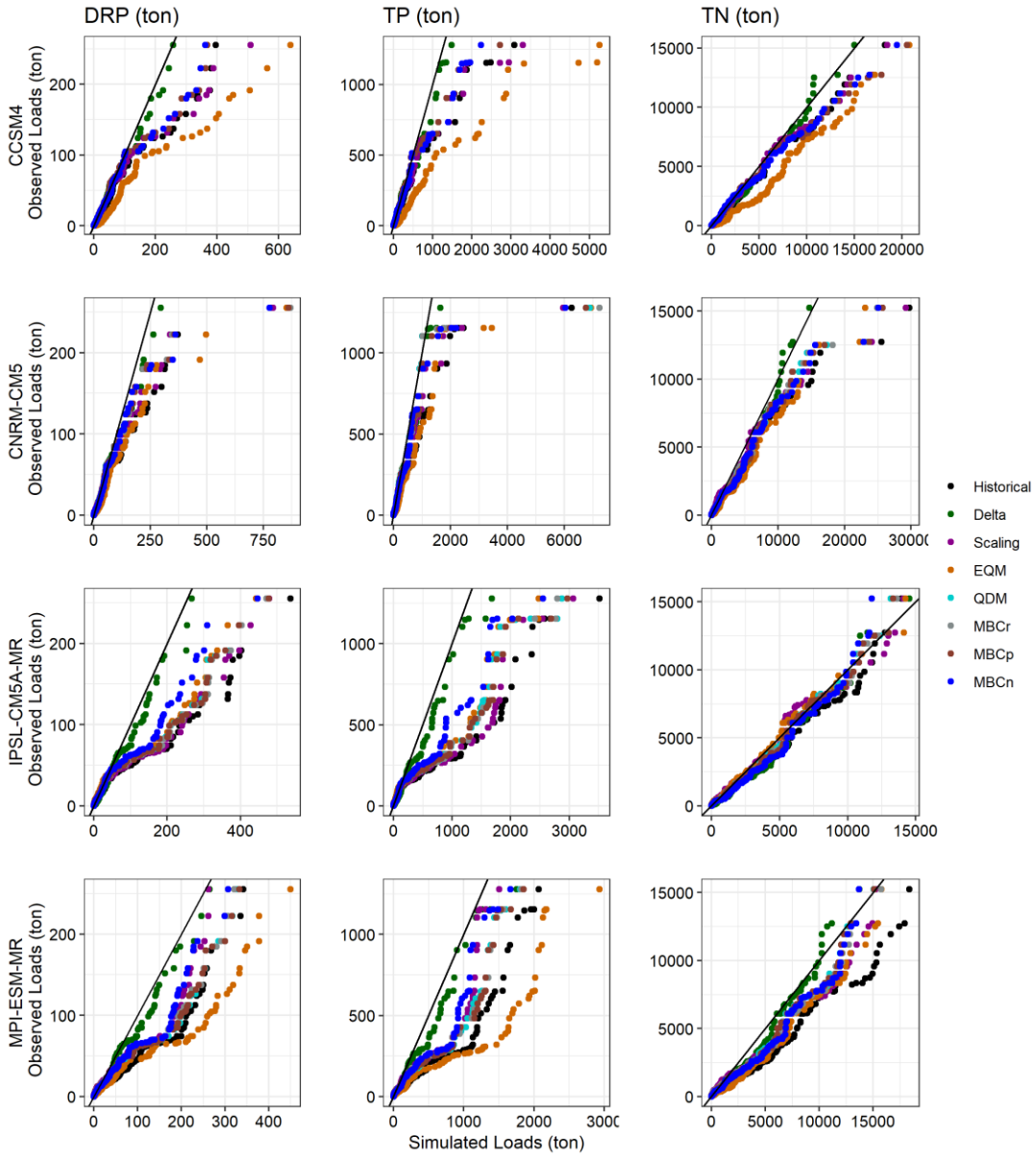


Figure 8. Quantile-Quantile (Q-Q) plots showing comparison between the DRP, TP, and TN loads from the SWAT driven by the 71 climate stations and the SWAT driven by bias-corrected precipitation (PCP), minimum temperature (TMIN), and maximum temperature (TMAX) climate model outputs.

Comparing the performance among climate models using the best BC methods (i.e. Delta, QDM, and MBCn), results showed that CCSM4 performed best for loads below 100

and 500 ton for DRP and TP, respectively (Figure 9). For TP, CCSM4 usually underestimated smaller loads and overestimated higher ones. When applying Delta, all models seem to perform similarly, with all models performing well when predicting smaller loads for DRP and TP while larger loads were best predicted by MPI and IPSL scenarios. When applying MBCn and QDM, CCSM4 and CNRM performed best principally when simulating smaller loads. CNRM-CM5 appeared to perform best in simulating higher DRP loads, but is poor in simulating TP and TN. MPI-ESM-MR simulated higher TP loads better than the other climate model scenarios, but it does not match as well as CCSM4 for the mid-range of observed loads.

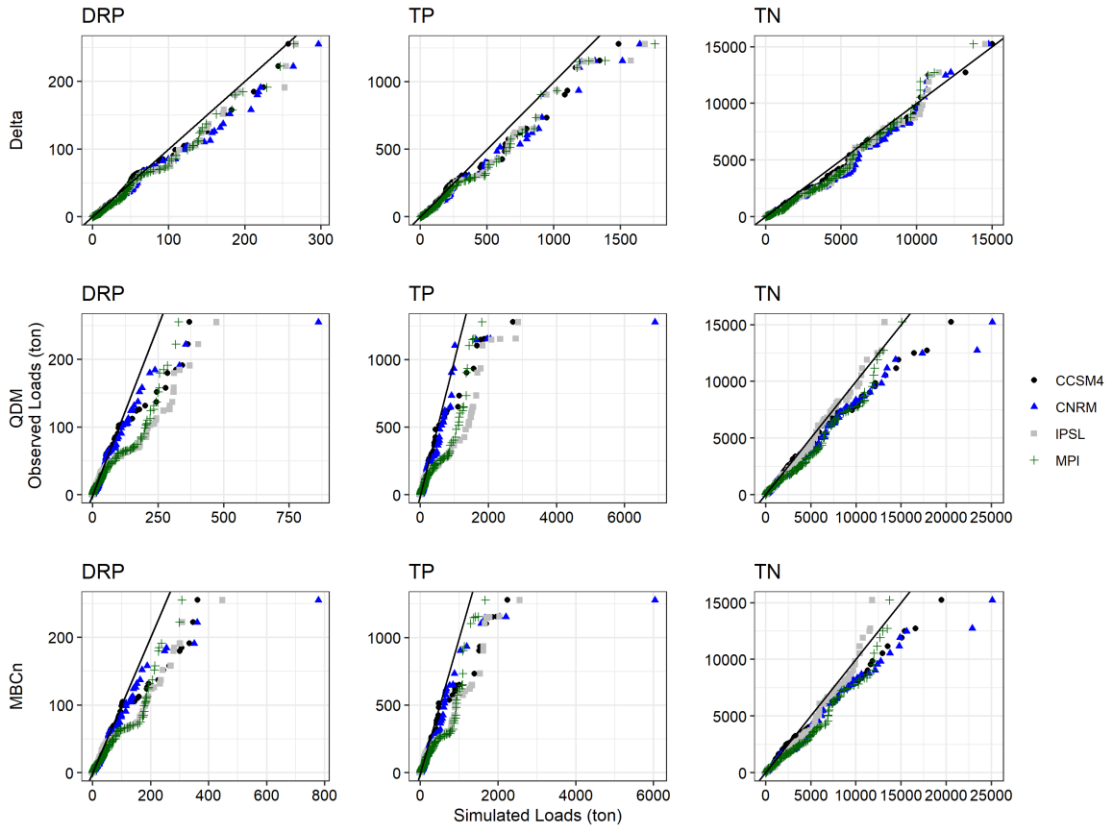


Figure 9. Quantile-Quantile (Q-Q) plots showing comparison between DRP, TP, and TN load from the SWAT driven by the 71 climate stations and the SWAT loads driven by bias-corrected climate model outputs using Delta, QDM, and MBCn.

2.2.5. Model Metrics Comparison

Based on fit statistics (Tables 2-4), Delta outperformed all the other methods, with MBCn ranking second most of the time. The choice of BC method also depends on the metric applied. MBCn ranked best after Delta when considering NSE and R^2, but based on PBIAS, Scaling and QDM rank best after Delta, with QDM ranking best at least once among all metrics when compared to Scaling.

Table 2. Percent Bias (PBIAS) result for model, bias correction method, and nutrient type. Calculations were based on monthly-averaged loads between loads from the SWAT driven by the observed climate data at the climate model resolution and the SWAT loads driven by the climate outputs used for each bias correction scenario. Best metric other than Delta is highlighted in each row.

			Bias correction						
Model	*Nutrient*	*no BC*	*Delta*	*Scaling*	*EQM*	*QDM*	*Mbc-N*	*Mbc-p*	*Mbc-r*
CCSM4	*TP*	41.30	0.00	31.10	173.40	**12.90**	16.80	**12.90**	**12.90**
	TN	10.90	0.00	**2.90**	53.50	5.10	4.70	5.10	5.10
	DRP	34.90	0.00	25.50	110.60	**12.40**	14.60	**12.40**	**12.40**
CNRM-CM5	*TP*	44.80	0.00	23.90	72.80	16.30	19.20	18.80	**15.10**
	TN	9.70	0.00	**0.20**	23.20	5.70	7.20	6.50	4.10
	DRP	30.00	0.00	14.60	45.70	7.20	9.10	7.90	**7.10**
MPI-ESM-MR	*TP*	91.80	0.00	**27.10**	129.70	42.10	28.50	43.50	40.60
	TN	41.40	0.00	14.30	31.00	**10.80**	14.40	11.30	11.90
	DRP	66.40	0.00	**22.80**	88.90	35.10	24.70	35.30	32.70
IPSL-CM5A-MR	*TP*	96.50	0.00	80.10	49.00	61.00	**34.20**	62.20	64.20
	TN	**0.30**	0.00	-13.20	-15.00	-8.70	-2.50	-8.00	-6.60
	DRP	66.70	0.00	54.00	31.90	43.40	**24.40**	43.60	45.40

Most of the *PBIAS* results were positive, indicating consistent overestimation of nutrient loads (Table 2). DRP and TP loads were highly and consistently overestimated among all models and BC methods. However, DRP overestimates were generally less than TP. *PBIAS* varied significantly among bias correction methods. For example, for CCSM4, EQM had *PBIAS* of 173 for TP, while Delta and QDM values were 0 and 12.9, respectively. TN was underestimated only for IPSL-CM5A-MR scenarios. However, it was overestimated by the other models (i.e. CCSM4, CNRM, and MPI). Based on NSE and R^2, MPI performed best for DRP and TP among the BC methods (Table 3-4). However, for TN, IPSL had the most

optimal NSE values, principally in the MBC-*N* scenario, while EQM had the best R^2 results. Compared to the cases without bias correction, all BC methods improved *PBIAS*, NSE, and R^2 results except for EQM. Delta showed best metric values. These analyses revealed that the choice of climate model and bias correction method will not only influence the nutrient model outcomes but also indicated that the choice of BC method varies by climate model, fit metric, and nutrient type.

Table 3. Nash-Sutcliffe Efficiency (NSE) results for model, bias correction method, and nutrient type. Calculations were based on monthly-averaged loads between loads from the SWAT driven by the observed climate data at the climate model resolution and the SWAT loads driven by the climate outputs used for each bias correction scenario. Best metric other than Delta is highlighted in each row.

			Bias correction						
Model	*Nutrient*	*no BC*	*Delta*	*Scaling*	*EQM*	*QDM*	*Mbc-N*	*Mbc-p*	*Mbc-r*
CCSM4	*TP*	-1.75	1.00	-1.93	-7.96	-1.07	**-0.99**	-1.07	-1.07
	TN	-0.74	1.00	-0.67	-1.13	-0.82	**-0.73**	-0.82	-0.82
	DRP	-1.48	1.00	-1.62	-3.91	-1.07	**-1.01**	-1.07	-1.07
CNRM-CM5	*TP*	-2.62	1.00	-2.04	-3.53	-2.29	**-1.73**	-2.27	-2.45
	TN	-0.97	1.00	-0.82	-0.64	-0.63	-0.61	-0.64	**-0.61**
	DRP	-1.32	1.00	-1.08	-1.58	-0.99	**-0.72**	-1.00	-1.00
MPI-ESM-MR	*TP*	-1.50	1.00	**-0.19**	-3.11	-0.45	-0.24	-0.50	-0.49
	TN	-1.17	1.00	-0.72	-0.60	**-0.41**	-0.44	-0.45	-0.50
	DRP	-1.27	1.00	**-0.41**	-2.39	-0.76	-0.45	-0.76	-0.69
IPSL-CM5A-MR	*TP*	-4.04	1.00	-3.36	-1.99	-2.43	**-1.34**	-2.49	-2.54
	TN	-0.47	1.00	-0.31	-0.31	-0.29	**-0.25**	-0.31	-0.32
	DRP	-3.26	1.00	-2.76	-1.84	-2.24	**-1.33**	-2.26	-2.32

Table 4. R^2 results for model, bias correction method, and nutrient type. Calculations were based on monthly-averaged loads between loads from the SWAT driven by the observed climate data at the climate model resolution and the SWAT loads driven by the climate outputs used for each bias correction scenario. The best metric other than Delta is highlighted in each row.

			Bias correction						
Model	*Nutrient*	*no BC*	*Delta*	*Scaling*	*EQM*	*QDM*	*Mbc-N*	*Mbc-p*	*Mbc-r*
CCSM4	*TP*	0.10	1.00	0.10	0.10	0.10	**0.11**	0.10	0.10
	TN	0.08	1.00	0.08	**0.12**	0.08	0.10	0.08	0.08
	DRP	0.08	1.00	0.09	0.09	0.09	**0.10**	0.09	0.09
CNRM-CM5	*TP*	0.22	1.00	0.21	0.21	0.19	**0.23**	0.19	0.19
	TN	0.14	1.00	0.13	**0.18**	0.15	0.16	0.15	0.15
	DRP	0.22	1.00	0.21	0.22	0.20	**0.25**	0.21	0.21
MPI-ESM-MR	*TP*	0.29	1.00	0.34	0.33	**0.36**	0.30	0.35	0.34
	TN	0.17	1.00	0.16	**0.20**	0.19	0.19	0.18	0.18
	DRP	0.27	1.00	0.31	0.31	0.31	0.28	0.31	**0.32**
IPSL-CM5A-MR	*TP*	0.12	1.00	0.10	**0.14**	0.14	0.13	0.14	0.14
	TN	0.12	1.00	0.13	0.12	0.12	**0.14**	0.12	0.12
	DRP	0.11	1.00	0.12	**0.14**	0.13	0.12	0.13	0.13

2.2.6. Changes in Future Nutrient Loads and Hydrological Responses

Based on the above results, the SWAT loads driven by the bias corrected CCSM4 model outputs performed generally best among all models, and Delta, QDM, and MBCn were the bias correction methods that best matched SWAT loads driven either by the 71 climate station observations or by the observed climate mapped to climate model resolution. Therefore, I explored how these combinations of model and BC method influenced mid-century (2051 to 2065) nutrient load projections (Figure 10). Using the Delta BC, DRP and TP loads are projected to increase in the future, except in the winter

and beginning of spring. TN loads are projected to decrease only in March and April on average. However, the QDM and MBCn correction methods lead to projected changes in the opposite direction where loads are projected to decrease, except in most winter seasons and June.

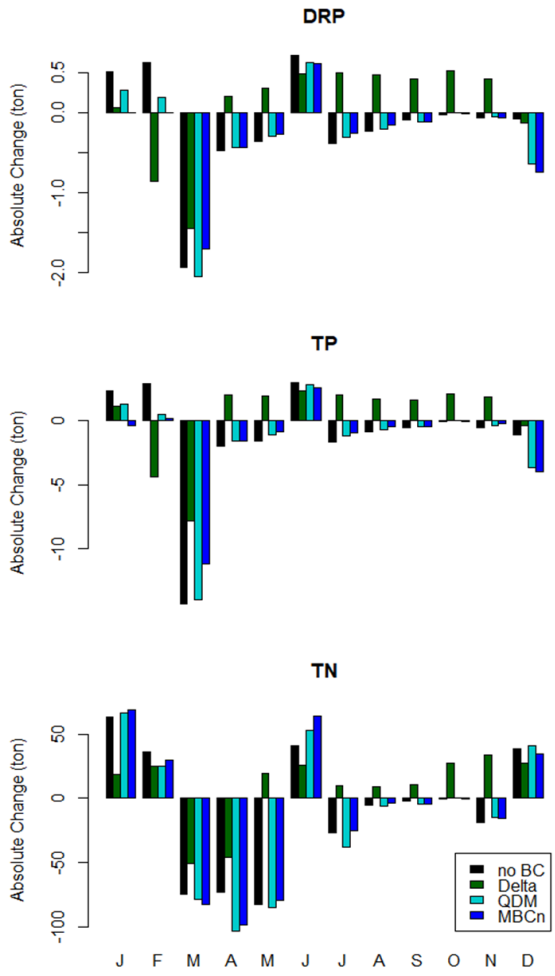


Figure 10. Monthly average absolute load change for the midcentury (2051-2065) period based on the change between the historical (1985-1999) and midcentury overall best SWAT results driven by no bias-corrected (no BC) as well as best bias corrected CCSM4 outputs for Dissolved Reactive Phosphorus (DRP), Total Phosphorus (TP), and Total Nitrogen (TN).

To understand these changes, I assessed changes in monthly evapotranspiration (ET), snowfall, flow, rainfall, TMIN, and TMAX for each BC method (Figure 11). When Delta was used, the simulated ET and flow differ significantly from the other BC methods. ET decreases consistently over all months (~ -10 mm/month), while flow mostly increases (~ 70 cms/month). For QDM and MBCn, ET mostly increased over the months (~ 5 mm/month), except in the end of summer and beginning of fall season, and flow overall decreased (~ - 50 cms/month), except in the winter, and July. In all BC methods, however, snowfall was consistently predicted to decrease in the future. Rainfall is expected to increase in the winter and in the first and third months of spring, while in other months are projected to decrease, which agrees with the increasing flow and decreasing snowfall patterns of QDM and MBCn scenarios. For temperature, I observed that both TMIN and TMAX are projected to increase in the future, which explains the increasing ET patterns in QDM and MBCn scenarios. These temperature patterns also explain the consistent decrease in snowfall for mid-century, because SWAT calculates snowfall by partitioning rainfall based on temperature values.

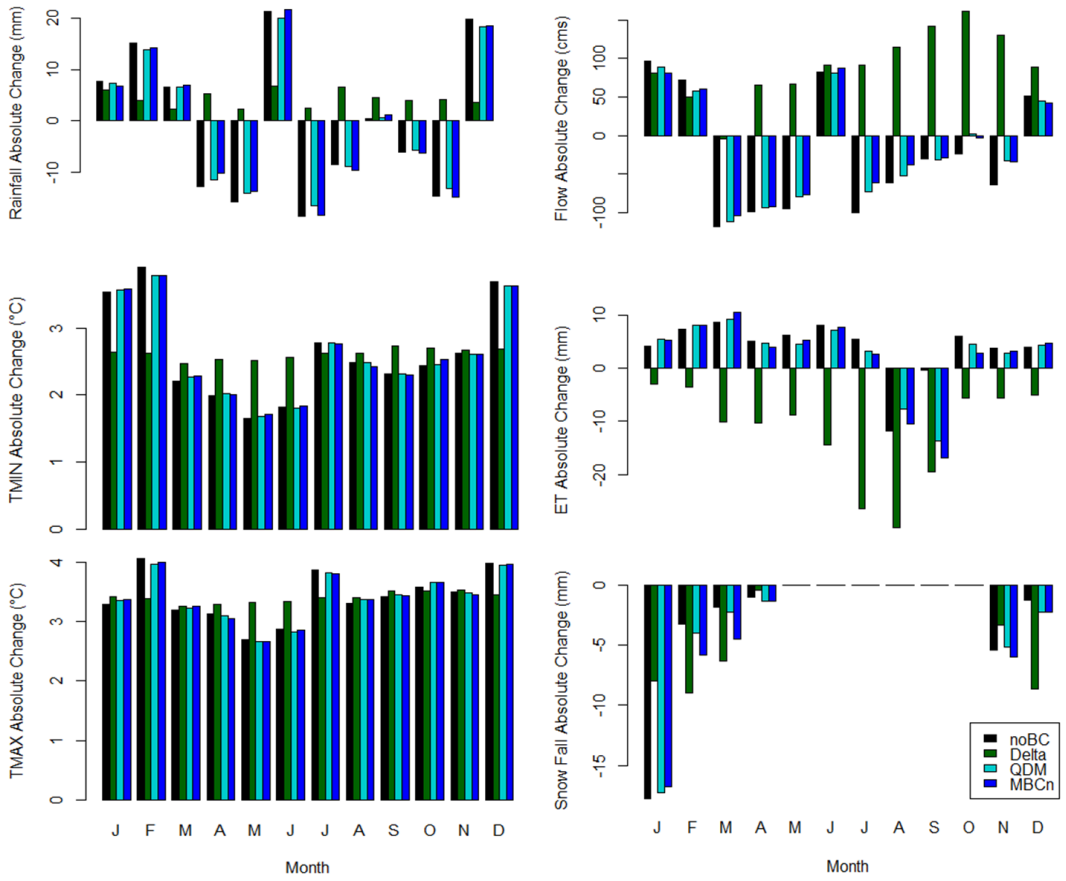


Figure 11. Monthly hydrological process changes for each bias correction method: rainfall (mm), maximum (TMAX) and minimum (TMIN) temperature (°C), evapotranspiration (ET, mm), snowfall (mm), and flow (cms). Change calculation was between the historical (1985-1999) and midcentury (2051-2065) overall best SWAT results driven by no bias-corrected (no BC) as well as best bias corrected CCSM4 outputs

Combining both assessments (Figures 10 and 11), March, April, and May DRP, TP and TN loads are projected to decrease as flow decreases and ET increases with potentially higher temperatures in the future. These results not only show how bias correction impacts

the nutrient modeling results, but also how it impacts the hydrological processes that can drive the nutrient load results.

2.3. DISCUSSION

I investigated the influence of bias correcting climate model precipitation and temperature output to be used as input to a watershed nutrient model. I compared the ability of the watershed model to simulate historical nutrient loads when driven by outputs (i.e. PCP, TMAX, and TMIN) from 4 climate models (CCSM4, CNRM-CM5, IPSL-CM5A-MR, and MPI-ESM-MR) without bias correction. I showed that model performance varied by nutrient type and magnitude, but that driving SWAT with CCSM4 led to the best fit to observations. Like other recent studies (Maraun et al., 2010; Ehret et al., 2012; Teutschbein & Seibert, 2013) I found that bias correcting climate model outputs, to correct for systematic and random errors in global and regional climate model outputs, can alter the findings of a study. Nutrient load studies that consider future climate should evaluate the impact of any bias correction methods applied across temporal scales (annual to monthly) and depending on nutrient of interest.

Our results reveal several inherent uncertainties persisted within simulations for the western Lake Erie basin, confirming behavior found in previous studies for regions outside of the Great Lakes. First, using climate model outputs as input to SWAT can increase uncertainty within loading predictions (Mehdi et al., 2015). A study addressing SWAT climate data input issues in the northeast of Brazil suggested that the choice of climatic inputs is critical for better representation of watershed processes, directly influencing

nutrient load forecasts (de Almeida Bressiani et al., 2015). In this study, for example, it was found that, in general, CCSM4 climate outputs performed best among the other models, but that MPI-ESM-MR performed better at higher DRP and TP loads, and IPSL-CM5A-MR performed best for TN. However, I also found that applying bias correction was necessary to improve the fit to model-observed loads.

Second, the way N and P processes are represented in the watershed model can influence how they respond to climate. For example, SWAT has been shown to routinely underestimate soil solution P, which likely leads to the underestimation of dissolved P (Vadas & White, 2010; Vadas et al., 2013), and climate change would impact particulate and dissolved P differently. While modeling P in SWAT will improve over time, studies generally report satisfactory performance in modeling N (Cakir et al., 2020; Logsdon & Chaubey, 2013). This explains why in our study the metrics were better for N when compared to TP and DRP. Additionally, Kujawa et al. (2020) showed that the main source of uncertainty when predicting N loads is related to the climate uncertainty, whereas uncertainty in P loads is mostly linked to uncertainty in the hydrologic model (SWAT).

Third, the uncertainties linked to nutrient prediction may also be attributed to the average based approach implemented by modelers when representing agricultural point and non-point sources of pollution in the basin (J. Wang & Baerenklau, 2015). For example, the use of rates to represent the chemical fertilizer application in crops and manure land application from intensified agricultural systems such as Concentrated Animal Feeding Operations (CAFOs) for the overall basin area (Kast et al., 2021). Information on how these operations may drive land use change and water quality conditions in their

surrounding may be crucial to improve our ability to represent the nutrient load reality in agricultural watersheds such as the western Lake Erie. One key concept currently ignored in watershed models is the spatial organization of these sources of pollution, specifically if their dispersion or clustering behavior can result in different environmental signatures. These components should be explored to improve our ability to model nutrient loads in these complex systems.

While several studies have shown that climate bias correction impacts streamflow, precipitation, and snowfall models (Teutschbein & Seibert, 2012; Bhowmik et al., 2017; Meyer et al., 2019; Wörner et al., 2019), I showed that BC selection also impacts watershed nutrient model performance. I also demonstrated that the impact of BC choice varies by type of nutrient modeled, climate model used, and the fit metric of interest. For example, I found that CCSM4 outputs led to the best watershed model performance for DRP and TP, principally after Delta, QDM, and MBCn bias correction. However, with the exception of Delta, based on NSE, MBCn is the most optimal BC method. On the contrary, if PBIAS is the metric of interest, QDM is the most proper BC choice, after Delta. Interestingly, for TN, the no-BC scenario appears to perform reasonably well in terms of PBIAS and NSE for the IPSL-CM5A-MR model. This variation among model metrics and nutrient load per climate model was also observed by Yuan et al. (2020), although they did not compare among bias correcting methods.

Although I found Delta to be the optimal BC method, hydrological modeling studies have suggested Delta has both advantages and disadvantages (Cannon et al., 2020). One advantage is that it bases future predictions on historical reality, which is already

understood, but it can be useful if the study seeks to understand the impact of current climate conditions on a future scenario. However, future interactions among climate, flow, or nutrient dynamics might differ from the historical period under non-stationary conditions (Lenderink et al., 2007). We showed that Delta bias corrections led to opposite directions of change when compared to other relevant univariate (QDM) and multivariate (MBCn) BC techniques, although QDM and MBCn reflected different magnitude changes. Delta has the potential to preserve the mean of the distribution in case of variables such as temperature as well as the mean and variance of bounded variables distribution such as precipitation. In the case of precipitation, Delta could improve the mean statistic but deteriorate the simulated variance principally when considering the extremes of a distribution (Cannon et al., 2020). This deterioration may explain the opposite direction of change when predicting monthly DRP, TP, and TN using Delta in comparison to QDM and MBCn. Therefore, Delta should be used with caution depending on the objective of the study, and because it assumes future climate, flow, and nutrient dynamics will function similarly to the present.

Studies that applied MBC to predict snow precipitation found MBC techniques, that account for the interdependencies among variables used as inputs, are superior to the univariate techniques such as QDM (Meyer et al., 2019). In this study, MBCn outperformed QDM in most of the cases, likely because in addition to correcting quantile dependent biases including frequencies and intensities of wet days, MBCn includes multivariate dependence structure among variables in the correction. However, this performance varied by climate model, metric of interest, and nutrient type. For CNRM-

CM5, QDM outperformed both MBCn and MBCp but not MBCr when considering PBIAS. In the MPI-ESM-MR model scenarios, QDM performed poorly in comparison to MBCs. However, the difference among metrics for QDM and MBCs is not high, which may be because we only used temperature data to correct precipitation and vice-versa. It was interesting that MBCs performed well even though only temperature and precipitation were used as inputs. If more variables such as solar radiation, wind speed, and relative humidity were used the MBC performance would probably improve.

Analyzing the impact of BC choice in future nutrient load changes using the overall best model-BC combinations (i.e CCSM4 no BC; CCSM4 Delta; CCSM4 QDM; CCSM4 MBCn), I found that BC choice impacts the direction of load. Kalcic et al (2019) compared their no-BC results with a Delta BC scenario and suggested that bias correcting can significantly impact nutrient load predictions. Our results expand and emphasize the importance of addressing uncertainty in the choice of climate models and BC methods when modeling nutrients. Furthermore, we found that loads, based on no-BC, QDM, and MBCn, are likely to decrease in the future spring period due to decreases in flow and snowfall, as well as increases in temperature and evapotranspiration. However, these hydrological changes were the opposite when considering Delta as the BC method, further illustrating the importance of this choice for future simulations. Future studies should consider a combination of bias correction methods among precipitation and temperature for modeling purposes.

