Land-to-Water Linkages: Nutrients Legacies and Water Quality Across Anthropogenic Landscapes

by

Danyka K. Byrnes

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Civil Engineering

Waterloo, Ontario, Canada, 2024

© Danyka K. Byrnes 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Jana Compton, PhD Senior Research Ecologist United States Environmental Protection Agency Western Ecology Division - Corvallis, OR
Supervisor	Nandita Basu Professor Department of Civil and Environmental Engineering Department of Earth and Environmental Sciences University of Waterloo
Supervisor	Kimberly Van Meter Professor Department of Geography The Pennsylvania State University
Internal Member	Bryan Tolson Professor Department of Civil and Environmental Engineering University of Waterloo
Internal-External Member	Andrea Brookfield Professor Department of Earth and Environmental Sciences University of Waterloo
Internal-External Member	Helen Jarvie Professor Department of Geography and Environmental Management University of Waterloo

Author’s Declaration

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

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

Statement of Contributions

The various chapters in my dissertation had contributions from different research collaborators and coauthors:

Chapter 2 and 3: Shuyu Chang, and Drs. Kimberly Van Meter and Nandita Basu contributed to the research of this chapter. Chapter 2 is submitted to Scientific Data and is in review and first authorship is shared between Danyka Byrnes and Shuyu Chang.

- **D.K.B**, K.J.V.M, and N.B.B. conceived the methods for the county-scale datasets.
- **D.K.B** and K.J.V.M created the model to generate the county-scale datasets.
- S.C, K.J.V.M, **D.K.B**, and N.B.B. conceived the methods of the downscaling.
- S.C and K.J.V.M created the model to downscale the county-scale datasets.

Chapter 4: The following chapter is based on a manuscript submitted for peer-review with my co-authors Drs. Kimberly Van Meter and Nandita Basu.

- **D.K.B**, N.B.B, K.J.V.M conceptualized the study
- **D.K.B**, N.B.B, K.J.V.M developed the methods.
- **D.K.B** conducted the data curation.
- **D.K.B** and N.B.B conducted the formal analysis and visualizations.

Chapter 5: The following chapter is written based on a prepared manuscript that has not been submitted for peer-review with my co-authors Drs. Kimberly Van Meter and Nandita Basu.

- **D.K.B**, N.B.B, and K.J.V.M conceptualized the study
- **D.K.B**, N.B.B, and K.J.V.M developed the methods.
- **D.K.B** completed the data curation.
- **D.K.B** conducted the formal analysis and visualizations.

Chapter 6: The following chapter is written based on a prepared manuscript that has not been submitted for peer-review with my co-authors Drs. Kimberly Van Meter and Nandita Basu.

- **D.K.B**, K.J.V.M and N.B.B conceptualized the study.
- **D.K.B** and K.J.V.M developed the methods.
- **D.K.B** completed the data curation.
- **D.K.B** and K.J.V.M conducted the formal analysis and visualizations.

Abstract

An increasing population and the intensification of agriculture has driven rapid changes in land use and increases in excess nutrients in the environment. Globally, excess nutrients in inland and coastal waters have led to persistent issues of eutrophication, ecosystem degradation, hypoxia, and drinking water toxicity. Over the past few decades, we have seen policies set to mitigate the degradation in water quality. The existing paradigm of water quality management is based on decades of research finding a linear relationship between the net nitrogen inputs to the landscape and stream nitrogen exports. For instance, in the U.S., in response to these nutrient problems, working groups have spent approximately a trillion dollars to improve water quality by upgrading wastewater treatment plants and implementing nutrient management plans to decrease watershed nitrogen and phosphorus inputs. Despite concerted efforts, in many cases we have not seen marked improvements in water quality. In cases where water quality has improved, it is frequently after decades of nutrient management. The lack of or delayed water quality improvement suggests the importance of other drivers in modulating the relationship between nutrient inputs and watershed exports. Indeed, watershed nutrient loads are not just a function of current-year nitrogen inputs but can also depend on the history of inputs to the watershed. However, we still have little understanding of the relationship between nutrient inputs related to exports and the extent that accumulated stores of nitrogen and phosphorus influence this relationship.

The central theme of my research has been an exploration of the history of anthropogenic nutrient use and the relationship between nutrient inputs and the response in water quality. Specifically, I have focused on the role of current nutrient inputs versus historical nutrient use in impacting water quality at the watershed scale, as well as the various landscape and climate controls that can mediate responses to changes in management. My research objectives will be to (1) develop a multi-decadal mass balance of nitrogen and phosphorus at the sub-watershed scale across the contiguous U.S. in order to investigate (2) the relationship between watershed nitrogen inputs and export and the drivers of changes in watershed nitrogen export, (3) the magnitude, spatial distribution, and drivers of nitrogen retention and legacy stores, and (4) the use and management of phosphorus in agricultural landscapes in the context of both food security and environmental health.

I began by developing county-scale nitrogen and phosphorus surplus datasets, TREND-N and TREND-P, for the contiguous U.S.—with surplus defined as the difference between anthropogenic inputs (fertilizer, manure, domestic inputs, biological nitrogen fixation, and atmospheric deposition) and non-hydrological export (crop and pasture uptake). In Chapter 2, I present the updates to a previously published TREND-N county-scale nitrogen mass balance dataset, improving crop and pasture uptake and livestock excretion methods. In Chapter 3, I develope new county-scale phosphorus surplus dataset, using similar methods. These datasets were then downscaled to a 250 m gridded-scale dataset, known as gTREND-Nitrogen and gTREND-Phosphorus, a step led by my

collaborator Shuyu Chang. These novel datasets serve as the foundational data for the subsequent chapters.

Next, in Chapter 4, I explored the relationship between net nitrogen inputs and nitrogen export for over 400 watersheds across the U.S. I used the newly developed nitrogen surplus dataset to understand how watershed-scale nitrogen surplus magnitudes and exports change over time and examine how the relationships are influenced by both natural and anthropogenic controls within watersheds. To achieve this, I used a set of 492 watersheds with nitrogen input and export data spanning from 1990 to 2017. We found that 284 watersheds had a significant ($p<0.1$) increasing or decreasing trend in both nitrogen surplus and nitrogen load. Of these watersheds, we identified 62 where both nitrogen surplus and export have been significantly increasing over the last two decades. These input-driven watersheds are characterized by high livestock density, agricultural area, and tile drainage. In contrast, nitrogen surplus and export have been decreasing in 127 "bright spot" watersheds, characterized by high population density and urban land use. Nitrogen surplus is also decreasing in 60 "transitioning" watersheds, but export is increasing as nitrogen surplus decreases. We argue that these watersheds are transitioning from agriculture to more urban areas, such that fertilizer inputs have decreased, but the higher nitrogen export is driven by legacy nitrogen stores. Finally, we found 35 watersheds demonstrating a delayed response, with nitrogen export decreasing despite an increase in nitrogen surplus. Climate appears to be the driver of response in these watersheds, with aridity likely driving lower nitrogen export, despite increasing inputs. The four typologies of nitrogen inputs and export relationships suggest that watersheds can act as filters and modulate the movement of nitrogen. Our results provide insights into the complex dynamics of nitrogen surplus and export relationships, as well as how the landscape, climate, and legacy nitrogen can influence these relationships.

In Chapter 4, I analyzed relationships between changes in nitrogen inputs and export, to understand what drives changes in watershed export, finding that legacy stores may be modulating the watershed response to changing net nitrogen inputs. However, we have limited knowledge of the magnitude and spatial distribution of legacy stores across North America. Therefore, in Chapter 5, we quantified how much nitrogen retention, which is the mass of nitrogen stored in legacy pools and nitrogen lost to denitrification, has accumulated in watersheds, and where it can be accumulating. To achieve this, we used existing datasets and machine learning algorithms to calculate the mass of ‘retained’ nitrogen in the landscape—defined as the nitrogen stored in the soil organic nitrogen pool, the groundwater pool, or lost through denitrification. Specifically, we built a random forest modeling framework trained on the watersheds’ nitrogen surplus and components, loads, and characteristics to predict nitrogen loads at the HUC8 scale across the U.S. We calculated retention for HUC8, which is the difference between nitrogen surplus and predicted loads, and found that nitrogen retention is highest in the Midwestern and Eastern U.S. because of low exports in regions with high agricultural inputs or high population density. Next, we used a data-driven approach to estimate legacy stores by

allocating retained nitrogen mass into their legacy pools. We partition nitrogen retention in the Upper Mississippi region HUC8 watersheds into the mass stored in the groundwater pool, soil organic nitrogen pools, and mass lost to denitrification. We found that, on average, 42% of the mass is stored in the soil organic nitrogen pool, 16.5% is stored in the groundwater pool, and 40% is lost to denitrification.

While these two chapters focused on nitrogen, in my final chapter we shifted to explore phosphorus use in agricultural systems. In my final chapter, we used the new gridded phosphorus surplus and components dataset to explore current and historical agricultural phosphorus use and management in landscapes within the context of both food security and environmental health. To characterize the extent of phosphorus depletion and excess, we employed indicators such as annual and cumulative phosphorus surplus and phosphorus use efficiency (PUE). We found that the evolution of agricultural phosphorus management is shaped by changing fertilizer management, the proliferation of concentrated animal operations, climate, and the landscape's memory of past phosphorus use. We further integrated both cumulative phosphorus surplus and PUE into a framework to quantify phosphorus sustainability in intensively managed landscapes. We found that in the 1980s, much of the agricultural land was undergoing ‘intensification,’ with positive and increasing cumulative stores because phosphorus inputs exceeded crop uptake (PUE < 1). By 2017, 29.5% of the agricultural land was undergoing ‘recovery’ and had positive cumulative phosphorus stores that were being depleted through improved phosphorus management (PUE > 1). However, 70% of the agricultural area in the U.S. is still undergoing ‘intensification,’ particularly in areas with more of their inputs from livestock manure, pointing to the need to treat manure as a resource instead of the current approach of treating it as a waste product.

By using novel datasets, we have been able to explore nutrient use across space and time and its impact on food security and environmental outcomes. I have made significant contributions towards expanding the discussion of nutrient us and fate, understanding the magnitude and distribution of cumulative net nutrient inputs stores in the landscape, as well as the ways in which intrinsic watershed properties, climate, land management, and historical nutrient use can modulate the relationship between inputs and export. Overall, my findings underscore the importance of nuanced, place-based, and context-dependent nutrient management strategies, with a focus on manure management, to address the diverse challenges of different agricultural systems and prevent unintended environmental consequences.

Acknowledgements

This research has been made possible with the financial support from the Odanak band, University of Waterloo, NSERC Alexander Graham Bell scholarship, the AGU Horton research grant, and various funding streams from my advisors including the Lake Futures grant.

I would like to thank my committee members, Bryan Tolson, Helen Jarvie, Andrea Brookfield, and Jana Compton for their willingness to engage with my research.

I would like to express my deepest gratitude to Dr. Nandita Basu and Dr. Kimberly Van Meter for their nine years of mentorship and support. Our discussions during work hours over zoom or at 5pm at a margarita bar have been instrumental in shaping me into the researcher I am today. I am also deeply thankful for the lab environment you both have fostered, a place where I have found not only excellent colleagues but lifelong friends. A special thanks to Tyler, Lamisa, Fred, Meghan, Simon, Caitlyn, Julianne, and Julia—your friendship and mentorship have been my anchor throughout this graduate school journey. Tyler and Caitlyn, I especially appreciate our extensive discussions that helped me refine my positionality within our work, allowing me to discover significant meaning beyond my specific niche.

Thank you to my coworking furry friends: Tookie, Hetfield, Ciuccio, Yuki, Meeko, Hazelnut, and Frankle for their companionship and occasional distractions that brought joy into my daily routine.

To my family—Al, Syl, Mercedes, Dylan, and David—thank you for your unwavering support and validation through the many ups and downs of the past few years. I deeply appreciate how you always took the time to understand my work so you could accurately and proudly share my achievements, all while reminding me of the importance of the bigger picture.

To my broader community, especially the all star line-up: Emilia, Garima, Tim, Michael, Laura, Kat, Alana, and Sarah—thank you for enthusiastically giving me the space to talk about my work when I needed to, but more importantly, for providing me with a rich life outside of it.

And finally, to Samuel—thank you for always grounding me, engaging in my curiosities, and encouraging me to find balance. But most importantly, thank you for reminding me of who I am and bringing me back home.

Dedication

This thesis is dedicated to my family. To the generations before me, who in the face of dispossession, displacement, hardship, and loss, worked to build a bright future. I stand on your shoulders and I too work with hope and purpose for those who come after me.

From earth, from water,
our people grew to love each other, in this manner.
For in all our languages, there is no he or she.
We are the children of the earth and of the sea.
– Alanis Obomsawin

Table of Contents

Examining Committee Membership ii
Author's Declaration iii
Statement of Contributions iv
Abstract v
Acknowledgements viii
Dedication ix
List of Figures xiii
List of Tables xv
Chapter 1 Introduction 1
1.1 Nitrogen in the anthropocene 1
1.2 Policy response to address water quality degradation 2
1.3 Fate of nutrients in the landscape 6
1.3.1 Nitrogen fluxes and legacies 6
1.3.2 Phosphorus fluxes and legacies 9
1.3.3 Quantifying and modeling nitrogen legacy and time lags 10
1.5 Knowledge gaps and research objectives 11
Chapter 2 Long-term nitrogen mass balance data for the contiguous United States (1930-2017) 12
Outline 12
2.1 Introduction 12
2.2 Methods 15
2.2.1 County-Scale Nitrogen Mass Balance Data 16
2.2.2 Creation of the gTREND-Nitrogen gridded dataset 27
2.3 Technical validation 30
2.3.1 Comparison of the county-scale TREND-Nitrogen v2 and v3 datasets 30
2.3.2 Validation of gTREND-Nitrogen with county-scale data 32
2.3.3 Use of gTREND vs county-scale TREND-Nitrogen for watershed-scale analysis 34
2.4 Limitations 35
2.5 Potential Use 36
2.6 Data Records 36
Chapter 3 Phosphorus gridded and county-scale TRajEctories Nutrient Dataset (TREND) for the contiguous United States (1930-2017) 40
Outline 40
3.1 Introduction 40
3.2 Methods 42
3.2.1 Dataset preparation: county-scale phosphorus mass balance 42
3.2.2 Creation of the gTREND-Phosphorus gridded dataset 52
3.3 Data Validation 55
3.3.1 Validation of gTREND-Phosphorus with county-scale data 55
3.5 Limitations 57
3.5 Data Records 57

3.4.1 TREND-Phosphorus dataset 57
3.4.2 gTREND-Phosphorus dataset 59
Chapter 4 Nitrogen Memoryscapes: Decoupling of Nitrogen Inputs and Riverine Export across the Continental United States 61
Outline 61
4.1 Introduction 61
4.2 Methods 65
4.2.1 Data sources and site selection for streamflow and nitrogen concentrations 65
4.2.1.2 Watershed nitrogen inputs and outputs 66
4.2.1.3 Spatial and temporal statistical analysis 66
4.3 Results and discussion 68
4.3.1 Spatial patterns in nitrogen surplus and nitrogen export 68
4.3.2 Temporal trends in nitrogen surplus and nitrogen export 69
4.3.3 Typologies of Watershed Responses to Nitrogen Dynamics: Decoupling and Landscape Drivers 72
4.3.4 Digging deeper: heterogeneity within quadrants provide insights into nutrient sources 76
4.4 Conclusion 77
Chapter 5 Nitrogen Retention across Continental US 79
Outline 79
5.1 Introduction 79
5.2 Methods 81
5.3.1 Model training data 83
5.3.2 Hydrological, climatic, and land management predictor dataset 83
5.3.3 Training model to estimate nitrogen loading in ungauged watersheds 84
5.3.4 Predicting nitrogen export, fate and legacy pools 86
5.3 Results and Discussion 89
5.3.1 USGS watershed model training data and results 89
5.3.2 Regional variability in nitrogen retention and the role of input sources. 93
5.3.3 Legacy pools in the Upper Mississippi region 101
5.4 Conclusion 105
Chapter 6 Phosphorus Sustainability in U.S. Agriculture: A Quadrant Framework Analysis 106
Abstract 106
6.1 Introduction 106
6.2 Methods 108
6.2.1 Data sources 108
6.2.2 Phosphorus sustainability metrics 108
6.3 Results and Discussion 110
6.3.1 Agricultural phosphorus usage across the contiguous United States 110
6.3.2 Toward a holistic approach for landscape socio-environmental evaluation 116
6.4 Conclusion 120
Chapter 7 Conclusions and Perspectives 122
7.1 Summary of major findings 122
7.2 Recommendations for future work 124

References 127
Appendix (Supplemental Figures and Tables) 150

List of Figures

Figure 1.1 Timelines of policy measures and hypoxic zone size reveal a lack of response to policy interventions across Europe and North America. 5
Figure 1.2 Transport and retention of anthropogenic nutrients across human-impacted landscapes..... 7
Figure 1.3 A conceptual diagram demonstrates the terminology used to describe the different how to group nitrogen pools in the subsurface. 8
Figure 2.1 N surplus magnitudes, as calculated from gTREND-Nitrogen data for the years (a) 1930, (b) 1960, (c) 1980, and (d) 2017 at a 250-m grid scale. 15
Figure 2.2 Data sources and workflow for creating the gTREND dataset..... 16
Figure 2.3 Agricultural and developed land use within the contiguous U.S..... 30
Figure 2.4 The TREND-Nitrogen county-scale nitrogen mass balance..... 31
Figure 2.5 Differences between the county-scale TREND-Nitrogen v2 and the new TREND-Nitrogen v3. 32
Figure 2.6 Relationship between county-scale TREND-Nitrogen v3.0 estimates of N mass balance components and the gridded gTREND-Nitrogen estimates, re-aggregated to the county scale. 33
Figure 2.7 Comparison of watershed N surplus magnitudes based on county-scale and gridded N surplus data..... 35
Figure 3.1 Data sources and workflow for creating the gTREND-Phosphorus dataset..... 43
Figure 3.2 Relationship between county-scale TREND-Phosphorus estimates of P mass balance components and the gridded gTREND-Phosphorus estimates, re-aggregated to the county scale..... 57
Figure 4.1. Nitrogen export ratio (E_R = Export/Surplus) as a function of climate, land use and management controls (2003-2017). 69
Figure 4.2 Watershed trends in nitrogen surplus (a-d) and export (e-h) over the last three decades (1990 - 2017). 71
Figure 4.3 Relationship between annual percent change of nitrogen surplus and export. 74
Figure 4.4 (a-d) Spider plot capturing the median characteristics of watersheds in each quadrant and (e) the distribution of select quadrant characteristics. 75
Figure 4.5 Comparison of the distribution of agricultural midwestern watershed attributes in Q1, Q2, and Q3 (agricultural land use > 25%). 77
Figure 5.1 Conceptual figure of our methods for estimating the nitrogen retained in HUC8 watersheds across the U.S..... 82
Figure 5.2 Summary of USGS riverine load training data. 90
Figure 5.3. Model predictor importance quantified through predictor importance values and SHapley Additive exPlanations (SHAP) values. 93
Figure 5.4 HUC2 regions input and fate of cumulative nitrogen (1958–2017). 94
Figure 5.5. Spatial distribution of cumulative nitrogen inputs and exports across HUC8 watersheds (1958–2017) measured in kg-N ha^{-1}. 96
Figure 5.6. Spatially varying relationship between crop use efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by the retention efficiency (%). 98
Figure 5.7 Relationships between input retention (%) and watershed properties across different HUC02 regions..... 100

Figure 5.8 Predicted mass of landscape retention, nitrogen stored in the soil organic and groundwater pool, and the nitrogen loss through denitrification. 103
Figure 6.1 Maps of the major components for the agricultural phosphorus mass balance: (a-c) fertilizer P, (d-f) manure P, and (g-i) crop phosphorus uptake, all in kg ha^{-1} y^{-1}. 111
Figure 6.2 Maps of (a-c) agricultural phosphorus surplus, and (d-f) phosphorus use efficiency. 113
Figure 6.3 Phosphorus use efficiency and input for three distinct hydrological regions across the U.S. . 115
Figure 6.4 National and regional cumulative phosphorus surplus from 1930 - 1980 and 1930 - 2017. ... 116
Figure 6.5 A composite diagram of the classification of agricultural land in the multi-indicator framework and how the quadrant classification changes over time. 117
Figure 6.6 Box plot of the percent of the total inputs are from manure for (a) 1980 and (2) 2017 for each quadrant. 120

List of Tables

Table 2.1 Livestock nitrogen consumption and excretion parameters. Consumption values are included for all grazing livestock. 21
Table 2.2 State-level confinement factors for beef cows, milk cows, and other cattle. 24
Table 2.3 Biological nitrogen fixation (BNF) parameters. 26
Table 2.4 Structure and File Names of the Gridded gTREND-N Dataset in the Data Repository. 38
Table 3.1 Livestock phosphorus consumption and excretion parameters. Consumption values are included for all grazing livestock. 47
Table 3.2 Structure and File Names of the TREND-P Dataset in the Data Repository. 58
Table 3.3 Structure and File Names of the Gridded gTREND-P Dataset in the Data Repository. In the file names, the YYYY values refer to the data year. 60

Chapter 1 Introduction

1.1 Nitrogen in the anthropocene

The increasing human population over the past 100 years has perturbed the phosphorus and nitrogen cycle in large part to meet the needs of a growing population. Globally, the fixation of anthropogenic nitrogen—through inorganic fertilizers, fossil fuel combustion, and nitrogen-fixing crops—has risen dramatically from 15 Tg N y^{-1} in 1890 to 210 Tg N y^{-1} in 2010 (Galloway & Cowling, 2002, 2021). Phosphorus fertilizer from phosphate rock has increased from 0.4 Tg P y^{-1} to 19 Tg P y^{-1} (Ashley et al., 2011). We see these global patterns reflected in the US, where both anthropogenic nitrogen and phosphorus inputs began to increase dramatically after the 1940s and have steadily increased since (Byrnes et al., 2020). Crop production has increased in tandem with inputs, but in many cases inputs still exceed crop needs. Across all regions in the U.S, nitrogen use efficiency (the ratio of crop uptake to total inputs) has decreased from 48% to 45% from the mid-1990s to mid-2000s, and phosphorus use efficiency has only slightly increased from 60 to 63% (Swaney et al., 2018b; Swaney & Howarth, 2019b). This leaves a large proportion of nitrogen and phosphorus to cycle through the environment.

Excess nitrogen and phosphorus in the environment contribute to widespread environmental degradation through the pollution of surface and groundwater resources. Nutrient runoff from the landscape can be transported to inland and coastal water bodies, where it promotes algal growth and can lead to cultural eutrophication and hypoxia (dissolved oxygen levels < 2 mg L^{-1}) (Malone & Newton, 2020). Over the past few decades, the incidence of nutrient-induced eutrophication and persistent hypoxic zones has increased largely due to rising nutrient loads. This has become a global issue, plaguing many coastal and lacustrine ecosystems (Beaulieu et al., 2019; Breitburg et al., 2018; Stoddard et al., 2016). We find eutrophication and hypoxia in each continent excluding Antarctica (Rabalais et al., 2009). This includes the Baltic Sea in Europe, Chesapeake Bay and the Gulf of Mexico in the U.S, Lake Erie in North America, and Lake Taihu in China (Conley et al., 2011; Hagy et al., 2004; Harding et al., 2019; Rabalais & Turner, 2019; Xu et al., 2015). This degradation poses a significant threat to a wide range of ecosystem services including tourism, recreation, and fisheries, as well as negatively impacting economic opportunities, human well-being, and ecosystem health (Conley et al., 2011; R. W. Howarth et al., 2011; Postel & Carpenter, 2012; Rabalais, Turner, & Scavia, 2002).

Nitrate is also a ubiquitous groundwater contaminant highlighted by the many cases globally of concentrations exceeding the World Health Organization's (WHO) maximum contaminant level (MCL) of 10 mg NO_3^--N L^{-1}. Nitrate concentrations above this limit have been documented in

40% of wells in Denmark, 28% in China, 25% in Zona Citricola, MX, and 90% in the Gaza Strip (Gu et al., 2013; Hansen et al., 2017; Pastén-Zapata et al., 2014; Shomar et al., 2008). In the U.S., nitrate is an ongoing and worsening issue, as demonstrated by a longitudinal study by Puckett et al. (2011), which found a significant increase in the percentage of sampled groundwater wells exceeding the MCL, rising from 2.2% in 1960 to 39% by 1980, with this trend anticipated to continue. Subsequent studies of drinking water wells capture the dire conditions of groundwater across the U.S. In 2007, approximately 45% of drinking water wells had concentrations exceeding the MCL (Toccalino & Hopple, 2010). Alarmingly, the population served by groundwater wells in violation of the MCL limits has remained relatively constant over the past two decades (Pennino et al., 2017) despite knowing that chronic exposure to elevated nitrate concentrations, even well below regulatory limits, has been linked to increased risk of cancer and is the known cause of methemoglobinemia in infants (Ward et al., 2018). This issue is particularly pressing given that approximately 43 million people rely on groundwater for drinking water, and an estimated 15 million people live in areas with groundwater nitrate concentrations exceeding natural levels (1 mg-N L^{-1}) (Dubrovsky et al., 2010). For people relying on private wells, a common situation for rural farmers, the options are limited and often costly: they must choose between investing in expensive filtration systems, purchasing bottled water, or consuming water with nitrate concentrations that can cause adverse health outcomes (Healy, 2018).

1.2 Policy response to address water quality degradation

Awareness surrounding the impacts of excess nutrients in the environment has spurred concerted efforts to reduce nutrient loads and address water quality (Backer et al., 2010; Diaz et al., 2017; Gren & Destouni, 2012; Mississippi River/Gulf of Mexico Watershed Nutrient Task Force, 2001; Van Meter et al., 2018). The 2001 Mississippi River/Gulf of Mexico Watershed Nutrient Task Force developed an action plan to manage point source and nonpoint source loading to reduce the 5-year average hypoxic zone extent to less than 5,000 km^2 by 2015 (USEPA, 2023). Earlier plans to address hypoxia and eutrophication were launched in other regions, including in the Chesapeake Bay, U.S., and the Baltic Sea, Europe (Destouni et al., 2017; Hagy et al., 2004; Lefcheck et al., 2018; Van Meter et al., 2018). In the 1980s, the Chesapeake Bay Program developed an action plan to implement a basin-wide strategy to reach 40% nitrogen and phosphorus load reduction by 2000 (Chesapeake Bay Program, 1988). The 1988 Helsinki Commission (HELCOM) mitigation plan, now included in the Baltic Sea Action Plan, outlined

nutrient loading targets for each country in the Baltic Sea drainage basin (Destouni et al., 2017; Vero et al., 2017).

Stakeholders across Europe and the U.S. have spent billions of dollars annually implementing point source controls and agricultural best management practices (BMP) (Hyytiäinen et al., 2014; Keiser & Shapiro, 2017). Despite such efforts, nutrient loads and hypoxic zone extents have been either increasing or have not recorded any significant decline in the last three decades, forcing target years to be reset for the future **(Figure 1.1)** (Destouni et al., 2017; L. Feng et al., 2024; Stackpoole et al., 2021; Van Meter et al., 2018; *Water Quality Standards Attainment and Monitoring - Chesapeake Progress*, 2021). One key reason for the apparent lack of success in meeting water quality goals is the time lags associated with legacy stores of nitrogen and phosphorus in highly managed watersheds, which act as a long-term source of nutrients to surface waters (Meals et al., 2010; Sharpley et al., 2013; Van Meter et al., 2018).

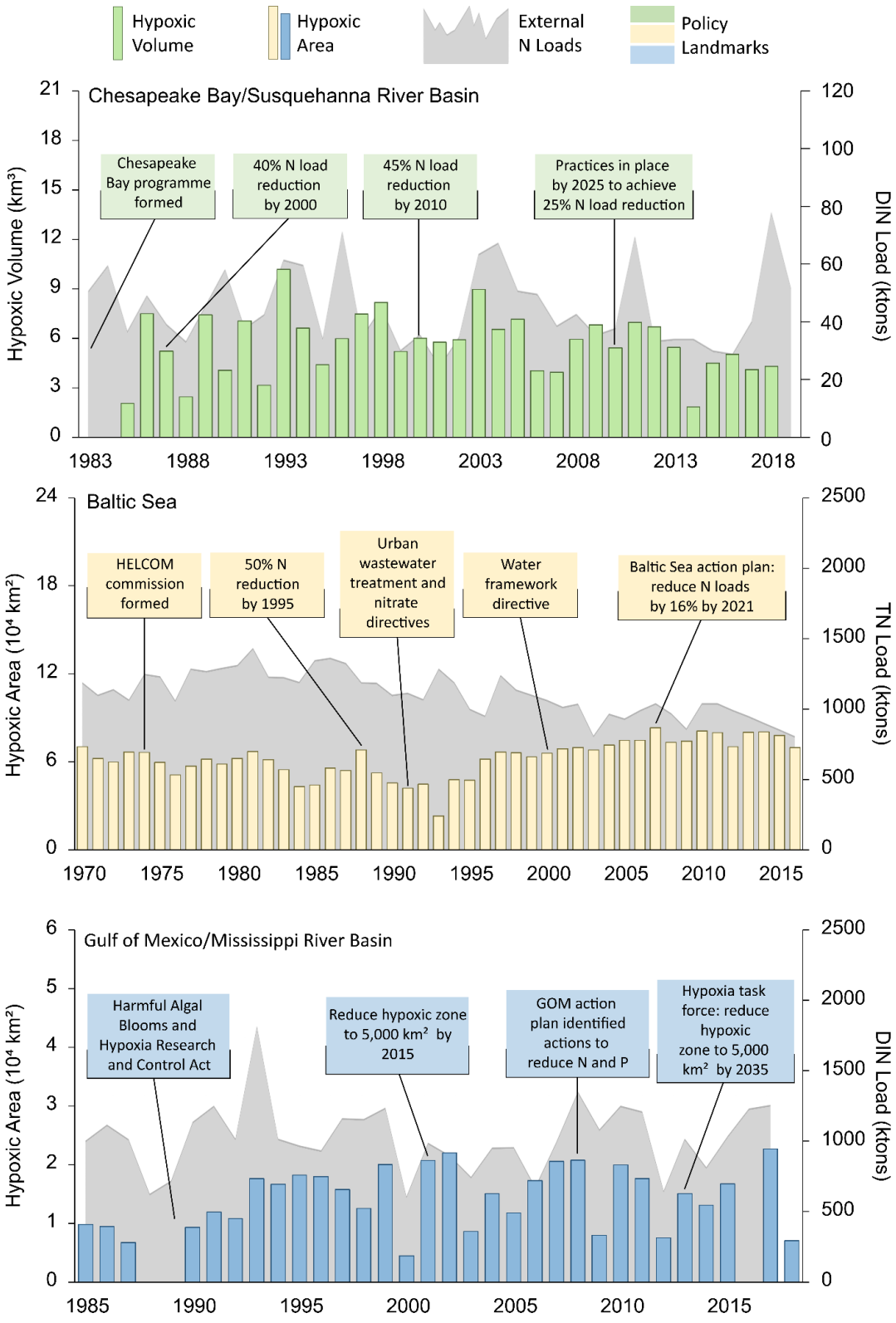

Hypoxic Volume
Hypoxic Area
External N Loads
Policy Landmarks
Chesapeake Bay/Susquehanna River Basin
Hypoxic Volume (km³)
DIN Load (ktons)
Chesapeake Bay programme formed
40% N load reduction by 2000
45% N load reduction by 2010
Practices in place by 2025 to achieve 25% N load reduction
Baltic Sea
Hypoxic Area (10⁴ km²)
TN Load (ktons)
HELCOM commission formed
50% N reduction by 1995
Urban wastewater treatment and nitrate directives
Water framework directive
Baltic Sea action plan: reduce N loads by 16% by 2021
Gulf of Mexico/Mississippi River Basin
Hypoxic Area (10⁴ km²)
DIN Load (ktons)
Harmful Algal Blooms and Hypoxia Research and Control Act
Reduce hypoxic zone to 5,000 km² by 2015
GOM action plan identified actions to reduce N and P
Hypoxia task force: reduce hypoxic zone to 5,000 km² by 2035

Figure 1.1 Timelines of policy measures and hypoxic zone size reveal a lack of response to policy interventions across Europe and North America.
Panels show time series for the hypoxic area or volume, the riverine nitrogen load, and selected policy landmarks and goals for the Chesapeake Bay, Baltic Sea, and Gulf of Mexico. External loads in (a) represent dissolved inorganic nitrogen (NO3-+NO2-) for the Susquehanna River, which accounts for more than half of the annual nitrogen load to Chesapeake Bay. Hypoxic volume data represents the early June volume in the Maryland mainstem of the Chesapeake Bay with dissolved oxygen (DO) concentrations < 2 mg/L. External loads in (b) represent an estimate of all total nitrogen inputs to the Baltic Sea, including riverine inputs, atmospheric deposition, direct point sources, and other diffuse inputs. Hypoxic area data represents the areal extent of the hypoxic area (DO < 2 mg/L) for the Baltic Sea's Bornholm basin. External loads in (c) represent dissolved inorganic nitrogen (NO3-+NO2-) for the Mississippi River, and the hypoxic area represents the mid-summer, bottom water hypoxic area (DO < 2 mg/L). Figure reproduced from Basu et al. (2022).

Although an understanding of nitrogen legacies and time lags has existed within the scientific community for decades, this mainly theoretical understanding has historically yet to be adequately considered when developing policy and targeted approaches. Stakeholders still expect short-term water quality improvement, and the lack of success in meeting goals also generates skepticism and disillusionment in policymakers and farmers on the efficacy of conservation measures adopted, jeopardizing future buy-in (Meals et al., 2010; Sharpley et al., 2013; Van Meter & Basu, 2015; Vero et al., 2017).

Only in the last half-decade have policymakers in heavily studied basins begun recognizing the role legacy plays in the efficacy of nutrient management and are now beginning to consider them in policy timelines (Ascott et al., 2021; Basu et al., 2022). For instance, in the Tulare Lake Basin in California, decades of nitrogen use in agriculture have contributed to high levels of nitrogen in groundwater, with some wells exceeding 20 mg-N L^{-1} (Harter et al., 2012). In 2008 the State Water Board prepared a report to the Legislature to delineate the cause and potential solution to California's poor groundwater quality. With the recognition of legacy nitrogen in the groundwater, remediation approaches aim to use legacy nitrogen for fertilization through irrigation water, offering a hopeful solution to the issue (California State Water Resources Control Board, 2017; Harter et al., 2012). Similarly, the Chesapeake Bay Program and the Mississippi River/Gulf of Mexico Hypoxia Task Force have worked to improve the understanding of the drivers of delays in nutrient load reductions and have acknowledged the importance of communicating lag times in water quality responses to stakeholders (Chesapeake Bay Program, 2020; Reckhow et al., 2011; USEPA, 2023).

Current knowledge regarding the magnitude and distribution of legacy nutrients across the U.S. and the timescales over which they may contribute to elevated nutrient concentrations remain inadequate, limiting the successful design of water quality solutions. To begin creating tangible and effective management decisions and water quality assessments for all nutrient-impacted watersheds across the U.S., we need a better understanding of the role of legacy nutrients in delays in water quality improvements. Specifically, we need to quantify nutrient stores in the landscape –in the soil organic pool, vadose zone, sediment, and groundwater – across the U.S. to begin understanding the role of legacy nutrients versus climate and land management on water quality across space and time.

1.3 Fate of nutrients in the landscape

1.3.1 Nitrogen fluxes and legacies

To effectively manage legacies and their impact on water quality, we need to quantify the stores and fluxes of nitrogen and phosphorus across regional and continental scales (Dupas et al., 2020; Goyette et al., 2016a; Hong et al., 2013; Van Breemen et al., 2002). The primary anthropogenic sources of nitrogen in the landscape are inorganic fertilizer, manure, human waste, biological nitrogen fixation, and atmospheric deposition. Byrnes et al. (2020) analyzed an 88-year trajectory of nitrogen inputs across the continental U.S. at the county level, finding that most counties receive inputs from all these sources in varying magnitudes. Human waste inputs predominantly influence urban counties, while atmospheric deposition is the dominant input in counties with less developed land use **(Figure 1.2)**. Crop uptake is a nitrogen sink, removing nitrogen through plant biomass harvesting. According to Swaney et al. (2018b), approximately 45% of nitrogen inputs are removed through crop uptake across the U.S., though this percentage varies depending on whether the agricultural system is dominated by livestock or fertilizers. We can calculate nitrogen surplus, which is the excess mass of nitrogen remaining after crop removal, by taking the difference between nitrogen inputs and crop uptake **(Figure 1.3)**. Across the U.S., Byrnes et al. (2020) found nitrogen surplus to vary from -8.6 to 1040 kg-N ha^{-1} y^{-1} with localized hotspots in urban counties. Additionally, large areas of high nitrogen inputs are evident in agricultural regions, where inorganic fertilizers contribute the largest portion of nitrogen mass, and hotspots are often found in areas with high-density livestock operations.