2.4. CONCLUSION

Bias correction of climate data is a common practice among watershed modelers across the globe, yet the impacts of different correction methods are rarely evaluated. We acknowledge that applying BC does not create climate variability, and so there is a stationarity of bias in simulations. However, BC is usually necessary (depending on the climate model choice and the objective of the study) to correct systematic errors of climate model outputs principally in the calibrating period (Hakala et al., 2019). Without this correction, the error propagation when forecasting could be substantial. The benefits of this study go beyond the Great Lake region in the USA, and our approach can be applied to any location in the world where watershed models have been developed and calibrated as well as where data exists. In places where there is limited water quality data to investigate the specific influence of BC on water quality predictions, this work can serve as a guide for which methods might be most appropriate.

CHAPTER 3

SPATIOTEMPORAL LAND USE CHANGE AND ENVIRONMENTAL DEGRADATION SURROUNDING CAFOS IN MICHIGAN AND NORTH CAROLINA

As discussed in Chapter 1 section 1.3, CAFO activities have the potential to change the landscape in their surroundings, which may lead to environmental degradation. In this chapter, I hypothesize that CAFO-impacted areas have a higher percentage of change over time from natural land cover (i.e., forest, wetland, and savanna) to anthropogenic land cover (i.e., cropland and shrublands). I expect that areas within 15 km of CAFOs (i.e., CAFO-impacted areas) experienced higher land cover change and environmental degradation compared to areas farther from CAFOs (i.e., control areas) (Sims et al., 2005; Centner, 2012; Furiness et al., 2019). I test these hypotheses examining patterns from 2000 to 2018 for two states in the U.S.—Michigan (MI) and North Carolina (NC)—where regulated CAFO location information is publicly available.

3.1. METHODS

3.1.1. Study Areas

This study focused on two states - Michigan (MI) and North Carolina (NC)- in the U.S. Like much of the U.S. and world (FAO, 2014), these two states have shown a clear shift from small to large animal farm operations in recent decades (NASS, 2017), affecting environmental quality and rural communities. Both states provide the locations of

permitted liquid waste CAFOs to the public; when this study was initiated, MI listed 328 CAFOs, showing a sparse spatial aggregation in comparison to NC, with approximately 2,600 CAFOs registered (Figure 12). Liquid waste CAFOs in these areas are primarily swine and dairy operations; NC has a growing poultry industry mainly using dry waste management, which is mostly unregulated under the National Pollutant Discharge Elimination System (NPDES), and the state does not disclose precise dry-waste poultry facility locations.

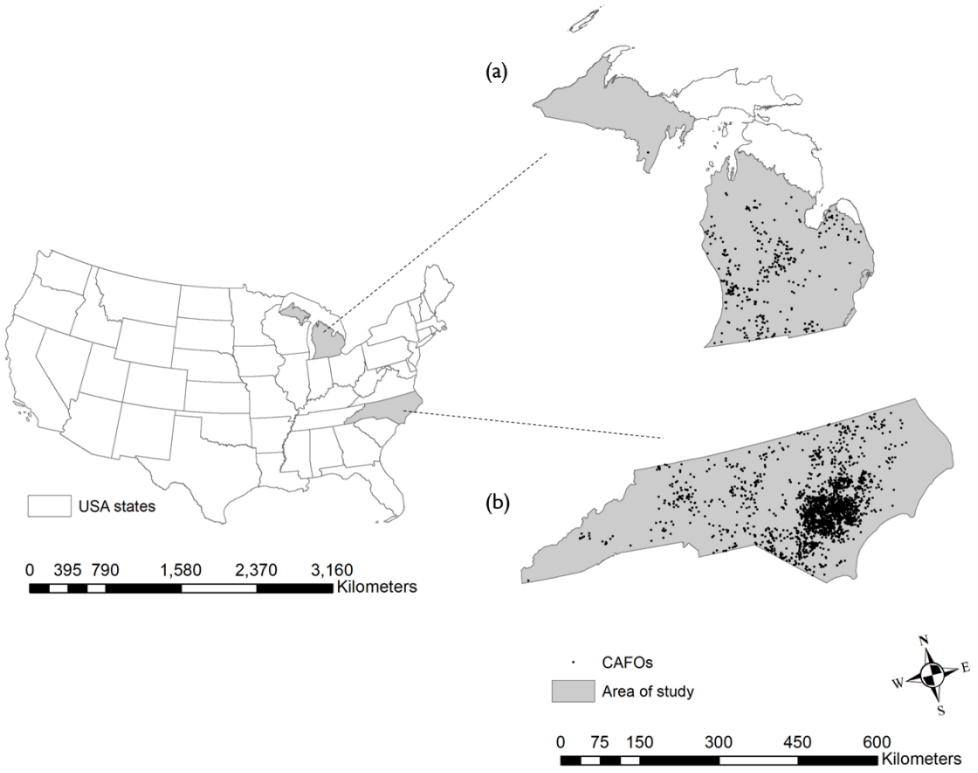


Figure 12. Study Areas located in the United States. Michigan (MI - a) with 328 CAFO (Concentrated Animal Feeding Operation) locations and North Carolina (NC - b) with 2,594 CAFOs. The gray area illustrates the analyses extent.

These states have distinct biophysical characteristics. Michigan (MI) has a humid continental climate with distinct summers and winters and a fairly even distribution of precipitation throughout the year (Peel et al., 2007). Its topography consists of flatlands, gentle and rolling hills, with average elevation of 270 meters. Summer temperatures can reach around 30 degrees C, while winters are cold and snowy with freezing temperatures. The average annual rainfall in MI is approximately of 800 mm, with 60% of the rainfall occurring in the growing season (NOAA, 2021). MI is divided into 5 main ecoregions: Northern Lakes and Forests, North Central Hardwood Forests, Eastern Corn Belt Plains, Drift Plains, and Huron/Erie Lake Plains (EPA, 1999). The Northern Lakes and Forests is dominated by nutrient poor glacial soils, forests, and extensive sandy outwash plains. The North Central Hardwood Forests consists of mosaic forests, wetlands and lakes, cropland agriculture, pasture, and dairy operations. This area formerly consisted of oak grassland areas and beech-sugar maple forest, but due to agriculture and urbanization, few remnants of native grassland remain. The Eastern Corn Belt Plains were primarily a rolling plain with local end moraines and more natural tree cover. Today, extensive crop and livestock production occur in these plains. The Drift Plains are characterized by marshes and dunes, while the Huron/Erie Lake Plain was originally covered by beech forests and elm-ash swaps but today most of the area has been cleared and artificially drained for agricultural practices.

North Carolina (NC) is a highly variable landscape. The state's western boundary is composed of the Southern Appalachian Mountains, including the tallest peak in the Eastern US (2037 m), sweeping across the Piedmont and low-lying Coastal Plain to the

Atlantic Ocean. NC is mainly characterized by a humid subtropical climate, with hot summer day time temperatures often reaching 32 degrees C, and mild winters with freezing temperatures and snow rarely occurring below the mountains (Peel et al., 2007). Precipitation is highly variable and occurs year-round (annual mean = 1143mm). Historically, native vegetation included extensive marshes, swamp forests, and long-leaf pine in the coastal plain, deciduous hardwood and evergreen forests in the piedmont, extending to the mountains which have some boreal conifer forests (Griffith et al., 2002). NC's economy is highly reliant on agriculture and forestry (UADA, 2018), which dominate the eastern portion of the state, and it also has among the fastest population growth in the country; the population has increased by 10% since the 2010 census and the state ranked 4th in population growth since 2018 (US Census Bureau, 2019).

Combined, MI and NC offer the opportunity to understand CAFO systems in different developmental stages in the U.S. While NC has a long-established CAFO system with a moratorium on new liquid waste operations, MI shows recent growth in the number of liquid-waste CAFOs (Walljasper, 2018). MI is an inland state, surrounded by 4 Great Lakes while NC is a coastal-influenced region, but both states have heavy animal and cropland agriculture.

3.1.2 Data

3.1.2.1 CAFOs Point Data

CAFO locations were assembled from the Michigan Department of Environmental Quality (EGLE, 2019) and North Carolina Department of Environmental Quality (NCDEQ, 2019). These datasets also provide information on the type of CAFOs registered,

the type of permits acquired, and the number of animals per farm. My focus was on liquid waste CAFOs, which are mainly cattle and swine operations in these two states.

3.1.2.2. MODIS Products

The Moderate Resolution Imaging Spectrometer (MODIS) images the same location at different times of the day and acquires complete coverage of Earth within 48 hours (Justice et al., 1998). A variety of MODIS products relevant to land use and environmental quality over time were used in this study (Appendix B – 2.1.7. and 2.2. Supplementary Table S-1 and Table S-2). Changes in land cover type detected from satellite imagery has been used extensively to detect land use changes and environmental degradation over time. For instance, conversion of forest to croplands or shrublands indicates anthropogenic activities that can impact environmental conditions (Jacquin et al., 2010; Lunetta et al., 2010; Eckert et al., 2015). Additionally, changes in evapotranspiration (ET) and the normalized difference vegetation index (NDVI) have been used to measure degradation in previous studies (Eckert et al., 2015; Alatorre et al., 2016a; C. Wang & Myint, 2016). Likewise, studies have associated the increase in land surface temperature (LST) and decline in both percent tree cover (PTC) and leaf area index (LAI) to shifts in land cover (i.e., forest to cropland) and environmental health impacts (Ishtiaque et al., 2016a; Estoque et al., 2017, 2018a). These land products are distributed by the Land Processes Distributed Active Archive Center (LP DAAC) (LP DAAC, 2019) and are available at various temporal resolutions: Daily, 8-Day, 16-Day, Monthly, Quarterly, and Yearly. The MODIS instrument also acquires data in four spatial resolutions: 250, 500, 1,000, and 5,600 meters. The temporal and spatial resolution of the products used in this

study varied depending on availability (Appendix B – 2.1.7. Supplementary Table S-1). I only used data representing the summer season (June 15th to September 15th) to avoid cloud and snow effects and to capture the highest intensity of the environmental parameters analyzed, particularly for vegetation indices which peak during the summer growing season in the study regions. More detail about each product and the indices used are provided in the Appendix B sections 2.1 and 2.2.

3.1.3. Data Analyses

This study focuses on a comparison of differences in the land use and environmental condition trends between 'CAFO-impacted' areas within 15 km of at least one regulated liquid waste CAFO and 'control' areas which were more than 15 km from a CAFO (Figure 13). The CAFO-impacted regions in NC and MI have 96,521 km^2 and 53,826 km^2 of area, respectively. The control area in MI has a total area of 66,106 km^2 and NC control area has 42,983 km^2. The 15 km distance radius represents the likely range of application of liquid manure from swine and dairy CAFOs (Freeze & Sommerfeldt, 1985; Kline et al., 2013; Long et al., 2018), which is a proxy for the extent of impacts from these operations. This approach allowed me to quantify the extent and nature of impacts likely related to permitted liquid waste CAFOs. Because most liquid manure is likely to be applied in closer proximity to CAFOs I also evaluated smaller distances (i.e., 3 km and 9 km radii around a given CAFO) to examine the sensitivity of the results to the chosen 15 km radius. To test my hypotheses, I first performed a yearly land use and land cover (LULC) change analysis using MODIS data between 2001 and 2017 to detect land use change in CAFO-impacted areas compared to control areas. I also compared environmental

conditions from 2000 to 2018 using a pixel-based trend assessment based on ET, PTC, NDVI, LAI and LST derived from MODIS.

3.1.3.1. Land Use Change Analysis

To investigate the first hypothesis that changes from natural to anthropogenic landscapes occurred more intensively in CAFO-impacted areas, LULC analysis was performed using the Land Change Modeler tool from the TerrSet Geospatial Monitoring and Modeling Software developed by Clark Labs (Clark Labs, 2019). LULC analysis is based on the net change area calculation in TerrSet, which takes the initial land cover areas (i.e., 2001 MODIS land cover product), adds the gains, and then subtracts the losses observed in the later land cover image (i.e., 2017 MODIS land cover product) (Clark Labs, 2019). I also performed a case study analysis mapping the transitions from each land use category to another using TerrSet. For each state, I selected units of two main types of CAFOs in the dataset, cattle and swine, to illustrate the trade-offs between land use types (i.e., savannas, wetlands, or forest losses versus gains in croplands or shrublands) as well as if these trade-offs were more likely to be driven by a certain type of CAFO (Figure 13). The identification of these trade-offs and the type of CAFO that is most likely to drive these changes may be crucial to develop future environmental solutions.

Although higher spatial resolution LULC products are available, I chose to use MODIS because it is likely to be a more reliable record of land cover change for our study areas. MODIS land cover does have some uncertainties in class accuracy, such as under-representation of wetlands and misclassification of grasslands (Sulla-Menashe & Friedl, 2018). However, while the National Land Cover Database (NLCD) has higher spatial

resolution, previous studies have identified inadequacies specifically with regard to CAFOs (Song et al., 2014; Martin et al., 2018). For instance, CAFOs in North Carolina were frequently misclassified as natural systems (often wetlands) by the NLCD (Martin et al., 2018). Among several global forest products available, a previous study found that the MODIS product was superior to the NLCD, with a lower root mean square error and the highest R^2 of all products evaluated, and that there was little difference in percentage among classes between the MODIS product and the NLCD (Song et al., 2014). Acknowledging the coarser spatial resolution of the MODIS product, I performed the same LULC analysis using the 30-m U.S. Department of Agriculture Cropland Data Layer (CDL). Comparable land use changes were identified using both MODIS and the CDL, therefore we retained the results of the MODIS-based LULC analysis for consistency with the MODIS products used in the environmental degradation analysis.

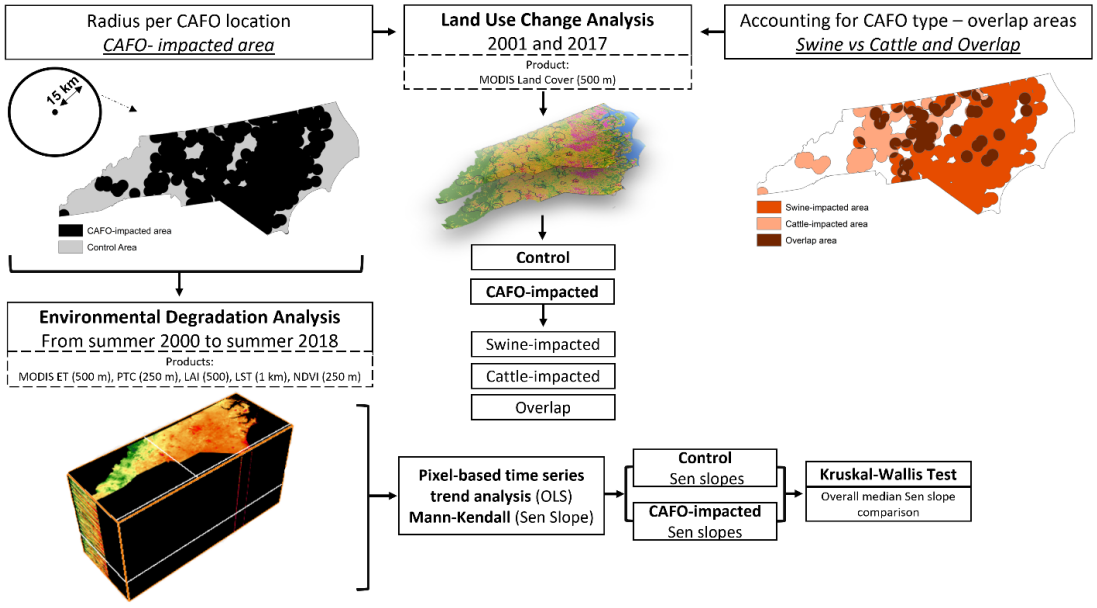


Figure 13. Methodology workflow of this study illustrating the Radius Assumption, Land

Use Change, and the Environmental Degradation Analyses in North Carolina; same was completed for Michigan.

3.1.3.2. Environmental Degradation Analysis

To assess the second hypothesis that CAFO-impacted areas experienced greater environmental degradation (i.e., declines in ET, PTC, NDVI, and LAI, and increases in LST) over time, I applied a temporal trend analysis based on the Mann-Kendall monotonic trend and the Theil-Sen median trend slope using the TerrSet Earth Trends Modeler. The Mann-Kendall test provides a nonlinear trend indicator that measures the degree to which a trend is consistently increasing or decreasing. It ranges from -1 to +1, where +1 indicates increasing trend, while -1 indicates a decreasing trend and 0 values indicate the lack of a consistent trend. This trend can be computed using a nonparametric technique developed by Sen (1968). Theil-Sen slope computes the median of slopes for the values observed at all time steps. This technique is robust against outliers and rejects values that are not reasonable without affecting slope estimations (Sen, 1968; Lamchin et al., 2018). One of the concerns with this technique is the potential false positive identification of trends in the presence of serial correlation and seasonality in the data (Yue & Wang, 2002; Bayazit & Önöz, 2007). Trend preserving pre-whitening is a pre-processing technique that can be apply before Mann-Kendall test to deal with positive and negative autocorrelation in the time series (Blain, 2015). However, for time series with a large number of observations as in our case, serial autocorrelation is not considered to be a major concern (Yue & Wang, 2002). In addition, specifically for image analysis, the original Mann-Kendall is recommended over modified tests (Militino et al., 2020).

I first constructed an image time series for each environmental parameter using all available summer season data (Appendix B – 2.1.7. Supplementary Table S-1). For each environmental parameter time series, I then generated a single summary image, where each pixel represented the slope value of the trend over time. I excluded urban pixels to avoid confusion with increasing LST trends due to urbanization (C. Wang & Myint, 2016). I also examined the trend preserving pre-whitening modification to the Mann-Kendall test, but this did not change the results substantially. Therefore, I retained the statistical results based on the standard Mann-Kendall trends I identified.

To compare CAFO-impacted areas to control areas, additional analyses were conducted. I defined the CAFO-impacted area in each state as all locations within 15 km of regulated liquid waste CAFOs; the remaining area was considered the control area (Figure 13). I calculated the Sen-Slope median within CAFO-impacted areas and control areas separately using geoprocessing tools in ArcMap version 10.4 from ESRI; I retained only the non-urban pixels with significant Mann-Kendall trends ($p \leq 0.1$). To determine if CAFOs were significantly related to detected environmental degradation, I applied a Kruskal-Wallis test. This test evaluates whether the overall median of the trend slopes from each environmental parameter in CAFO-impacted areas were significantly different from the trend slopes in control areas. I chose Kruskal-Wallis because it is a robust nonparametric test (i.e. it does not rely on a specific data distribution and has higher statistical power when applied to normal or short-tailed data distributions), while it also accommodates comparison among groups with unequal sample sizes (Siegel, 1957; Mahoney & Magel, 1996). I expected that the slope median of ET, LAI, PTC, and NDVI

trends in the CAFO-impacted areas would be lower than the slope median in the control areas, while the slope median of LST trends would be higher in the CAFO-impacted area in comparison to the control area.

3.2. RESULTS

3.2.1. Land Use Change

Major net changes in land use occurred in CAFO-impacted areas (i.e., areas within 15 km of CAFOs) in both MI and NC between 2001 and 2017 (Figure 14). In MI, the extent of cropland increased in CAFO-impacted areas, mainly at the expense of savannas; ~141,640 ha of savanna were lost, while croplands increased by ~182,100 ha. In MI CAFO-impacted areas, savanna loss was nearly 40,460 ha greater than compared to control areas. Similarly, in control areas, cropland increases were approximately four times smaller than in the CAFO-impacted areas. While significant forest losses were not detected in MI near CAFOs, an increase in forest was detected in control areas. Although loss of savannas also occurred in NC, I observed an increase in forest and shrubland in NC CAFO-impacted areas, rather than cropland expansion. The results show that ~81,000 ha of savanna loss and ~ 48,600 ha of shrubland expansion occurred in NC CAFO-impacted areas. Surprisingly, NC control areas experienced a loss of ~35,600 ha of forest. These results indicate potential land degradation in CAFO-impacted areas and reveal different land change histories between MI and NC between 2001 and 2017. Note that these changes were similar when using 3 km and 9 km radii, with NC showing stronger land use changes, principally within 3 km radius (i.e., increase in croplands at the expense of forest),

confirming our hypothesis of land degradation near CAFOs (Appendix B – 2.3. Supplementary Figure S-1).

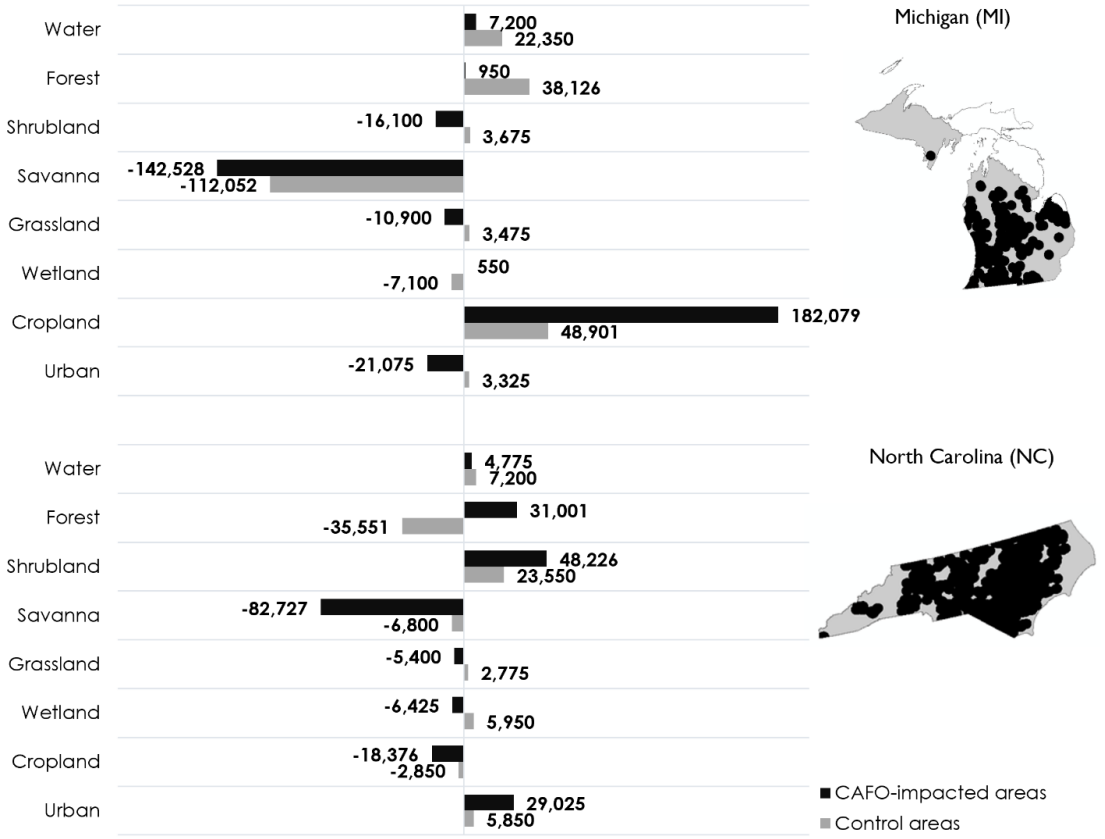


Figure 14. Land use change detected in hectares in Michigan and North Carolina using 2001 and 2017 MODIS data. Black bars illustrate the gains and losses in CAFO-impacted areas, while the gray bars indicate gains or losses observed in control areas.

Considering the type of CAFOs, and 'overlap' regions impacted by both swine and dairy CAFOs, revealed additional land use change patterns that were not detected when considering all CAFOs regardless of animal type (Figure 15). Swine operations were more strongly associated with changes in land use than dairy operations. For instance, shrublands replaced savannas primarily near swine CAFOs. Additionally, while savanna area was

converted to croplands (in MI) and shrublands (in NC), wetlands decreased in favor exclusively of croplands in swine CAFO-impacted areas in NC. In MI grassland conversion to croplands occurred in swine CAFO-impacted areas, as well as in overlap impact areas. In contrast, in NC, this transition was observed primarily in dairy CAFO-impacted areas. While forest extent increased mostly in the vicinity of swine CAFOs, forest loss was observed in both states exclusively in dairy CAFO-impacted areas; shrublands mainly replaced these forests in NC, while in MI they were replaced by croplands. The overlap impact area accounted for significant land conversion in both states, including increases in cropland extent at the expenses of savannas, grasslands, shrublands, and wetlands. In general, while croplands and shrublands most commonly replaced savannas in CAFO-impacted areas, cropland replacement of wetland and grassland extent was concentrated in the overlap region.

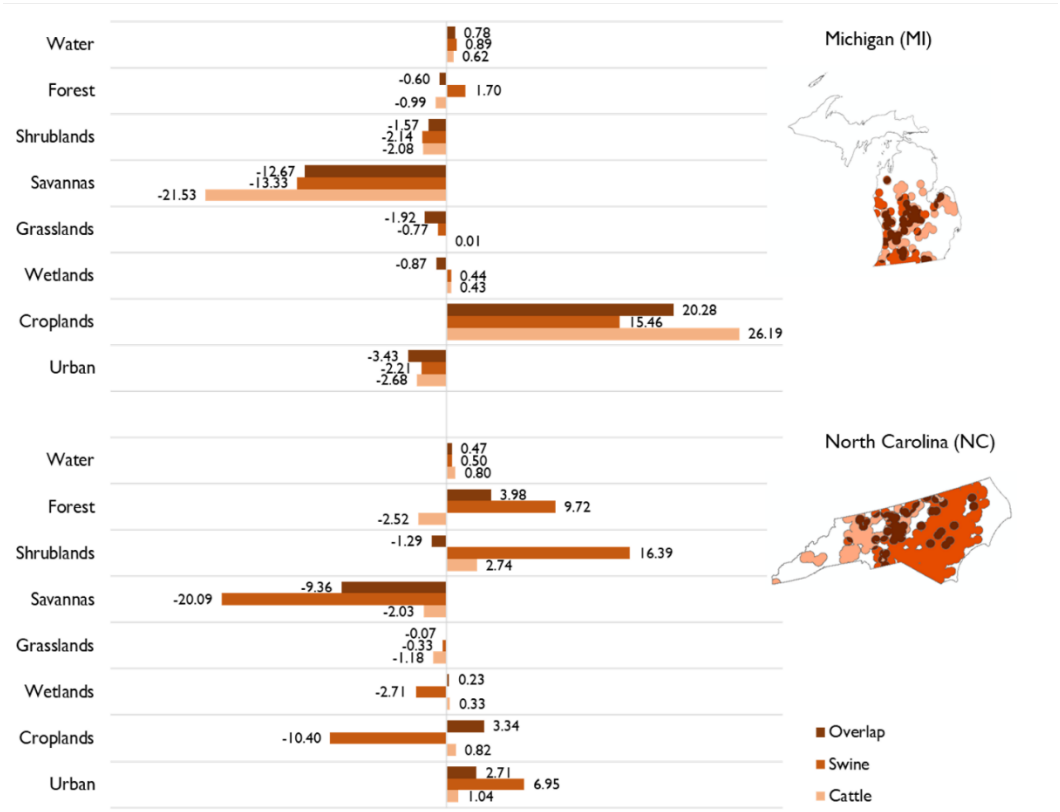
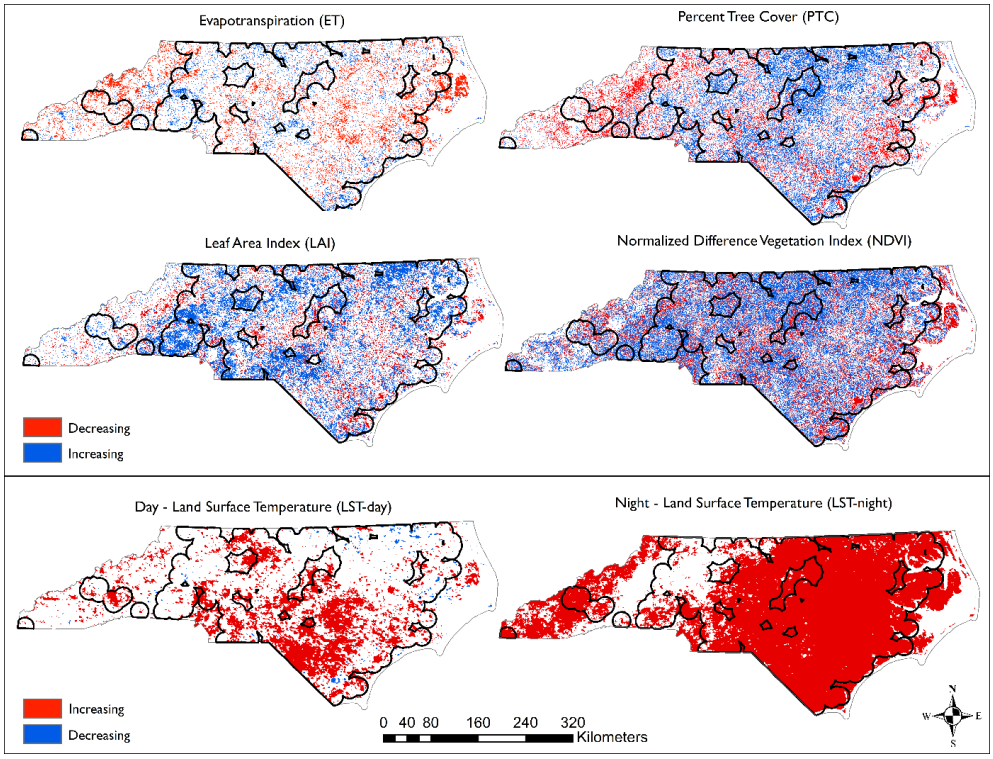


Figure 15. Percentage of total losses surrounding swine CAFOs and dairy CAFOs in both Michigan (MI) and North Carolina (NC). Areas impacted by both dairy and swine CAFOs (i.e., overlap) are shown in dark brown.