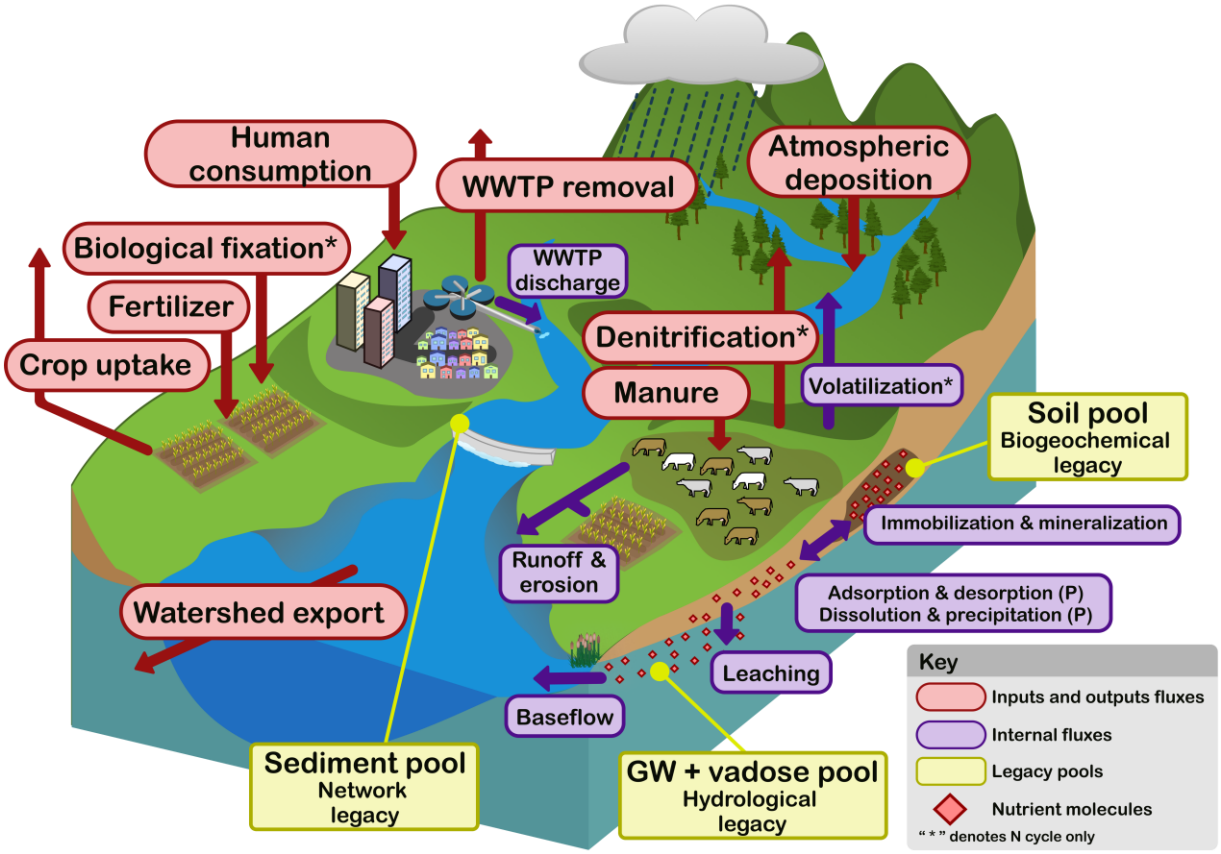


Figure 1.2 Transport and retention of anthropogenic nutrients across human-impacted landscapes.

Agricultural areas frequently act as sources of nutrients due to decades of over-application of commercial fertilizers and livestock manure. Nitrogen and phosphorus can be retained as biogeochemical legacies in agricultural soils in organic forms or, for phosphorus, sorbed to soil particles. Nitrogen and phosphorus can also be retained in a dissolved form in groundwater as hydrologic legacies, although this makes up a much smaller fraction of the phosphorus legacies. Nutrients' current and legacy sources are transported from the landscape to lakes and coastal zones via erosion and overland flow or through baseflow fluxes. Red arrows are for inputs and outputs from the watershed mass balance, purple arrows are internal fluxes that move nitrogen between pools that do not remove nitrogen from the watershed, and yellow boxes are the biological and hydrological legacy pools.

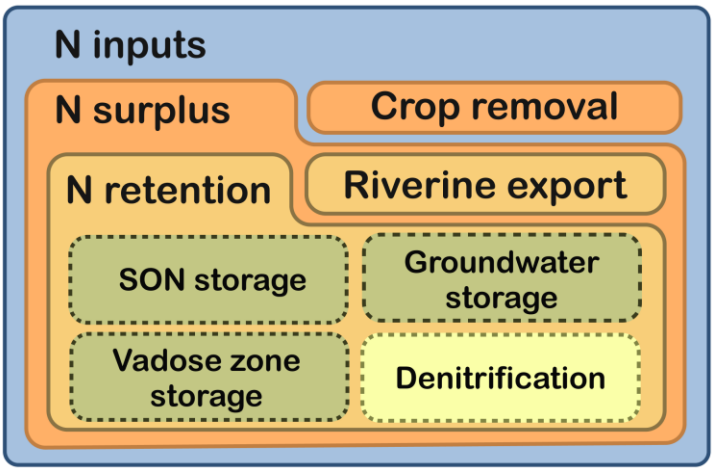


Figure 1.3 A conceptual diagram demonstrates the terminology used to describe the different how to group nitrogen pools in the subsurface. Nested boxes represent how the terms relate to one another. Nitrogen surplus is the difference between nitrogen inputs and non-hydrological export. Nested in the nitrogen surplus term is nitrogen that is retained and nitrogen exported through riverine pathways. Within nitrogen retention, nitrogen could be stored in the SON pool, vadose zone, groundwater pool, or denitrified. Legacy nitrogen includes nitrogen in the SON pool, groundwater, and vadose zone storage. Boxes are not to scale.

The surplus nitrogen is temporarily stored within the landscape until it is exported from the watershed through riverine export or denitrification **(Figure 1.3**). This stored nitrogen, or legacy nitrogen, is found as biogeochemical legacies stored as soil organic nitrogen (SON) within the vegetation root zone and as hydrological legacies stored as dissolved inorganic nitrogen in the vadose zone or groundwater pool (Van Meter et al. 2017). Thus, quantification of legacy nitrogen requires estimation of riverine export and denitrification.

Legacy nitrogen stored in the subsurface will eventually reach the surface water systems through erosion, runoff, or baseflow fluxes decades after initial application to the watershed. The time elapsed between when the nitrogen molecule enters the watersheds and when it leaves is commonly referred to as the *time lag* (D. Chen et al., 2018; Meals et al., 2010). Time lags at the watershed scale, conceptualized by the sum of (1) biogeochemical time lag and (2) hydrologic time lag (**Figure 1.3**), have been found to range from a single year to multiple decades (D. Chen et al., 2018; Dupas et al., 2018; Van Meter et al., 2018; Van Meter & Basu, 2017). Biogeochemical time lags arise from the delay in nitrogen transport from microbial cycling between inorganic and organic nitrogen forms. Indeed, isotope studies from the past two decades were among the first to find evidence of these biogeochemical time lags, finding that 12-15% of ^{15}N fertilizer remained in the root zone decades after application (Jenkinson et al., 2004; Sebilo

et al., 2013). The biogeochemical legacy stores in the crop root zones can have time lags ranging from under ten years to over 70 years (Hamilton, 2012). Van Meter et al. (2016) were among the first to estimate the regional rate of 25-70 kg-N ha y^{-1} of nitrogen accumulation in agricultural soil across the Mississippi River Basin.

Organic nitrogen can be mineralized from the SON pool and transformed into inorganic nitrogen (NH_4^+, NO_3^-, NO_2^-), a form of nitrogen that is more bioavailable for crops but also more susceptible to leaching and volatilization. Nitrate, and nitrite to a much lesser degree, is highly mobile and leaches from the root zone to the vadose zone and groundwater pool. Nitrate can then accumulate as hydrological legacies in the vadose zone dissolved in pore water. Hydrological time lags, which can range from days to decades and are generally longer than biogeochemical time lags (McDowell et al., 2021), are influenced by precipitation-driven flushing, as well as the hydraulic conductivity and connectivity of the soils (Meals et al., 2010; Tesoriero et al., 2013).

1.3.2 Phosphorus fluxes and legacies

The primary anthropogenic sources of phosphorus entering watersheds are agricultural inputs of fertilizer, manure, and atmospheric deposition, along with domestic sources such as human and food waste and detergents (Van Meter et al., 2021) (**Figure 1.2**). The most significant input into U.S. watersheds comes from agricultural sources (Sabo et al., 2021). Historically, phosphorus from human waste and detergents has contributed a large phosphorus load to rivers. However, policies aimed at reducing point source pollution, including upgrades to wastewater treatment and limits on phosphorus content in detergents, have led to a substantial reduction in domestic phosphorus loads (Litke, 1999; Sabo et al., 2021; Scavia et al., 2014). As a result, much of the point source phosphorus is now removed and stored in landfills (network legacies) (Van Meter et al., 2021). Atmospheric deposition, while historically a minor contributor, remains below 0.5 kg-P ha^{-1} y^{-1} across the U.S. (Sabo et al., 2021).

Once phosphorus enters the watersheds, it can accumulate at various points along the land–freshwater continuum. Phosphorus can be stored in the soil profile in organic and inorganic forms, cycling through these phases via microbial processes. Phosphorus can also accumulate in the soil profile as biogeochemical legacies through precipitation with other metals and adsorbing to clay particles, metal oxides, and carbonates (Schlesinger & Bernhardt, 2020). A smaller proportion of excess phosphorus, compared to nitrogen, is stored as hydrologic legacies in groundwater (Holman et al., 2008; Van Meter et al., 2021). Crops remove soluble phosphorus from the soil profile to integrate into their biomass. The use efficiency of phosphorus has

improved significantly from the mid-1980s because of improved nutrient management (Q. Zhang et al., 2023). In some cases crop uptake will exceed the annual inputs, leading to negative phosphorus surpluses and soil phosphorus mining in highly cultivated areas (Sabo et al., 2021; Swaney & Howarth, 2019b). The critical pathway for phosphorus export from the soil is through erosion from upland areas and subsurface runoff through agricultural field tile drains (Sharpley et al., 2013; Smith et al., 2015). In the U.S., it is estimated that up to 5.4 kg ha^{-1} yr^{-1} of phosphorus is lost from agricultural soils annually, a flux magnitude that exceeds current annual surplus estimates in many areas (Sabo et al., 2021). However, typically we find that between 3-30% of annual phosphorus surplus is exported to rivers (H. Han et al., 2011; Hong et al., 2012; Van Meter et al., 2021). Once phosphorus reaches the river network, it can accumulate in riparian areas and reservoirs (network legacies) (Van Meter et al., 2021).

1.3.3 Quantifying and modeling nitrogen legacy and time lags

Only in the past two decades has the accumulation of legacy nitrogen stores in the subsurface been considered in models and mass balance studies. Progress has been made in developing both data-driven statistical and process-based modeling approaches to better characterize these legacy stores. For example, the parsimonious process-based ELeMENT model was developed to quantify the legacy stores in the Mississippi and Susquehanna River Basin. From the 1960s, soil organic nitrogen in agricultural soils accumulated 6.9 and 0.35 kg-N ha^{-1} $y^{-1,}$ respectively, while groundwater accumulated nitrogen at a rate of 3.1 and 9.3 kg-N ha^{-1} y^{-1} (Van Meter et al., 2017). Ascott et al. (2017) estimated the mass of nitrogen accumulated in the vadose zone globally and found that arid regions with deep groundwater can have stores of dissolved nitrogen exceeding 4,000 kg-N ha^{-1} More recently, Van Meter et al. (2023) used a machine learning model and groundwater well concentration data and estimated that approximately 15 Tg of N (328 ± 167 kg-N ha^{-1}) of nitrogen has accumulated in the groundwater over the last 50 years.

There has also been development in modeling approaches to capture the effects of legacy nutrients on water quality. For example, McIsaac et al. (2001) developed a regression equation using NANI from two past time periods (previous 2-5 years and 6-9 years) to predict the nitrogen exported to the Gulf of Mexico from the Mississippi River Basin. Chen et al. (2014) developed a similar model using NANI but included a term to account for the previous year's storage of legacy nitrogen mass. Chen et al. (2014) and Van Meter et al. (2017) used a cross-correlation analysis to estimate the lag time between changes in nitrogen input trajectory and subsequent changes in stream nitrogen concentration trajectory in the Grand River watershed in Canada and Yongan Watershed in China. They found the lag time to range between 1 to 34 years. McDowell

et al. (2021) also used cross-correlation and a generalized additive model (GAM) to model lag times in watersheds across New Zealand, finding time lags ranging from 1 year to 9.5 years. Dupas et al. (2020) followed a travel time convolution approach to estimate lag times in 16 watersheds in western France, finding lag times to range between 2 and 14 years. Studies have linked shorter lag times with proximal flow paths, such as riparian soil and tile drainage (Ehrhardt et al., 2019; Van Meter & Basu, 2017), while longer lag times have been associated with higher evapotranspiration (McDowell et al., 2021), geologies characterized by low hydraulic conductivity and low-porosity geologic classes (Dupas et al., 2020; McDowell et al., 2021), as well as larger and steeper catchments due to a diversity of flow paths and increased flow path lengths (McDowell et al., 2021; Van Meter & Basu, 2017). Although we have an improved understanding of the importance of nitrogen and phosphorus legacies in delaying water quality improvement, we still need a comprehensive characterization of legacy stores' magnitude and spatial distribution.

1.5 Knowledge gaps and research objectives

The key to adequately characterizing the extent and impacts of legacy nitrogen and phosphorus requires understanding the linkages between landscape, groundwater, and surface water and their associated nutrient pools and transport pathways. While recent research has begun highlighting the role of legacies in water quality delays, the work has been limited in spatial and temporal scale. Large-spatial scale studies rarely have data from before the mid-1980s and rarely consider temporal trends (Hong et al., 2012, 2017; Swaney et al., 2012). Many existing studies have focused on single watersheds or drainage regions (Boyer et al., 2002; Dupas et al., 2018; Goyette et al., 2016a; H. Han et al., 2011; Metson et al., 2020; Van Breemen et al., 2002; Van Meter et al., 2018; Van Meter & Basu, 2017) or exclusively on surface water or groundwater (Hansen et al., 2019; Kim et al., 2020).

In my thesis, I aim to address these gaps by investigating: (**RO1**) the multi-decadal inputs and crop export of nitrogen and phosphorus at the sub-watershed scale across the contiguous U.S., (**RO2**) the relationship between watershed nitrogen inputs and export, (**RO3**) the drivers of changes in watershed nitrogen export, (**RO4**) the magnitude, spatial distribution, and drivers of nitrogen retention and legacy stores, and (**RO5**) the use and management of phosphorus in agricultural landscapes in the context of both food security and environmental health.

Chapter 2 Long-term nitrogen mass balance data for the contiguous United States (1930-2017)

Outline

Excess nitrogen (N) within the landscape can lead to environmental issues such as water pollution and downstream eutrophication. The quantification of landscape N fluxes can help identify areas of excess N and can improve our understanding of the sources, sinks, and transport of N within ecosystems. The TREND-N and gTREND-Nitrogen dataset provides a comprehensive, long-term (1930-2017) nitrogen (N) mass balance for the contiguous United States at a spatial resolution of county and 250 meters. The downscaled dataset integrates county-scale estimates of N fluxes with gridded land use and population data to estimate grid-scale surface fluxes of N, including fertilizer, atmospheric deposition, manure inputs, biological N fixation, crop N uptake, and population-based human waste. The downscaled gTREND-N data will allow for fine-scale insights into both historical and current N dynamics, addressing the limitations of previous datasets developed at coarser spatial scales. The data is openly available and will help to inform the development of effective policy and management strategies to mitigate the negative impacts of excess N and to promote sustainable nutrient management practices. This work has been submitted for publication in Scientific Data, a Nature journal. The first authorship of this work is shared between Danyka Byrnes and Shuyu Chang. Please see *Statement of Contributions* for more details.

2.1 Introduction

Fluxes of reactive N in the environment have increased over the past century due to population growth, shifts in diet, agricultural intensification, and increasing reliance on biofuels (Galloway & Cowling, 2021; Glibert, 2020; Van Meter et al., 2018). Anthropogenic N enters the environment through inorganic fertilizer, livestock manure, biological N fixation from leguminous crops, atmospheric deposition, and human waste. While N is frequently applied to agricultural land for crop production, it is estimated globally that only 40-50% of agricultural N is used by crops, leaving a significant portion of excess N to flow to the environment (Conant et al., 2013; Swaney et al., 2018b).

Increasing flows of N to the environment threaten groundwater and surface waters as well as the coastal ocean. Excess N has contributed to eutrophication of fresh and coastal waters, triggering the appearance of hypoxic zones and having devastating effects on aquatic ecosystems, including

fish kills and loss of biodiversity (Rabalais, Turner, & Wiseman, 2002). An estimated two-thirds of coastal systems in the U.S. alone are moderately to severely impacted by excess nutrient loading (Davidson et al., 2011). Excess N also negatively affects drinking water quality. Drinking water wells in rural areas are unregulated and are at high risk of elevated nitrate concentrations that can lead to a variety of health effects, including an increased risk of colorectal cancer (Knobeloch et al., 2000; Picetti et al., 2022; Van Meter, Schultz, et al., 2023). Surface waters have also threatened drinking water sources through the persistence of toxins from algal bloom after drinking water treatment (Davidson et al., 2011; Glibert, 2020; M. M. Steffen et al., 2017).

Fossil fuel combustion as well as the production and overuse of N fertilizer and manure have also contributed to increased fluxes of atmospheric reactive N (Chang et al., 2021; Y. Chen et al., 2020; Gao & Cabrera Serrenho, 2023). While stringent air pollution policies in the U.S. have decreased deposition of oxidized forms of N (Lloret & Valiela, 2016), the emissions from manure and fertilizer have yet to be mitigated on a large scale. These emissions contribute to acidification of terrestrial and aquatic ecosystems and the generation of greenhouse gasses (Y. Chen et al., 2020; Gao & Cabrera Serrenho, 2023). Excess N in soils and surface water from agriculture accounts for approximately 63% of all anthropogenic emissions of N_2O – among the most potent of the greenhouse gasses – and are continuing to increase (Glibert, 2020; Syakila & Kroeze, 2011).

In response to these increases in N fluxes and increased threats to water quality, policymakers have set policy goals to manage excess N. However, despite efforts to manage nutrients within watersheds, decreases in N loading to rivers, lakes, and coastal zones have been slow to occur, and many have yet to see marked improvements in water quality (Basu et al., 2022; Chang et al., 2021; Van Meter, Byrnes, et al., 2023). It is increasingly understood that N emissions from watersheds are a function of not only the current-year N inputs, but also of legacy stores of N in the soil profile and in the groundwater pool (Ascott et al., 2021; Van Meter, Basu, Veenstra, et al., 2016). Thus, managing impaired water quality requires an understanding of both past and present fluxes of N across all land uses. Accordingly, to better understand long-term N dynamics and the distribution of legacy N stores, we need long-term spatially explicit N mass balances datasets (Batool et al., 2022).

Datasets estimating the various components of the N mass balance have been compiled at multiple spatial and temporal scales across the U.S. Early regional mass balance studies were conducted by Howarth et al. (1996), who developed static nitrogen budgets for major Atlantic

Ocean drainage basins. Boyer et al. (2002) followed a similar approach to estimate the net nitrogen inputs in the Northeastern U.S. These publications, in many ways, paved the way of using livestock inventories and crops harvested from agricultural census data to estimate the nitrogen from manure production, from biological fixation, and nitrogen removed through biomass removal. Over time, these methods were further refined by updating empirical parameter values, extending temporal coverage beyond single-year snapshots to span multiple decades, and expanding the range of inputs and outputs considered.

The first large-scale estimate was produced by Hong, Swaney and Howarth (2011; 2019a) who compiled a county-scale data set of anthropogenic N inputs and outputs covering the contiguous U.S. at 5-year intervals for the period 1987 to 2017. Building on this, Sabo et al. (2019) expanded the scope of previous studies by using watershed boundaries instead of county-scale ones, making it more applicable for hydrology studies. They included a more comprehensive inventory of N inputs for the years 2002, 2007, and 2012, however, excluded estimates of nutrient uptake from pasture, a key component in nutrient cycling, especially in the western U.S. Hamlin et al. (2020) have developed estimates of landscape N inputs across the Great Lakes region at a 30-m spatial resolution for years 2008–2015. Yang et al. (2016) have developed a county-scale data set covering the period 1930–2012, but their data are limited to manure nutrient production. Goyette et al. (2016b) extended previous methods to create a transboundary nitrogen budget for the Saint Lawrence drainage area from 1901 to 2011, to our knowledge, is the longest budget timespan. Finally, Byrnes et al.(2020, 2020d) developed a more complete, long-term, county-scale nitrogen mass balance dataset for the years 1930-2017 that also increased the number of crops considered from the 16 crops in Swaney et al.(2018a) and the 21 crop types in Sabo et al.(2019) to 63 different crops, which includes fruit and nut trees, fruits, and vegetables as well as field crops.