3.2.2. Environmental Degradation

Changes in environmental quality were investigated through a Mann-Kendall pixel-based monotonic trend analysis of a variety of indicators in MI and NC (Figures 16 and 17). We observed significant increases in both day and nighttime LST in NC, mostly in CAFO-impacted areas (Figure 16 and Table 5). Additionally, most indications of degradation (i.e., significant decreases in ET, LAI, and NDVI) occurred near CAFOs in NC, while improvements in environmental quality were mainly detected in control areas.

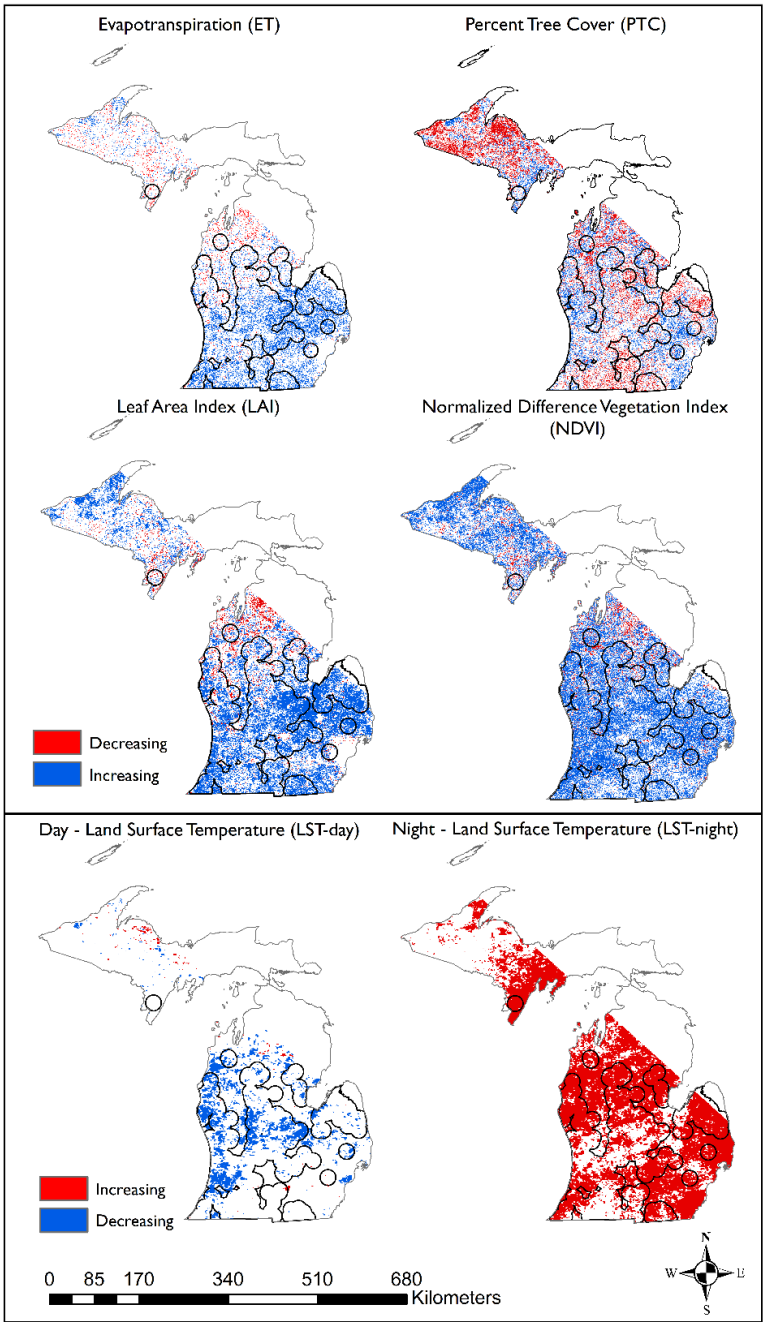
One environmental quality indicator showed the opposite pattern; surprisingly, PTC generally increased over time in NC CAFO-impacted areas, while decreases in PTC were observed in control areas.



For indicators ET, LAI, NDVI, and PTC: blue = increasing trends; red= decreasing trends; LST – day and LST – night: blue = decreasing trends, red = increasing trends

Figure 16. Changes in environmental quality indicators in North Carolina from 2000 to 2018 indicated by Mann-Kendall trend analysis across multiple parameters from MODIS. Significant ($p < 0.1$) Thiel-Sen slopes are shown for each parameter.

Distinct from NC, significant decreasing PTC trends were observed in the Lower Peninsula of MI, where regulated CAFOs are located, and in upper MI (Figure 17). Overall, across MI, significant increases in ET, LAI, and NDVI were observed. Although no significant positive day time LST trend was observed, nighttime LST did significantly increase over time across MI, principally in CAFO-impacted areas.



For indicators ET, LAI, NDVI, and PTC: blue = increasing trends; red= decreasing trends; LST – day and LST – night: blue = decreasing trends, red = increasing trends

Figure 17. Changes in environmental quality in Michigan (MI) from 2000 to 2018 detected by Mann-Kendall trend analysis across multiple parameters from MODIS. Significant ($p < 0.1$) Thiel-Sen slopes are shown for each parameter.

I identified significant differences in environmental quality trends over time between CAFO-impacted areas and control regions across multiple indicators (Table 5). In Michigan (MI), for each parameter, the overall slope median in the CAFO-impacted areas was significantly different from control areas (i.e., Kruskal-Wallis tests $p < 0.001$), yet the overall environmental quality trends did not consistently follow our expectations. In North Carolina (NC), median ET significantly decreased over time in CAFO-impacted areas, while no trend was observed in control areas. Additionally, increasing median day and nighttime LST trends in NC CAFO-impacted areas were stronger and significantly different than control areas and had greater magnitude. These findings confirmed our hypothesis that CAFO-impacted areas experienced environmental degradation in NC.

Table 5. Kruskal-Wallis test results evaluating the differences in environmental quality trends over time in CAFO-impacted areas compared to control areas.

	Michigan (MI)			North Carolina (NC)		
	Sen Slope Median		**Kruskal-Wallis test**	**Sen Slope Median**		**Kruskal-Wallis test**
Parameters	*CAFOs*	*No CAFOs*	*p*	*CAFOs*	*No CAFOs*	*p*
ET	0.302	0.245	**<0.001***	-0.268[a]	0	**<0.001***
LAI	0.018	0.015	**<0.001***	0.014	0.013	**<0.001***
LST-day	-0.291	-0.273	**<0.001***	0.258[b]	0.224	**<0.001***
LST-night	0.261	0.273	**<0.001***	0.315[b]	0.24	**<0.001***
NDVI	4.839	2.952	**<0.001***	3.269	2.135	**<0.001***
PTC	0.125	-0.30	**<0.001***	0.414	-0.286	**<0.001***

[a] Environmental quality parameter that met the hypothesis of stronger decreasing trends in areas with CAFOs than without CAFOs.
[b] Environmental quality parameter that met the hypothesis of stronger increasing trends in areas with CAFOs than without CAFOs.
*Significance level of 0.05

3.3. DISCUSSION

Cropland and shrubland expansion observed in CAFO-impacted areas in both MI and NC suggests that CAFOs may be a driver of land conversion in forest, wetland,

grassland, and savanna systems. While other studies have found similar results, such as an increase in shrublands due to livestock presence (J. R. Brown & Carter, 1998; Rundel et al., 2014; Schreiner-McGraw et al., 2020), this study is the first to examine this issue around CAFOs. This finding has important implications for future policies regarding land acquisition and management by these entities. It is possible that other drivers also contributed to observed LULC; for example, development of large-scale solar projects (SEIA, 2019) especially in NC may have also contributed to observed forest loss. Additionally, commercial poultry operations occurring in proximity to swine CAFOs in NC may have also contributed to the observed environmental changes (EWG, 2019). Dry-waste poultry operations are not subject to the same waste distribution constraints, regulatory permitting process, or location disclosure requirements as the liquid waste CAFOs that we examined in this study (Copeland, 2010). Considering these and additional drivers and assessing their individual and cumulative impacts is warranted in future work.

The CAFO-associated land conversions detected could impact environmental health. For instance, the conversion from savannas to croplands and shrublands has been linked to losses of organic matter and biodiversity (Kronberg & Ryschawy, 2019). Additionally, excess nutrients in water bodies in NC, which can impact water quality conditions, has been shown to be influenced by increases in animal agriculture and the decline of wetland cover (Rothenberger et al., 2009). A study tracing nutrient pollution from CAFOs in NC pointed out that stream of closer proximity to CAFOs presented heavier pollution signatures when compared to unimpacted streams (C. N. Brown et al., 2020). This pattern has also been recently observed in Wisconsin, U.S. (Raff & Meyer,

2021a). A recent study reported impacts to the natural ecosystem in MI, where conversion of ~37% of grassland to croplands occurred between 2008 and 2013, which reduced biomass availability (Mladenoff et al., 2016). This conversion may be link to the increase of CAFO operations in the region. Furthermore, deforestation is responsible for a variety of negative impacts globally. On average, the clearing of Amazonian forests to create pastures for livestock is responsible for 56% of losses in abundance, biomass, richness, and diversity of soil macrofauna and microbial communities in Brazil (Franco et al., 2019). In analyzing environmental degradation, these findings indicate that NC ecosystem condition was potentially impacted by its long history of CAFO activities. Previous studies on the relationship between cash crop expansion and loss of natural habitats nearby swine CAFOs (Warner, 1994; Thu & Durrenberger, 1998), and the increase in swine operations between 2011 and 2017 in NC (Walljasper, 2018) support the findings of decreasing ET and increasing LST trends. Surprisingly, the increase of both day and nighttime LST in the vicinity of CAFOs reveal patterns reminiscent of urban heat island effects (Estoque et al., 2017), which warrants further study.

Government initiatives may be able to mitigate the impacts that CAFO activities can have on the environment. In this study, I observed that PTC generally increased over time in NC potentially due to conservation strategies. One example is the program to protect riparian buffers and conserve or restore land in selected river basins in NC including, the Neuse River, Tar-Pamlico River, and Catawba River (NCDEQ, 2020). Most of the increase in forest occurred in the headwater areas of these basins, which can help prevent pollutants and nutrients from different sources, including CAFOs, from reaching

the streams. Interestingly, this conservation strategy program is not implemented statewide. The Cape Fear River basin, where most of the liquid-waste CAFO facilities are located, is an example of one of the basins in NC where the riparian protection program has not yet come into force. This lack of environmental regulation enforcement is also seen in Asia. A survey in the 2000s found that over 90% of the CAFOs in China were established without conducting proper environmental impact assessments (EIA) (Hu et al., 2017). The assessment of the environment prior to the establishment of these operations could help to better site and design these facilities so that impacts of land conversion for feeding, improper manure handling and disposal, and potential nearby water body contamination from overflow of storage lagoons can be avoided or mitigated. This study suggests that CAFO-impacted areas, specifically those within 15 km radius from these operations, may be critical zones for conducting environmental impact assessment prior to CAFO establishment and for implementing mitigation strategies for impacts linked to changes in land use and environmental quality.

Two key factors may have affected detection of the CAFO-associated impacts observed. First, I only considered regulated CAFOs with publicly disclosed location information. While recent advances are making it easier to identify all large operations (Handan-Nader & Ho, 2019a; Prenafeta-Boldú & Kamilaris, 2019), location information for all animal operations is not usually accessible in most states. This is also an issue in other countries. The rapid geographic expansion of these operations in the recent decades has been largely unnoticed due to the lack of data in Brazil (Vale et al., 2019). Due to lack of accountability, unregulated facilities that are often not included in the publicly available

data may also have a higher environmental impact than their regulated counterparts and should therefore be considered in future studies. Secondly, this study exclusively used satellite image datasets representing the months of June through September, the growing season for most crops in the U.S. and certainly for these study areas. The literature has shown that croplands are positively related to ET, NDVI, and LAI parameters used to examine degradation trends in this study (Asner et al., 2003a; Duchemin et al., 2006; Tesemma et al., 2014). In fact, a study identified NDVI increases over time (normally associated with improved environmental quality) due to the increase in croplands principally near CAFOs (Qi et al., 2017). This relationship may have affected the ability to effectively track environmental quality changes in MI where the increase in cropland extent over time was much greater than that observed in NC during the study period. These nuances should be examined further in future research.

New environmental and agricultural management policies should consider the type of CAFO being regulated, in addition to whether a CAFO uses a liquid waste management system. For instance, NC has permits for distinct CAFO types, while MI CAFO facilities still operate under one general permit. Our study revealed that the type of animal production may be an important determinant of the scope of environmental impacts. We found that distinct CAFO types were associated with unique signatures of land use change. In NC, losses of savanna, wetland, and grassland were more frequent near swine CAFOs, while forest loss was more common near dairy CAFOs. Beef and dairy CAFOs have expanded to the Amazon biome in the recent decades as an alternative to reduce deforestation from implementation of pasture-based systems (Vale et al., 2019). The

argument of reduced in-property deforestation rates of beef and dairy CAFOs has allowed faster expansion of these operations while the off-site effects were still not accounted for. However, this study reveals that the off-property effects of dairy CAFOs can be substantial and lead to additional off-site deforestation within 15 km radius. Additionally, MI results suggest that the early stage of CAFO establishment as well as the type of CAFO can explain differences in signatures of land use change and environmental quality. While NC had 0% increase in the establishment of these operations in the recent decade, MI is among the states with highest increase in dairy operations in the U.S., with approximately 41% increase from 2011 to 2017 in the establishment of CAFOs statewide (Walljasper, 2018). Changes in land use policy, design and permitting could help mitigate future loss of environmental quality in MI and beyond. The intensified beef production and its inefficient manure management system led to the degradation of 15 million ha of grasslands in Brazil (FAO, 2013). This study found that grassland losses were 50% greater in MI (in agreement with the increase in CAFO operations in the region) compared to NC. Changes in policies or new initiatives are necessary to reverse the impact of these losses. These results advance our knowledge of the spatial extent and intensity patterns of impacts from liquid-waste CAFOs. These methods and findings can support new strategies and policies as well as help address the UN sustainable development goals, such as goals 6 (ensure water quality), 12 (sustainable production patterns) and 15 (protect natural resources) (United Nations, 2015).

The spatial distribution of CAFOs should be considered and monitored, as part of future policies and regulations when considering new permits. The number of CAFOs

operating in both states differed significantly; MI has approximately 13% of the total number of permitted CAFOs operating in NC. Furthermore, NC has a longer history of CAFO development than MI. This may explain why the land cover change and the environmental degradation analysis differed between the two states (i.e., the systems may be at different stages of transformation) and could be considered for future management of CAFOs. MI CAFO locations are more spatially dispersed compared to NC where CAFOs occur at high densities, primarily in the eastern Coastal Plain region (Figure 12). This spatial aggregation may lead to increased impact per unit of land area, exacerbating individual impacts (Carrel et al., 2016; Miralha & Kim, 2018). We found that certain impacts were concentrated within areas where swine and dairy CAFO-impacted areas overlapped spatially. This adds support to a growing body of evidence that the density of CAFOs may affect the nature and severity of impacts (Son et al., 2021).

3.4. CONCLUSION

This study presents previously unquantified spatiotemporal impacts from regulated CAFOs and significantly advances our understanding of previously unknown consequences of CAFO management and expansion. However, more information is needed to properly manage these systems. For example, in the U.S. there is no public database of where regulated and unregulated animal operations are, making it difficult to properly measure impacts and develop new solutions (Handan-Nader & Ho, 2019a). This study highlights the need for improved strategies to manage these complex networks of individually operating, yet collectively impactful CAFOs. From this work it is possible to infer that the associated land use change and environmental degradation observed is likely

a function of the type of CAFO, length of time CAFOs have been present, number and density of CAFOs present, and regulatory influences. While this study focused on land use change and environmental impacts, more work is needed to expand these methods to consider other drivers of change, quantify the specific effects of CAFO clustering on impacts, and gather more data regarding manure management (e.g., timing and application methods, transfer off site). Finally, given that many of the changes identified here may be driven by land use change occurring on lands where manure is applied, more work is needed to understand the barriers to economically viable use of manure as a fertilizer. The methods used here for both land use change and environmental degradation analyses are well-established in research and can be applied to explore degradation from CAFOs in other regions, and to explore other potential drivers of changes. Based on the results and literature, it is likely that similar trends could be observed in other areas as CAFOs continue to expand in the U.S. and globally. New techniques that help uncover unregulated CAFO facilities will enable consideration of these facilities in future analyses (Handan-Nader & Ho, 2019a). As information regarding CAFOs continues to expand (NRDC, 2019), this work can be leveraged to help to manage these food production systems for environmental sustainability.

CHAPTER 4

THE SPATIAL ORGANIZATION OF CAFOS AND ITS RELATIONSHIP TO WATER QUALITY IN THE U.S.

As discussed in chapter 1 section 1.4., the spatial clustering of pollution sources may exacerbate environmental impacts. However, this spatial component has yet to be evaluated for the industrialized animal agriculture sector. In this chapter, I hypothesize that CAFO-clustered watersheds are likely to present higher concentrations of total phosphorus (TP) and total nitrogen (TN) than CAFO-dispersed basins. I test this hypothesis by examining the relationship between the spatial pattern (i.e., clustering or dispersion) of CAFOs in 15 states across the U.S. and the TP and TN flow-weighted mean concentrations per watershed present in these states. To distinguish the impact of clustering from the number of CAFO present in a watershed, I also investigated the linear relationship between TN and TP concentrations, the number of CAFOs per watershed, and each watershed CAFO spatial pattern. I also gathered the states that provided animal number information per CAFO and evaluated the relationship between the spatial cluster of the number of animals per CAFO and TP and TN concentrations. This study brings a comprehensive understanding of the relationship between the spatial organization of CAFOs and water quality, provides the basis for further regulation of livestock production, and calls for spatially explicit policy strategies to prevent environmental impacts associated with intensified animal agriculture in the U.S.

4.1. METHODS

4.1.1. CAFO locations database

The 2003 CAFO rule stated that CAFOs had the duty to apply for NPDES permits, which requires information about the location of the facility as well as requires a nutrient management plan. However, the 2008 revision of this rule established that a CAFO operator has no longer the duty to apply for the permit if it was determined that the facility has no potential to discharge manure, litter, or process wastewater to waters of the U.S. (WOTUS) (Rosov et al., 2020). This revision complicates our ability to identify where these large animal farms are and where they are potentially land-applying manure. Adding to this change in regulation, additional rules protect the farmer's privacy right to not disclose their location (Steinzor & Huang, 2012). This explains the lack of information about the location of these operations for all the states in the U.S. as well as the unknown locations of unregulated farms (i.e., operations that do not meet the threshold of animals required to apply for a permit or do not discharge in waterways) which tend to operate without much regulation. Attempting to ameliorate this gap, I gathered and digitized CAFO locations in 15 U.S. states (Table 6). Most of the data were publicly available, expect for 3 states (Mississippi, Wisconsin, and Oregon) in which the data was provided by the state agency responsible via public record request. Each state data required georeferencing, processing, or digitizing techniques to meet the general format in this study. Each state data required georeferencing, processing, or digitizing techniques to standardize to a useable format in this study. This data handling process resulted in an ArcGIS point shapefile database with attributes of longitude, latitude, state, and number of animals. The animal

numbers were converted to animal units based on the definition established by the Code of Federal Regulations (1979) (Appendix C –3.1. Section 1).

Table 6. Number of CAFOs per state and the source of each data that composed the database developed for this study.

State	Acronym	Source	Date Accessed	# *CAFOs	# Animal	Animal Type
Alabama	AL	Alabama Department of Environmental Management (ADEM)	Sep 22, 2020	220	Y	N
Arizona	AZ	Arizona Department of Environmental Quality (ADEQ)	Mar 1, 2020	117	N	N
Florida	FL	Florida Department of Environmental Protection (FDEP)	Aug 20, 2020	416	N	N
Indiana	IN	Indiana Department of Environmental Management (IDEM), Office of Land Quality (OLQ)	Sep 21, 2020	1784	Y	Y
Iowa	IA	Iowa Department of Natural Resources (IDNR)	Nov 8, 2019	12367	Y	Y
Michigan	MI	Michigan Department of Environment, Great Lakes, and Energy (EGLE- MiWaters)	May 3, 2020	288	N	N
Minnesota	MN	Minnesota Pollution Control Agency (MPCA)	Sep 23, 2020	824	Y	Y
Mississippi	MS	Mississippi Department of Environmental Quality (MDEQ), OPC Environmental Permits Division	Apr 15, 2021	1490	N	Y
Missouri	MO	Missouri Department of Natural Resources (MDNR)	Sep 21, 2020	655	Y	Y
North Carolina	NC	The North Carolina Department of Environmental Quality (DEQ)	Sep 4, 2020	2526	Y	Y
Ohio	OH	Ohio Environmental Protection Agency (OEPA)	Sep 21, 2020	31	N	N
Oregon	OR	Oregon Department of Agriculture (ODA)	Oct 1, 2020	517	Y	Y
Pennsylvania	PA	Department of Environmental Protection (DEP)	Oct 8, 2020	413	Y	N
Tennessee	TN	TN Department of Environment & Conservation	Sep 28, 2020	51	N	N
Texas	TX	Texas Commission on Environmental Quality (TCEQ)	Apr 29, 2021	428	N	N
Wisconsin	WI	Wisconsin Department of Natural Resources (WDNR)	Sep 4, 2020	311	Y	Y

*CAFOs in this study does not align with the federal definition. The definition of these operations per states tend to vary due to regulatory differences.

Regulations are inconsistent among states (Rosov et al., 2020) and some states provide information on more than just regulated operations, which complicates the understanding of this compiled database. States like Oregon and Iowa, maintain a database

of small AFOs and CAFOs that do not discharge to waters of the U.S. These small AFOs when clustered may pose a risk to their surrounding environment. Indiana has chosen to regulate large animal farms under their own state-level permit, rather than the NPDES system, so its database includes more operations than would be regulated by NPDES (US EPA, 2015). Therefore, our CAFO database is composed of both large permitted and some unpermitted CAFOs and AFO facility locations when available (Figure 18). Some of the states also provided animal number and animal type information seen in Table 6. Note that most of the states in this study only provided the location of large operations, because of that I use the terminology CAFO to refer to all farms in this database.

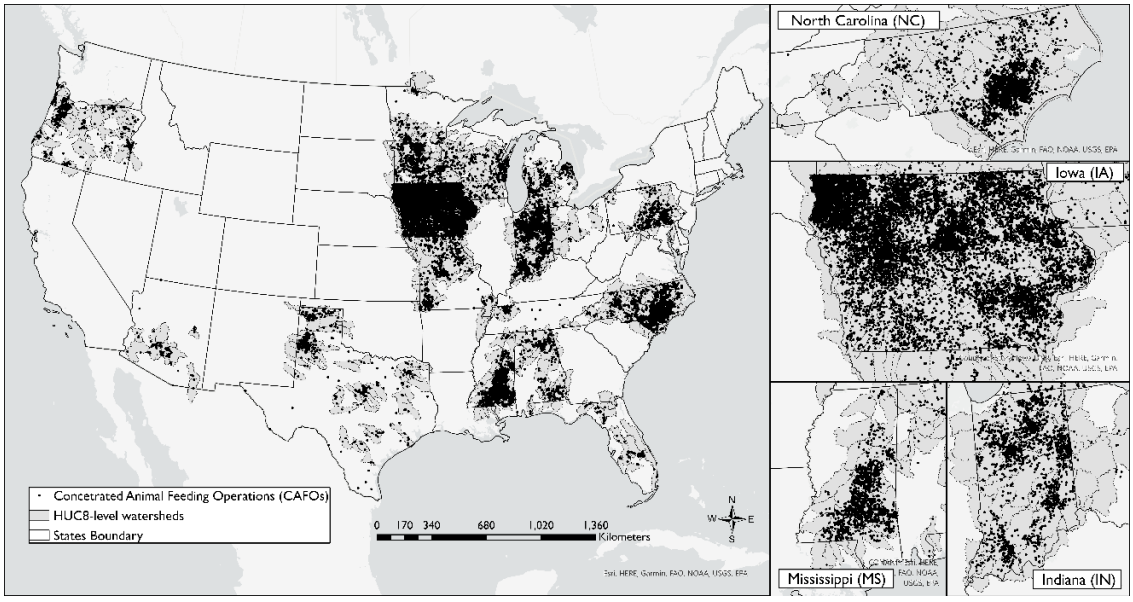


Figure 18. Geographical location of Concentrated Animal Feeding Operations (CAFOs) used in this study and their respective hydrologic units (HUs – code 8).

4.1.2. Water quality data

The U.S. Geological Survey released the water-quality concentration and streamflow data inputs used to develop the Spatially References Regressions on Watershed Attributes (SPARROW) models (Saad et al., 2019). They screened approximately 5,200 streamflow and 3,300 nutrient sites sampled by 137 agencies and organizations in the U.S from 1999 to 2014. I evaluated these inputs seasonally and selected stations with streamflow (m^3/s), TP (mg/l), and TN (mg/l) data available from 2005 to 2014 (Appendix C 3.2. Section 2). I selected this temporal range due to the intensification of animal agriculture in the U.S. in this period (von Keyserlingk et al., 2013; Key et al., 2017; Walljasper, 2018; McDonald et al., 2018). I matched the closest streamflow stations (based on the Euclidian distance criteria) to each TP and TN station to calculate the seasonal average flow-weighted mean concentration (FWMC) of TN and TP (Meals et al., 2011, 2013; Richards & Baker, 1993). I calculated FWMC by dividing the total load over the estimation period by the total streamflow of the same period (Equation 4.1):

$$FWMC = \frac{\sum_1^n c_i * t_i * q_i}{\sum_1^n t_i * q_i}$$

Where c_i represents the i^{th} sample concentration in mg/l; $t_i$represents the time window for the i^{th} sample (day is the time window in this study); and q_i is the flow registered by the station closest to the station that registered the i^{th} sample concentration (NCWQR, 2005). Streamflow, TP, and TN were matched daily from 2005 to 2014, then averaged by season.

4.1.3. Nearest Neighbor Index (NNI)

The Nearest Neighbor Index (NNI) or the Average Nearest Neighbor analysis package in ArcGIS pro 2.8 calculates the division between the observed distance among the point features and the expected distance based on the total area within which the point features are located. The observed distance is the average distance between each point and its nearest neighbor (Ebdon, 1985). I used the *arcpy* package to perform the NNI calculation among CAFOs within each HUC8 in this study (Figure 18). If the observed mean distance is less than the expected, NNI will be less than 1 and the CAFOs within the HUC8 area are considered clustered (i.e., CAFO-clustering). If the observed mean distance is greater than the expected, NNI is greater than 1 and a dispersion pattern is observed among the CAFOs (i.e., CAFO-dispersion) in a certain watershed. This pattern is validated based on a *z*-score calculation which is converted to *p*-value, allowing us to tell if the clustering or dispersion pattern observed is significant or not. In this study, I set a significance level of 0.1 ($p < 0.1$). I also scaled the NNI from most negative (dispersion) to most positive or zero (clustering) to observe the linear association between water quality and the spatial pattern detected. For the analysis, I only selected watersheds with more than 2 animal feeding operations. This analysis was used to evaluate if the significant clustering and dispersion of CAFO displays different water quality signals within the 15 states selected.

4.1.4. Local Moran's *I* index

I gathered the states that provided information on type and number of animals per CAFO and calculated their corresponding Animal Units (AUs) following the federal code

AU definition (Appendix C - 3.1. Section 1; *Code of Federal Regulations*, 1979) to scale animal numbers across animal types. I used the AU of each CAFO to investigate if the clustering of AUs influenced TP and TN concentrations. To identify the local spatial patterns associated with animal numbers, we applied a local indicator of spatial association (LISA) (Anselin, 1995). CAFOs with a certain number of animals can be clustered (spatial clusters) or exist individually (spatial outliers). In this study, a spatial cluster is associated with a CAFO with high AUs surrounded by other CAFOs with high AUs (i.e., High-High cluster; HH). In contrast, a spatial outlier is detected when a CAFO with high AUs is surrounded by CAFOs with normal or low AUs (i.e., High-Low outlier; HL). If a CAFO with low AU is in the vicinity of CAFOs with high AUs, this pattern is considered a Low-High outlier (LH). A spatial cluster can also occur when a CAFO with low AUs is surrounded by other CAFOs with low AUs (i.e., a Low-Low cluster; LL). I identified these patterns using the Local Moran's *I* index calculated in ArcGIS Pro 2.7 (Equation 2; Anselin, 1995; Zhang et al., 2008) :

$$I_i = \frac{x_i - \bar{x}}{\sigma^2} \sum_{j=1, j \neq i}^{n} [w_{ij} (x_j - \bar{x})]$$

where x_i is the value of the variable x at location i; $\bar{x}$ is the average value of x with the sample number of n; x_j is the value of the variable x at all the other locations (where $j \neq i$); σ^2 is the variance of variable x; and w_{ij} is a spatial weight matrix which can be defined as the inverse of the distance d_{ij} among locations i and j. In this study, the weight w_{ij} is determined using a 15 km fixed distance: samples within 15 km are given the same weight, while those outside the distance band are given weight of 0. This fixed distance was

established based on studies that investigated manure hauling distances and its impact to the environment (Long et al., 2018; Furiness et al., 2019; Miralha, Muenich, Schaffer-Smith, et al., 2021). A high positive Local Moran's *I* value implies that a certain location has similarly high or low values as its neighbors (spatial cluster) while a high negative *I* indicates that a certain location has distinct values as its neighbors (spatial outlier). Spatial outliers are an indication of individual hotspots, while spatial clusters indicate regional hotspots. Note that to perform this analysis, I first projected the data to avoid Chordal distance calculations since the area of analysis is larger than 30 degrees. We performed a conditional permutation with Local Moran's *I* to avoid the assumption of normal distribution when calculating the *p*-value. In this study, all the Local Moran's *I* indices were tested using 999 permutations, and the significance level was set at < 0.05.

4.1.5. Analysis of the spatial organization of CAFOs and water quality relationship

NNI: I calculated the seasonal FWMC for TP and TN per HUC8 (Figure 18 and 19) and compared these concentrations between watersheds that presented overall CAFO-clustering or CAFO-dispersion patterns. I also compared the number of CAFOs per HUC8, their respective TP and TN seasonal FWMC, and their NNI spatial pattern detected to better distinguish if the water quality conditions are more likely a function of the number of CAFOs or the way CAFOs are spatially organized within a watershed.

Local Moran's *I*: To compare the local spatial association of AUs to water quality, I first detected the closest water quality station to each CAFO with AU information

available. I used each individual station to calculate the respective TP and TN seasonal FWMC for each CAFO and aggregated the concentrations based on the local spatial pattern detected (High-High cluster (HH); Low-Low cluster (LL); Low-High outlier (LH); High-Low outlier (HL)). I performed a simple median difference statistic test (Kruskal-Wallis) to evaluate if there was a statistically significant difference among the seasonal median of TP and TN concentrations per scenario.

4.2. RESULTS

4.2.1. NNI and water quality

From the 443 HUC8 watersheds, 355 presented CAFO-clustering patterns while CAFO-dispersion was detected in only 88 watersheds (Figure 19). By excluding watersheds without significant patterns ($p< 0.1$), I had a total of 249 CAFO-clustering watersheds and 31 CAFO-dispersion watersheds subsequently available to compare with TP and TN concentrations. States such as IA, NC, MO, IN, and FL presented overall clustered patterns, while CAFO-dispersion watersheds were detected in states in the west (OR AZ,), south (AL), and northern (WI, MN) parts of the U.S. Although the number of CAFOs in IA basins was higher than the number of operations in the other 14 states in this study and 50% of the watersheds within IA state had non-significant ($p > 0.1$) clustering or dispersion patterns and were excluded from the water quality scenarios.

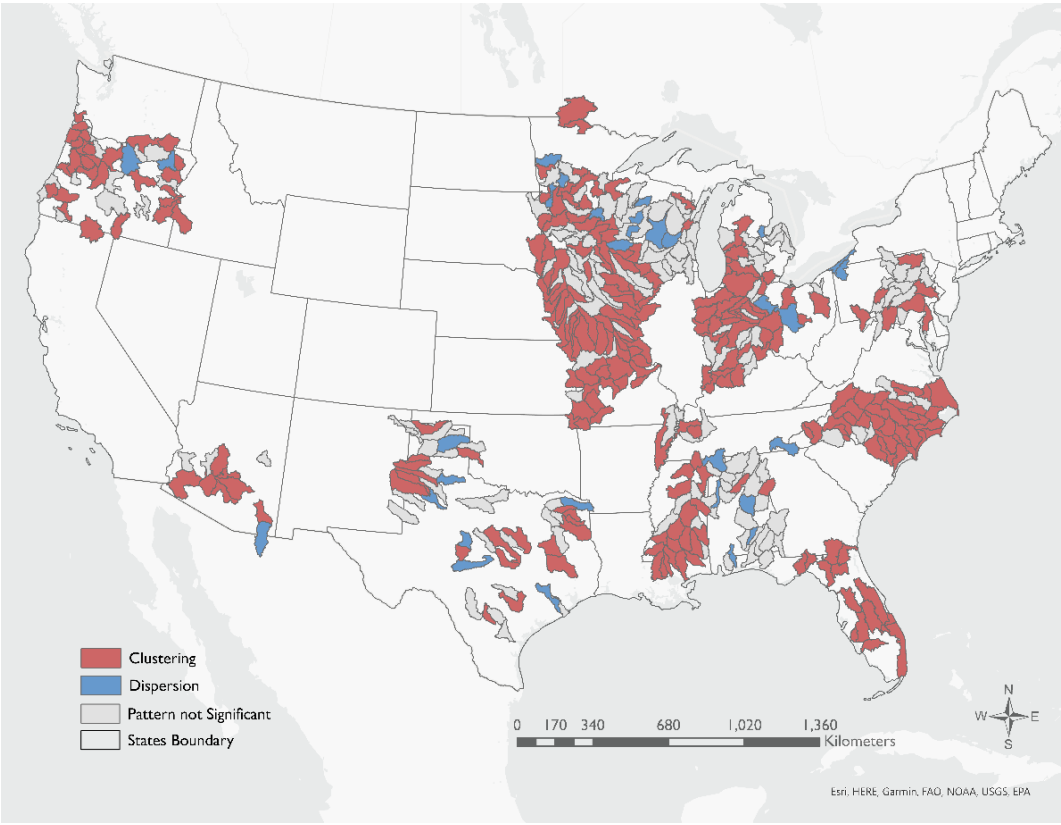


Figure 19. Significant ($p<0.1$) clustering and dispersion patterns of CAFOs in 15 U.S. states.

I compared both clustering and dispersion distributions to TN and TP concentrations and we found that overall TN and TP concentrations are significantly higher in watersheds where CAFOs are clustered than in CAFO-dispersion watersheds *(*Figure 20; *Kruskal-Wallis p-value <0.001*). The median TN and TP concentration in CAFO-clustered watersheds was approximately 0.1 mg/l and 0.6 mg/l higher than in CAFO-dispersed basins. The highest TP and TN concentrations were overall associated with CAFO-clustering hydrologic units. Median FWMC TP concentrations in CAFO-clustered sites was above 0.1 mg/L, while CAFO-dispersed FWMC TP concentrations were below this value. While there are no federal regulations for stream TP concentrations, states have

begun adopting TP criteria (US EPA, 2016), and some like Wisconsin set TP criteria for streams to be less than 0.1 mg/L (Wisconsin Legislature: NR 102.06, 2020). Seasonally, these concentrations patterns for both TP and TN also held, principally during spring and summer seasons when manure is usually land-applied in nearby areas (Appendix C – 3.2.1. and 3.2.2. SI – Figure 1 and 2).

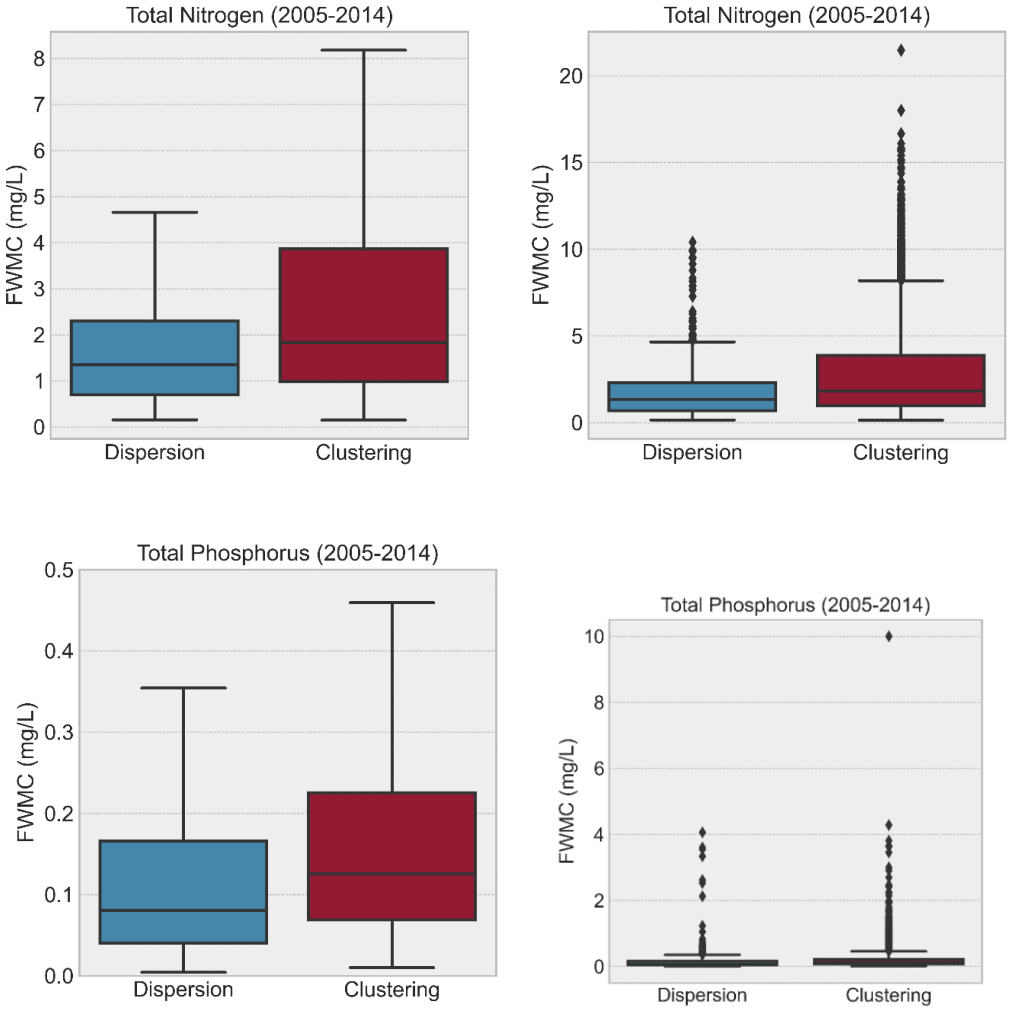


Figure 20. *Left* - Overall Total Phosphorus (TP) and Total Nitrogen (TN) flow-weighted mean concentration (FWMC) distribution per watershed's spatial pattern. *Right*- Same overall pattern FWMC distribution of TP and TN but displaying outliers.

I compared the clustering and dispersion patterns with the number of CAFOs per watershed and their respective TP and TN concentrations (Figure 21). The results suggest that watersheds tend to exhibit clustering patterns as the number of CAFOs increase, which is intuitive. For TP, the linear relationship between the number of CAFOs and the FWMC was significant and moderately positive in clustered watersheds (r = 0.3; p <0.001) while for the CAFO dispersed basins this relationship was only significant at 0.1 level (r =0.42; p =0.059).

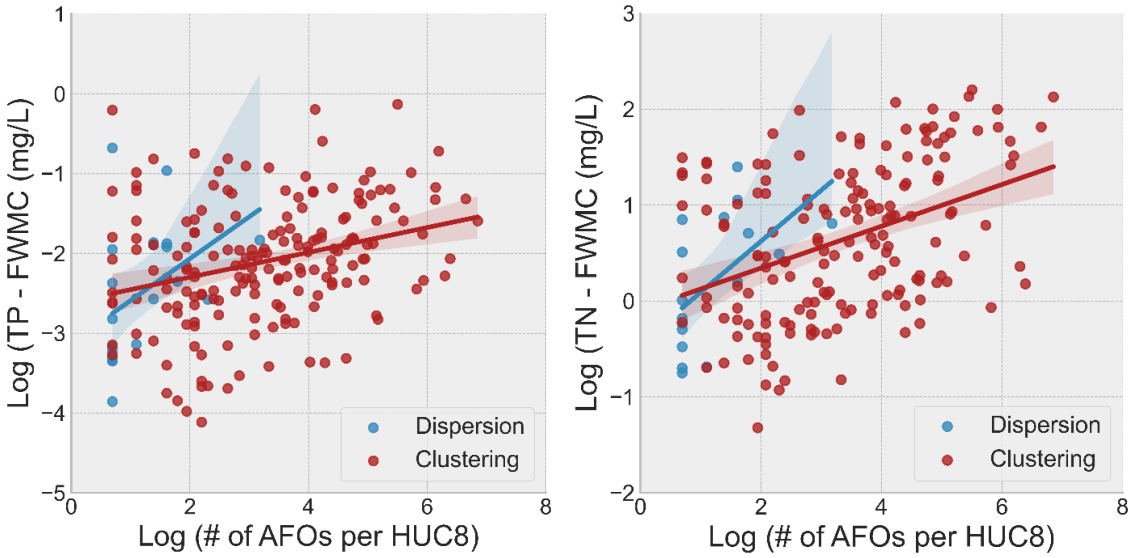


Figure 21. Linear relationship between number of CAFOs and flow-weighted mean concentration (FWMC - mg/l) per nutrient and significant spatial pattern (Clustering or Dispersion).

To better understand the relationship between spatial clustering of CAFOs and water quality, I selected watersheds with less than 20 CAFOs that were also classified as CAFO-clustered and CAFO-dispersed, because CAFO clustering increases with the number of CAFOs in a given watershed, restricting available spatial patterns. This analysis

restated what we previously found in figure 20. Higher concentrations of TP (i.e., 10 mg/l) and TN (i.e., above 15 mg/l) were found in CAFO-clustering watersheds when compared to the ones classified as CAFO-dispersed (Appendix C – 3.2.3. SI – Figure 3). However, only the median TP concentration difference between the 2 scenarios was statistically significant at a level of 0.05 (TP *Kruskal-Wallis p <0.001*; TN *Kruskal-Wallis p =0.11*)

4.2.2. Local Moran's *I*, animal units per CAFO, and water quality

From the 15 states analyzed in this study, only 9 had data available on animal number and type per CAFO. The number of AUs per farm ranged between 0.06 to 50,000, with a median of 528, and a standard deviation of 1.672 AUs. Using the AUs, we performed the local spatial autocorrelation analysis (Local Moran's *I*).The indices of Local Moran's *I* ranged from -10.40 to 476.15. CAFOs with statistically significant Local Moran's *I* indices ($p< 0.05$) are displayed according to their spatial cluster category while non-significant patterns were excluded and displayed in gray (Figure 22). Among the 19,084 CAFOs with AU information, 23.5% presented non-significant spatial clustering patterns. Approximately 54% of the CAFOs were categorized as LL clusters, which means that these CAFOs have low or normal values of AUs and so does their neighboring farms. Most of these LL-cluster farms were in IA, MN, and west NC. This revealed that ~13% of CAFOs were part of the HH cluster category (i.e., both centroid and neighboring operations with high AUs). HH clusters were mostly detected in states such as WI, IN, east NC, and MO. About 9.5 % of the total number of CAFOs were detected as spatial outliers (i.e., HL and LH outliers). While I found several LH outliers in OR, PA, east NC, and IN, most of the

HL outliers were in IA and MN. Spatial patterns in AL were generally not significant when compared to the other 8 states in this analysis.

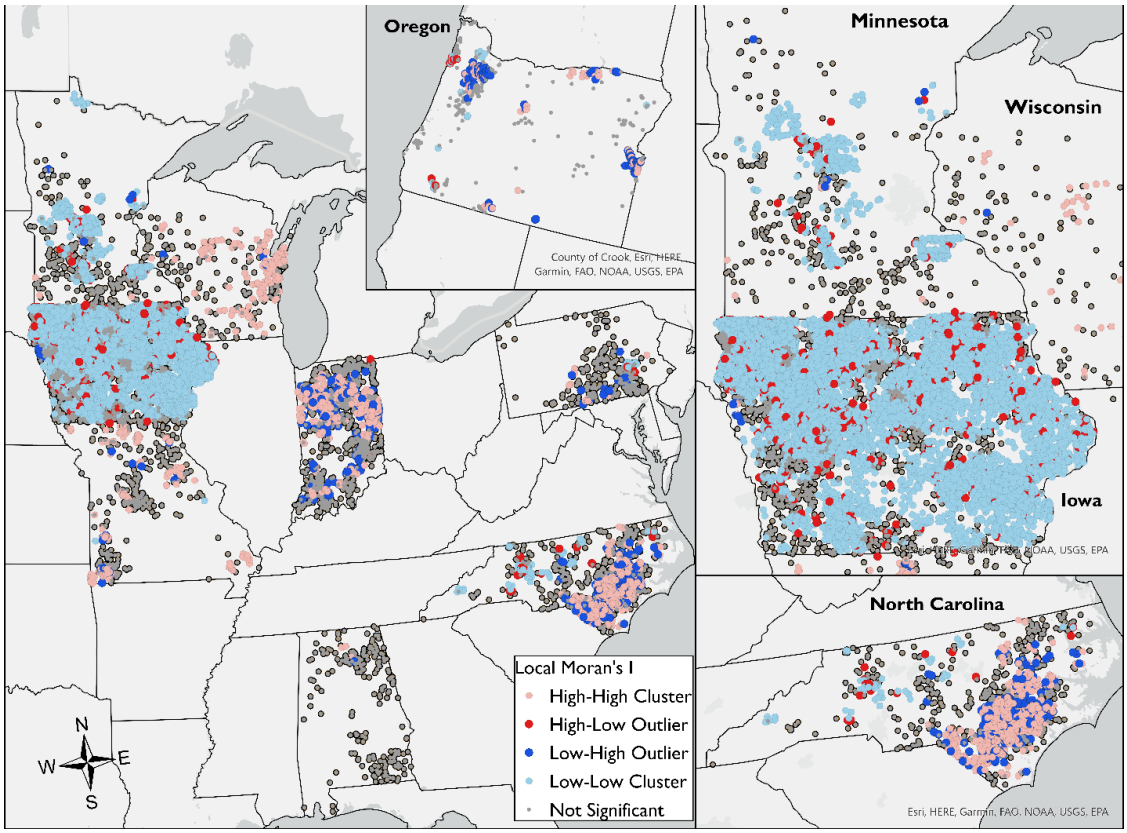


Figure 22. Local Moran's *I* result displayed by spatial cluster and outlier at 0.05 significance level.

Comparing the spatial cluster and outlier types with their closest station's nutrient concentrations, I found that both LL (i.e., light blue) and HL (i.e., red) clusters are associated with higher concentrations of TP and TN (Figure 23). First, local Moran's *I* was also able to detect a regulatory threshold of 1000 animal units (i.e., log (7)) when differentiating the low from the high spatial clusters and outliers (Figure 23 – scatter plots). Although farms with the largest AUs and with neighbors of similar high AU values (HH cluster; light pink) eventually produce large amounts of manure, their surrounding TP and

TN concentrations were significantly lower than other CAFOs with different spatial cluster and outlier classification. LH outlier CAFOs (i.e., blue) also displayed an overall lower TP and TN rates. The average TP concentrations for each cluster (HH, LL) and outlier (HL, LH) group were 0.16, 0.34, 0.36, and 0.16 mg/l, respectively. The maximum TP concentration for HH and LL clusters was 4.86 mg/l while for spatial outliers this concentration reached 3.9 mg/l. For TN, the average concentration per scenario was 2.1 (HH), 7.2 (LL), 7.4 (HL), and 2.40 (LH) mg/l. I tested if both TP and TN distributions were significantly different per scenario, and all scenarios were different at the 0.05 significance level (*Kruskal-Wallis p <0.0001*). Although the relationship between TN/TP concentrations and AUs was not very clear, when classified by cluster and outlier type, I was able to observe that lowest TP and TN concentrations were associated with scenarios when large CAFOs are clustered (HH) and when a small CAFO is surrounded by large operations (LH). These results have the potential to change current CAFO policies in place.

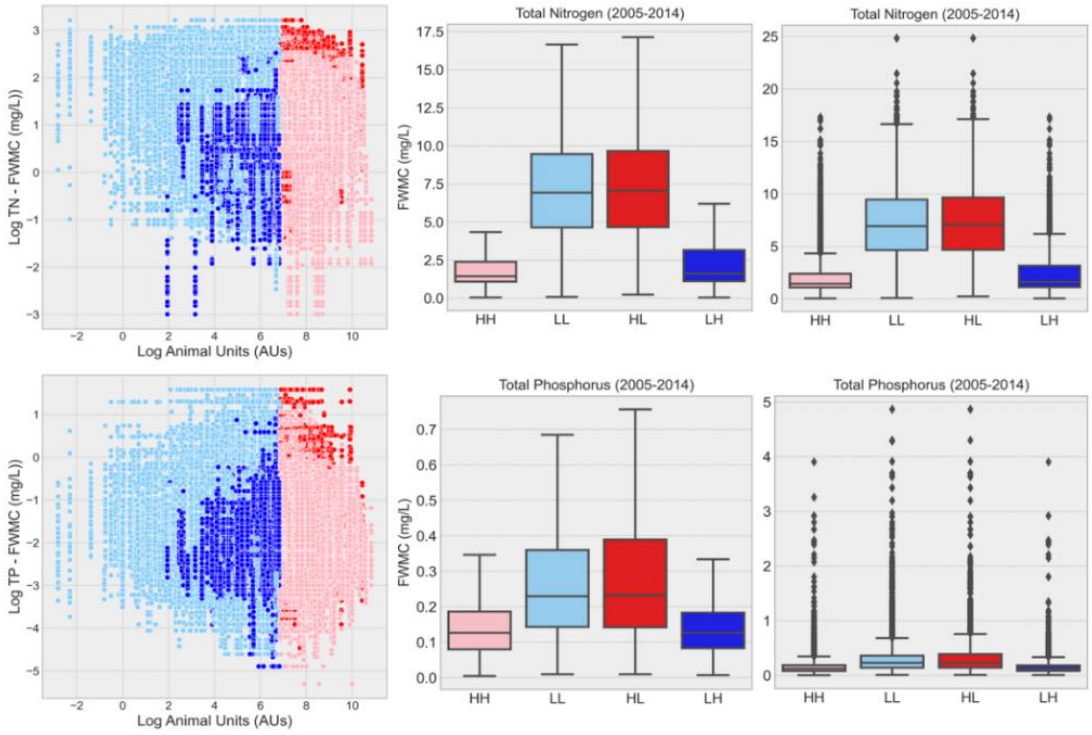


Figure 23**.** Total Phosphorus (TP) and Total Nitrogen (TN) flow-weighted mean concentrations (FWMC -mg/l) per spatial cluster (High-High – HH = a large CAFO surrounded by Large CAFOs; Low-Low – LL = a small CAFO surrounded by small CAFOs) and outlier (High-Low – HL= a large CAFO surrounded by small CAFOs; Low-High – LH = a small CAFO surrounded by large CAFOs) scenario as well as animal units (AUs).

I also draw the relationship between positive Local Moran's *I* indeces of spatial clusters and water quality (Appendix C – 3.3.1. SI- Figure 6). I fit a linear model and found that as *I* becomes larger (HH cluster CAFOs) TP and TN concentrations tend to be lower, TP and TN are relatively higher when *I* is significant and closer to zero (LL cluster CAFOs). The animal type in each CAFO may also play a role in the concentrations

observed. Regions where LL clusters dominated presented different animal type distribution. IA had predominantly swine operations, while Cattle operations predominated in MN and east NC (Appendix C – 3.4.1. SI-Table 1). The way these states regulate these operations and how they consider the animal type in these policies may play a role in the water quality outcomes observed.

4.3. DISCUSSION

While studies have mentioned that the spatial aggregation of CAFOs is likely to exacerbate environmental impacts (Thurow & Thompson, 1998; Yang et al., 2016; Martin et al., 2018; Miralha, Muenich, Schaffer-Smith, et al., 2021), this study is the first to investigate this spatial component and reveal that the clustering of these operations by itself leads to stronger negative environmental outcomes. Specifically, I found that CAFO-clustered watersheds were negatively impacted in terms of water quality. This study revealed that HUCs with clustered CAFOs presented overall higher concentration of TN and TP than CAFO-dispersed HUCs. States in part of the U.S. corn belt such as IA, IN, MO, and OH, responsible for major applications of N from manure and chemical fertilizers and major drivers of the largest hypoxic zone in the U.S. (i.e., dead zone in the Gulf of Mexico) (Scavia et al., 2003; Turner et al., 2008; Glibert, 2020), were also states where CAFOs in general presented clustering patterns (Appendix C – 3.2.4. SI-Figure 4). A previous study investigated the water quality conditions in two watersheds also included in this study - the South Fork basin in IA and the Little Cobb basin in MN (Kalkhoff et al., 2016). They found that these two basins, with growing intensified livestock agriculture,

presented overall lower N and P concentrations, and in our study, we detected a random spatial organization of CAFOs in these basins (i.e., neither clustering nor dispersion of CAFOs). Additionally, this same basin in MN presented no significant spatial cluster or outlier when AUs and Moran's *I* was considered. These facilities may be sparse in space, managing their manure properly or beyond the watershed boundary, which warrants further study.

Based on the TN and TP concentration limits established by the U.S. Environmental Protection Agency (EPA) per ecoregions for rivers and streams (US EPA, 2013) as well as lakes and reservoirs (US EPA, 2020), the highest concentrations of TN and TP found in CAFO-clustering watersheds are likely to drive large algae blooms or other impacts from excess nutrients. The concentrations found in this study also exceed safe drinking water limits for N and limits beginning to be set by states for P (US EPA, 2016). Seasonally this finding also holds, principally during spring and summer, when manure is most likely to be land-applied (Appendix C – 3.2.1 and 3.2.2. SI – Figures 1 and 2).

The role of livestock in pollution has been ignored by nutrient reduction strategies in 12 states of the Mississippi River basin. These states, such as IA, MS, MO, and MN have the authority to monitor and implement strategies to reduce nutrient pollution. However, they have no resources to implement these strategies, because the subsidies directed to nutrient pollution reduction are held in a federal level, which potentially leads to the neglection of the role of CAFOs and manure hauling in water pollution (Secchi & Mcdonald, 2019). Regulators could use the concept of clustering or target CAFO-clustered watersheds in nutrient reduction strategies. The usefulness of spatial aggregation of point

sources of pollution will depend on the purpose of the assessment as well as on the spatial resolution of the monitoring program implemented (Schuwirth, 2020). This study demonstrated that the spatial arrangement of these operations matter when it comes to water pollution. Schuwirth (2020) also suggested that environmental assessments must be done at a catchment scale, because it is usually a manageable scale by many states (Secchi & Mcdonald, 2019). This study successfully demonstrated the efficiency of considering HUC 8-level basins as the targeted boundary area for the assessment of the spatial arrangements of CAFOs and its relationship to water quality. Note that other anthropogenic sources of pollution were not considered in this study, which warrants further study as most total maximum daily load (TMDL) implementations do not consider the spatial *arrangement* of emitters during their process.

An aggregation criterion was recently adopted by the Farm System Reform Act of 2021. It states that a CAFO should not be established within 3 miles from another CAFO, and the neighboring facility should not be under common membership or control. However, this criterion may be not enough to prevent pollution and further degradation. Previous research has established that liquid-waste CAFOs tend to haul manure within ~ 15 km radius from the operation facilities (Centner, 2012; Long et al., 2018; Furiness et al., 2019), but this distance increases when it comes to dry-waste CAFOs such as many poultry operations (Ribaudo et al., 2003). Due to manure-hauling distances and the fact most CAFOs tend to cluster near land resources to avoid high transportation costs (Ribaudo et al., 2003), operators tend to apply manure over crop N:P requirements (Kellogg et al., 2000; Long et al., 2018) likely leading to environmental degradation (Miralha, Muenich,

Schaffer-Smith, et al., 2021). In this study, I established that neighboring farms are usually 15 km apart because of manure hauling distances, and this criterion allowed us to detect several spatial clusters associated with worse water quality conditions (i.e., Local Moran's *I* and water quality analysis). For the clustering of enterprises to work in favor of the environment, there should exist diversity among the type of entities within the cluster (Deutz & Gibbs, 2008). Essentially, in a clustered industrial ecosystem, one facility's by-products or surplus should become another's raw material or energy (Garner & Keoleian, 1995). Strategies incentivizing exchange between cropland owners and CAFO operators should be considered to enhance nutrient recycling from these operations in the U.S. Another form of incentive is the concept of manuresheds currently implemented by the Arkansas nutrient surplus program (Spiegal et al., 2020). Additional regulatory incentives may help such as the dairy farms in MN that are required to have adequate land base for manure application prior to the permit acquisition (MPCA, 2016). The clustering of animals is corollary to the clustering of humans in cities. In cities we are able to economically and efficiently treat human waste products due to economies of scale. This same approach could be applied in animal cluster "cities", especially as advanced technologies like anaerobic digesters evolve. Regulatory incentivization and the clustering between CAFOs and other industries could be considered to promote environmental quality and foster the sustainability of these agricultural systems.