While several datasets have been developed, most either lack the temporal depth, spatial resolution, or comprehensive inclusion of diverse crops, making it difficult to fully understand long-term nutrient dynamics. In the present work, we address the above limitations through the updating of TREND-N and the construction of a gridded Trajectories Nutrient Dataset for nitrogen (gTREND-N). gTREND-N is a long-term (1930-2017), 250-m grid scale N mass budget, with estimates of all individual components of the budget for the contiguous U.S. (**Figure 2.1**). In developing this dataset, we have updated the county-scale TREND-N dataset, refining the crop and pasture uptake estimates as well as livestock manure excretion estimates. We then combined county-scale data from the TREND-Nitrogen dataset with gridded land use and population data to provide fine-scale estimates of both surface fluxes of N (fertilizer,

atmospheric deposition, manure inputs, biological N fixation, crop N uptake) as well as population-based fluxes of human waste. We demonstrate herein the consistency of the new gridded dataset with the county-scale data that serves as its foundation and also quantify improvements in estimates of watershed-scale N fluxes that can be obtained with the downscaled, gridded data in comparison with the coarser county-scale estimates.

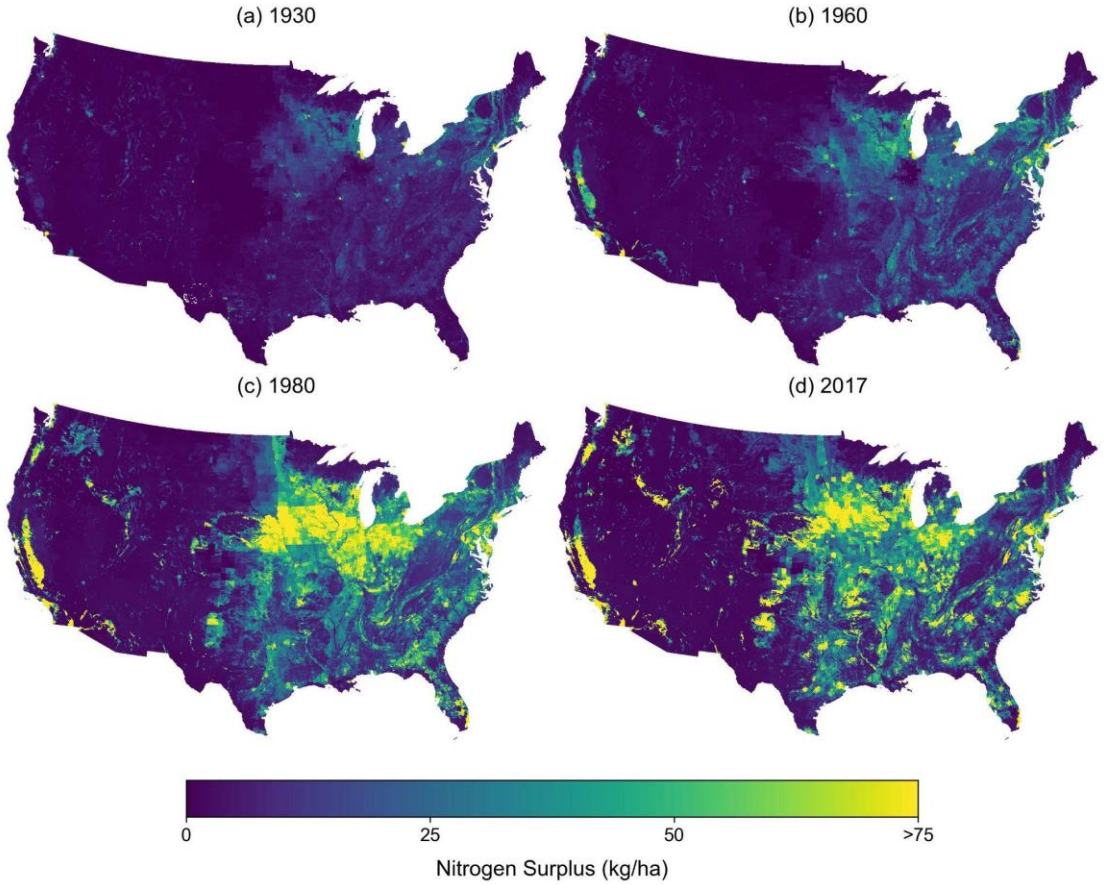


Figure 2.1 N surplus magnitudes, as calculated from gTREND-Nitrogen data for the years (a) 1930, (b) 1960, (c) 1980, and (d) 2017 at a 250-m grid scale.

2.2 Methods

The gTREND-N dataset provides 250-m grid scale estimates of N in components of the N mass balance (fertilizer, livestock manure, biological N fixation, crop N uptake, N in human waste) for the period 1930-2017. As shown in **Figure 2.2**, grid-scale estimates of these N fluxes were downscaled from a new update (version 3) of the county-scale TREND-Nitrogen dataset (Byrnes et al., 2020, 2020d), together with publicly available gridded land use/land cover data (Multi-Resolution Land Characteristics Consortium, 2016; T. L. Sohl et al., 2014), gridded population

data (Fang & Jawitz, 2018), and gridded atmospheric N deposition data (National Atmospheric Deposition Program, 2018). Details regarding the county-scale TREND-Nitrogen dataset, the steps required to facilitate the downscaling process, and additional preparation steps for gridded datasets for human population and atmospheric deposition are described in more detail below. The first authorship of this work is shared between Danyka Byrnes and Shuyu Chang. Please see *Statement of Contributions* for further details.

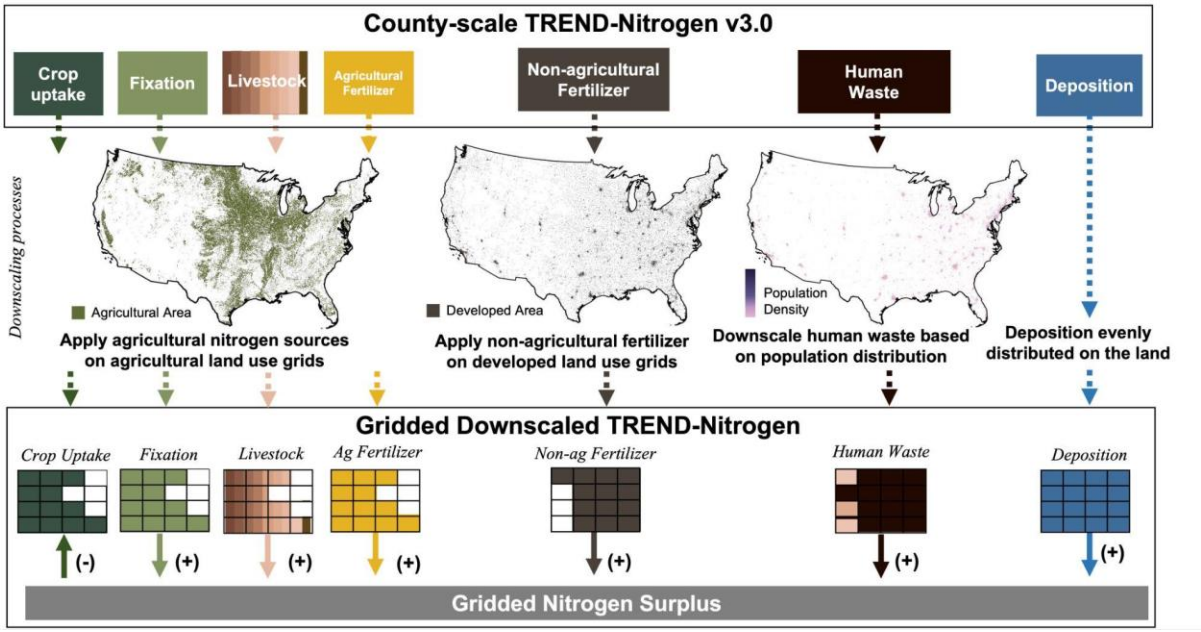


Figure 2.2 Data sources and workflow for creating the gTREND dataset.
(1) Agricultural components of the N mass balance such as crop uptake, fixation, manure, and agricultural fertilizer were downscaled to areas with agricultural land use; (2) Non-agricultural fertilizer was downscaled to areas with developed land use; (3) Atmospheric N deposition was distributed evenly across all land uses; (4) Human waste was downscaled based on gridded population data.

2.2.1 County-Scale Nitrogen Mass Balance Data

The components of the nitrogen mass balance were calculated every year from 1930–2017 at the county scale, as quantified in the newly updated TREND-Nitrogen v3 dataset, available through the Figshare data repository, doi.org/10.6084/m9.figshare.20915989.v1 (Byrnes et al., 2020b, 2022). Here, we describe our methods for estimating the different components at county scale, with specific focus on updates in methodology used to create version 3 of the TREND-N dataset. For additional details, see the original Byrnes et al. publication and the methods for the TREND-N v2 dataset available in the readme file on Figshare (Byrnes et al., 2022).

2.2.1.1 Atmospheric deposition

The county-scale dataset includes estimates of wet and dry deposition for both oxidized (NO_x) and reduced (NH_x) forms of N, $DEP_{wet,ox,}$ $DEP_{dry,ox,}$ $DEP_{wet,red,}$ and $DEP_{dry,red,}$ all in kg-N y^{-1}. Recent estimates of wet and dry oxidized N deposition (1986-2017) were obtained from modeled deposition data from the National Atmospheric Deposition Program (NADP) at a 4 x 4 km grid cell size (NADP, n.d.). Prior to 1986 (1930 -1985), current estimates were scaled backwards based on historical estimates of CO_2 emissions from fossil fuel combustion (Oak Ridge National Laboratory, 2017). Scaling was achieved by taking the ratio of N deposition to CO_2 emissions for years with available data to create a scaling factor for each individual county, i, and then multiplying the scaling factor, $S_{dep}(i,t)$, by the annual CO_2 emission data.

$$DEP(i,t) = CO2(t) * S_{dep}(i,t) \quad \text{(Eq. 2.1)}$$

CO_2 emission magnitudes are assumed to be a reasonable proxy for N deposition based on the strong correlation ($R^2 = 0.98$) between 1940-1986 N emission rates (Houlton et al., 2013) and reported CO_2 emissions.

For the period 1930-2014, county-scale deposition of reduced wet and dry N ($DEP_{wet,red,}$ and $DEP_{dry,red}$) was calculated by aggregating the modeled grid-scale estimates of reduced wet and dry N deposition (Durack & Taylor, 2019) to the county scale, and then summing the monthly values to obtain annual estimates. For the period 2015-2017, deposition magnitudes were assumed to remain constant.

2.2.1.2 Fertilizer

County-scale estimates of both farm and non-farm N fertilizer use, $FERT_f$ and $FERT_{nf}$ (kg-N y^{-1}), were obtained from several different sources. From 1945–1985, data was obtained from Alexander and Smith (1990). From 1987-2012, data was obtained from Brakebill and Gronberg (2017). We interpolated the county fertilizer to gap filling missing fertilizer in 1986. For the period 1930–1944, gridded (5 km × 5 km resolution), estimates of N fertilizer application were obtained from Cao et al. (2018) and aggregated to the county scale. In the Cao et al. dataset, U.S. fertilizer use is provided for the period 1850 to 2015. Due to differences in methodology, the magnitudes of fertilizer use in the Cao et al. were generally lower than those available in the USGS dataset in overlapping years. Thus, to fuse the pre-1945 Cao et al. data with the USGS data post-1945, USGS county fertilizer use data was scaled backwards based on trends in the Cao et al. gridded dataset.

To obtain estimates of non-farm fertilizer for the period 1930-1986, we first calculated the mean ratio of non-farm to total N fertilizer use, as reported by Brakebill & Gronberg for individual counties (Brakebill & Gronberg, 2017). Then, for the years prior to 1987, we used those county-scale ratios to partition total N fertilizer use into estimates of farm and non-farm fertilizer use.

For the period 2013–2017, national-scale estimates of N fertilizer use (Food and Agriculture Organization of the United Nations (FAO), 2022) were used to scale the county-scale estimates of both farm and non-farm N fertilizer use forward in time. Scaling was achieved by taking the ratio of mean county-scale fertilizer use (farm and non-farm) to mean national fertilizer use for the years 1987-2012 to create scaling factors for each individual county and then multiplying the scaling factors, S_f and S_{nf}, by the annual estimates of N fertilizer use (2013-2017):

$$FERT_f(i,t) = FERT_{natl}(t) * S_f(i,t) \quad \text{(Eq. 2.2)}$$

$$FERT_{nf}(i,t) = FERT_{natl}(t) * S_{nf}(i,t) \quad \text{(Eq. 2.3)}$$

where $FERT_f(i,t)$ and $FERT_{nf}(i,t)$ are the total farm and non-farm N fertilizer use, respectively, for individual counties and $FERT_{natl}$ is the FAO estimate of national N fertilizer use.

2.2.1.3 Livestock manure

Manure N inputs, MAN (kg-N y^{-1}), are defined herein as the N from livestock excretion (urine and manure). For the current dataset, county-scale manure N, by livestock type, was calculated as described in Byrnes et al.,(2020, 2022), with the following changes in methodology: (1) use of a USGS dataset(Falcone, 2021) as a source for U.S. Census of Agriculture livestock inventory numbers for the period (1950-2017); (2) improved methods for handling undisclosed livestock inventory numbers from the Census of Agriculture, resulting in fewer data gaps; (3) use of a data-fusion approach to supplement the 5-year U.S. Census of Agriculture data with annual Agricultural Survey data; and (4) use of time-varying livestock weights in calculating manure N values. Details regarding these changes in methodology are provided below.

(1) All livestock inventory data originates from the U.S. Census of Agriculture. Livestock inventory numbers for cattle, hogs, poultry, sheep, and horses for the period 1950-2017 (reported at 5-year intervals) were obtained from a USGS dataset by Falcone (2021) that includes not only Census of Agriculture numbers but also estimates of livestock inventory numbers for counties with censored data in some census years. For the period 1930-1945, and for livestock categories not explicitly included in the USGS dataset (pullets, mules, donkeys, goats), data was obtained from the ICPSR data archive (1930–

2012) (Haines et al., 2018) and NASS Quick Stats (USDA National Agricultural Statistics Service, 2019).

(2) For years and for counties for which livestock inventory data from the Census of Agriculture is "undisclosed" due to privacy concerns, data gaps are filled by using the data from the most recent year for which there is existing data for that county and livestock type (Falcone, 2021). If there is undisclosed data for more than six census time periods (~30 years), the inventory is assumed to be zero.

(3) A data-fusion approach was used to integrate the 5-year U.S. Census of Agriculture data with annual Agricultural Survey data. Using this approach, the Census of Agriculture livestock inventory data was used as reported and the annual Agricultural Survey data was scaled to the Census of Agriculture data. Scaling was achieved by taking a ±10-year window around each year in which Agricultural Survey data was to be used. For these 20 years, we created a scaling factor, $S_{lvstk}(i,j,t)$, for county i, livestock type j, in year t, calculated by taking the ratios of census data to survey data for years where both were available, and then taking the mean of those ratio values. Then, for each year with available survey data, but no census data, crop production or crop area values were calculated by multiplying the scaling factor by the reported survey data:

$$INV(i,j,t) = INV_{surv}(i,j,t) * S_{lvstk}(i,j,t) \quad \text{(Eq. 2.4)}$$

where $INV(i,j,t)$ is the estimate of the number of heads of a specific livestock type j in county i in year t, and $INV_{surv}(i,j,t)$ is the number of heads of that livestock reported in the Agricultural Survey.

The magnitude of N excretion per animal varies with animal weight, and livestock weights have increased over time due to both genetic evolution and improvements in feeding (Falcone, 2021). Due to the importance of these changes, we have adopted, for the current dataset, a new methodology for scaling livestock weights over time based on USGS methods (Falcone, 2021). In the current methodology, we use time-varying livestock weights based on use of annual slaughtered livestock weight data (U.S. Department of Agriculture, n.d.) Reported weights were available for the time period 1958-2017 (U.S. Department of Agriculture, n.d.). Any data gaps across this period were filled via linear interpolation. For the period 1930-1957, we used the reported 1958 weights, assuming no change in livestock weights for that period. No time-varying weights were available for horses, so we assumed weight to remain constant across the

study period. In this new approach, 2008 was chosen as a base year due to the availability of N excretion parameters for 2008 livestock weights (USDA NRCS, 2008). Temporally varying, livestock type-specific scaling factors, $S_{wt}(i,t)$, were then developed based on the ratio of the 2008 weight to the reported weight for any given year (**Supplemental Table 2.1**).

With the resulting county-scale livestock inventory trajectories, N in livestock manure was calculated as a function of these inventory numbers and the estimated N content of the manure produced per animal per year (Byrnes et al., 2020, 2022):

$$MAN(i,j,t) = INV(i,j,t) * p(j)_{lvstk,excr} * S_{wt}(i,j,t) \quad \text{(Eq. 2.5)}$$

where MAN(i,j,t) is the county-scale, time-varying estimated N in manure from livestock type j, $p(j)_{lvstk,excr}$ is the estimated N content in manure from that livestock type, and $S_{wt}(i,j,t)$ is the time-varying scaling factor for livestock weights for county i and livestock type j. N excretion magnitudes for all of the livestock types were then aggregated to provide one annual, county-scale value:

$$MAN(i,t) = \sum_{j=1}^{n} MAN(i,j,t) \quad \text{(Eq. 2.6)}$$

where *n* is the number of different livestock types. The parameter values used for 2008 livestock N consumption and excretion are provided in **Table 2.1**. Excretion values were obtained from the USDA Agricultural Waste Management Field Handbook (USDA NRCS, 2008). Consumption values for cattle were obtained based on a calculated ratio of consumption to excretion (Byrnes et al., 2022; Hong et al., 2011)**.**

Table 2.1 Livestock nitrogen consumption and excretion parameters. Consumption values are included for all grazing livestock.