This study found that not only the spatial organization of these facilities matter, but also the number of animals they hold. Clustering of CAFOs has found to exist between enterprises of similar size classification in MO (Matisziw & Hipple, 2001). Specifically,

the only operations that underwent clustering were the largest ones (class IA in Missouri). Although the clustering of these operations was previously investigated, no previous study has ever linked their spatial organization to environmental conditions. Besides demonstrating that clustering is common among CAFOs with high AUs, we found that the spatial cluster most threatening to water quality in the U.S. is the cluster of CAFOs with low AUs (i.e., LL cluster – Figure 23). This finding has the potential to change CAFO regulations in the U.S. Most of the regulatory agencies focus on controlling large operations due to the amount of waste they produce, ignoring entities that do not meet the threshold of animals to be considered as a large CAFO or do not have the duty to apply for a permit because there is no potential of manure discharge to the waters of the U.S. (Ritzel, 2014). This individual-based regulation approach may fail when multiple entities add up to produce the same amount of waste as one large entity. As demonstrated in this study, the spatial autocorrelation of small CAFOs (i.e., lower AUs) was associated to higher concentrations of TP and TN while the cluster of large operations (i.e., HH) was in general within safe drinking water TP and TN concentration limits (Figure 23). This spatial autocorrelation phenomenon is also a characteristic of water quality (Miralha & Kim, 2018), meaning that nearby water quality stations upstream and downstream of the river network tend to resemble approximately same TP and TN concentrations. This raises a concern on the pollution of tributaries or main rivers closer to areas or waterways impacted by these small CAFOs, indicating not only a threat to the aquatic ecosystem but also to the health of nearby communities.

To account for the nutrient contribution from manure produced by these animal operations in watershed models, modelers often rely on average manure application rates to represent manure management practices in the whole watershed area in study, neglecting the spatial dynamics of animal agriculture and changes in management practices other than manure application and crop patterns (J. Wang & Baerenklau, 2015). When solely relying on application rates, models have not shown much significant improvement when modeling observed TP and TN loads (Apostel et al., 2021; Kast et al., 2021). More information on the location of these animal facilities, their distance to water bodies, and the number of animals they hold may improve the representation of these point and non-point sources of nutrient pollution in models. Most importantly, considering their spatial heterogeneity via metrics such as NNI and Local Moran's *I* have the potential to reveal the spatial component of these sources inherent in water quality that is still neglected in models.

I acknowledge the limitations of this study which warrant further research. First, other sources of pollution (e.g., human wastewater, fertilizer applications) were not accounted for in this study which may have influenced the water quality patterns identified. However, it does not disregard our findings since both NNI and Local Moran's *I* analysis illustrated similar regional clustering patterns, although these two techniques are based on different approaches. Second, I did not fully account for the clustering of animal types. Although I have provided the counts of CAFOs per animal type (Appendix C – 3.4.1 SI- Table 1) and broadly discussed the dominance of an animal type in states where high clustering was identified, the point analysis per animal type was not further investigated but it has the

potential to facilitate the elaboration of policies per type of CAFO. I also acknowledge that a lack of water quality data in the western U.S. (e.g., AZ, OR), principally for TN concentrations, could influence observed patterns. (i.e., AZ, and OR). However, by evaluating the patterns, I was still able to associate higher concentrations of TP for the CAFO-clustered watersheds in these regions (Appendix C – 3.2.5. SI- Figure 5). The lack of information on the location of these operations as well as the type and number of animals they hold also may have impaired our results. Although efforts to identify the location of these operations have been made (Chugg et al., 2021; Handan-Nader & Ho, 2019a), more information is needed to properly manage these systems across larger scales. Federal and state agencies along with researchers should come together to create a database of small and large animal farms to improve environmental assessment and monitoring of nutrient pollution.

4.4. CONCLUSION

This study reveals spatial patterns of CAFOs and its relationship to water quality previously unaccounted for in the literature. With nearest neighbor index (NNI), it was revealed that the spatial organization of CAFOs matters when it comes to the surrounding water quality conditions. Most importantly, I found that the spatial autocorrelation of these farms with respect to their animal units (Local Moran's *I* analysis) relates to water quality differently. The results show that although larger CAFOs tend to cluster over space (i.e., HH cluster), they may also better handle their excess manure as expected by federal and state environmental regulations. It also shows that it is the clustering of CAFOs with low

animal units (i.e., LL cluster) that may impact water quality the most. These patterns detected are likely to be associated with the *individualistic* approach to regulation that we have in the U.S. These results indicate the importance of gathering more information about the location of all animal agriculture entities (principally for better representing these entities in nutrient load modeling), as well as for new regulations or incentives considering the clustering pattern of CAFOs and other emitters prior to the issuance of permits.

CHAPTER 5

FINAL REMARKS AND FUTURE WORK

5.1. Takeaways

This dissertation explored novel approaches to better investigate uncertainties associated with the excess of nutrients in the U.S., bringing insights on nutrient management strategies principally when it comes to climate uncertainty and underregulated point sources of pollution. Our results suggest that policymakers, watershed managers, and environmental agencies should consider: (1) a robust climate change scenario analysis when it comes to the prediction of nutrient loads principally in watersheds with intensified animal agriculture; (2) implementing environmental impact assessments within 15 km from a CAFO facility, prior to the issuing of a permit, as an attempt to avoid extreme land use changes and subsequent environmental degradation driven by these operations; (3) regulating CAFOs (principally small unregulated facilities) based on their spatial organization principally when it comes to manure hauling distances, as this study has shown that the clustering of unregulated small CAFOs may be more detrimental to the water quality conditions in their surroundings than the clustering of larger, regulated CAFOs. Considering both the intensity of land use changes and spatial aggregation metrics of these animal operations can also improve watershed model nutrient load predictions under climate uncertainty. As these intensified food production systems grow with the population, these research findings advance modeling approaches, current environmental policies, and guide future decisions towards the reduction of nutrient loads and sustainable development in the U.S. and worldwide.

5.2. Mapping AFOs in the U.S.

This dissertation work, especially chapters 3 and 4 could have been greatly improved by having access to locations of all animal operations in the U.S. While few states provide data on regulated CAFOs, there are still many farms that operate just under the threshold to not be obligated to hold permits, which is the case of AFOs. AFOs tend to operate with a number of animals close to the regulatory thresholds, tend to cluster over space, and are significantly unmonitored by agencies (EWG, 2019). A complete database of AFO locations would allow for more accurate evaluation of the social and environmental impacts of these animal operations, which have yet to be rigorously quantified despite advancements in the detection of these facilities (Handan-Nader & Ho, 2019b). The advancements in mapping CAFOs have been based on object detection and pixel-based algorithms, but no study has considered developing a model that includes social, economic, and environmental metrics as predictors. I compared the performance of several machine learning algorithms based on socio-economic and environmental metrics associated with CAFO regions and results reached 99% of accuracy when detecting and distinguishing the type of these operations. The next step in this research is to scale this analysis to provide the first and most needed comprehensive database of AFO locations in the U.S. Limitations such as training and test sets for the model were overcome (Figure 24), but there is still a need to account for the irregular land parcel sizes at a national level to decrease the uncertainties associated with the implementation

of a grid-based model.

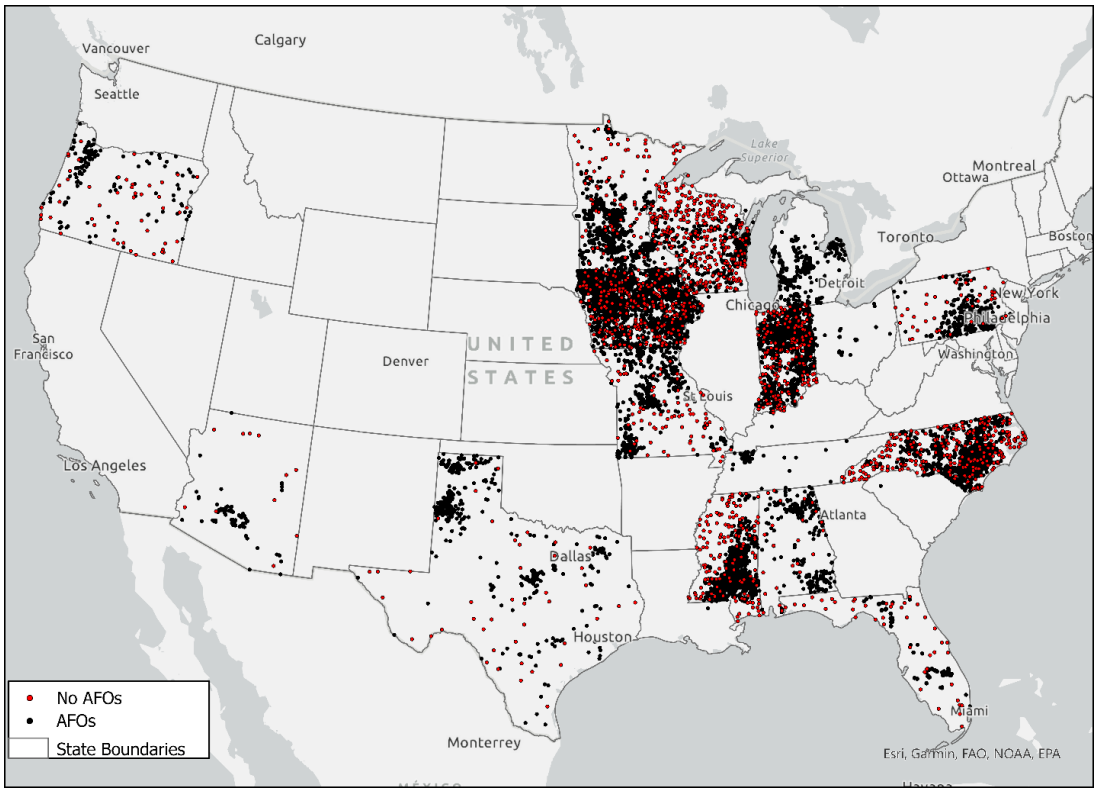


Figure 24. AFO locations and verified “no AFO” locations dataset developed for the training and test set for the modeling of these animal production systems over space

REFERENCES

Alatorre, L. C., Sánchez-Carrillo, S., Miramontes-Beltrán, S., Medina, R. J., Torres-Olave, M. E., Bravo, L. C., Wiebe, L. C., Granados, A., Adams, D. K., Sánchez, E., & Uc, M. (2016a). Temporal changes of NDVI for qualitative environmental assessment of mangroves: Shrimp farming impact on the health decline of the arid mangroves in the Gulf of California (1990–2010). *Journal of Arid Environments*, *125*, 98–109. https://doi.org/10.1016/j.jaridenv.2015.10.010

Alatorre, L. C., Sánchez-Carrillo, S., Miramontes-Beltrán, S., Medina, R. J., Torres-Olave, M. E., Bravo, L. C., Wiebe, L. C., Granados, A., Adams, D. K., Sánchez, E., & Uc, M. (2016b). Temporal changes of NDVI for qualitative environmental assessment of mangroves: Shrimp farming impact on the health decline of the arid mangroves in the Gulf of California (1990–2010). *Journal of Arid Environments*, *125*, 98–109. https://doi.org/10.1016/j.jaridenv.2015.10.010

Anh, P. T., Dieu, T. T. M., Mol, A. P., Kroeze, C., & Bush, S. R. (2011). Towards eco-agro industrial clusters in aquatic production: The case of shrimp processing industry in Vietnam. *Journal of Cleaner Production*, *19*(17–18), 2107–2118.

Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, *27*(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Apostel, A., Kalcic, M., Dagnew, A., Evenson, G., Kast, J., King, K., Martin, J., Muenich, R. L., & Scavia, D. (2021). Simulating internal watershed processes using multiple SWAT models. *Science of The Total Environment*, *759*, 143920. https://doi.org/10.1016/j.scitotenv.2020.143920

Asner, G. P., Scurlock, J. M. O., & Hicke, J. A. (2003a). Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, *12*(3), 191–205. https://doi.org/10.1046/j.1466-822X.2003.00026.x

Asner, G. P., Scurlock, J. M. O., & Hicke, J. A. (2003b). Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, *12*(3), 191–205. https://doi.org/10.1046/j.1466-822X.2003.00026.x

Austin, K. G., González-Roglich, M., Schaffer-Smith, D., Schwantes, A. M., & Swenson, J. J. (2017). Trends in size of tropical deforestation events signal increasing dominance of industrial-scale drivers. *Environmental Research Letters*, *12*(5), 054009. https://doi.org/10.1088/1748-9326/aa6a88

Ayres, R., & Ayres, L. (2002). *A Handbook of Industrial Ecology*. Edward Elgar Publishing. https://doi.org/10.4337/9781843765479

Baek, S., & Smith, C. D. (2019). Potential contaminant runoff from Californias dairy concentrated animal feeding operations (CAFOs): A geospatial analysis. *International Journal of Water Resources and Environmental Engineering*, *11*(1), 1–13. https://doi.org/10.5897/IJWREE2018.0803

Bayazit, M., & Önöz, B. (2007). To prewhiten or not to prewhiten in trend analysis? *Hydrological Sciences Journal*, *52*(4), 611–624. https://doi.org/10.1623/hysj.52.4.611

Bhowmik, R. D., Sankarasubramanian, A., Sinha, T., Patskoski, J., Mahinthakumar, G., & Kunkel, K. E. (2017). Multivariate Downscaling Approach Preserving Cross Correlations across Climate Variables for Projecting Hydrologic Fluxes. *Journal of Hydrometeorology*, *18*(8), 2187–2205. https://doi.org/10.1175/JHM-D-16-0160.1

Blain, G. C. (2015). The influence of nonlinear trends on the power of the trend-free pre-whitening approach. *Acta Scientiarum. Agronomy*, *37*, 21–28. https://doi.org/10.4025/actasciagron.v37i1.18199

Brands, E. (2014). *Siting restrictions and proximity of Concentrated Animal Feeding Operations to surface water*. https://doi.org/10.1016/j.envsci.2014.01.006

Brown, C., Mallin, M., & Loh, A. N. (2020). Tracing nutrient pollution from industrialized animal production in a large coastal watershed. *Environmental Monitoring and Assessment*, *192*. https://doi.org/10.1007/s10661-020-08433-9

Brown, C. N., Mallin, M. A., & Loh, A. N. (2020). Tracing nutrient pollution from industrialized animal production in a large coastal watershed. *Environmental Monitoring and Assessment*, *192*(8), 515. https://doi.org/10.1007/s10661-020-08433-9

Brown, J. R., & Carter, J. (1998). *Spatial and temporal patterns of exotic shrub invasion in an Australian tropical grassland*. 10.

Burkholder, J., Libra, B., Weyer, P., Heathcote, S., Kolpin, D., Thorne, P. S., & Wichman, M. (2007). Impacts of Waste from Concentrated Animal Feeding Operations on Water Quality. *Environmental Health Perspectives*, *115*(2), 308–312. https://doi.org/10.1289/ehp.8839

Cakir, R., Sauvage, S., Gerino, M., Volk, M., & Sánchez-Pérez, J. M. (2020). Assessment of ecological function indicators related to nitrate under multiple human stressors in a large watershed. *Ecological Indicators*, *111*, 106016. https://doi.org/10.1016/j.ecolind.2019.106016

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: An N-dimensional probability density function transform for climate model simulations of multiple

variables. *Climate Dynamics*, *50*(1), 31–49. https://doi.org/10.1007/s00382-017-3580-6

Cannon, A. J., Piani, C., & Sippel, S. (2020). Bias correction of climate model output for impact models. In *Climate Extremes and Their Implications for Impact and Risk Assessment* (pp. 77–104). Elsevier. https://doi.org/10.1016/B978-0-12-814895-2.00005-7

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, *28*(17), 6938–6959. https://doi.org/10.1175/JCLI-D-14-00754.1

Carrel, M., Young, S. G., & Tate, E. (2016). Pigs in Space: Determining the Environmental Justice Landscape of Swine Concentrated Animal Feeding Operations (CAFOs) in Iowa. *International Journal of Environmental Research and Public Health*, *13*(9), 849. https://doi.org/10.3390/ijerph13090849

Centner, T. J. (2011). Addressing water contamination from concentrated animal feeding operations. *Land Use Policy*, *28*(4), 706–711. https://doi.org/10.1016/J.LANDUSEPOL.2010.12.007

Centner, T. J. (2012). Regulating the land application of manure from animal production facilities in the USA. *Water Policy; Oxford*, *14*(2), 319–335. http://dx.doi.org.ezproxy1.lib.asu.edu/10.2166/wp.2011.086

Chapin, F. S., Matson, P. A., & Mooney, H. A. (2002). *Principles of terrestrial ecosystem ecology*. Springer.

Chapra, S. C., Boehlert, B., Fant, C., Bierman, V. J., Henderson, J., Mills, D., Mas, D. M. L., Rennels, L., Jantarasami, L., Martinich, J., Strzepek, K. M., & Paerl, H. W. (2017). Climate Change Impacts on Harmful Algal Blooms in U.S. Freshwaters: A Screening-Level Assessment. *Environmental Science & Technology*, *51*(16), 8933–8943. https://doi.org/10.1021/acs.est.7b01498

Chugg, B., Anderson, B., Eicher, S., Lee, S., & Ho, D. E. (2021). Enhancing Environmental Enforcement with Near Real-Time Monitoring: Likelihood-Based Detection of Structural Expansion of Intensive Livestock Farms. *ArXiv Preprint ArXiv:2105.14159*.

Clark Labs. (2019). *TerrSet Geospatial Monitoring and Modeling Software*. Clark Labs. https://clarklabs.org/terrset/

Code of Federal Regulations: 1949-1984. (1979). U.S. General Services Administration, National Archives and Records Service, Office of the Federal Register.

Copeland, C. (2010). *Animal Waste and Water Quality: EPA Regulation of Concentrated Animal Feeding Operations (CAFOs)*. 24.

Culbertson, A. M., Martin, J. F., Aloysius, N., & Ludsin, S. A. (2016). Anticipated impacts of climate change on 21st century Maumee River discharge and nutrient loads. *Journal of Great Lakes Research*, *42*(6), 1332–1342. https://doi.org/10.1016/j.jglr.2016.08.008

Dagnew, A., Scavia, D., Wang, Y., Muenich, R., Long, C., & Kalcic, M. (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, I., Cho, K. H., & Scavia, D. (2012). Evaluating Causes of Trends in Long-Term Dissolved Reactive Phosphorus Loads to Lake Erie. *Environmental Science & Technology*, *46*(19), 10660–10666. https://doi.org/10.1021/es302315d

Daniels, T. L. (1997). Where Does Cluster Zoning Fit in Farmland Protection? *Journal of the American Planning Association*, *63*(1), 129–137. https://doi.org/10.1080/01944369708975730

de Almeida Bressiani, D., Srinivasan, R., Jones, C. A., & Mendiondo, E. M. (2015). Effects of spatial and temporal weather data resolutions on streamflow modeling of a semi-arid basin, Northeast Brazil. *International Journal of Agricultural and Biological Engineering*, *8*(3), 125–139.

De Vries, J. W., Corre, W. J., & Van Dooren, H. J. C. (2010). *Environmental assessment of untreated manure use, manure digestion and codigestion with silage maize: Deliverable for the'EU-AGRO-BIOGAS'project*. Wageningen UR Livestock Research.

Delwiche, C. C. (1970). The nitrogen cycle. *Scientific American*, *223*(3), 136–147.

Deutz, P., & Gibbs, D. (2008). Industrial Ecology and Regional Development: Eco-Industrial Development as Cluster Policy. *Regional Studies*, *42*(10), 1313–1328. https://doi.org/10.1080/00343400802195121

Donner, S. D., & Scavia, D. (2007). How climate controls the flux of nitrogen by the Mississippi River and the development of hypoxia in the Gulf of Mexico. *Limnology and Oceanography*, *52*(2), 856–861. https://doi.org/10.4319/lo.2007.52.2.0856

Douglas-Mankin, K. R., Srinivasan, R., & Arnold, J. G. (2010). Soil and Water Assessment Tool (SWAT) model: Current developments and applications. *Transactions of the ASABE*, *53*(5), 1423–1431.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Ezzahar, J., Hoedjes, J. C. B., Kharrou, M. H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.-C., & Simonneaux, V. (2006). Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. *Agricultural Water Management*, *79*(1), 1–27. https://doi.org/10.1016/j.agwat.2005.02.013

Duke Bass Connections. (2016). *Animal Waste Management and Global Health (2016-2017) | Duke Bass Connections*. https://bassconnections.duke.edu/project-teams/animal-waste-management-and-global-health-2016-2017

Ebdon, D. (1985). *Statistics in geography* (Vol. 754). Blackwell Oxford.

Eckert, S., Hüsler, F., Liniger, H., & Hodel, E. (2015). Trend analysis of MODIS NDVI time series for detecting land degradation and regeneration in Mongolia. *Journal of Arid Environments*, *113*, 16–28. https://doi.org/10.1016/j.jaridenv.2014.09.001

EGLE. (2019). *Michigan Department of Environment, Great Lakes, and Energy*. MiWaters – Water Resources Information and Forms. https://miwaters.deq.state.mi.us/nsite/map/help

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., & Liebert, J. (2012). HESS Opinions" Should we apply bias correction to global and regional climate model data?". *Hydrology & Earth System Sciences Discussions*, *9*(4).

EPA. (1999). *Level III Ecoregions of Michigan*. https://hort.purdue.edu/newcrop/cropmap/michigan/maps/MIeco3.html

Estoque, R. C., Murayama, Y., & Myint, S. W. (2017). Effects of landscape composition and pattern on land surface temperature: An urban heat island study in the megacities of Southeast Asia. *Science of The Total Environment*, *577*, 349–359. https://doi.org/10.1016/j.scitotenv.2016.10.195

Estoque, R. C., Myint, S. W., Wang, C., Ishtiaque, A., Aung, T. T., Emerton, L., Ooba, M., Hijioka, Y., Mon, M. S., Wang, Z., & Fan, C. (2018a). Assessing environmental impacts and change in Myanmar's mangrove ecosystem service value due to deforestation (2000–2014). *Global Change Biology*, *24*(11), 5391–5410. https://doi.org/10.1111/gcb.14409

Estoque, R. C., Myint, S. W., Wang, C., Ishtiaque, A., Aung, T. T., Emerton, L., Ooba, M., Hijioka, Y., Mon, M. S., Wang, Z., & Fan, C. (2018b). Assessing environmental impacts and change in Myanmar's mangrove ecosystem service value due to deforestation (2000–2014). *Global Change Biology*, *24*(11), 5391–5410. https://doi.org/10.1111/gcb.14409

EWG. (2019). *Under the Radar: New Data Reveals N.C. Regulators Ignored Decade-Long Exploding Growth of Poultry CAFOs*. Environmental Working Group (EWG). https://www.ewg.org/research/under-radar

Fang, G. H., Yang, J., Chen, Y. N., & Zammit, C. (2015). Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China. *Hydrology and Earth System Sciences*, *19*(6), 2547–2559. https://doi.org/10.5194/hess-19-2547-2015

FAO (Ed.). (2013). *Tackling climate change through livestock: A global assessment of emissions and mitigation opportunities*. Food and Agriculture Organization of the United Nations.

FAO. (2017). *Food and Agriculture Organization of The United Nations—Water pollution from agriculture: A global review—Executive summary*. 35.

FAO, F. A. A. O. O. T. U. N. (2014). *World mapping of animal feeding systems in the dairy sector*. Food & Agriculture Org.

Farm System Reform Act of 2021. (2021). *The Senate of the United States—Farm System Reform Act of 2021'—A moratorium on large concentrated animal feeding operations*. 35.

Franco, A. L. C., Sobral, B. W., Silva, A. L. C., & Wall, D. H. (2019). Amazonian deforestation and soil biodiversity. *Conservation Biology*, *33*(3), 590–600. https://doi.org/10.1111/cobi.13234

Freeze, B. S., & Sommerfeldt, T. G. (1985). Breakeven Hauling Distances for Beef Feedlot Manure in Southern Alberta. *Canadian Journal of Soil Science*, *65*(4), 687–693. https://doi.org/10.4141/cjss85-074

Furiness, C., Cowling, E., Allen, L., Abt, R., Frederick, D., Zering, K., & Campbell, R. (2019). *Forests as an Alternative for Swine Manure Application*. NC State Extension Pubblications. https://content.ces.ncsu.edu/forests-as-an-alternative-for-swine-manure-application

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., & Sutton, M. A. (2008). Transformation of the Nitrogen Cycle: Recent Trends, Questions, and Potential Solutions. *Science*, *320*(5878), 889–892. https://doi.org/10.1126/science.1136674

Garcia, D. J., Lovett, B. M., & You, F. (2019). Considering agricultural wastes and ecosystem services in Food-Energy-Water-Waste Nexus system design. *Journal of Cleaner Production*, *228*, 941–955. https://doi.org/10.1016/j.jclepro.2019.04.314

Garner, A., & Keoleian, G. A. (1995). *Industrial Ecology: An Introduction*. 32.

Gassman, P. W., Reyes, M. R., Green, C. H., & Arnold, J. G. (2007). The soil and water assessment tool: Historical development, applications, and future research directions. *Transactions of the ASABE*, *50*(4), 1211–1250.

Gellings, C. W., & Parmenter, K. E. (2016). *Energy Efficiency in Fertilizer Production and Use*. 15.

Gill, D., Rowe, M., & Joshi, S. J. (2018). Fishing in greener waters: Understanding the impact of harmful algal blooms on Lake Erie anglers and the potential for adoption of a forecast model. *Journal of Environmental Management*, *227*, 248–255. https://doi.org/10.1016/j.jenvman.2018.08.074

Glasgow, H. B., & Burkholder, J. M. (2000). Water Quality Trends and Management Implications from a Five-Year Study of a Eutrophic Estuary. *Ecological Applications*, *10*(4), 1024–1046. https://doi.org/10.2307/2641015

Glibert, P. (2020). From hogs to HABs: Impacts of industrial farming in the US on nitrogen and phosphorus and greenhouse gas pollution. *Biogeochemistry*, *150*. https://doi.org/10.1007/s10533-020-00691-6

Godfray, H. C. J., Aveyard, P., Garnett, T., Hall, J. W., Key, T. J., Lorimer, J., Pierrehumbert, R. T., Scarborough, P., Springmann, M., & Jebb, S. A. (2018). Meat consumption, health, and the environment. *Science*, *361*(6399). https://doi.org/10.1126/science.aam5324

Griffith, G., Omernick, J., Comstock, J., Schfale, M., McNab, W., Lenat, D., Glover, J., & Shelburne, V. (2002). *Ecoregions of North Carolina and South Carolina*.

Guidry, V. T., Rhodes, S. M., Woods, C. G., Hall, D. J., & Rinsky, J. L. (2018). Connecting Environmental Justice and Community Health Effects of Hog Production in North Carolina. *North Carolina Medical Journal*, *79*(5), 324–328. https://doi.org/10.18043/ncm.79.5.324

Guignard, M. S., Leitch, A. R., Acquisti, C., Eizaguirre, C., Elser, J. J., Hessen, D. O., Jeyasingh, P. D., Neiman, M., Richardson, A. E., Soltis, P. S., Soltis, D. E., Stevens, C. J., Trimmer, M., Weider, L. J., Woodward, G., & Leitch, I. J. (2017). Impacts of Nitrogen and Phosphorus: From Genomes to Natural Ecosystems and Agriculture. *Frontiers in Ecology and Evolution*, *5*. https://doi.org/10.3389/fevo.2017.00070

Gutiérrez, J. M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martín, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., Dubrovsky, M., Ribalaygua, J., … Pagé, C. (2019). An intercomparison of a large ensemble of

statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. *International Journal of Climatology*, *39*(9), 3750–3785. https://doi.org/10.1002/joc.5462

Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., Seibert, J., & Maurice, P. (2019). *Hydrological Modeling of Climate Change Impacts*.

Handan-Nader, C., & Ho, D. E. (2019a). Deep learning to map concentrated animal feeding operations. *Nature Sustainability*, *2*(4), 298–306. https://doi.org/10.1038/s41893-019-0246-x

Handan-Nader, C., & Ho, D. E. (2019b). Deep learning to map concentrated animal feeding operations. *Nature Sustainability*, *2*(4), 298. https://doi.org/10.1038/s41893-019-0246-x

Harden, S. L. (2015). *Surface-water quality in agricultural watersheds of the North Carolina Coastal Plain associated with concentrated animal feeding operations* (USGS Numbered Series No. 2015–5080; Scientific Investigations Report, p. 70). U.S. Geological Survey. https://doi.org/10.3133/sir20155080

Harun, S. M., & Ogneva-Himmelberger, Y. (2013). Distribution of industrial farms in the United States and socioeconomic, health, and environmental characteristics of counties. *Geography Journal*, *2013*.