Livestock Type	Livestock N Consumption, 2008 (kg-N head^{-1} yr^{-1})	Livestock N Excretion, 2008 (kg-N head^{-1} yr^{-1})
Cattle, Beef Cows	79.33	69.54
Cattle, Milk Cows	195.62	151.73
Cattle, Other (Heifers, bulls, steers, calves)	70.32	61.64
Goats	8.90	7.45
Sheep and Lambs	8.90	7.45
Horses	50.06	44.70
Hogs and Pigs	-	13.80
Poultry, Turkey	-	1.24
Poultry, Broilers	-	0.41
Poultry, Other Chickens	-	0.58

2.2.1.4 Crop uptake

County-scale crop N uptake in county i and year t was estimated as the sum of crop-specific uptake values CROP(i,j,t) (kg-N y^{-1}) for all crop types j in that county (field crops, fruit and nut trees, fruits, and vegetables). The crop-specific, county-scale N uptake values CROP(i,j,t) were calculated on the basis of either mass or harvested area by multiplying the harvested crop magnitude, HARV(i,j,t) in census-reported units, e.g. bushels, hundredweight, acres, by a crop-specific conversion factor, c(j), to convert to kg crop, by the crop-specific percent dry matter, d(j), where applicable, and then by the the crop-specific N uptake values, $p_{crop}(j,t)$ obtained from literature (Byrnes et al., 2020, 2022):

$$CROP(i,j,t) = HARV(i,j,t) * c(j) * d(j) * p_{crop}(j,t) \quad \text{(Eq. 2.7)}$$

Total county-scale crop N uptake, CROP(i,t) was then calculated by summing the uptake values for all of the 63 included crops.

$$CROP(i,t) = \sum_{j=1}^{63} CROP(i,j,t) \quad \text{(Eq. 2.8)}$$

The 63 different crops included here were confirmed to cover ≥97% of the county-scale cropped area See **Supplemental Table 2.2-2.4** for conversion factors, dry matter values, and N uptake parameters for field crops, fruits and vegetables, and orchard crops, respectively.

Annual crop production data, $HARV(i,j,t)$, was obtained using a data-fusion method that integrates data from the Census of Agriculture, which reports data every 5 years, with the U.S. Agricultural Survey, which provides projected annual estimates. Census data was obtained from the ICPSR data archive (1930–2012) (Haines et al., 2018) and NASS Quick Stats (USDA, 2023) and Agricultural Survey data was obtained from NASS Quick Stats (USDA, 2023). Using this approach, and to ensure the compatibility of the datasets, the Census of Agriculture crop data was used as reported, and the annual Agricultural Survey data, $HARV_{surv}(i,j,t)$ was scaled to the Census of Agriculture data for the years when census data was not available (Eq. 9 and 10). Scaling was achieved by taking a ±10-year window around each year in which the Agricultural Survey was to be used. For these 20 years, we created a county-specific scaling factor, S_{crop} (i,j,t), calculated by taking the ratios of census data to survey data for years where both were available, and then taking the mean of those ratio values across all census years within the specified time window. Then, for each year with available survey data, but no census data, crop production or crop area values were calculated by multiplying the scaling factor by the reported survey data:

$$HARV(i,j,t) = HARV_{surv}(i,j,t) * S_{crop}(i,j,t) \quad \text{(Eq. 2.9)}$$

where $HARV_{surv}$(i,j,t) represents the estimates of crop production or harvested crop area in county i for crop type j in year t from the Agricultural Survey.

Censored county-scale data from the Census of Agriculture was gap-filled by calculating the difference between reported state-level production values and the sum of all reported county-level values. This difference was then apportioned equally to all of the counties with censored data as a function of the ratio between county-scale and state-scale values from the nearest census years with reported data.

One notable change between the county-scale crop N uptake data used here and that included in TREND-Nitrogen v2 is the parameter values used for three major U.S. crops: corn, soybeans, and wheat. For the period 1930-1989, we assume constant N uptake values (kg-N kg-crop^{-1}) based on parameter values from the International Plant Nutrition Institute (IPNI) and other

literature sources (**Supplemental Table 2.2**). Then, for the period 2013-2017, we use newly estimated parameter values that reflect current uptake magnitudes (Ciampitti & Vyn, 2012; G. McIsaac, personal communication, 2023; Scott et al., 2006; Villamil et al., 2019). For the period 1990-2013, we then linearly interpolate the N uptake parameters between the two time periods. Both the pre-1990 and current crop N uptake parameters for corn, soy, and wheat are provided in **Supplemental Table 2.2**.

2.2.1.5 Pasture uptake

County-scale pasture N uptake, PAST(i,t) (kg-N y^{-1}), was calculated as the N content of pasture grasses consumed by grazing cattle in county i per year t, which requires both county-scale livestock data and area estimates of both cropland and non-cropland pasture. Estimates of pasture area for the period 1945-2017 were obtained from the Census of Agriculture.(Haines et al., 2018; *USDA - National Agricultural Statistics Service - State Agricultural Overview*, n.d.) For the period 1930-1944, gridded estimates of pasture area from (Ramankutty & Foley, 1999) were aggregated to the county scale. To ensure continuity between the two periods and data sources, we used the coarser 1930-1944 estimates of pasture area to scale the census-based estimates backward in time. The backward scaling was achieved as follows:

$$AREA_{past}(i, t-1) = AREA_{past}(i,t) + (\Delta P(i,t) * AREA_{past}(i,t)) \quad \text{(Eq. 2.10)}$$

where $AREA_{past}$ is the sum of cropland and non-cropland pasture areas of individual counties, i, t is the annual timestep, and ΔP is the fractional annual change in pasture area estimated from the gridded dataset, calculated as $PAST_g(t) - PAST_g(t-1)$. The maximum possible N uptake from pasture, $PAST_{max,i}$ was calculated as follows:

$$PAST_{max}(i,t) = AREA_{past}(i,t) * p_{past} \quad \text{(Eq. 2.11)}$$

where p_{past} is the annual-scale area–based N content of pasture grasses (kg ha^{-1}).

Livestock inventory numbers were obtained from the Census of Agriculture (Falcone, 2021; Haines et al., 2018; *USDA - National Agricultural Statistics Service - State Agricultural Overview*, n.d.). For cattle, to account for some animals being raised in confinement while others are allowed to graze, we developed estimates of the number of grazing cattle using state-scale estimates of the fraction of different types of non-confined cattle, e.g., beef cows, dairy cows, other cattle (**Table 2.2**) (Kellogg et al., 2000a). For horses, goats, and sheep, we assumed that all animals were allowed to graze, and for hogs and poultry, we assumed no grazing.

Table 2.2 State-level confinement factors for beef cows, milk cows, and other cattle. The confinement factor refers to the fraction of each indicated livestock category that does not have access to pasture.

State	Beef Cattle	Milk Cattle	Other Cattle	State	Beef Cattle	Milk Cattle	Other Cattle
Minnesota	0.15	0.90	0.58	Washington	0.05	0.80	0.44
Arkansas	0.10	0.50	0.24	New Hampshire	0.10	0.80	0.34
Connecticut	0.10	0.90	0.42	New York	0.10	0.80	0.39
Delaware	0.10	0.80	0.39	North Carolina	0.00	0.59	0.00
Illinois	0.10	0.80	0.38	Oregon	0.05	0.60	0.28
Indiana	0.10	0.60	0.33	Pennsylvania	0.05	0.80	0.39
Iowa	0.10	0.87	0.46	Rhode Island	0.10	0.80	0.29
Maine	0.10	0.80	0.37	Texas	0.05	0.75	0.34
Maryland	0.10	0.80	0.37	Utah	0.05	0.80	0.31
Massachusetts	0.10	0.80	0.35	Vermont	0.20	0.90	0.48
Mississippi	0.10	0.60	0.23	West Virginia	0.00	0.80	0.25
Missouri	0.10	0.65	0.30	Wisconsin	0.08	0.90	0.45
Kentucky	0.08	0.70	0.30	South Dakota	0.10	0.80	0.34
Michigan	0.08	0.90	0.50	Virginia	0.10	0.60	0.34
Arizona	0.05	0.80	0.52	Nevada	0.05	0.80	0.26
California	0.05	0.80	0.43	New Jersey	0.10	0.80	0.35
Colorado	0.05	0.80	0.59	New Mexico	0.00	0.85	0.31
Kansas	0.05	0.85	0.48	South Carolina	0.00	0.59	0.17
Montana	0.01	0.75	0.22	Wyoming	0.05	0.80	0.28
Alabama	0.00	0.40	0.02	Nebraska	0.08	0.80	0.63
Florida	0.00	0.50	0.05	North Dakota	0.00	0.80	0.22
Georgia	0.00	0.70	0.10	Ohio	0.10	0.90	0.45
Idaho	0.00	0.95	0.42	Oklahoma	0.00	0.65	0.32
Louisiana	0.00	0.50	0.03	Tennessee	0.10	0.60	0.23

Potential N uptake from consumption of pasture grasses by livestock, $CONS_{max}(i,j,t)$, was first calculated by multiplying the county inventory numbers for the different livestock types, $INV(i,j,t)$, by both the fraction of unconfined animals, $(1-f_{conf}(j,t))$, and by literature estimate of N consumption for the different livestock types (**Table 2.1**) (Byrnes et al., 2022; Swaney et al., 2018a; USDA NRCS, 2008).

$$CONS_{max}(i,j,t) = INV(i,j,t) * (1 - f_{conf}(i,j,t)) * p_{cons}(j) * S_{wt}(i,j,t) \quad \text{(Eq. 2.12)}$$

where $CONS_{max}(i,j,t)$ is the potential N consumption for livestock type j in county i, year t, $p_{cons}(j)$ is the N intake parameter for that cattle type, and $S_{wt}(i,j,t)$ is the time-varying scaling factor for livestock weights. The county-scale estimates of total potential pasture N consumption, $CONS_{max}(i,t)$, were then calculated as the sum of potential intake values for all cattle types:

$$CONS(i,t) = \sum_{j=1}^{n} CONS(i,j,t) \quad \text{(Eq. 2.13)}$$

If potential consumption values exceeded the calculated maximum N uptake for pasture area, pasture N uptake values, $PAST(i,t)$ were capped at the area-based P uptake value:

$$PAST(i,t) = \begin{cases} CONS_{max}(i,t), & CONS_{max}(i,t) < PAST_{\max}(i,t) \\ PAST_{max}(i,t), & CONS_{max}(i,t) \geq PAST_{\max}(i,t) \end{cases} \quad \text{(Eq. 2.14)}$$

For additional information on calculation of pasture N uptake, see Byrnes et al. (2022).

2.2.1.6 Biological nitrogen fixation

County-scale biological N fixation, $BNF(i,t)$ (kg-N y^{-1}), was calculated for N-fixing crops using either mass-based (Eq. 14) or area-based (Eq. 15) methods (Bouwman et al., 2005; Byrnes et al., 2020; H. Han & Allan, 2008; Hong et al., 2013). Accordingly, calculations were based on either crop production data (mass in kg) or county-scale estimates of crop area (ha) obtained from the U.S. Census of Agriculture (Haines et al., 2018; *USDA - National Agricultural Statistics Service - State Agricultural Overview*, n.d.):

$$BNF(i,j,t) = PROD(i,j,t) * p_{crop,m}(j,t) * f_{bnf}(j) \quad \text{(Eq. 2.15)}$$

$$BNF(i,j,t) = AREA_{crop}(i,j,t) * p_{bnf,a}(j) \quad \text{(Eq. 2.16)}$$

where $BNF(i,j,t)$ is the estimate of county-scale biological N fixation for in county i, *for* crop j, in year t, $PROD(i,j,t)$ is the mass (kg) of harvested crop j, calculated as shown in Eq. 12, $p_{crop,m}$ (j,t) is the mass-based (kg-N kg-crop^{-1}) N content of crop j, f_{bnf} (j) is the fraction of crop N uptake

considered to be derived from N fixation in crop j, and $p_{bnf,a}(j)$ is the area-based fixation rate (kg-N ha^{-1}) (Byrnes et al., 2022; H. Han & Allan, 2008; Hong et al., 2013).

The mass-based approach was used for soybeans, peanuts, alfalfa and non-alfalfa hay, lentils, and beans. The area-based approach was used for snap beans and cropland pasture. Total county-scale biological N fixation, BNF(i,t) was then calculated by summing the BNF values for all of the N-fixing crops:

$$BNF(i,t) = \sum_{j=1}^{n} BNF(i,j,t) \quad \text{(Eq. 2.17)}$$

where *n* is the number of N-fixing crops. For a listing of all BNF parameter values, see **Table 2.3**.

Table 2.3 Biological nitrogen fixation (BNF) parameters. Parameter values for nitrogen fixing crops, including harvested crop units, crop-unit conversion values, dry matter percentages, fraction N in crop dry matter, and the fraction of harvested N derived from nitrogen fixation.

Crop Type	Reporting Unit	Conversion Value (kg-crop/ reporting unit)	Fraction Dry Matter (kg-$crop_{dry}$/kg-crop)	Fraction N in Dry Matter (kg-N /kg-$crop_{dry}$)	Fraction of Harvested N from Fixation (kg-N_{fix} /kg-N_{harv})	Citation
Alfalfa hay	dry tons	907.18	0.904	0.026	1.23	(Hong et al., 2013; *IPNI Estimates of Nutrient Uptake and Removal*, n.d.)
Beans	bushels	27.22	0.900	0.046	0.395	(Hossain et al., 2016; Lander et al., 1998a)
Beans	hundredweight	45.36	0.900	0.046	0.395	(Hossain et al., 2016; Lander et al., 1998a)
Cropland pasture	acres	907.18	1.000	0.025	0.331	(Hong et al., 2013; Meisinger & Randall, 1991)
Hay, (excl alfalfa)	dry tons	907.18	0.867	0.020	0.308	(Hong et al., 2013; *IPNI Estimates of Nutrient Uptake and Removal*, n.d.; Meisinger & Randall, 1991)
Lentils	pounds	0.45	0.887	0.044	0.803	(Hossain et al., 2016; Lander et al., 1998a)
Lentils	hundredweight	45.36	0.887	0.044	0.803	(Hossain et al., 2016; Lander et al., 1998a)
Peanuts	bushels	9.98	0.940	0.043	1.110	(Hong et al., 2013; *Standard Weights per Bushel for Agricultural Commodities*, n.d.)
Peanuts	pounds	0.45	0.940	0.043	1.110	(Hong et al., 2013)
Snap beans	acres	3872.6	0.130	0.030	0.675	(Hong et al., 2013; Meisinger & Randall, 1991; *USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey)*, n.d.; Westermann et al., 1981)
Soybeans[A]	bushels	27.22	0.906	0.065	1.110	(Hong et al., 2013)

Soybeans[B]	bushels	27.22	0.870	0.062	1.110	(Hong et al., 2013; Villamil et al., 2019)

[A] Value before 1980 [B] Value after 2014

2.2.1.7 Human waste

In the county-scale TREND-Nitrogen dataset, human N inputs were estimated using county-scale estimates of human population and literature estimates of annual human N excretion (5 kg-N $capita^{-1}$ y^{-1}) (Byrnes et al., 2020; Hong et al., 2011). Population data at the county scale was obtained at decadal time increments for the period 1930-1990 and at annual time increments for the period 1991-2017 (U.S. Census Bureau, n.d., 2007, 2016). For the period 1930-1990, the decadal data was interpolated to obtain annual estimates of population.

Note here that a large fraction of human waste is treated in wastewater treatment plants or septic systems before being discharged to downstream waters. In addition, the efficiencies of these treatment systems vary spatially and have changed over time. Quantifying these treatment trajectories is beyond the scope of the current work.

2.2.2 Creation of the gTREND-Nitrogen gridded dataset

County-scale N fluxes were downscaled to the 250-m grid scale according to land use/land cover and population densities, as described below. For the downscaling, N fluxes were divided into three different land-use categories: (1) agricultural, (2) developed/urban, and (3) all land (**Figure 2.3**). Agricultural N fluxes were considered to include four major landscape flux categories: (1) N fertilizer, (2) N in livestock manure N, (3) biological N fixation in cropland and pastureland, and (4) N uptake by crops and pasture. Developed/urban fluxes were considered to include (1) N in human waste and (2) non-farm N fertilizer. Finally, atmospheric N deposition was considered to be an all-land flux. While natural land use including wetlands and water areas received atmospheric deposition, they were not assigned any agricultural or urban N fluxes.

Land-use trajectories were established based on two different land use and land cover (LULC) datasets: (1) a modeled historical LULC dataset created at a resolution of 250 m (hereafter referred to as the Sohl LULC dataset) (T. Sohl et al., 2016b; T. L. Sohl et al., 2014), and (2) the widely used National Land Cover Database (NLCD), available at a resolution of 30 m (T. L. Sohl et al., 2018). These two datasets were integrated and harmonized to reconstruct temporally continuous and spatially explicit land-use trajectories from 1930-2017 at the 250-m grid scale for the contiguous United States (CONUS). Specifically, we use here the annual 250-m Sohl LULC data from 1938 to 2005, and then use 30-m NLCD data for the years 2006, 2008, 2011, 2013,

2016. The NLCD data was resampled from 30 m to 250 m using the QGIS-python Nearest Neighbor Function, version 3.28 (*QGIS*, n.d.). Consistent land use was assumed for the years 1930-1937 and in the years between the years of NLCD availability after 2005.

Gridded population-density data was obtained from the 1-km gridded, decadal population maps developed for the contiguous U.S. by Fang & Jawitz (2018, 2019). These maps were created using a parsimonious modeling approach, in which decennial U.S. census data was brought together with a variety of other factors, including socioeconomic desirability, suitability, and inhabitability. While five different population models of varying complexity were created in the Fang & Jawitz study,(2018) we use here the results of the M3 model, which was of intermediate complexity but found to be the most accurate based on validation using county subdivision population data. To provide annual estimates of population for the gTREND-Nitrogen downscaling work, we used linear interpolation between the modeled years, and then assumed a constant population for the years 2011-2017. The annual 1-km maps of human population were resampled to a 250-m resolution using the QGIS-python Nearest Neighbor Function, version 3.28 to align with other gridded components of the gTREND-Nitrogen dataset.

2.2.2.1 Agricultural nitrogen fluxes

To prepare for downscaling of the agricultural N fluxes, annual, gridded binary land-use maps were created to classify 250-m grid cells as either agricultural or non-agricultural, g$LU_{ag}(k,t)$ = 0/1, where *k* refers to the individual grid cell and *t* to year. Agricultural land use is defined here as areas used to produce crops, areas being actively tilled, and areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay. Specifically, note that agricultural land was considered to include both cultivated cropland and hay/pasture in both the Sohl LULC dataset and the NLCD. Land not falling into either the developed or agricultural land classes was considered to be non-agricultural land, $gLU_{ag}(k,t) = 0$.

Total agricultural area in each county *i* in each year *t*, $AREA_{ag}(i,t)$ (ha), was then calculated as the sum of the gridded agricultural land. For each agricultural component of the N mass balance, *j*, county-scale N flux values normalized by agricultural area, $N_{ag,cnty}(i,j,t)$ (kg ha-agland^{-1} y^{-1}) were calculated by dividing the total county N flux by the time-varying agricultural area of each county, i. County-scale vector maps of $N_{ag,cnty}$ values were then rasterized to a 250-m grid scale. To appropriately assign agricultural N fluxes only to grid cells with agricultural land use, the county-scale flux values were multiplied by the 0 or 1 values in the binary land-use grid.

2.2.2.2 Nitrogen fluxes on developed land

As noted above, N fluxes on developed land were considered to include non-farm N fertilizer and human waste. For the non-farm fertilizer, the county-scale estimates of non-farm fertilizer, $FERT_{nf}$ ($kg\text{-}N\ y^{-1}$), were downscaled based on developed land use, assuming that homes and businesses using fertilizer would not occur on either agricultural or "other" land. Using a method similar to that developed for the agricultural N fluxes, we created binary maps to classify 250-m grid cells as either developed or non-developed, $gLU_{dev}(k,t)$ = 0/1, where *k* refers to the individual grid cell and *t* to year. Land not falling into either the developed or agricultural land classes was considered to be non-developed, $gLU_{dev}(k,t)$ = 0.

Total developed area in each county in each year, $AREA_{dev}(i,t)$ (ha), was then calculated as the sum of the gridded developed land in that county in year *t*. County-scale vector maps of $FERT_{nf,dev}$ values were then rasterized to a 250-m grid scale. Next, to assign the non-farm N fertilizer fluxes only to grid cells with developed land use, the county-scale flux values were multiplied by the 0 or 1 values in the binary land-use grid.

For human waste, annual, grid-scale estimates of N from human waste, $gHUM(k,t)$ ($kg\text{-}N\ ha^{-1}\ y^{-1}$) were developed using gridded estimates of human population(Fang & Jawitz, 2018) and literature-based estimates of per capita annual human N excretion (5 kg-N $pp^{-1}\ y^{-1}$).(Hong et al., 2011)

2.2.2.3 Atmospheric deposition

The county-scale atmospheric N deposition magnitudes, $DEP_{ox}(i,t)$ and $DEP_{red}(i,t)$, were normalized by county area, $AREA(i,t)$:

$$DEP_{ox,norm}(i,t) = \frac{DEP_{ox}(i,t)}{AREA_{cnty}(i,t)} \quad \text{(Eq. 2.18)}$$

$$DEP_{red,norm}(i,t) = \frac{DEP_{red}(i,t)}{AREA_{cnty}(i,t)} \quad \text{(Eq. 2.19)}$$

For the all-land fluxes, all grid cells of our 250-m gridded map of the U.S. were assigned 1 values, $gLU_{all}(k,t)$ = 1, where *k* refers to the individual grid cell and *t* to year. The normalized deposition flux values were then multiplied by the 1 values in the land-use grid.

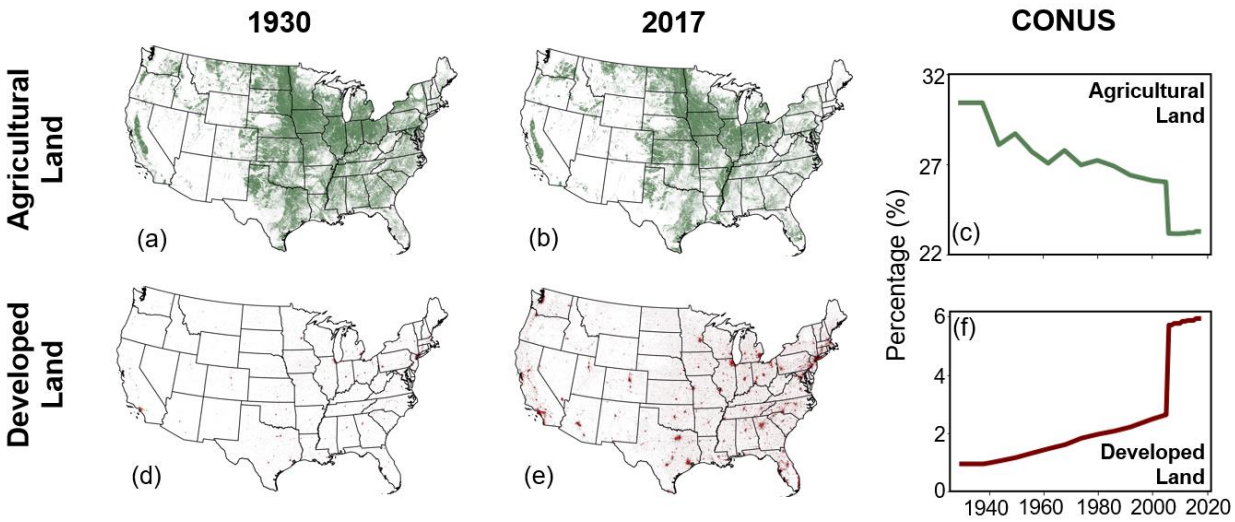


Figure 2.3 Agricultural and developed land use within the contiguous U.S. Maps (a,b,d,e) show the agricultural and developed land use in 1930 and 2017. Land-use categories are based on a simplification of categories in the NLCD and Sohl LULC datasets. Land-use trajectories for each land category are shown in (c) and (f).

2.2.2.4 Gridded nitrogen surplus

The annual grid-scale estimates of the N surplus, gNS, were calculated by raster summation of the gridded components of the N mass balance. Here, we define the N surplus as a sum of N inputs (fertilizer N, biological N fixation, atmospheric N deposition, manure N, and human N) and non-hydrologic N outputs (crop N uptake) (Byrnes et al., 2020):

$$gNS(k,t) = gfERT(k,t) + gBNF(k,t) + gDEP_{ox}(k,t) + gMAN(k,t) + gHUM(k,t) - gCROP(k,t) - gPAST(k,t) \quad \text{(Eq. 2.20)}$$

where gFERT is the sum of farm and non-farm N fertilizer, gBNF is biological N fixation, gDEP is oxidized N deposition (both wet and dry forms), gMAN is N in livestock manure, gHUM is the N in human waste, gCROP is N uptake in crops, and gPAST is N uptake by pasture grasses consumed by livestock, all in units of kg ha^{-1} y^{-1}.

2.3 Technical validation

2.3.1 Comparison of the county-scale TREND-Nitrogen v2 and v3 datasets

As described in the Methods, we first created a new version (v3.0) of the county-scale TREND-Nitrogen dataset for the contiguous U.S. (**Figure 2.4**). One major difference between the v2.0 and v3.0 datasets is the new uptake parameters for corn, soybeans, and wheat, three of the most common field crops grown in the contiguous U.S. It is increasingly acknowledged that protein

and N concentrations have decreased in corn and other grains with the development of new higher-yielding hybrids (Ciampitti & Vyn, 2012; G. McIsaac, personal communication, 2023; Scott et al., 2006; Villamil et al., 2019). To reflect current knowledge regarding changes in grain N content over time, we now use time-varying parameter values for these three crops. Based on the adopted methodology, uptake magnitudes remain the same for the period 1930-1989 but then decrease over the period 1990-2017. In 2017, these updated parameter values led to a 16% decrease in estimates of corn N uptake, a 4% decrease for soybeans, and a 19% decrease for wheat compared with the previously used parameter values.

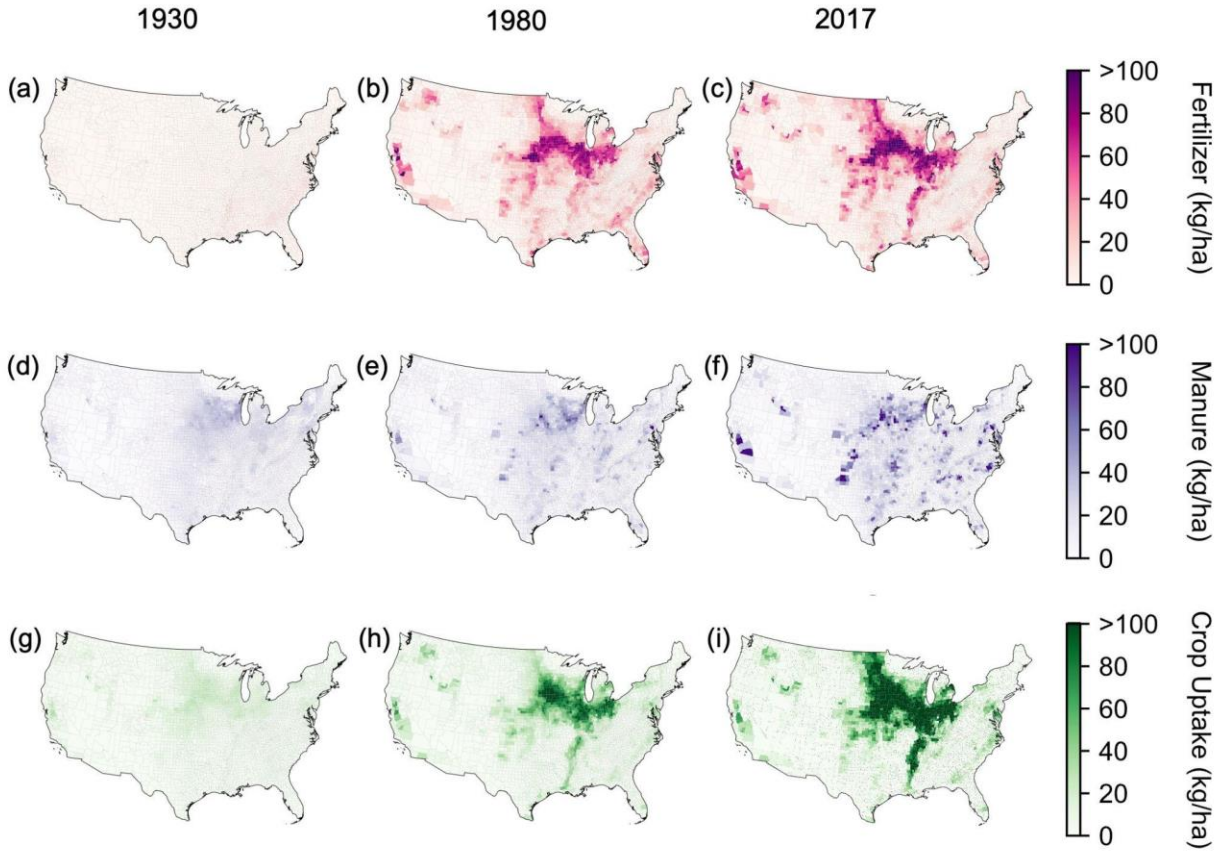


Figure 2.4 The TREND-Nitrogen county-scale nitrogen mass balance. Shown here are county-scale N fertilizer inputs (a,b,c), manure N inputs (d,e,f), and crop N uptake (g,h,i), which are three major components of the overall N mass balance. Values are shown as kg ha^{-1} county land.

Due to these decreases in crop uptake N uptake for the three largest field crops, we also see some differences in the overall N surplus, especially in the midwestern states. As shown in **Figure 2.5a**, the 2017 median N surplus magnitude for the U.S. midwest is approximately 4 kg/ha greater than that predicted in the previous TREND-N v2 dataset. In contrast, in other areas of the country, which are not as dominated by the top three field crops, N surplus magnitudes are similar to those previously predicted by the earlier dataset. In **Figure 2.5e**, the blue areas of the map demonstrate that in some areas of the country, N surplus magnitudes are actually somewhat lower than predicted by the TREND-N v2 dataset. These small decreases, typically less than 1 kg ha^{-1} at the county scale, are largely due to version 3 differences in methodology for predicting N in livestock manure as well as pasture N uptake.

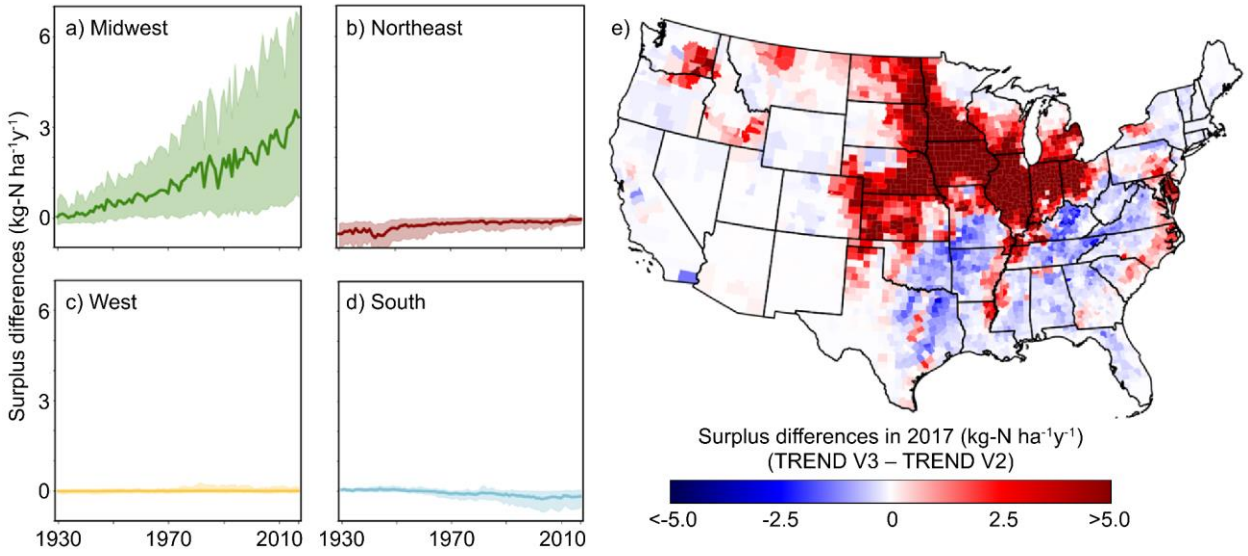


Figure 2.5 Differences between the county-scale TREND-Nitrogen v2 and the new TREND-Nitrogen v3. Due to changes in estimates of crop N uptake in corn, soybeans, and wheat, we see decreases in crop N uptake and thus increases in the overall N surplus compared with the older TREND-Nitrogen v2 dataset (a,b,c,d). These changes are especially noticeable in the U.S. midwestern states (a,e).

2.3.2 Validation of gTREND-Nitrogen with county-scale data

An important validation step for the gridded gTREND-Nitrogen dataset was to ensure conservation of mass between the estimates of annual N fluxes in the TREND-Nitrogen v3 county-scale dataset and those in gTREND-Nitrogen. For this step, the gridded N estimates for each component of the N mass balance were re-aggregated to the original county scale, and mass values were compared via linear regression analysis (**Figure 2.6)**.

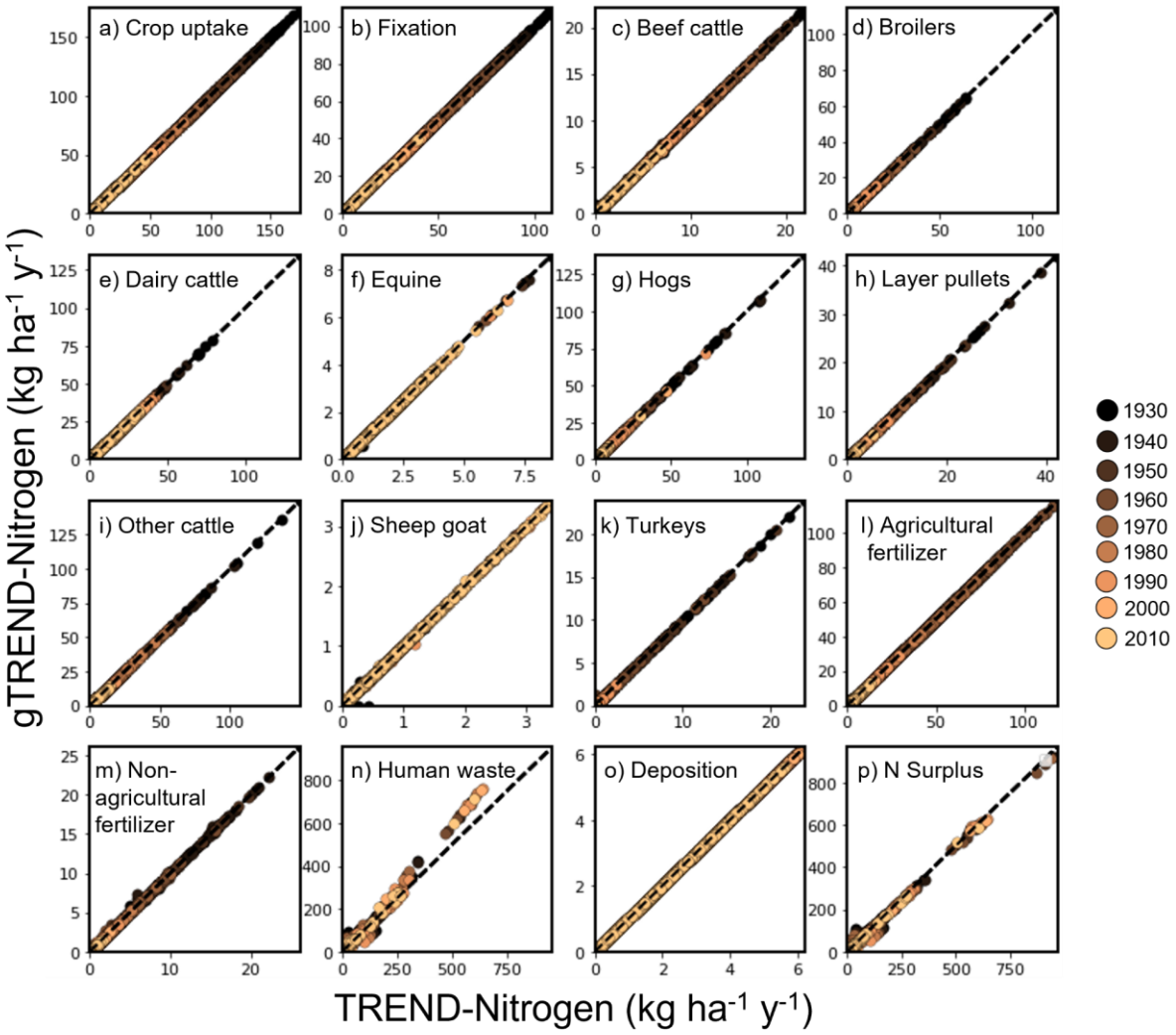


Figure 2.6 Relationship between county-scale TREND-Nitrogen v3.0 estimates of N mass balance components and the gridded gTREND-Nitrogen estimates, re-aggregated to the county scale. Data points represent estimates for each county for each decadal year from 1930 to 2010. The dashed lines represent a 1:1 relationship.

As shown in **Figure 2.6**, we found a strong correlation between the county-scale estimates from TREND-Nitrogen v3 and the gTREND-Nitrogen data re-aggregated to the county scale (R^2 > 0.99). In this context, the greatest deviation from a 1:1 relationship between the two datasets was found for the human waste component of the mass balance (**Figure 2.6n**). The difference is due to differences between the population datasets used for TREND-Nitrogen (county) and gTREND-Nitrogen. Even in this case, however, we see a strong correlation between estimates of N inputs for the two datasets (R^2 = 0.99).

2.3.3 Use of gTREND vs county-scale TREND-Nitrogen for watershed-scale analysis

When developing this down-scaled data product, we predicted that use of the gridded gTREND nitrogen data would be of particular value when estimating N fluxes for smaller watersheds. To test this hypothesis, we calculated N surplus magnitudes for more than 1,000 watersheds of varying size using both the county-scale TREND-Nitrogen dataset and the gridded gTREND-Nitrogen dataset. For this analysis, we included 1,285 watersheds from the U.S. Geological Survey (USGS) National Water Quality Program (NWQP) Surface Water Trends project.