Hribar, C. (2010). *Understanding concentrated animal feeding operations and their impact on communities*. https://stacks.cdc.gov/view/cdc/59792

Hu, Y., Cheng, H., & Tao, S. (2017). Environmental and human health challenges of industrial livestock and poultry farming in China and their mitigation. *Environment International*, *107*, 111–130. https://doi.org/10.1016/j.envint.2017.07.003

Ishtiaque, A., Myint, S. W., & Wang, C. (2016a). Examining the ecosystem health and sustainability of the world's largest mangrove forest using multi-temporal MODIS products. *Science of The Total Environment*, *569–570*, 1241–1254. https://doi.org/10.1016/j.scitotenv.2016.06.200

Ishtiaque, A., Myint, S. W., & Wang, C. (2016b). Examining the ecosystem health and sustainability of the world's largest mangrove forest using multi-temporal MODIS products. *Science of The Total Environment*, *569–570*, 1241–1254. https://doi.org/10.1016/j.scitotenv.2016.06.200

Jacquin, A., Sheeren, D., & Lacombe, J.-P. (2010). Vegetation cover degradation assessment in Madagascar savanna based on trend analysis of MODIS NDVI time series. *International Journal of Applied Earth Observation and Geoinformation*, *12*, S3–S10. https://doi.org/10.1016/j.jag.2009.11.004

Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., Strahler, A., Lucht, W., Myneni, R. B., Knyazikhin, Y., Running, S. W., Nemani, R. R., Wan, Z., Huete, A. R., van Leeuwen, W., Wolfe, R. E., … Barnsley, M. J. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1228–1249. https://doi.org/10.1109/36.701075

Kalcic, M. M., Chaubey, I., & Frankenberger, J. (2015). Defining Soil and Water Assessment Tool (SWAT) hydrologic response units (HRUs) by field boundaries. *International Journal of Agricultural and Biological Engineering*, *8*(3), 69–80.

Kalcic, M. M., Frankenberger, J., & Chaubey, I. (2015). Spatial Optimization of Six Conservation Practices Using Swat in Tile-Drained Agricultural Watersheds. *JAWRA Journal of the American Water Resources Association*, *51*(4), 956–972.

Kalcic, M. M., Muenich, R. L., Basile, S., Steiner, A. L., Kirchhoff, C., & Scavia, D. (2019). Climate Change and Nutrient Loading in the Western Lake Erie Basin: Warming Can Counteract a Wetter Future. *Environmental Science & Technology*, *53*(13), 7543–7550. https://doi.org/10.1021/acs.est.9b01274

Kalkhoff, S. J., Hubbard, L. E., Tomer, M. D., & James, D. E. (2016). Effect of variable annual precipitation and nutrient input on nitrogen and phosphorus transport from two Midwestern agricultural watersheds. *Science of The Total Environment*, *559*, 53–62. https://doi.org/10.1016/j.scitotenv.2016.03.127

Kast, J. B., Apostel, A. M., Kalcic, M. M., Muenich, R. L., Dagnew, A., Long, C. M., Evenson, G., & Martin, J. F. (2021). Source contribution to phosphorus loads from the Maumee River watershed to Lake Erie. *Journal of Environmental Management*, *279*, 111803. https://doi.org/10.1016/j.jenvman.2020.111803

Kellogg, R. L., States, U., Lander, C. H., Moffitt, D. C., & D, N. G. P. (2000). *Manure Nutrients Relative to the Capacity of Cropland and Pastureland to Assimilate Nutrients: Spatial and Temporal Trends for the United*.

Kennedy, L. (1999). Cooperating for Survival: Tannery Pollution and Joint Action in the Palar Valley (India). *World Development*, *27*(9), 1673–1691. https://doi.org/10.1016/S0305-750X(99)00080-7

Keplinger, K. O., & Hauck, L. M. (2006). The Economics of Manure Utilization: Model and Application. *Journal of Agricultural and Resource Economics*, *31*(2), 414–440. JSTOR.

Key, N., McBride, W. D., Ribaudo, M., & Sneeringer, S. (2017). *Trends and Developments in Hog Manure Management: 1998-2009* (SSRN Scholarly Paper

ID 2981722). Social Science Research Network. https://papers.ssrn.com/abstract=2981722

Kline, K. L., Singh, N., & Dale, V. H. (2013). Cultivated hay and fallow/idle cropland confound analysis of grassland conversion in the Western Corn Belt. *Proceedings of the National Academy of Sciences*, *110*(31), E2863–E2863.

Kondraju, T. T., & Rajan, K. S. (2019). Excessive Fertilizer Usage Drives Agriculture Growth but Depletes Water Quality. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-3/W1*, 17–23. https://doi.org/10.5194/isprs-annals-IV-3-W1-17-2019

Kosten, S., Huszar, V. L. M., Bécares, E., Costa, L. S., Donk, E. van, Hansson, L.-A., Jeppesen, E., Kruk, C., Lacerot, G., Mazzeo, N., Meester, L. D., Moss, B., Lürling, M., Nõges, T., Romo, S., & Scheffer, M. (2012). Warmer climates boost cyanobacterial dominance in shallow lakes. *Global Change Biology*, *18*(1), 118–126. https://doi.org/10.1111/j.1365-2486.2011.02488.x

Kronberg, S. L., & Ryschawy, J. (2019). Chapter 5—Negative Impacts on the Environment and People From Simplification of Crop and Livestock Production∗. In G. Lemaire, P. C. D. F. Carvalho, S. Kronberg, & S. Recous (Eds.), *Agroecosystem Diversity* (pp. 75–90). Academic Press. https://doi.org/10.1016/B978-0-12-811050-8.00005-4

Kujawa, H., Kalcic, M., Martin, J., Aloysius, N., Apostel, A., Kast, J., Murumkar, A., Evenson, G., Becker, R., Boles, C., Confesor, R., Dagnew, A., Guo, T., Logsdon Muenich, R., Redder, T., Scavia, D., & Wang, Y.-C. (2020). The hydrologic model as a source of nutrient loading uncertainty in a future climate. *Science of The Total Environment*, *724*, 138004. https://doi.org/10.1016/j.scitotenv.2020.138004

Lall, S. V., & Mengistae, T. (2005). *Business Environment, Clustering, and Industry Location: Evidence from Indian Cities*. World Bank Publications.

Lamchin, M., Lee, W.-K., Jeon, S. W., Wang, S. W., Lim, C. H., Song, C., & Sung, M. (2018). Long-term trend and correlation between vegetation greenness and climate variables in Asia based on satellite data. *Science of The Total Environment*, *618*, 1089–1095. https://doi.org/10.1016/j.scitotenv.2017.09.145

Lenderink, G., Buishand, A., & van Deursen, W. (2007). Estimates of future discharges of the river Rhine using two scenario methodologies: Direct versus delta approach. *Hydrology and Earth System Sciences*, *11*(3), 1145–1159. https://doi.org/10.5194/hess-11-1145-2007

Leung, L. R., Mearns, L. O., Giorgi, F., & Wilby, R. L. (2003). Regional climate research: Needs and opportunities. *Bulletin of the American Meteorological Society*, *84*(1), 89–95.

Lifset, R., & Graedel, T. E. (2002). Industrial ecology: Goals and definitions. *A Handbook of Industrial Ecology*, 3–15.

Liu, J., Kleinman, P. J. A., Aronsson, H., Flaten, D., McDowell, R. W., Bechmann, M., Beegle, D. B., Robinson, T. P., Bryant, R. B., Liu, H., Sharpley, A. N., & Veith, T. L. (2018). A review of regulations and guidelines related to winter manure application. *Ambio*, *47*(6), 657–670. https://doi.org/10.1007/s13280-018-1012-4

Logsdon, R. A., & Chaubey, I. (2013). A quantitative approach to evaluating ecosystem services. *Ecological Modelling*, *257*, 57–65. https://doi.org/10.1016/j.ecolmodel.2013.02.009

Long, C. M., Muenich, R. L., Kalcic, M. M., & Scavia, D. (2018). Use of manure nutrients from concentrated animal feeding operations. *Journal of Great Lakes Research*, *44*(2), 245–252. https://doi.org/10.1016/j.jglr.2018.01.006

LP DAAC, N. (2019). *The Land Processes Distributed Active Archive Center (LP DAAC)—Homepage*. https://lpdaac.usgs.gov/

Lunetta, R. S., Shao, Y., Ediriwickrema, J., & Lyon, J. G. (2010). Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data. *International Journal of Applied Earth Observation and Geoinformation*, *12*(2), 81–88. https://doi.org/10.1016/j.jag.2009.11.005

Mahoney, M., & Magel, R. (1996). Estimation of the Power of the Kruskal-Wallis Test. *Biometrical Journal*, *38*(5), 613–630. https://doi.org/10.1002/bimj.4710380510

Mallin, M. A., & Cahoon, L. B. (2003). Industrialized animal production—A major source of nutrient and microbial pollution to aquatic ecosystems. *Population and Environment*, *24*(5), 369–385.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., & Thiele-Eich, I. (2010). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, *48*(3). https://doi.org/10.1029/2009RG000314

Martin, K. L., Emanuel, R. E., & Vose, J. M. (2018). Terra incognita: The unknown risks to environmental quality posed by the spatial distribution and abundance of concentrated animal feeding operations. *Science of The Total Environment*, *642*, 887–893. https://doi.org/10.1016/j.scitotenv.2018.06.072

Matisziw, T. C., & Hipple, J. D. (2001). Spatial Clustering and State/County Legislation: The Case of Hog Production in Missouri. *Regional Studies*, *35*(8), 719–730. https://doi.org/10.1080/00343400120084704

McDonald, J. M., Hoppe, R. A., & Newton, D. (2018). *Three decades of consolidation in US agriculture*.

McDonald, J. M., Law, J., & Mosheim, R. (2020). *Consolidation in US Dairy Farming*.

Meals, D. W., Richards, R. P., & Dressing, S. A. (2013). Pollutant load estimation for water quality monitoring projects. *Tech Notes*, *8*, 1–21.

Meals, D. W., Spooner, J., Dressing, S. A., & Harcum, J. B. (2011). Statistical analysis for monotonic trends. *Tech Notes*, *6*, 23.

Mehan, S., Aggarwal, R., Gitau, M. W., Flanagan, D. C., Wallace, C. W., & Frankenberger, J. R. (2019). Assessment of hydrology and nutrient losses in a changing climate in a subsurface-drained watershed. *Science of The Total Environment*, *688*, 1236–1251. https://doi.org/10.1016/j.scitotenv.2019.06.314

Mehdi, B., Ludwig, R., & Lehner, B. (2015). Evaluating the impacts of climate change and crop land use change on streamflow, nitrates and phosphorus: A modeling study in Bavaria. *Journal of Hydrology: Regional Studies*, *4*, 60–90. https://doi.org/10.1016/j.ejrh.2015.04.009

Meyer, J., Kohn, I., Stahl, K., Hakala, K., Seibert, J., & Cannon, A. J. (2019). Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments. *Hydrology and Earth System Sciences*, *23*(3), 1339–1354. https://doi.org/10.5194/hess-23-1339-2019

Michalak, A. M. (2016). Study role of climate change in extreme threats to water quality. *Nature*, *535*(7612), 349–350. https://doi.org/10.1038/535349a

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., Chaffin, J. D., Cho, K., Confesor, R., Daloğlu, I., DePinto, J. V., Evans, M. A., Fahnenstiel, G. L., He, L., Ho, J. C., Jenkins, L., Johengen, T. H., Kuo, K. C., LaPorte, E., … Zagorski, M. A. (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*, *110*(16), 6448–6452. https://doi.org/10.1073/pnas.1216006110

Militino, A. F., Moradi, M., & Ugarte, M. D. (2020). On the Performances of Trend and Change-Point Detection Methods for Remote Sensing Data. *Remote Sensing*, *12*(6), 1008. https://doi.org/10.3390/rs12061008

Miralha, L., & Kim, D. (2018). Accounting for and Predicting the Influence of Spatial Autocorrelation in Water Quality Modeling. *ISPRS International Journal of Geo-Information*, *7*(2), 64. https://doi.org/10.3390/ijgi7020064

Miralha, L., Muenich, R. L., Scavia, D., Wells, K., Steiner, A. L., Kalcic, M., Apostel, A., Basile, S., & Kirchhoff, C. J. (2021). Bias correction of climate model outputs influences watershed model nutrient load predictions. *Science of The Total Environment*, *759*, 143039. https://doi.org/10.1016/j.scitotenv.2020.143039

Miralha, L., Muenich, R. L., Schaffer-Smith, D., & Myint, S. W. (2021). Spatiotemporal land use change and environmental degradation surrounding CAFOs in Michigan and North Carolina. *Science of The Total Environment*, *800*, 149391. https://doi.org/10.1016/j.scitotenv.2021.149391

Mladenoff, D. J., Sahajpal, R., Johnson, C. P., & Rothstein, D. E. (2016). Recent Land Use Change to Agriculture in the U.S. Lake States: Impacts on Cellulosic Biomass Potential and Natural Lands. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0148566

Mole, B. (2013). *Farming up Trouble*. *499*, 3.

Moore, S. K., Trainer, V. L., Mantua, N. J., Parker, M. S., Laws, E. A., Backer, L. C., & Fleming, L. E. (2008). Impacts of climate variability and future climate change on harmful algal blooms and human health. *Environmental Health*, *7*(2), S4. https://doi.org/10.1186/1476-069X-7-S2-S4

MPCA, M. P. C. A. (2016, October 26). *Construction, operation, and technical requirements*. Minnesota Pollution Control Agency. https://www.pca.state.mn.us/water/construction-operation-and-technical-requirements

Muenich, R. L., Kalcic, M., & Scavia, D. (2016). Evaluating the Impact of Legacy P and Agricultural Conservation Practices on Nutrient Loads from the Maumee River Watershed. *Environmental Science & Technology*, *50*(15), 8146–8154. https://doi.org/10.1021/acs.est.6b01421

NASS, U. (2017). *USDA - National Agricultural Statistics Service—Census of Agriculture*. https://www.nass.usda.gov/AgCensus/

NCDEQ. (2019). *North Carolina—Department of Environmenta Qualtity*. https://deq.nc.gov/cafo-map

NCDEQ. (2020). *Riparian Buffer Protection Program*. https://deq.nc.gov/about/divisions/water-resources/water-quality-permitting/401-buffer-authorization/riparian-buffer

NCWQR, N. C. for W. Q. R. (2005). *Time weighted and flow weighted mean concentrations*. https://ncwqr.files.wordpress.com/2017/06/d-time-weighted-and-flow-weighted-mean-concentrations.pdf

NOAA. (2021). *U.S. Climate Atlas | National Centers for Environmental Information (NCEI)*. https://www.ncdc.noaa.gov/climateatlas/

NRCS. (2020). *Watershed Boundary Dataset (WBD) Overview | NRCS*. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/water/watersheds/dataset/?cid=nrcs143_021623

NRDC. (2019). *CAFOs: What We Don't Know Is Hurting Us*. NRDC. https://www.nrdc.org/resources/cafos-what-we-dont-know-hurting-us

Ogneva-Himmelberger, Y., Huang, L., & Xin, H. (2015). CALPUFF and CAFOs: Air Pollution Modeling and Environmental Justice Analysis in the North Carolina Hog Industry. *ISPRS International Journal of Geo-Information*, *4*(1), 150–171. https://doi.org/10.3390/ijgi4010150

O'Neil, J. M., Davis, T. W., Burford, M. A., & Gobler, C. J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae*, *14*, 313–334. https://doi.org/10.1016/j.hal.2011.10.027

Oun, A., Kumar, A., Harrigan, T., Angelakis, A., & Xagoraraki, I. (2014). Effects of Biosolids and Manure Application on Microbial Water Quality in Rural Areas in the US. *Water*, *6*(12), 3701–3723. https://doi.org/10.3390/w6123701

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., & Dasgupta, P. (2014). *Climate change 2014: Synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Ipcc.

Paerl, H. W., Otten, T. G., & Kudela, R. (2018). Mitigating the Expansion of Harmful Algal Blooms Across the Freshwater-to-Marine Continuum. *Environmental Science & Technology*, *52*(10), 5519–5529. https://doi.org/10.1021/acs.est.7b05950

Peel, M. C., Finlayson, B. L., & McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, *11*(5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007

Pepper, I. L., Brooks, J. P., & Gerba, C. P. (2019). Chapter 23 - Land Application of Organic Residuals: Municipal Biosolids and Animal Manures. In M. L. Brusseau, I. L. Pepper, & C. P. Gerba (Eds.), *Environmental and Pollution Science (Third Edition)* (pp. 419–434). Academic Press. https://doi.org/10.1016/B978-0-12-814719-1.00023-9

Porter, M. E. (1998). *Clusters and the new economics of competition* (Vol. 76). Harvard Business Review Boston.

Prenafeta-Boldú, F. X., & Kamilaris, A. (2019). AI assists in locating hidden farms. *Nature Sustainability*, *2*(4), 262. https://doi.org/10.1038/s41893-019-0264-8

Qi, J., Xin, X., John, R., Groisman, P., & Chen, J. (2017). Understanding livestock production and sustainability of grassland ecosystems in the Asian Dryland Belt. *Ecological Processes*, *6*(1), 22. https://doi.org/10.1186/s13717-017-0087-3

Rabalais, N. N., Turner, R. E., Díaz, R. J., & Justić, D. (2009). Global change and eutrophication of coastal waters. *ICES Journal of Marine Science*, *66*(7), 1528–1537. https://doi.org/10.1093/icesjms/fsp047

Raff, Z., & Meyer, A. (2021a). CAFOs and Surface Water Quality: Evidence from Wisconsin. *American Journal of Agricultural Economics*. https://doi.org/10.1111/ajae.12222

Raff, Z., & Meyer, A. (2021b). CAFOs and Surface Water Quality: Evidence from Wisconsin. *American Journal of Agricultural Economics*. https://doi.org/10.1111/ajae.12222

Randad, P. R., Dillen, C. A., Ortines, R. V., Mohr, D., Aziz, M., Price, L. B., Kaya, H., Larsen, J., Carroll, K. C., Smith, T. C., Miller, L. S., & Heaney, C. D. (2019). Comparison of livestock-associated and community-associated Staphylococcus aureus pathogenicity in a mouse model of skin and soft tissue infection. *Scientific Reports*, *9*(1), 6774. https://doi.org/10.1038/s41598-019-42919-y

Razaq, M., Zhang, P., Shen, H., & Salahuddin. (2017). Influence of nitrogen and phosphorous on the growth and root morphology of Acer mono. *PLOS ONE*, *12*(2), e0171321. https://doi.org/10.1371/journal.pone.0171321

Ribaudo, M., Kaplan, J. D., Christensen, L. A., Gollehon, N., Johansson, R., Breneman, V. E., Aillery, M., Agapoff, J., & Peters, M. (2003). Manure management for water quality costs to animal feeding operations of applying manure nutrients to land. *USDA-ERS Agricultural Economic Report*, *824*.

Richards, R. P., & Baker, D. B. (1993). Trends in Nutrient and Suspended Sediment Concentrations in Lake Erie Tributaries, 1975–1990. *Journal of Great Lakes Research*, *19*(2), 200–211. https://doi.org/10.1016/S0380-1330(93)71211-3

Ritzel, B. (2014). *US EPA's Efforts to Regulate CAFOs*.

Rosov, K. A., Mallin, M. A., & Cahoon, L. B. (2020). Waste nutrients from U.S. animal feeding operations: Regulations are inconsistent across states and inadequately

assess nutrient export risk. *Journal of Environmental Management*, *269*, 110738. https://doi.org/10.1016/j.jenvman.2020.110738

Rothenberger, M. B., Burkholder, J. M., & Brownie, C. (2009). Long-Term Effects of Changing Land Use Practices on Surface Water Quality in a Coastal River and Lagoonal Estuary. *Environmental Management*, *44*(3), 505–523. https://doi.org/10.1007/s00267-009-9330-8

Rundel, P. W., Dickie, I. A., & Richardson, D. M. (2014). Tree invasions into treeless areas: Mechanisms and ecosystem processes. *Biological Invasions*, *16*(3), 663–675. https://doi.org/10.1007/s10530-013-0614-9

Ruttenberg, K. C. (2003). 8.13—The Global Phosphorus Cycle. In H. D. Holland & K. K. Turekian (Eds.), *Treatise on Geochemistry* (pp. 585–643). Pergamon. https://doi.org/10.1016/B0-08-043751-6/08153-6

Saad, D. A., Argue, D. M., Schwarz, G. E., Anning, D. W., Ator, S. W., Hoos, A. B., Preston, S. D., Robertson, D. M., & Wise, D. A. (2019). *Water-quality and streamflow datasets used for estimating long-term mean daily streamflow and annual loads to be considered for use in regional streamflow, nutrient and sediment SPARROW models, United States, 1999-2014* [Data set]. U.S. Geological Survey. https://doi.org/10.5066/F7DN436B

Scanes, C. G. (2018). Chapter 18—Impact of Agricultural Animals on the Environment. In C. G. Scanes & S. R. Toukhsati (Eds.), *Animals and Human Society* (pp. 427–449). Academic Press. https://doi.org/10.1016/B978-0-12-805247-1.00025-3

Scavia, D., David Allan, J., Arend, K. K., Bartell, S., Beletsky, D., Bosch, N. S., Brandt, S. B., Briland, R. D., Daloğlu, I., DePinto, J. V., Dolan, D. M., Evans, M. A., Farmer, T. M., Goto, D., Han, H., Höök, T. O., Knight, R., Ludsin, S. A., Mason, D., … Zhou, Y. (2014). Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. *Journal of Great Lakes Research*, *40*(2), 226–246. https://doi.org/10.1016/j.jglr.2014.02.004

Scavia, D., Kalcic, M., Muenich, R. L., Read, J., Aloysius, N., Bertani, I., Boles, C., Confesor, R., DePinto, J., Gildow, M., Martin, J., Redder, T., Robertson, D., Sowa, S., Wang, Y.-C., & Yen, H. (2017). Multiple models guide strategies for agricultural nutrient reductions. *Frontiers in Ecology and the Environment*, *15*(3), 126–132. https://doi.org/10.1002/fee.1472

Scavia, D., Rabalais, N. N., Turner, R. E., Justić, D., & Wiseman Jr, W. J. (2003). Predicting the response of Gulf of Mexico hypoxia to variations in Mississippi River nitrogen load. *Limnology and Oceanography*, *48*(3), 951–956.

Schreiner-McGraw, A. P., Vivoni, E. R., Ajami, H., Sala, O. E., Throop, H. L., & Peters, D. P. C. (2020). Woody Plant Encroachment has a Larger Impact than Climate

Change on Dryland Water Budgets. *Scientific Reports*, *10*(1), 8112. https://doi.org/10.1038/s41598-020-65094-x

Schuwirth, N. (2020). Towards an integrated surface water quality assessment: Aggregation over multiple pollutants and time. *Water Research*, *186*, 116330. https://doi.org/10.1016/j.watres.2020.116330

Secchi, S., & Mcdonald, M. (2019). The state of water quality strategies in the Mississippi River Basin: Is cooperative federalism working? *Science of The Total Environment*, *677*, 241–249. https://doi.org/10.1016/j.scitotenv.2019.04.381

SEIA. (2019). *North Carolina SEIA - SOlar Energy Industries Association*. SEIA. /state-solar-policy/north-carolina-solar

Sen, P. K. (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, *63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

Siegel, S. (1957). Nonparametric Statistics. *The American Statistician*, *11*(3), 13–19. https://doi.org/10.1080/00031305.1957.10501091

Sims, J. T., Bergström, L., Bowman, B. T., & Oenema, O. (2005). Nutrient management for intensive animal agriculture: Policies and practices for sustainability. *Soil Use and Management*, *21*(1), 141–151. https://doi.org/10.1111/j.1475-2743.2005.tb00118.x

Son, J.-Y., Muenich, R. L., Schaffer-Smith, D., Miranda, M. L., & Bell, M. L. (2021). Distribution of environmental justice metrics for exposure to CAFOs in North Carolina, USA. *Environmental Research*, *195*, 110862. https://doi.org/10.1016/j.envres.2021.110862

Song, X.-P., Huang, C., Feng, M., Sexton, J. O., Channan, S., & Townshend, J. R. (2014). Integrating global land cover products for improved forest cover characterization: An application in North America. *International Journal of Digital Earth*, *7*(9), 709–724. https://doi.org/10.1080/17538947.2013.856959

Sonne, C., Ok, Y. S., Dietz, R., & Alstrup, A. K. O. (2019). Pig slurry needs modifications to be a sustainable fertilizer in crop production. *Environmental Research*, *178*, 108718. https://doi.org/10.1016/j.envres.2019.108718

Sousan, S., Iverson, G., Humphrey, C., Lewis, A., Streuber, D., & Richardson, L. (2021). High-frequency assessment of air and water quality at a concentration animal feeding operation during wastewater application to spray fields. *Environmental Pollution*, *288*, 117801.

Spiegal, S., Kleinman, P. J. A., Endale, D. M., Bryant, R. B., Dell, C., Goslee, S., Meinen, R. J., Flynn, K. C., Baker, J. M., Browning, D. M., McCarty, G., Bittman, S., Carter, J., Cavigelli, M., Duncan, E., Gowda, P., Li, X., Ponce-Campos, G. E., Cibin, R., … Yang, Q. (2020). Manuresheds: Advancing nutrient recycling in US agriculture. *Agricultural Systems*, *182*, 102813. https://doi.org/10.1016/j.agsy.2020.102813

Stein, L. Y., & Klotz, M. G. (2016). The nitrogen cycle. *Current Biology*, *26*(3), R94–R98. https://doi.org/10.1016/j.cub.2015.12.021

Steinzor, R. I., & Huang, L.-Y. (2012). Agricultural Secrecy: Going Dark Down on the Farm: How Legalized Secrecy Gives Agribusiness a Federally Funded Free Ride. *Center for Progressive Reform Briefing Paper*, *1213*.

Stewart, J. S., Schwarz, G. E., Brakebill, J. W., & Preston, S. D. (2019). Catchment-level estimates of nitrogen and phosphorus agricultural use from commercial fertilizer sales for the conterminous United States, 2012. In *Catchment-level estimates of nitrogen and phosphorus agricultural use from commercial fertilizer sales for the conterminous United States, 2012* (USGS Numbered Series No. 2018–5145; Scientific Investigations Report, Vols. 2018–5145). U.S. Geological Survey. https://doi.org/10.3133/sir20185145

Stow, C. A., Glassner-Shwayder, K., Lee, D., Wang, L., Arhonditsis, G., DePinto, J. V., & Twiss, M. R. (2020). Lake Erie phosphorus targets: An imperative for active adaptive management. *Journal of Great Lakes Research*, S038013302030040X. https://doi.org/10.1016/j.jglr.2020.02.005

Sulla-Menashe, D., & Friedl, M. A. (2018). *User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product*. 18.

Tan, M. L., Gassman, P., Yang, X., & Haywood, J. (2020). A Review of SWAT Applications, Performance and Future Needs for Simulation of Hydro-Climatic Extremes. *Advances in Water Resources*, 103662.

Taranu, Z. E., Gregory-Eaves, I., Leavitt, P. R., Bunting, L., Buchaca, T., Catalan, J., Domaizon, I., Guilizzoni, P., Lami, A., McGowan, S., Moorhouse, H., Morabito, G., Pick, F. R., Stevenson, M. A., Thompson, P. L., & Vinebrooke, R. D. (2015). Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecology Letters*, *18*(4), 375–384. https://doi.org/10.1111/ele.12420

Tesemma, Z. K., Wei, Y., Western, A. W., & Peel, M. C. (2014). Leaf Area Index Variation for Crop, Pasture, and Tree in Response to Climatic Variation in the Goulburn–Broken Catchment, Australia. *Journal of Hydrometeorology*, *15*(4), 1592–1606. https://doi.org/10.1175/JHM-D-13-0108.1

Teutschbein, C., & Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, *456–457*, 12–29. https://doi.org/10.1016/j.jhydrol.2012.05.052

Teutschbein, C., & Seibert, J. (2013). Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions? *Hydrology and Earth System Sciences*, *17*(12), 5061–5077. https://doi.org/10.5194/hess-17-5061-2013

Teutschbein, C., Sponseller, R. A., Grabs, T., Blackburn, M., Boyer, E. W., Hytteborn, J. K., & Bishop, K. (2017). Future Riverine Inorganic Nitrogen Load to the Baltic Sea From Sweden: An Ensemble Approach to Assessing Climate Change Effects. *Global Biogeochemical Cycles*, *31*(11), 1674–1701. https://doi.org/10.1002/2016GB005598

Thu, K. M., & Durrenberger, E. P. (1998). *Pigs, Profits, and Rural Communities*. State University of New York Press. http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=5535&site=ehost-live

Thurow, A. P., & Thompson, P. B. (1998). *Toward an Augmented Theory of Cooperative Behavior: The Case of Clustering in Animal Agriculture*.