The results of our analysis first show a strong correlation between the county-based and the grid-based watershed-scale estimates of N surplus magnitudes ($R^2 = 0.71$) (**Figure 2.7a**). This strong correlation is expected, as the grid-based data is, of course, derived from the county-scale data. There is, however, a notable amount of scatter away from the 1:1 line. To further explore this scatter further, we show in **Figure 2.7b** the ratios of county-based N surplus estimates to grid-based estimates ($R = SURP_{county}$ (i)/$SURP_{gridded}$ (i)) for all 1,285 study watersheds. In the figure, R values > 1 indicate a higher N surplus estimate with the county-scale data, and R values < 1 indicate a higher estimate from the grid-scale data. These results are shown across a range of watershed sizes, extending from a minimum of <1 km^2 to a maximum of more than 2,000,000 km^2. As we hypothesized, for smaller watersheds (those less than 50,000 km^2), it can make a large difference whether the county-based or grid-based data is used for the estimation of N fluxes (R = 1.04, IQR = 0.48) (**Figure 2.7c**). For larger watersheds, however, there is less variability around the 1:1 line (R = 1.18, IQR = 0.28). This greater variability for the smaller watersheds suggests that use of coarser, county-scale estimates for smaller watersheds can potentially lead to large over- or under-estimates of N fluxes. These results clearly demonstrate the value of using down-scaled N mass balance data when attempting to understand and quantify watershed-scale N dynamics.

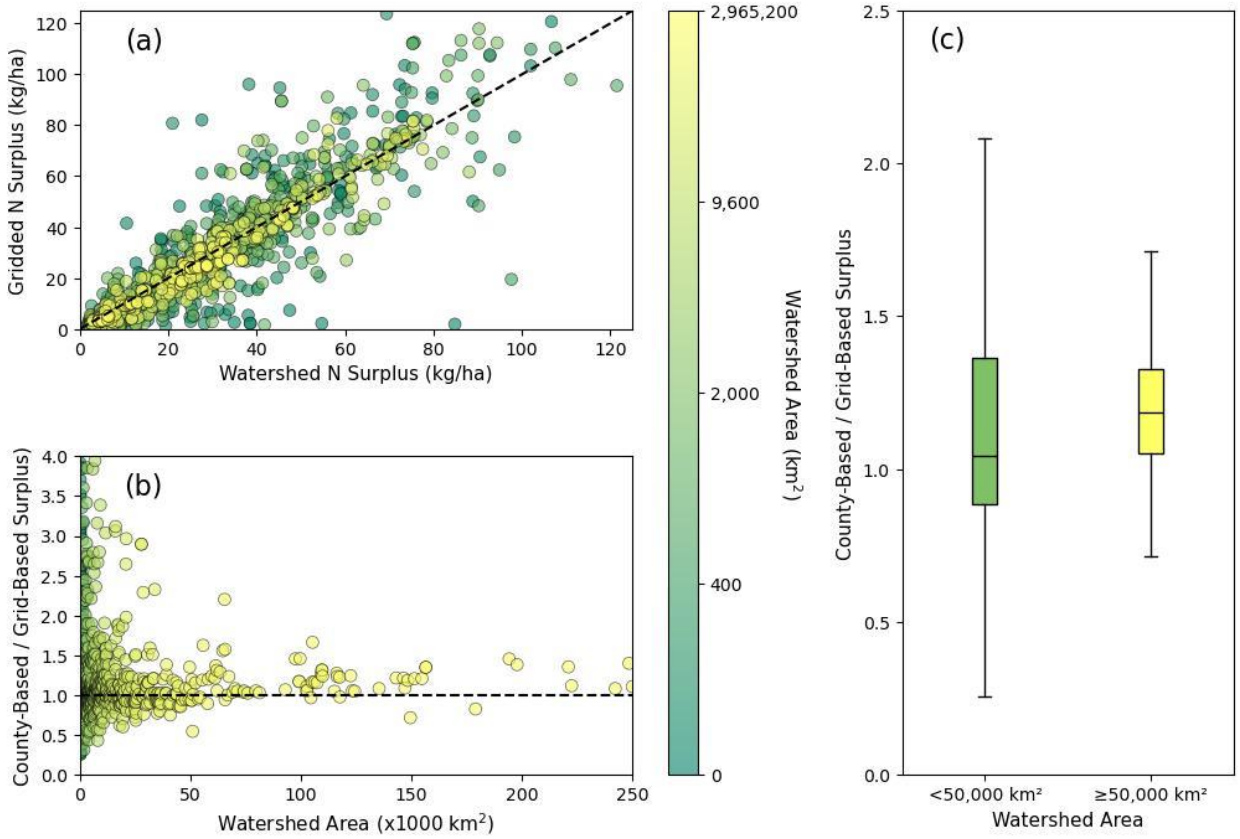


Figure 2.7 Comparison of watershed N surplus magnitudes based on county-scale and gridded N surplus data. (a) Overall, we see a 1:1 relationship between estimates of N surplus magnitudes from the county-scale and gridded data ($R^2 = 0.71$). (b,c) For large watersheds, >50,000 km^2, it makes less difference which dataset is used. For smaller watersheds, however, use of county-scale data can lead to large overestimates or underestimates of N fluxes.

2.4 Limitations

The TREND datasets have limitations regarding the scope of nitrogen sources considered. While we account for major anthropogenic inputs (synthetic fertilizer, manure, biological nitrogen fixation in croplands, human waste, and atmospheric deposition), several other potential nitrogen sources and fates were not included. Background biological nitrogen fixation from non-agricultural ecosystems, nitrogen released through rock weathering, volatilization and denitrification losses, and riverine export were excluded from the calculations. Though these sources contribute to the overall nitrogen cycle, at large scales they represent smaller fluxes compared to the major agricultural and anthropogenic inputs considered (Compton et al., 2003; Galloway et al., 1995; Lin et al., 2020; Sabo et al., 2019). While these sources have been quantified in field to regional-scale budgets, data at the national scale remains limited. For

instance, Compton et al. (2003) found that biologically fixed nitrogen from Red Alder trees in watersheds dominated by this species can contribute up to 35 kg-N ha^{-1} yr^{-1}. However, there are no existing datasets of Red Alder tree species distributions across the US spanning 1930 to 2017. Similarly, riverine nitrogen export at the county or gridded scale is highly uncertain because export is frequently modeled at the watershed level. These data limitations result in a high degree of uncertainty when estimating these processes at relevant spatial and temporal scales (Bouwman et al., 2013; Davidson, 2009), leading us to exclude them from the surplus calculations to be later estimated through modeling exercises.

2.5 Potential Use

The expectation for the gTREND-Nitrogen dataset is that it will be widely used for environmental monitoring and assessment and that it will be of particular value for water quality model development. Increasingly, data-driven water quality models are being developed in the domain of large-sample hydrology, which typically requires data for hundreds or even thousands of catchments (Kratzert et al., 2023). To ensure accessibility of the gTREND-Nitrogen data for such watershed-scale applications, we provide a script that can be used to aggregate the gridded data to desired watershed, administrative, or other type of boundary. The script is available through the project homepage at https://github.com/vanmeter-kvm/gTREND-Nitrogen/.

2.6 Data Records

The gTREND-Nitrogen dataset covers the contiguous United States and includes 22 N mass balance components for the years 1930-2017. The current version of the gTREND-Nitrogen dataset is available in figshare (doi forthcoming). In the figshare repository, the gTREND-Nitrogen data is zipped into subfolders, organized around different components of the N mass balance. Inside the zipped folders, data is available in the Tag Image File Format (TIFF). Each TIFF file represents an annual N mass balance component with the unit of kg-N $ha^{-1}y^{-1}$ at 250-m grid scale. We also provide aggregated N surplus estimates at hydrologic unit code 8 (HUC8) level from 1930 to 2017 in comma-separated value (.csv) format.

The dataset is organized into multiple zipped subfolders, with 22 GRID_scale subfolders and one HUC8_scale subfolder, as follows:

- The zipped GRID_scale folders include one subfolder per N mass balance component. Each zipped folder then contains 88 TIFF files corresponding to the years 1930-2017. A detailed description of the data structure and files provided in the data repository is provided in **Table 2.4**.

- The zipped HUC_scale folder contains one csv file (*Nsurplus_HUC8.csv*) and one shapefile (*HUC8.gpkg*). The HUC8.shp contains the HUC8 boundaries for the continental United States, which is used to extract basin-averaged N surplus values. In *Nsurplus_HUC8.csv,* the first column is named "HUC8," which contains the unique HUC8 basin identifiers. The csv also includes one column for each year of the dataset, with the naming convention "YYYY_Nsurplus," with YYYY values referring to the data year. The data in these columns corresponds to aggregated basin-averaged N surplus values from 1930 to 2017.

Code for the TREND-Nitrogen county-scale data is available upon request and will be available on a later date at github.com/danykakbyrnes/TREND_Nutrients.git.

Table 2.4 Structure and File Names of the Gridded gTREND-N Dataset in the Data Repository. In the file names, the YYYY values refer to the data year.

Zipped Subfolder Name	File name	Zipped Folder Content	
		File Naming Convention	Number of Files
Oxidized Atmospheric N Deposition	Atmospheric_Oxidized.zip	Atmospheric_Oxidized_YYYY.tif	88
Reduced Atmospheric N Deposition	Atmospheric_Reduced.zip	Atmospheric_Reduced_YYYY.tif	88
Crop and Pasture N Uptake	CropUptake_Ag.zip	CropUptake_Ag_YYYY.tif	88
Crop N Uptake	CropUptake_Cropland.zip	CropUptake_Cropland_YYYY.tif	88
Pasture N Uptake	CropUptake_Pasture.zip	CropUptake_Pasture_YYYY.tif	88
Farm Fertilizer	Fertilizer_Ag.zip	Fertilizer_Ag_YYYY.tif	88
non-Farm Fertilizer	Fertilizer_Dev.zip	Fertilizer_Dev_YYYY.tif	88
Biological N Fixation (all-agriculture)	Fix_Ag.zip	Fix_Ag_YYYY.tif	88
Cropland N Fixation	Fix_Cropland.zip	Fix_Cropland_YYYY.tif	88
Pasture N Fixation	Fix_Pasture.zip	Fix_Pasture_YYYY.tif	88
Human Waste N	Human.zip	Human_YYYY.tif	88
Livestock Waste N - Milk Cows	MilkCow.zip	MilkCow_YYYY.tif	88
Livestock Waste N - Beef Cattle	BeefCow.zip	BeefCow_YYYY.tif	88
Livestock Waste N - Other Cattle	OtherCattle.zip	OtherCattle_YYYY.tif	88
Livestock Waste N - Horses	Equine.zip	Equine_YYYY.tif	88
Livestock Waste N - Hogs	Hogs.zip	Hogs_YYYY.tif	88
Livestock Waste N - Layers & Pullets	LayerPullets.zip	LayerPullets_YYYY.tif	88

Livestock Waste N - Broilers	Broilers.zip	Broilers_YYYY.tif	88
Livestock Waste N - Turkeys	Turkeys.zip	Turkeys_YYYY.tif	88
Livestock Waste N - Sheep & Goats	SheepGoat.zip	SheepGoat_YYYY.tif	88
Livestock Waste N - All	Livestock.zip	Livestock_YYYY.tif	88
N Surplus	Surplus_N.zip	Surplus_N_YYYY.tif	88
HUC8-scale N surplus magnitudes	HUC8_scale.zip	Nsurplus_HUC8.csv HUC8.gpkg	2

Chapter 3 Phosphorus gridded and county-scale TRajEctories Nutrient Dataset (TREND) for the contiguous United States (1930-2017)

Outline

Phosphorus (P) is essential for the U.S. food system, but its overuse has accelerated the cultural eutrophication and hypoxia of surface waters. Despite numerous policies and substantial investments aimed at reducing phosphorus loads, nutrient-induced algal blooms continue to occur each summer across the U.S., threatening drinking water security and limiting the economic and social use of surface waters. It is now understood that persistent eutrophication is partly due to the long shadow of decades of phosphorus overuse: accumulated stores of legacy phosphorus. Quantifying landscape P fluxes can help better understand key sources, sinks, and regions at risk of large P stores. To this end, we created two long-term (1930-2017) P mass balance datasets at the county and gridded scale, providing unprecedented temporal resolution and spatial coverage. These datasets offer opportunities for fine-scale insights into both historical and current P dynamics. Our new, open-source dataset can provide opportunities for fine-scale insights into both historical and current P dynamics. The first authorship of this work is shared between Danyka Byrnes and Shuyu Chang. Please see *Statement of Contributions* for more details.

3.1 Introduction

The discovery of phosphorus by a 17th-century alchemist marked the beginning of its recognition as a vital component for improving soil fertility (Ashley et al., 2011). By the early 1800s, ground bones were used as a source of phosphorus, leading to the mining of phosphate rock and the eventual creation of "superphosphate" fertilizers (Cordell et al., 2009). Phosphorus has since become a crucial input for modern agriculture, playing a key role in bolstering global food systems and enhancing food security. This development was pivotal in addressing food insecurity by allowing farmers to supplement natural soil phosphorus levels with external sources rather than relying solely on manure and crop residues.

The increased use of phosphorus in agriculture over the past century has been driven by population growth, dietary changes, agricultural intensification, and increasing reliance on biofuels (Galloway & Cowling, 2021; Glibert, 2020; Van Meter et al., 2018). After World War II, the mechanization of farming practices and the intensification of the agricultural system led to a rapid increase in phosphorus use across the U.S. (Glibert, 2020). This intensification also contributed to the consolidation of small farms into larger operations, resulting in a system with

significant nutrient surpluses and decoupling of nutrient sources and sinks (Byrnes et al., 2020; Glibert, 2020). In U.S. agricultural systems, phosphorus uptake by crops has consistently remained below 80%, leaving a substantial portion of excess nutrients available to enter the environment (Conant et al., 2013; Swaney et al., 2018b; Swaney & Howarth, 2019b).

Phosphorus from the landscape contributes to cultural eutrophication and hypoxia in bottom waters in both fresh and coastal waters (Carpenter et al., 1998; Rabalais, Turner, & Wiseman, 2002; Schindler et al., 2016). The consequences of eutrophication and associated hypoxic zones are multi-faceted. It can threaten drinking water quality sources with elevated algal bloom microcystin toxins, change the species composition, and degrade the habitat of major fisheries, with conditions favoring some species at the loss of others (Carpenter et al., 1998; Le Moal et al., 2019; Rabalais, Turner, & Wiseman, 2002). Furthermore, the excess growth of algae also restricts the use of water for recreation and industry (Carpenter et al., 1998; Malone & Newton, 2020; Sharpley & Wang, 2014).

Policymakers have set goals to manage excess nutrients in response to freshwater and coastal ecosystem degradation. In the Chesapeake Bay and Lake Erie drainage areas, nutrient reduction goals were set to reduce phosphorus loads by up to 40% and 50%, respectively (Chesapeake Bay Program, 1988; Great Lakes Water Quality Agreement Nutrient Annex Subcommittee, 2015). However, despite efforts to manage nutrients within the contributing watersheds, in many cases, there has been little improvement in downstream water quality, and we continue to see persistent eutrophication and late-summer hypoxia (Jarvie et al., 2013; Van Meter et al., 2021). It is increasingly recognized that phosphorus exports from watersheds are influenced not only by current-year inputs but also by legacy stores of phosphorus in the landscape (Jarvie et al., 2013; Müller-Karulis et al., 2024; Sharpley et al., 2013; Van Meter et al., 2021). The residence time of phosphorus means it can threaten water quality decades after application via soil erosion and remobilization from sediments along the land-river continuum (Alewell et al., 2020; Sharpley et al., 2013). Indeed, recent work in the Baltic Sea found that the residence time of phosphorus across the drainage basins ranged from 4.8 to 30 years. Similarly, a modeling study in the Grand River watersheds in the Lake Erie drainage basin found that even under aggressive phosphorus management scenarios, the 40% phosphorus load reduction target would not be met within 25 years (Van Meter et al., 2021). There is now unequivocal evidence that legacy phosphorus must be considered in management plans to achieve desired water quality outcomes. Managing impaired water quality requires understanding past and present fluxes across all land uses (Haygarth et al., 2014); however, we cannot manage what we do not measure. Therefore, to achieve this goal, long-term and spatially explicit phosphorus mass balance datasets are essential

for understanding nutrient dynamics and the distribution of legacy stores (Batool et al., 2022; Byrnes et al., 2020).

Mass balance datasets have been developed across the U.S. to capture phosphorus fluxes in the landscape and understand the impact of phosphorus on water quality. Howarth et al. (1996) were among the first to develop a regional mass balance study to quantify nitrogen and phosphorus fluxes to the Atlantic Ocean. Other regional mass balances were developed, including work from David and Gentry (2000), who developed a phosphorus mass balance for Illinois from 1945 to 1998. Goyette et al. (2016b) developed a mass balance of the transnational St. Lawrence River Basin from 1901-2012. The first large-scale mass balance, the NAPI (net anthropogenic phosphorus input) dataset, was developed by Swaney et al. (2018a), covering the contiguous U.S. from 1987 to 2017. Sabo et al. (2021) also developed a U.S-scale mass balance from 2002-2017 and expanded on the fluxes commonly included in previous datasets. They also used HUC8 watershed boundaries instead of county-scale ones, making it more applicable for hydrology studies. These datasets have their applications but are still limited in spatial and temporal scale to understand the long-term effects of phosphorus fluxes on environmental outcomes across the U.S.

Our current work aims to develop a long-term phosphorus mass balance dataset for quantifying phosphorus fluxes and surplus across space and time. We developed TREND-Phosphorus, a county-scale phosphorus mass budget spanning 1930-2017, with estimates of all individual components of the budget for the contiguous U.S. We then downscaled our county data to a 250 m grid to produce gTREND-Phosphorus, the first 88-year, 250 m gridded phosphorus mass budget dataset.

3.2 Methods

3.2.1 Dataset preparation: county-scale phosphorus mass balance

The components of the phosphorus mass balance were calculated annually at the county-scale from 1930–2017. We also downscaled the county-scale phosphorus fluxes (fertilizer, livestock manure phosphorus, crop uptake, atmospheric deposition, rock weathering, detergent, human waste, and food waste), to a 250-m grid scale with publicly available gridded land use/land cover data (Multi-Resolution Land Characteristics Consortium, 2016; T. L. Sohl et al., 2014) and gridded population data (Fang & Jawitz, 2018). Compiling such a large dataset requires harmonizing datasets, often created for different purposes, for different end users, and covering different spatial and temporal scales. The steps taken to harmonize disparate datasets and

downscale the county dataset are described below. The first authorship of this work is shared between Danyka Byrnes and Shuyu Chang. Please see Statement of Contributions for further details.