Tullo, E., Finzi, A., & Guarino, M. (2019). Review: Environmental impact of livestock farming and Precision Livestock Farming as a mitigation strategy. *Science of The Total Environment*, *650*, 2751–2760. https://doi.org/10.1016/j.scitotenv.2018.10.018

Turner, R. E., Rabalais, N. N., & Justic, D. (2008). Gulf of Mexico Hypoxia: Alternate States and a Legacy. *Environmental Science & Technology*, *42*(7), 2323–2327. https://doi.org/10.1021/es071617k

UADA. (2018). *North Carolina | The Economic Contributions and Impacts of U.S. Food, Fiber, and Forest Industries—University of Arkansas System—Division of Agriculture (Research & Extension)*. https://economic-impact-of-ag.uada.edu/north-carolina/

United Nations. (2015). *Transforming our World: The 2030 Agenda for Sustainable Development | Department of Economic and Social Affairs*. https://sdgs.un.org/publications/transforming-our-world-2030-agenda-sustainable-development-17981

United Nations. (2016). Water and Sanitation. *United Nations Sustainable Development*. https://www.un.org/sustainabledevelopment/water-and-sanitation/

US Census Bureau. (2019). *Population Estimates Continue to Show the Nation's Growth Is Slowing*. Census. https://www.census.gov/newsroom/press-releases/2019/popest-nation.html

US EPA, O. (2013, March 19). *Ecoregional Nutrient Criteria for Rivers and Streams* [Data and Tools]. https://www.epa.gov/nutrient-policy-data/ecoregional-nutrient-criteria-rivers-and-streams

US EPA, O. (2015, August 25). *NPDES CAFO Regulations Implementation Status Reports* [Reports and Assessments]. https://www.epa.gov/npdes/npdes-cafo-regulations-implementation-status-reports

US EPA, O. (2016, January 6). *State Progress Toward Developing Numeric Nutrient Water Quality Criteria for Nitrogen and Phosphorus* (States, Territories) [Data and Tools]. https://www.epa.gov/nutrient-policy-data/state-progress-toward-developing-numeric-nutrient-water-quality-criteria

US EPA, O. (2020, December 9). *Ambient Water Quality Criteria to Address Nutrient Pollution in Lakes and Reservoirs* [Reports and Assessments]. https://www.epa.gov/nutrient-policy-data/ambient-water-quality-criteria-address-nutrient-pollution-lakes-and-reservoirs

USEPA. (2003). National Pollutant Discharge Elimination System Permit Regulation and Effluent Limitation Guidelines and Standards for Concentrated Animal Feeding Operations (CAFOs). *Federal Register*, *68*(29), 100.

USEPA. (2004). *Risk Assessment Evaluation for Concentrated Animal Feeding Operations*. https://nepis.epa.gov

USEPA. (2020). *National Summary of State Information | Water Quality Assessment and TMDL Information | US EPA*. https://ofmpub.epa.gov/waters10/attains_nation_cy.control#status_of_data

USEPA, O. (2018). *Summary of the Clean Water Act 1972—2018 version of the Clean Water Act CWA from the U.S. Code* [Overviews and Factsheets]. https://www.epa.gov/laws-regulations/summary-clean-water-act

Vadas, P. A., Bolster, C. H., & Good, L. W. (2013). Critical evaluation of models used to study agricultural phosphorus and water quality. *Soil Use and Management*, *29*(s1), 36–44. https://doi.org/10.1111/j.1475-2743.2012.00431.x

Vadas, P. A., & White, M. J. (2010). Validating soil phosphorus routines in the SWAT model. *Transactions of the ASABE*, *53*(5), 1469–1476.

Vale, P., Gibbs, H., Vale, R., Christie, M., Florence, E., Munger, J., & Sabaini, D. (2019). The Expansion of Intensive Beef Farming to the Brazilian Amazon. *Global*

Environmental Change, *57*, 101922. https://doi.org/10.1016/j.gloenvcha.2019.05.006

Verma, S., Bhattarai, R., Bosch, N. S., Cooke, R. C., Kalita, P. K., & Markus, M. (2015). Climate Change Impacts on Flow, Sediment and Nutrient Export in a Great Lakes Watershed Using SWAT. *CLEAN – Soil, Air, Water*, *43*(11), 1464–1474. https://doi.org/10.1002/clen.201400724

von Keyserlingk, M. A. G., Martin, N. P., Kebreab, E., Knowlton, K. F., Grant, R. J., Stephenson, M., Sniffen, C. J., Harner, J. P., Wright, A. D., & Smith, S. I. (2013). Invited review: Sustainability of the US dairy industry. *Journal of Dairy Science*, *96*(9), 5405–5425. https://doi.org/10.3168/jds.2012-6354

Walljasper, C. (2018, June 7). *Large animal feeding operations on the rise*. Investigate Midwest. https://investigatemidwest.org/2018/06/07/large-animal-feeding-operations-on-the-rise/

Wang, C., & Myint, S. W. (2016). Environmental Concerns of Deforestation in Myanmar 2001–2010. *Remote Sensing*, *8*(9), 728. https://doi.org/10.3390/rs8090728

Wang, J., & Baerenklau, K. A. (2015). How Inefficient Are Nutrient Application Limits? A Dynamic Analysis of Groundwater Nitrate Pollution from Concentrated Animal Feeding Operations. *Applied Economic Perspectives and Policy*, *37*(1), 130–150.

Warner, R. E. (1994). Agricultural Land Use and Grassland Habitat in Illinois: Future Shock for Midwestern Birds? *Conservation Biology*, *8*(1), 147–156. JSTOR.

Williams, M. R., & King, K. W. (2020). Changing Rainfall Patterns Over the Western Lake Erie Basin (1975–2017): Effects on Tributary Discharge and Phosphorus Load. *Water Resources Research*, *56*(3), e2019WR025985. https://doi.org/10.1029/2019WR025985

Wilson, S. M., & Serre, M. L. (2007). Examination of atmospheric ammonia levels near hog CAFOs, homes, and schools in Eastern North Carolina. *Atmospheric Environment*, *41*(23), 4977–4987. https://doi.org/10.1016/j.atmosenv.2006.12.055

Wisconsin Legislature: NR 102.06. (2020). https://docs.legis.wisconsin.gov/code/admin_code/nr/100/102/i/06

Withers, P. J. A., & Jarvie, H. P. (2008). Delivery and cycling of phosphorus in rivers: A review. *Science of The Total Environment*, *400*(1), 379–395. https://doi.org/10.1016/j.scitotenv.2008.08.002

Wörner, V., Kreye, P., & Meon, G. (2019). Effects of Bias-Correcting Climate Model Data on the Projection of Future Changes in High Flows. *Hydrology*, *6*(2), 46. https://doi.org/10.3390/hydrology6020046

Wurtsbaugh, W. A., Paerl, H. W., & Dodds, W. K. (2019). Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *WIREs Water*, *6*(5), e1373. https://doi.org/10.1002/wat2.1373

Xu, H., Brown, D. G., & Steiner, A. L. (2018). Sensitivity to climate change of land use and management patterns optimized for efficient mitigation of nutrient pollution. *Climatic Change*, *147*(3), 647–662. https://doi.org/10.1007/s10584-018-2159-5

Xu, Y. (2020). *Hydrology and Climate Forecasting R package—Hyfo* [R]. https://github.com/Yuanchao-Xu/hyfo

Yan, B., Shi, W., Yan, J., & Chun, K. P. (2017). Spatial distribution of livestock and poultry farm based on livestock manure nitrogen load on farmland and suitability evaluation. *Computers and Electronics in Agriculture*, *139*, 180–186. https://doi.org/10.1016/j.compag.2017.05.013

Yang, Q., Tian, H., Li, X., Ren, W., Zhang, B., Zhang, X., & Wolf, J. (2016). Spatiotemporal patterns of livestock manure nutrient production in the conterminous United States from 1930 to 2012. *Science of The Total Environment*, *541*, 1592–1602. https://doi.org/10.1016/j.scitotenv.2015.10.044

Yoon, S., & Nadvi, K. (2018). Industrial clusters and industrial ecology: Building 'eco-collective efficiency' in a South Korean cluster. *Geoforum*, *90*, 159–173. https://doi.org/10.1016/j.geoforum.2018.01.013

Yu, B., Shang, S., Zhu, W., Gentine, P., & Cheng, Y. (2019). Mapping daily evapotranspiration over a large irrigation district from MODIS data using a novel hybrid dual-source coupling model. *Agricultural and Forest Meteorology*, *276–277*, 107612. https://doi.org/10.1016/j.agrformet.2019.06.011

Yuan, S., Quiring, S. M., Kalcic, M. M., Apostel, A. M., Evenson, G. R., & Kujawa, H. A. (2020). Optimizing climate model selection for hydrological modeling: A case study in the Maumee River Basin using the SWAT. *Journal of Hydrology*, 125064.

Yue, S., & Wang, C. Y. (2002). Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resources Research*, *38*(6), 4-1-4–7. https://doi.org/10.1029/2001WR000861

Zhang, C., Luo, L., Xu, W., & Ledwith, V. (2008). Use of local Moran's I and GIS to identify pollution hotspots of Pb in urban soils of Galway, Ireland. *Science of The Total Environment*, *398*(1), 212–221. https://doi.org/10.1016/j.scitotenv.2008.03.011

APPENDIX A

SUPPLEMENTARY INFORMATION – BIAS CORRECTION OF CLIMATE MODEL OUTPUTS INFLUENCES WATERSHED MODEL NUTRIENT LOAD PREDICTIONS

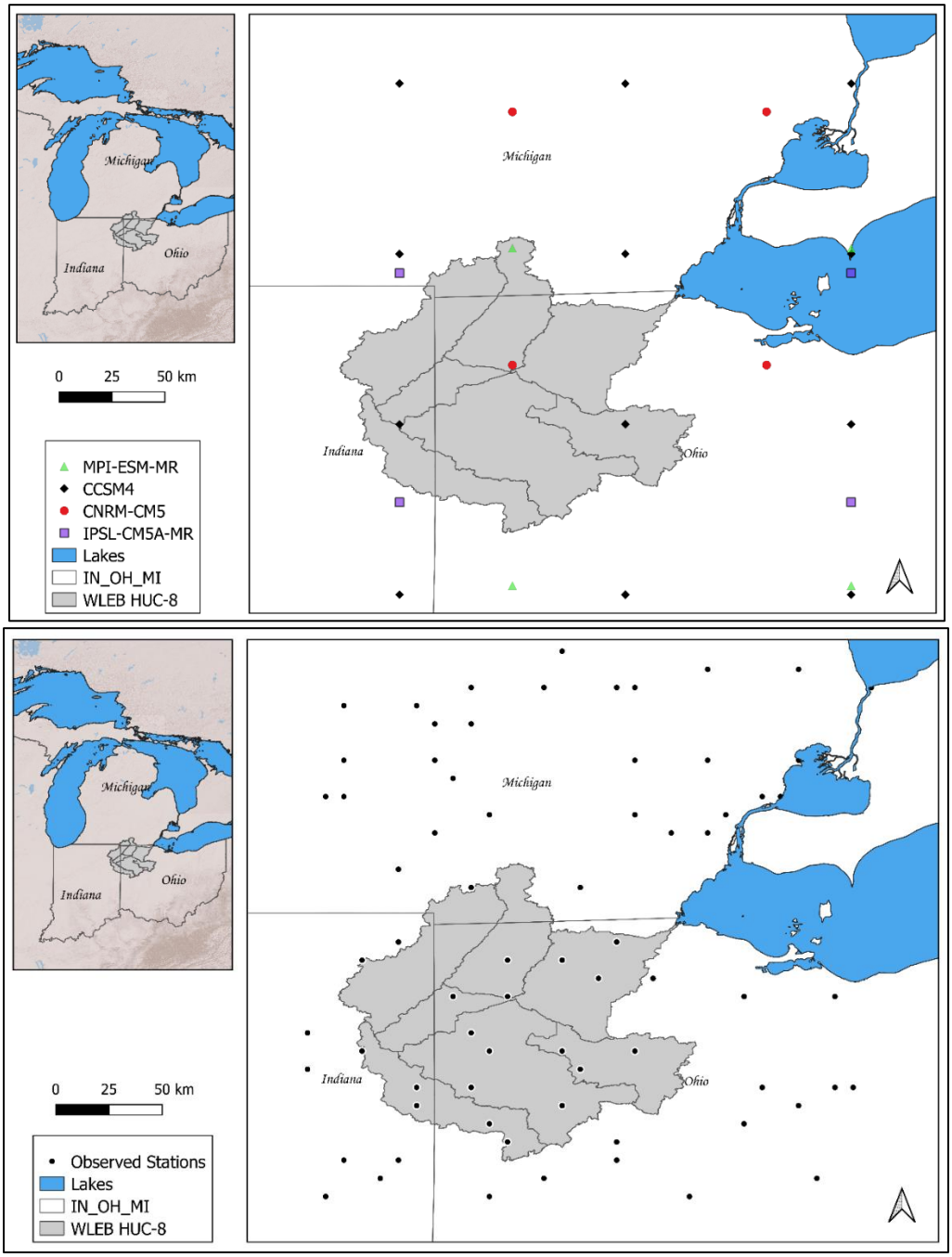
1.1. Table 1. Daily and monthly Apostel et al. (in review) SWAT model calibration and validation statistics for nutrient loads.

		Calibration (2005-2015)		**Validation (2000-2004)**	
	Statistic	**Daily**	**Monthly**	**Daily**	**Monthly**
TP	R^2	0.60	0.61	0.47	0.51
	NSE	0.58	0.52	0.46	0.44
	PBIAS	-3.76	-3.23	-18.53	-18.35
DRP	R^2	0.63	0.68	0.63	0.74
	NSE	0.62	0.67	0.63	0.73
	PBIAS	2.03	1.51	-9.89	-10.22
TN	R^2	0.63	0.78	0.75	0.82
	NSE	0.55	0.69	0.68	0.71
	PBIAS	-0.40	-1.24	-6.44	-6.73

1.2. Table 2. Seasonal Bias per climate model for temperature in K and Total precipitation in millimeters for the Western Lake Erie Basin. Models in bold are the ones used in the study.

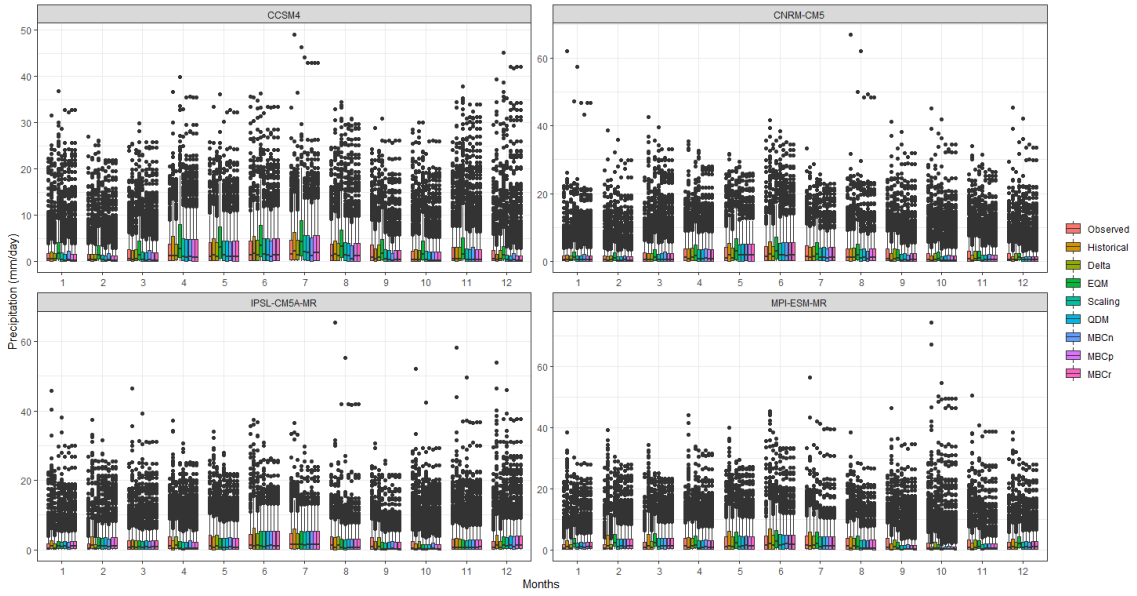
	Temperature (K)				**Total Precipitation (mm)**			
Models	SON bias	DJF bias	MAM bias	JJA bias	SON bias	DJF bias	MAM bias	JJA bias
ACCESS1-3	1.12	1.72	1.67	3.18	21.50	33.49	36.44	8.24
BCC-CSM1-1	0.97	-1.06	0.31	3.03	-4.46	14.76	10.42	-30.62
CanESM2	3.62	6.57	4.88	6.21	-14.27	17.69	28.95	-27.12
CCSM4	0.68	-0.82	1.93	1.45	-2.98	3.44	10.28	15.77
CESM1-CAM5	-0.60	-1.47	1.39	-0.82	-	-	-	-
CNRM-CM5	-1.18	-1.41	-1.00	-0.01	-10.09	-1.16	17.46	7.95
CSIRO-Mk3-6-0	0.71	-0.93	-1.68	3.18	-3.67	6.04	13.37	-27.23
FIO-ESM	0.26	2.98	2.86	0.84	-1.88	19.54	9.68	6.46
GFDL-ESM2M	-1.72	1.26	-1.89	-0.41	-5.62	23.75	25.18	16.52
GISS-E2-R	-1.29	-2.57	-0.59	-1.02	21.99	4.17	14.51	54.45
HadGEM2-AO	0.49	-2.72	1.64	3.17	1.90	17.07	17.54	-13.44
INMCM4	-1.03	1.56	0.56	-2.00	-3.81	23.35	34.70	8.49
IPSL-CM5A-MR	0.78	0.17	0.58	1.70	1.51	44.05	2.85	4.72
MIROC5	2.27	1.64	0.21	3.17	26.00	14.28	32.21	15.42
MIROC-ESM	1.62	3.91	3.49	3.58	11.05	2.63	-2.24	24.94
MPI-ESM-MR	0.76	0.71	0.27	-0.15	-1.94	45.10	39.83	24.46
MRI-CGCM3	-2.92	-2.05	-1.44	-1.54	21.95	14.18	29.40	2.92
NorESM1	-1.92	-1.65	1.17	-1.45	-9.36	-1.01	2.65	24.00

1.3.Figure 1. Spatial grid of each climate model (top) and the 71 weather stations (bottom) of the observed climate data used in this study.

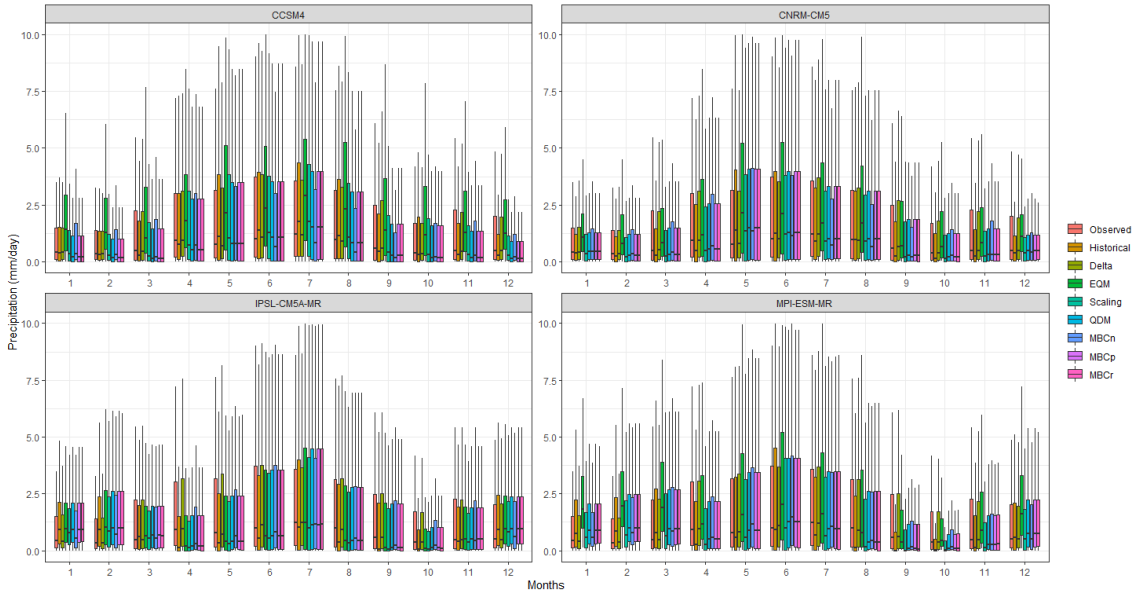


1.4.Figure 2. Plots of average monthly-sum precipitation in mm with (a.1) and without (a2) outliers; (b) monthly average of maximum temperature (TMAX) in Celsius; and (c) monthly average of minimum temperature (TMIN) in Celsius for the historical period (1980-1999) per each climate model and scenario (observed, historical (climate output not bias corrected), and bias corrected).

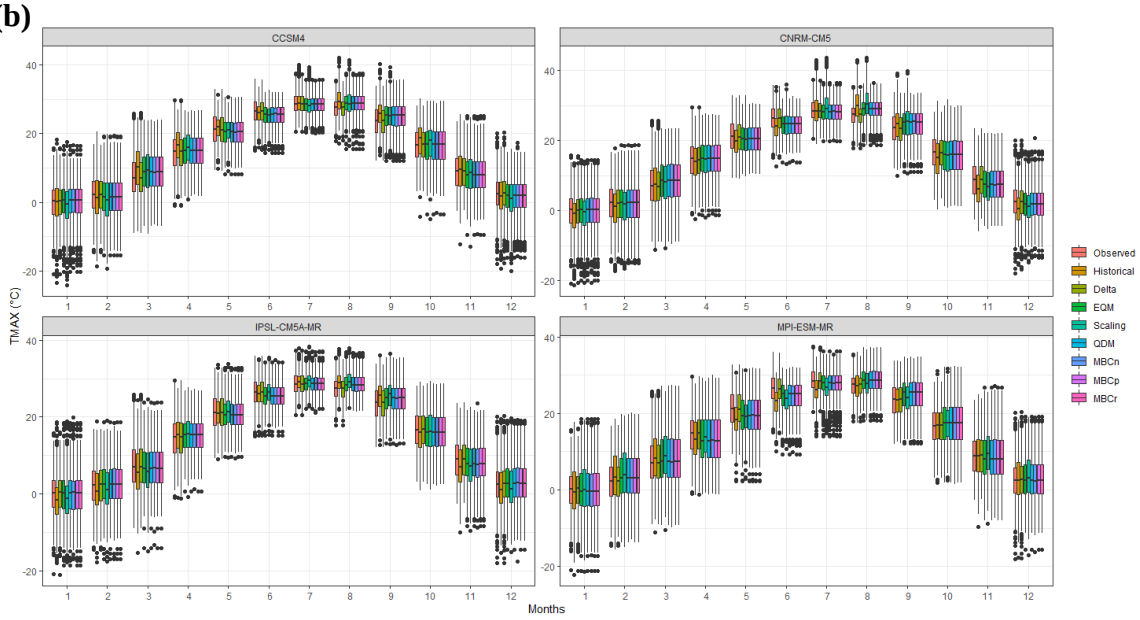
(a.1)



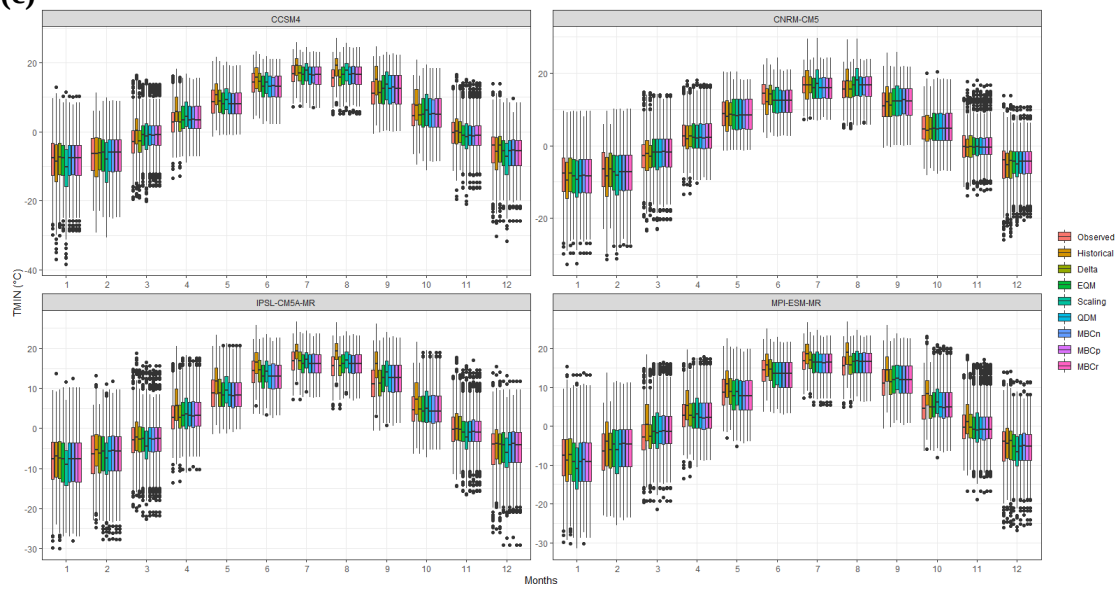
(a.2)



(b)


CCSM4
CNRM-CM5
IPSL-CM5A-MR
MPI-ESM-MR
TMAX (°C)
Months
Observed
Historical
Delta
EQM
Scaling
QDM
MBCn
MBCp
MBCr

(c)


CCSM4
CNRM-CM5
IPSL-CM5A-MR
MPI-ESM-MR
TMIN (°C)
Months
Observed
Historical
Delta
EQM
Scaling
QDM
MBCn
MBCp
MBCr

APPENDIX B

SUPPLEMENTARY INFORMATION – SPATIOTEMPORAL LAND USE CHANGE AND ENVIRONMENTAL DEGRADATION SURROUNDING CAFOS IN MICHIGAN AND NORTH CAROLINA

2.1. MODIS Product Information

2.1.1. Land Use and Land Cover (LULC)

For this study, LULC analysis was used to examine annual changes from 2001 to 2017 using the MODIS (Moderate Resolution Imaging Spectroradiometer) Land Cover Type 2 yearly 500-m resolution product (MCD12Q1 v006). We used the University of Maryland (UMD) classification with 16 distinct classes, which we further simplified to 9 classes: water, forest, shrubland, savannas, grassland, wetland, cropland, urban, and non-vegetated.

2.1.2. Percent Tree Cover (PTC)

We would expect a reduction in forested areas over time to indicate a decrease in ecosystem health (Estoque et al., 2018b). To analyze tree cover changes over time, we used PTC image layers from the Vegetation Continuous Fields yearly 250-m product (MOD44B v006). This product is a sub-pixel-level representation of tree cover, and the pixel value indicates the percent area covered by tree canopy with pixel values between 0 and 100 percent. A total of 17 PTC images were used in this study, each representing a single year from 2000 to 2016. As the image of 2017 to 2018 was not available at the time the analysis was performed, our analysis only examined changes from 2000 through 2016.

2.1.3. Leaf Area Index (LAI)

LAI is a key characteristic used in vegetation and ecosystems studies as a basic description of vegetation conditions (Asner et al., 2003b). Its measurement is based on the leaf area per unit ground area in broadleaf species, while LAI of conifers is determined as one-half the total needle surface area per unit ground area. For our study of changes over time, we examined the MODIS LAI image layer (MOD15A2H v006), an 8-day composite dataset at 500-m resolution, including images from 2000 to 2018. Each image uses the best pixel available from all image acquisitions within the 8-day period.

2.1.4. Normalized Difference Vegetated Index (NDVI)

NDVI is a well- known index used by numerous scientific studies as a measure of ecosystem health. It is mainly correlated with biophysical properties such as forest cover,

LAI, and biomass (Eckert et al., 2015; Ishtiaque et al., 2016b; Alatorre et al., 2016b). Therefore, NDVI is useful for the investigation of the environmental conditions surrounding CAFO facilities. We used the NDVI image layer in the Vegetation Indices (VI) MODIS product (MYD13Q1 v006) for this study. Each 250-m pixel in a single image contains the NDVI value acquired by the sensor within a 16-day period. The 16-day composite is generated from two 8-day composite Surface Reflectance granules in the 16-day period. Our analysis considered images for the summer period from 2002 to 2018.

2.1.5. Land Surface Temperature (LST)

Changes in LST are useful for studying ecosystem responses (Estoque et al., 2018b). Increases in LST can be associated with forest degradation, urbanization, and other anthropogenic activities driving changes in the environment. Here, we examined the 8-day per-pixel average of daily MODIS Land Surface Temperature and Emissivity images at 1000-m resolution (MOD11A2 v006). We used both daytime (LST-day) and nighttime (LST-night) surface temperature bands for the summer season (13 images per year) from 2000 to 2018.

2.1.6. Evapotranspiration (ET)

ET represents the water loss by plant transpiration and surface evaporation, and as fewer vegetated features are present in the environment, ET is likely to decrease (Ishtiaque et al., 2016b; Yu et al., 2019). For instance, a study found that changes in Myanmar's mangrove ecosystem were linked to a decrease in ET over time (Estoque et al., 2018b). Therefore, temporal measurements of ET may be useful in evaluating the health of CAFO-impacted areas. ET is estimated based on the Penman-Monteith equation that requires land cover, LAI/fPAR (fraction of Photosynthetically Active Radiation), and albedo inputs which are available in other MODIS products. For this study, we used the 8-day composite ET data from the MODIS Evapotranspiration/Latent Heat Flux 500-m product (MOD16A2 v006). The pixel values represent kilograms per square meter of ET across 8 days. For our study, we analyzed summer season images from 2001 and 2018.

2.1.7. Supplementary Table S-1. MODIS product characteristics and number of images used per state.

Variables*	MODIS product	Years	Spatial Resolution (m)	Temporal Resolution	Number of images
LC	MCD12Q1	2001 and 2017	500	Yearly	2
PTC	MOD44B	2000-2017	250	Yearly	17
LAI	MOD15A2H	2000-2018	500	8-day	247
NDVI	MYD13Q1	2002-2018	250	16-day	119
LST	MOD11A2	2000-2018	1000	8-day	247
ET	MOD16A2	2001-2018	500	8-day	234

* LC – Land Cover product; PTC – Percent Tree Cover; LAI – Leaf Area Index; NDVI- Normalized Difference Vegetated Index; LST – Land Surface Temperature; ET – Evapotranspiration
Data can be accessed via: NASA EARTHDATA - https://search.earthdata.nasa.gov/search

2.2. Land Use product reclassification

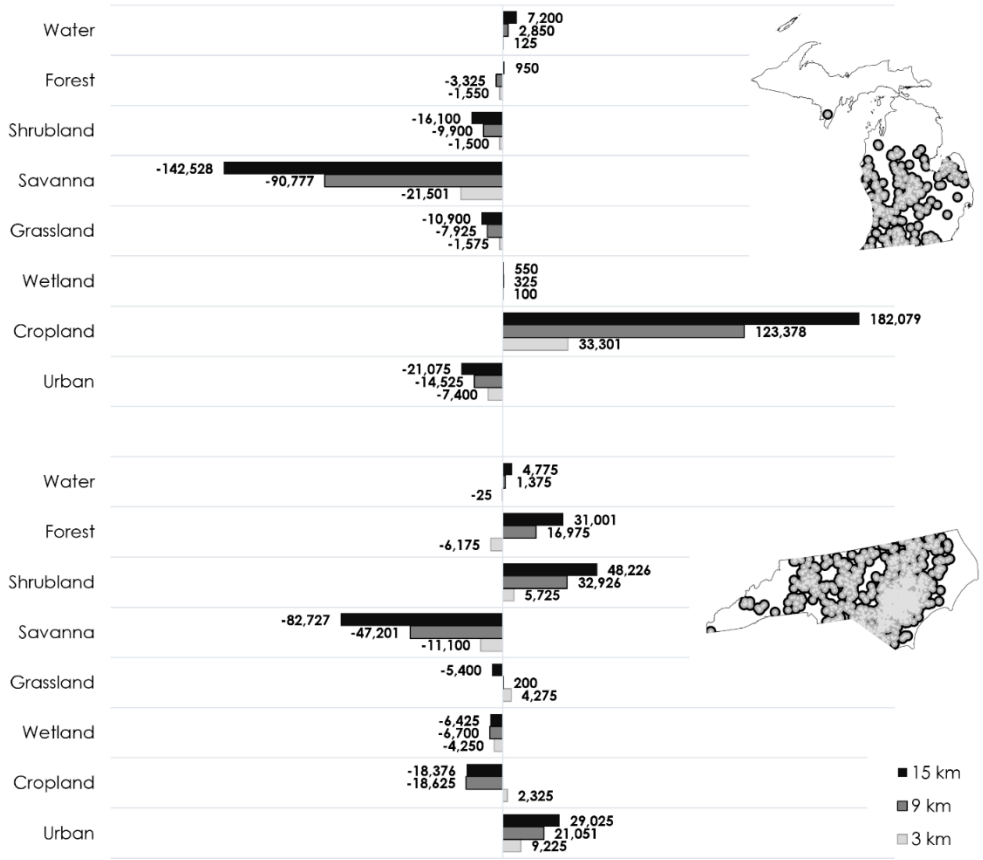
2.2.1. Supplementary Table S-2. MODIS Land cover product reclassification description used in this study.

MODIS Land Cover Class	Description*	Class in this study
Water bodies	At least 60% of area is covered by permanent water bodies.	*Water*
Evergreen Needleleaf Forests	Dominated by evergreen conifer trees (canopy >2m). Tree cover >60%.	*Forest*
Evergreen Broadleaf Forests	Dominated by evergreen broadleaf and palmate trees (canopy >2m). Tree cover >60%.	
Deciduous Needleleaf Forests	Dominated by deciduous needleleaf (larch) trees (canopy >2m). Tree cover >60%.	
Deciduous Broadleaf Forests	Dominated by deciduous broadleaf trees (canopy >2m). Tree cover >60%.	
Mixed Forests	Dominated by neither deciduous nor evergreen (40-60% of each) tree type (canopy >2m). Tree cover >60%.	
Closed Shrublands	Dominated by woody perennials (1-2m height) >60% cover.	*Shrublands*
Open Shrublands	Dominated by woody perennials (1-2m height) 10-60% cover.	
Woody Savannas	Tree cover 30-60% (canopy >2m).	*Savannas*
Savannas	Tree cover 10-30% (canopy >2m).	
Grasslands	Dominated by herbaceous annuals (<2m).	*Grasslands*
Permanent Wetlands	Permanently inundated lands with 30-60% water cover and >10% vegetated cover.	*Wetlands*
Croplands	At least 60% of area is cultivated cropland.	*Croplands*
Cropland/Natural Vegetation Mosaics	Mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation	
Non-Vegetated Lands	At least 60% of area is non-vegetated barren (sand, rock, soil) or permanent snow and ice with less than 10% vegetation.	*Non-vegetated*
Urban and Built-up Lands	At least 30% impervious surface area including building materials, asphalt, and vehicles.	*Urban*

*Description from Sulla-Menashe and Friedl, 2018. User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product: https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf

2.3. Radii distance sensitivity analysis

To evaluate the sensitivity of land use land cover change analysis to use of a 15km distance from CAFOs (i.e., 'CAFO-impacted' areas), we compared results using 3 km, 9km, and 15 km distances. The results remained generally constant in both study areas (Supplementary Figure S-2). Michigan (MI) land use change results were consistent among all three radii. Results for North Carolina were consistent in the 15 km and 9 km radii, while changes within the 3 km radius changes were even stronger and in agreement with our hypothesis and with the changes observed in MI (e.g., forest losses and gain in cropland area).



2.3.1. Supplementary Figure S-1. Sensitivity of land use land cover change analysis to radius distance used to define CAFO-impacted areas and control areas, illustrating changes detected in the vicinity of CAFOs in acres.

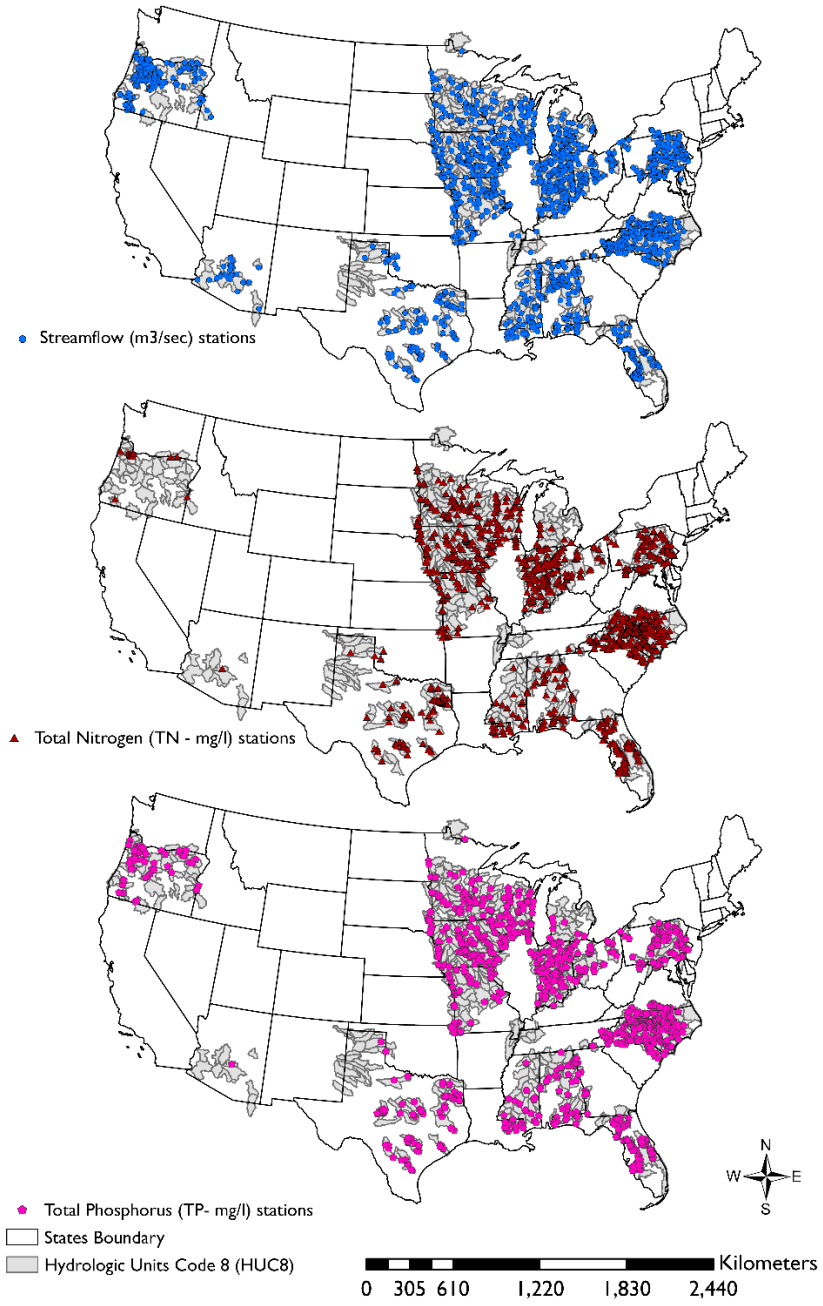
APPENDIX C

SUPPLEMENTARY INFORMATION –THE SPATIAL ORGANIZATION OF CAFOS AND ITS RELATIONSHIP TO WATER QUALITY IN THE U.S.

3.1. Section 1: Animal Units (AUs) calculation based on the Code of Federal Regulations, 1949-1984

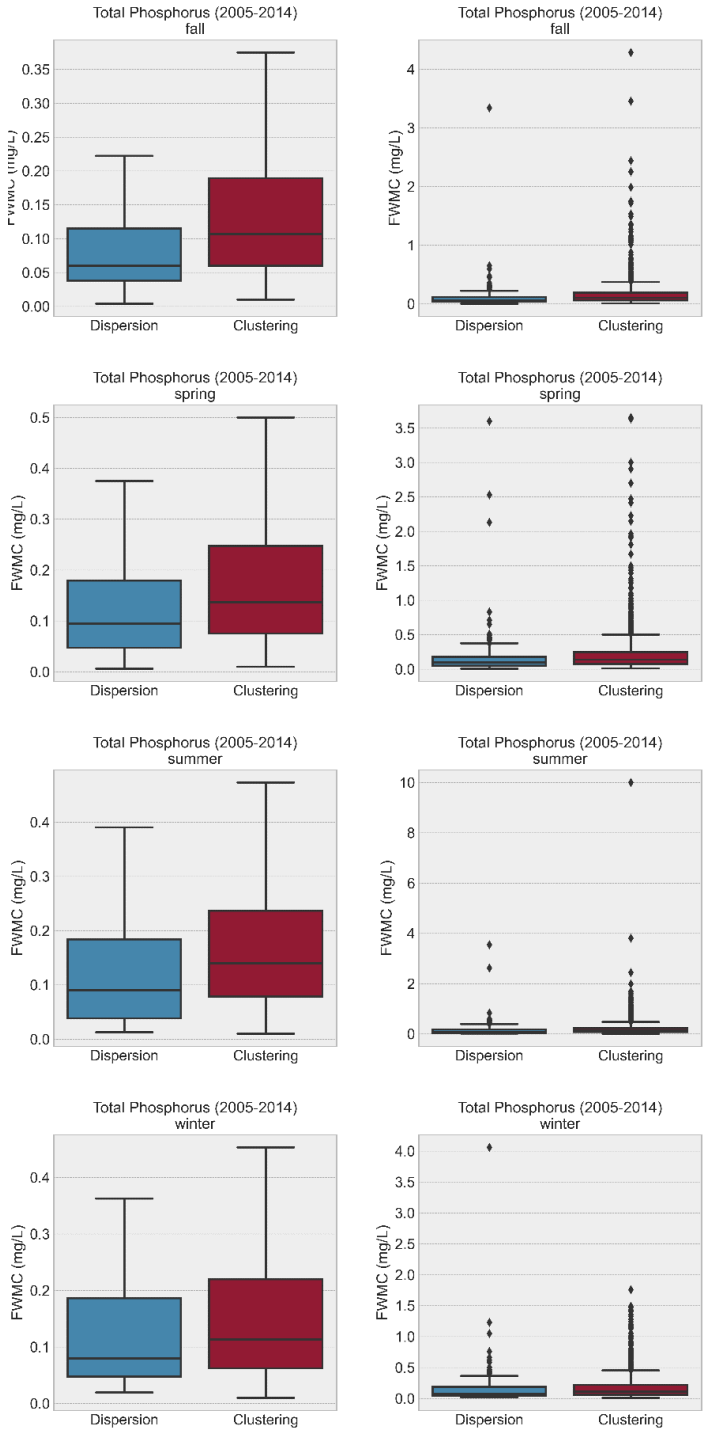
We based our AU calculation on the following regulation: *"003.03 ANIMAL UNIT means a unit of measurement for any animal feeding operation calculated by adding the following numbers: The number of slaughter and feeder cattle multiplied by 1.0, plus the number of cow/calf pairs multiplied by 1.2, plus the number of mature dairy cattle multiplied by 1.4, plus the number of swine weighing 55 pounds or more multiplied by 0.4, plus the number of weaned pigs weighing less than 55 pounds multiplied by 0.04, plus the number of sheep multiplied by 0.1, plus the number of horses multiplied by 2.0, plus the number of chickens multiplied by 0.01, plus the number of turkeys multiplied by 0.02, plus the number of ducks multiplied by 0.2. For immature dairy cattle or species not listed, the number of animal units shall be calculated as the average weight of the animals, divided by 1,000 pounds, multiplied by the number of animals."* If an animal type was not included in this description, we assigned the multipliers that closely matched the animals not stated by the regulation. For instance, the regulation does not mention lamb, but in general lams are closely related to sheep, so we assigned the multiplier 0.1 to lamb farms. The same was done to different types of chicken operations, we applied 0.2 as the multiplier for all of them across the states.

3.2. Section 2: Location of Total Phosphorus (TP), Total Nitrogen (TN), and Streamflow stations.

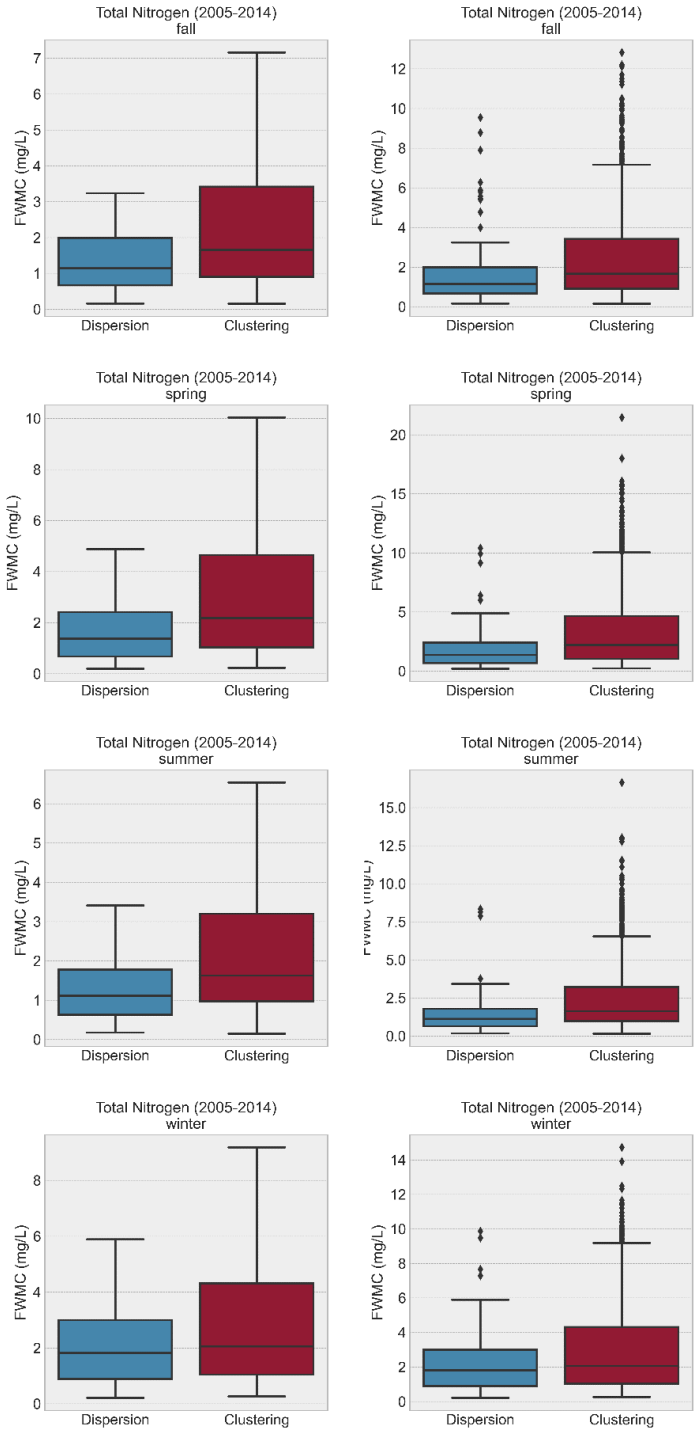


3.3. Section 3: Clustering and Dispersion watershed patterns and Total Phosphorus (TP) and Nitrogen (TN) concentrations.

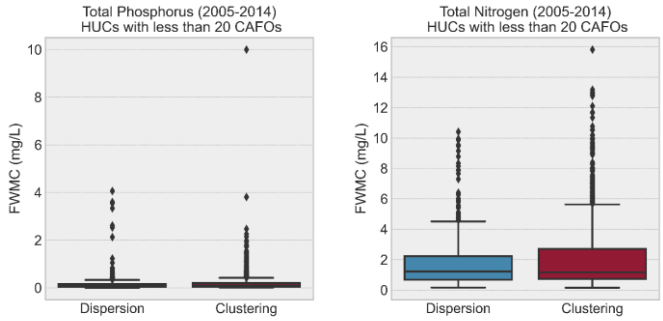
This section highlights seasonal TP and TN flow-weighted mean concentrations per spatial pattern identified. We observed that in general the highest concentrations of TP and TN occurred during the seasons in which manure tends to be land-applied (Figure 1 and 2). When comparing the watersheds with 20 CAFOs or less, we found that CAFO-clustered watersheds tend to be associated with the highest levels of TN and TP (Figure 3). However, TN median difference was not statistically significant at 0.05 level ($p=0.11$). We also compared the spatial patterns per state in the Corn Bel for TN and per west states for TP (Figure 4).



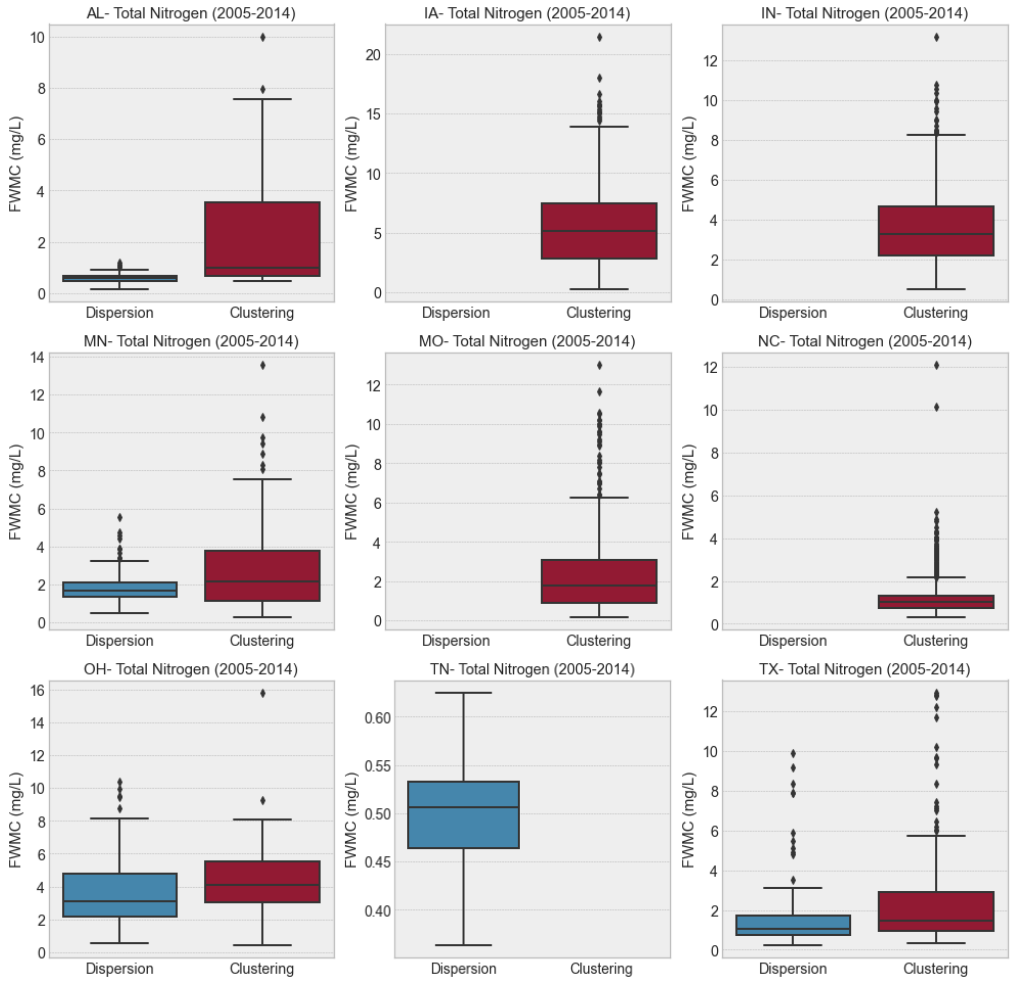
3.3.1. SI- Figure 1- Seasonal TP concentration with (right) and without outliers (left) per spatial pattern observed.



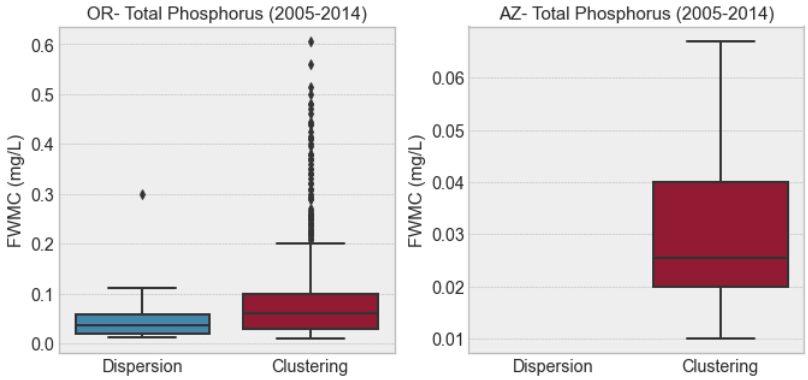
3.3.2. SI- Figure 2. Seasonal TN concentration with (right) and without outliers (left) per spatial pattern observed.



3.2.3. SI- Figure 3. Total Phosphorus and Total Nitrogen flow-weighed mean concentration per spatial pattern when hydrologic units (HUC8) had less than or equal to 20 CAFOs within their boundaries.



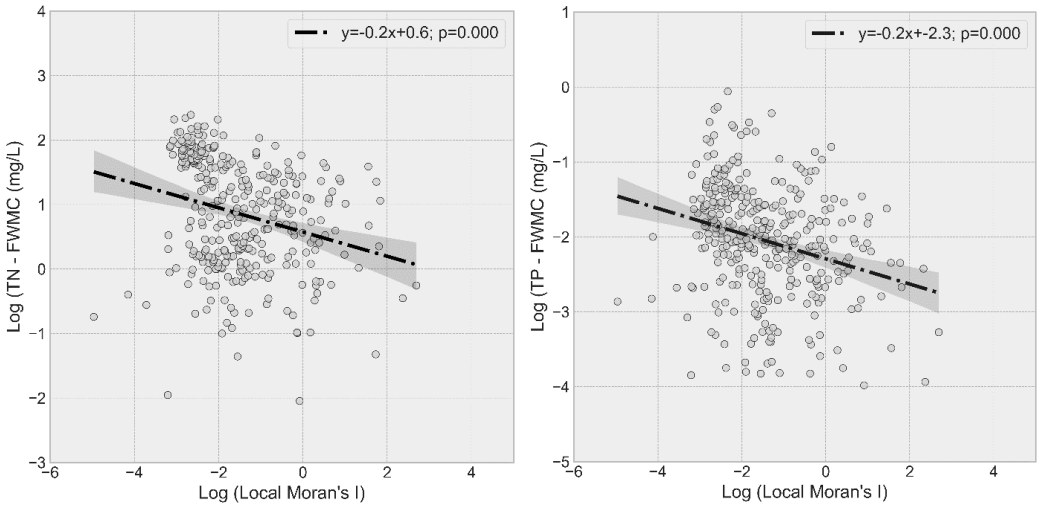
3.3.4. SI – Figure 4. Total Nitrogen flow-weighted mean concentration (FWMC – mg/l) per state in the Corn Belt and North Carolina, and per spatial pattern.



3.3.5. SI – Figure 5. Total Phosphorus (TP) flow-weighted mean concentration (FWMC – mg/l) per US west state and spatial pattern.

3.4. Section 4: The relationship between Local Moran's I indices of spatial clusters and water quality

We fit a linear model to detect what is the influence of Moran's I index in the prediction of TP and TN concentrations. We found that a 0.2 decrease in I would result in a unit increase in TP or TN. This parameter can be used in future water quality modeling studies.



3.4.1. SI-Figure 6. Average of flow-weighed mean concentration (FWMC – mg/l) of Total Phosphorus (TP) and Total Nitrogen (TN) per positive Local Moran's *I* indeces per spatial clusters (HH and LL)

3.5. Section 5: Counts of CAFOs per animal type per State

3.5.1. SI -Table 1. Number of CAFOs per state per animal type. Few CAFOs grow more than one type of animal in their facilities.

Iowa

Animal type	*# CAFOs*
Swine	9118
Beef	2821
Dairy	407
Chicken	206
Turkeys	157
Sheep	77
Horses	55

Missouri

Animal Type	*# CAFOs*
Swine > 55	379
Broilers	227
Swine < 55	95
Dairy	26
Layers	8
Beef	8
Turkeys	5
Goat	1
Pullets	1

North Carolina

Animal Type	*# CAFOs*
Swine	2248
Cattle	250
Wet Poultry	19
Horses	5
Other Animals	2
Dry Poultry	1

Indiana

Animal type	*# CAFOs*
Finishers_pig	1200
Nursery_pig	537
Sows	308
Beef	118
Turkeys	108
Dairy	92
Layers	79
Dairy_calve	76
Broilers	57
Dairy_Heif	56
Horses	54
Pullets	54
Beef_calve	44
Boars	43
Veal_calve	34
Poultry	13
Sheep	10
Ducks	3

Oregon

Animal Type	*# CAFOs*
Dairy cattle	248
Cattle Feedlots	137
Beef Cattle	34
Broiler chicken	25
All other animal production	25
Goat farming	14
Chicken egg	12
Hog and Pig Farming	10
Sheep and Lambs	7
Horses and Other Equine	5

Minnesota

Animal Type	*# CAFOs*
Swine 55-300 lbs	180
Dairy Cattle >1000 lbs	127
Beef Cattle - Cow & calf pair	124
Beef Cattle - Slaughter/Stock	122
Beef Cattle - Feeder/heifer	66
Turkeys >5 lbs	32
Swine >300 lbs	31
Horses	26
Dairy Cattle - Heifer	22
Swine <55 lbs	12
Beef Cattle - Calf	11
Turkeys <5 lbs	8
Chickens - Layers <5 lbs	7
Dairy Cattle - Calf	5
Dairy Cattle <1000 lbs	5
Chickens - Broilers <5 lbs	4
Chickens - Broilers >5 lbs	4
Goats	4
Chickens - liquid manure	3
Elk	3
Sheep, Lambs	3
Bison, Buffalo	2
Chickens - Layers >5 lbs	1
Foxes	1
Ducks - dry manure	1
Other	1

Wisconsin

Animal Type	*# CAFOs*
Dairy	287
Swine	12
Beef	7
Chickens	6
Turkeys	2
Ducks	1