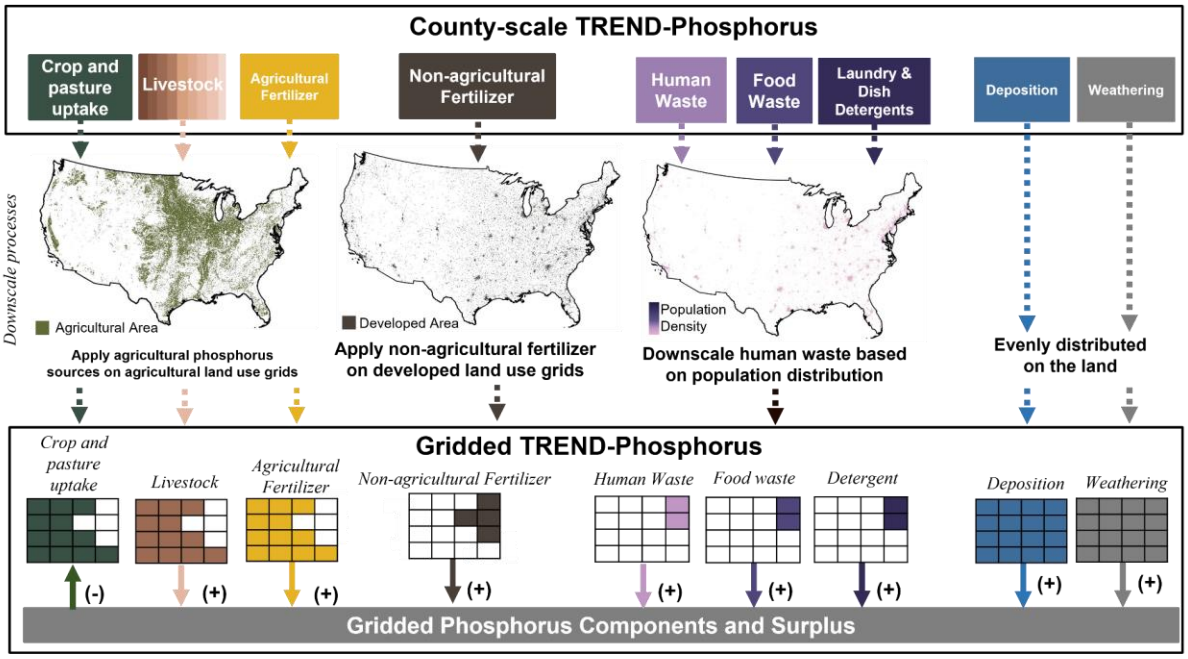


Figure 3.1 Data sources and workflow for creating the gTREND-Phosphorus dataset. (1) Agricultural components of the phosphorus mass balance such as crop uptake, manure, and agricultural fertilizer are downscaled to areas with agricultural land use; (2) Non-agricultural fertilizer was downscaled to areas with developed land use; (3) Atmospheric deposition and phosphorus rock weathering are distributed evenly across all land uses; (4) Human waste, food waste, and detergent is downscaled based on gridded population data.

3.2.1.1 Inorganic phosphorus fertilizer

County-scale estimates of both farm and non-farm phosphorus (P) fertilizer use, $FERT_f$ and $FERT_{nf}$ (kg-P y^{-1}), were obtained from several different sources. From 1945-1985, data was obtained from Alexander and Smith (1990). From 1987-2012, data was obtained from Brakebill and Gronberg (2017). For the period 1930-2014, we created county-specific scaling factors by taking the mean ratio of each county's fertilizer use to the total of U.S. phosphate rock fertilizer used or consumed for the period 1945-1949 (*Phosphate Rock - Historical Statistics (Data Series 140) 2018 Update*, 2018). The scaling factors were then multiplied by the mass of the U.S. phosphate rock fertilizer used or consumed for each year of the period 1930-1944, as follows:

$$FERT_{tot}(i,t) = FERT_{rock}(t) * S(i) \quad \text{(Eq. 3.1)}$$

where $FERT_{tot}(i,t)$ $(kg\text{-}P\ y^{-1})$ is the total phosphorus fertilizer use for county *i* in year *t*, $FERT_{rock}(t)$ is the U.S. phosphate rock fertilizer used or consumed ($kg\text{-}P\ y^{-1}$) in year *t*, and $S(i)$ is the county-scaling factor. Some counties are not included in the 1945-1986 dataset; in this case, we used the same approach starting in 1987 instead of 1945, using the 1987-1991 period to create the scaling factors.

To obtain estimates of non-farm fertilizer for 1930-1986, we first calculated the mean ratio of non-farm to total P fertilizer use, as reported by Brakebill & Gronberg (2017) for individual counties. Then, for the years prior to 1987 (1930-1986), we used those county-scale ratios to partition total P fertilizer use into estimates of farm and non-farm fertilizer use.

For 2013–2017, national-scale estimates of P fertilizer use(Food and Agriculture Organization of the United Nations (FAO), 2022) were used to scale the county-scale estimates of both farm and non-farm P fertilizer use forward in time. Scaling was achieved by taking the ratio of mean county-scale fertilizer use (farm and non-farm) to mean national fertilizer use for the years 1987-2012 to create scaling factors for each county and then multiplying the scaling factors, S_f and S_{nf}, by the annual estimates of P fertilizer use (2013-2017):

$$FERT_f(i,t) = FERT_{natl}(t) * S_f(i) \quad \text{(Eq. 3.2)}$$

$$FERT_{nf}(i,t) = FERT_{natl}(t) * S_{nf}(i) \quad \text{(Eq. 3.3)}$$

where $FERT_f(i,t)$ and $FERT_{nf}(i,t)$ are the total farm and non-farm P fertilizer use, respectively, for individual counties and $FERT_{natl}$ is the FAO estimate of national P fertilizer use.

3.2.1.2 Phosphorus in livestock excretion

Manure phosphorus inputs, *MAN*, defined here as waste excretion (urine and manure) from ten livestock inventories including dairy cows, beef cows, and other cattle (calves, heifers that have not calved, steers, bulls), broilers, other chickens (laying chickens and pullets), turkeys, hogs and pigs, goats, sheep and lambs, and equine (horses, mules, donkeys and ponies). All livestock inventory data originates from the U.S. Census of Agriculture. Livestock inventory numbers for cattle, hogs, poultry, sheep, and horses for the period 1950-2017 (reported at 5-year intervals) were obtained from a USGS dataset that includes not only Census of Agriculture numbers but also estimates of livestock inventory numbers for counties with censored data in some census years (Falcone, 2021). For the period 1930-1945 and livestock categories not explicitly included in the USGS dataset (pullets, mules, donkeys, goats), data was obtained from the ICPSR data archive (1930–2012) (Haines et al., 2018) and NASS Quick Stats (USDA National Agricultural

Statistics Service, 2019). Livestock census category changed over time and we consolidated reported categories to maintain in livestock groups (**Supplemental Table 3.1**)

For years and for counties for which livestock inventory data from the Census of Agriculture is "undisclosed" due to privacy concerns, data gaps are filled by using the data from the most recent year for which there is existing data for that county and livestock type ((Falcone, 2021) If there is undisclosed data for more than six census time periods (~30 years), the inventory is assumed to be zero.

A data-fusion approach was used to integrate the 5-year U.S. Census of Agriculture data with annual Agricultural Survey data. Using this approach, the Census of Agriculture livestock inventory data was used as reported, and the annual Agricultural Survey data was scaled to the Census of Agriculture data. Scaling was achieved by taking a ±10-year window around each year in which the Agricultural Survey was to be used. For these 20 years, we created a scaling factor, $S_{lvstk}(i,j,t)$, for county i, livestock type j, in year t, calculated by taking the ratios of census data to survey data for years where both were available, and then taking the mean of those ratio values. Then, for each year with available survey data, but no census data, crop production or crop area values were calculated by multiplying the scaling factor by the reported survey data:

$$INV(i,j,t) = INV_{surv}(i,j,t) * S_{lvstk}(i,j,t) \quad \text{(Eq. 3.4)}$$

where $INV(i,j,t)$ is the estimate of the number of heads of a specific livestock type j in county i in year t, and $INV_{surv}(i,j,t)$ is the number of heads of that livestock reported in the Agricultural Survey.

The magnitude of P excretion per animal varies with animal weight, and livestock weights have increased over time due to both genetic evolution and improvements in feeding (Falcone, 2021). We use time-varying livestock weights based on use of annual slaughtered livestock weight data (U.S. Department of Agriculture, n.d.). Reported weights were available for the time period 1958-2017 (U.S. Department of Agriculture, n.d.). Any data gaps across this period were filled via linear interpolation. For the period 1930-1957, we used the reported 1958 weights, assuming no change in livestock weights for that period. No time-varying weights were available for horses, so we assumed weight to remain constant across the study period. In this new approach, 2008 was chosen as a base year due to the availability of P excretion parameters for 2008 livestock weights (USDA NRCS, 2008). Temporally varying, livestock type-specific scaling factors, $S_{wt}(i,t)$, were then developed based on the ratio of the 2008 weight to the reported weight for any given year (**Supplemental Table 2.1**).

With the resulting county-scale livestock inventory trajectories, P in livestock manure was calculated as a function of these inventory numbers and the estimated P content of the manure produced per animal per year (Byrnes et al., 2020, 2020d, 2022):

$$MAN(i,j,t) = INV(i,j,t) * p(j)_{lvstk,excr} * S_{wt}(i,j,t) \quad \text{(Eq. 3.5)}$$

where MAN(i,j,t) is the county-scale, time-varying estimated P in manure from livestock type j, $p(j)_{lvstk,excr}$ is the estimated P content in manure from that livestock type, and $S_{wt}(i,j,t)$ is the time-varying scaling factor for livestock weights for county i and livestock type j. The parameter values used for 2008 livestock P consumption and excretion are provided in **Table 3.1**. Excretion values were obtained from the USDA Agricultural Waste Management Field Handbook (USDA NRCS, 2008). Consumption values for cattle were obtained based on a calculated ratio of consumption to excretion (Byrnes et al., 2022; Hong et al., 2011).

Table 3.1 Livestock phosphorus consumption and excretion parameters. Consumption values are included for all grazing livestock.

Livestock Type	Livestock P Consumption, 2008 (kg-P head^{-1} y^{-1})	Livestock P Excretion, 2008 (kg-P head^{-1} y^{-1})
Cattle, Beef	18.32	16.06
Cattle, Milk	33.02	25.61
Cattle, Other (heifers, bulls, steers, calves)	10.89	9.55
Goats	1.39	1.16
Sheep and Lambs	1.39	1.16
Horses	9.45	8.44
Hogs and Pigs	-	2.35
Poultry, Turkey	-	0.36
Poultry, Broilers	-	0.12
Poultry, Other Chickens	-	0.18

3.2.1.3 Crop uptake

County-scale crop phosphorus uptake in county i and year t was estimated as the sum of crop-specific uptake values CROP(i,j,t) (kg- P y^{-1}) for all crop types j in that county (field crops, fruits and vegetables, and orchard crops). The crop-specific, county-scale P uptake values CROP(i,j,t) were calculated on the basis of either mass, harvested area, or number of trees by multiplying, where applicable, the harvested crop magnitude, HARV(i,j,t) in census-reported units, e.g. bushels, hundredweight, acres, by a crop-specific conversion factor, c(j), to convert to kg crop, by the crop-specific percent dry matter, d(j). Then multiply by the crop-specific phosphorus uptake values, p_{crop}(j,t):

$$CROP(i,j,t) = HARV(i,j,t) * c(j) * d(j) * p_{crop}(j,t) \quad \text{(Eq. 3.6)}$$

Total county-scale crop P uptake, CROP(i,t) was then calculated by summing the uptake values for all of the included crops. The 63 different crops included here were confirmed to cover ≥97% of the county-scale cropped area. See **Supplemental Table 3.2-3.4** for conversion factors, dry matter values, and P uptake parameters field crops, fruits and vegetables, and orchard crops, respectively. Parameter values are derived from the literature. Field crop parameter values are available from the NAPI dataset, IPNI, and NRCS (*IPNI Estimates of Nutrient Uptake and Removal*, n.d.; Lander et al., 1998b; Swaney & Howarth, 2019a). For crops not included in those datasets, we converted the TREND-Nitrogen parameters values using a ratio of the nitrogen and phosphorus content of the crops from the USDA FoodData Central nutrient database (United States Department of Agriculture, n.d.).

Annual crop production data, *HARV(i,j,t)*, was obtained using a data-fusion method that integrates data from the Census of Agriculture, which reports data every 5 years, with the U.S. Agricultural Survey, which provides projected annual estimates. Census data was obtained from the ICPSR data archive (1930–2012) (Haines et al., 2018) and NASS Quick Stats (USDA, 2023), and Agricultural Survey data was obtained from NASS Quick Stats (USDA, 2023). Using this approach, and to ensure the compatibility of the datasets, the Census of Agriculture crop data was used as reported, and the annual Agricultural Survey data, $HARV_{surv}(i,j,t)$ was scaled to the Census of Agriculture data for the years when census data was not available (Eq. 3.7). Scaling was achieved by taking a ±10-year window around each year in which the Agricultural Survey was to be used. For these 20 years, we created a county-specific scaling factor, S_{crop} (i,j,t), calculated by taking the ratios of census data to survey data for years where both were available, and then taking the mean of those ratio values across all census years within the specified time window. Then, for each year with available survey data, but no census data, crop production or crop area values were calculated by multiplying the scaling factor by the reported survey data:

$$HARV(i,j,t) = HARV_{surv}(i,j,t) * S_{crop}(i,j,t) \quad \text{(Eq. 3.7)}$$

where $HARV_{surv}(i,j,t)$ represents the estimates of crop production or harvested crop area in county i for crop type j in year t from the Agricultural Survey.

In the census, harvest data for fruit and nut trees were reported, depending on the year, as either the *num*ber of trees or as acres of trees. To estimate phosphorus removal from tree uptake, $CROP_{tree}$ (kg-P y^{-1}), we first needed to standardize census inventories between the years. We converted acres of trees to the number of trees using county-specific orchard density (den_{orch}; *trees acre*$^{-1}$) from the nearest available census year with available data. When county data was

not available, we used state-level orchard density. We used literature values to estimate average, non-time varying annual mass of phosphorus removal through tree uptake (p_{tree}) in county *i*, tree type *j*, in year *t*:

$$HARV_{tree}(i,j,t) = HARV_{acres}(i,j,t) * den_{orch} \quad \text{(Eq. 3.8)}$$

where $HARV_{tree}(i,j,t)$ in number of trees after using county or state-level average orchard density (den_{orch}) to convert acres of tree ($HARV_{acres}$) *for* county *i*, crop type *j*, in year *t* (**Supplemental Table 3.4**).

The nutrient concentrations of grain corn, wheat, and soybeans have evolved with the development of new, higher-yielding hybrids (Ciampitti & Vyn, 2012; G. McIsaac, personal communication, 2023; Scott et al., 2006; Villamil et al., 2019). To accurately reflect current knowledge regarding changes in grain phosphorus content over time, we now use time-varying parameter values for these three common field crops in the contiguous U.S. For the period 1930-1989, we assume constant P uptake values (kg-N kg-crop^{-1}) based on parameter values from the International Plant Nutrition Institute (IPNI) and other literature sources. Then, for the period 2013-2017, we use newly estimated parameter values that reflect current uptake magnitudes (Ciampitti & Vyn, 2012; G. McIsaac, personal communication, 2023; Scott et al., 2006; Villamil et al., 2019). For the period 1990-2013, we then linearly interpolate the P uptake parameters between the two time periods. Both the pre-1990 and current crop P uptake parameters for corn, soy, and wheat are provided in **Supplemental Table 3.2**.

Censored county-scale data from the Census of Agriculture was gap-filled by calculating the difference between reported state-level production values and the sum of all reported county-level values. This difference was then apportioned equally to all of the counties with censored data as a function of the ratio between county-scale and state-scale values from the nearest census years with reported data.

3.2.1.4 Pasture uptake

County-scale pasture P uptake, PAST (*kg-P y^{-1}*), estimates the phosphorus content removed from grazing livestock (cattle, horse, goats, and sheep), based on both the total available pasture area and the number of grazing livestock in the county. To account for different yields, for our calculations, pasture were separated into two categories: “non-cropland pasture” and “cropland pasture”, based on Hong et al. (2011) Pasture categories are from the Agricultural Census for the period 1945-2017. Prior to 1945, we used aggregated fraction of county area considered cropland or pastureland for each county from Ramankutty and Foley (1999) and used it to backward scale

the non-cropland ($AREA_{ncp}$) and cropland pasture area ($AREA_{cp}$) in each county for each census year. Non-cropland pasture is scaled as follows:

$$AREA_{ncp}(i,t-1) = AREA_{ncp}(i,t) + \Delta P_{ncp}(i,t) * AREA_{ncp}(i,t) \quad \text{(Eq. 3.9)}$$

$$\Delta P_{ncp}(i,t) = (FRAC_{past}(i,t-1) - FRAC_{past}(i,t)) \,/\, FRAC_{past}(i,t) \quad \text{(Eq. 3.10)}$$

where AREAncp (acres) is the area of non-cropland pasture in county i, year t, $\Delta P_{ncp}(i,t)$ is the scaling factor for pasture in county i, year t, and FRACpast is the fraction of the total area that is considered pasture in county i in year t. If all non-cropland pasture were maximally grazed, $PAST_{nc}$, the max phosphorus uptake (kg-P y^{-1}), was calculated as follows:

$$PAST_{nc,max}(i,t) = AREA_{past}(i,t) * y_{ncp} * p_{ncp} \quad \text{(Eq. 3.11)}$$

where y_{ncp} is the yield for the pasture category (kg $acres^{-1}$) and p_{ncp} is the content in pasture grasses (kg-P kg-$pasture^{-1}$). The same equations (Eq. 5.12 and 5.13) are used for non-cropland pasture $PAST_{c,\,max}$ using the fraction of the total area that is considered cropland ($FRAC_{crop}$) to backward scale cropland pasture. Non-cropland and cropland pasture phosphorus uptake are then summed together as total pasture, $PAST_{max}(i,t)$.

Livestock inventory numbers were obtained from the Census of Agriculture (Falcone, 2021; Haines et al., 2018; *USDA - National Agricultural Statistics Service - State Agricultural Overview*, n.d.). For cattle, to account for some animals being raised in confinement while others are allowed to graze, we developed estimates of the number of grazing cattle using state-scale estimates of the fraction of different types of non-confined cattle for the year 2002, e.g. beef cows, dairy cows, other cattle (**Table 2.2**) (Kellogg et al., 2000b). For horses, goats, and sheep, we assumed that all animals were allowed to graze, and for hogs and poultry, we assumed no grazing. Since large-scale confinement is a product of the last 20-years (Glibert, 2020; J. M. MacDonald & McBride, 2009), we scale the confinement factors over time using a national cattle-to-farm ratio (CF_R), normalized to 1 in 2002, the census year closest to the publication of the confinement fractions (USDA National Agricultural Statistics Service, 2019).

Potential P uptake from consumption of pasture grasses by livestock, $CONS_{max}(i,j,t)$, was first calculated by multiplying the county inventory numbers for the different livestock types, $INV(i,j,t)$, by both the fraction of unconfined animals, $(1-f_{conf}(j,t))$, and by literature estimate of P consumption for the different livestock types (**Table 3.1**) (Byrnes et al., 2022; Swaney et al., 2018a; USDA NRCS, 2008).

$$CONS_{max}(i,j,t) = INV(i,j,t) * (1 - f_{conf}(i,j) * CF_R(t)) * p_{cons}(j) * S_{wt}(i,j,t) \quad \text{(Eq. 3.12)}$$

where $CONS_{max}(i,j,t)$ is the potential P consumption for livestock type *j* in county *i*, year *t*, $p_{cons}(j)$ is the N intake parameter for that cattle type, and $CF_R(t)$ is the cattle-to-farm ratio, normalized to be equal to 1 in 2002, and $S_{wt}(i,j,t)$ is the time-varying scaling factor for livestock weights. The county-scale estimates of total potential pasture P consumption, $CONS_{max}(i,t)$, were then calculated as the sum of potential intake values for all livestock types:

$$CONS_{max}(i,t) = \sum_{j=1}^{n} CONS_{max}(i,j,t) \quad \text{(Eq. 3.13)}$$

If potential consumption values exceeded the calculated maximum N uptake for pasture area, pasture N uptake values, $PAST(i,t)$ were capped at the area-based P uptake value:

$$PAST(i,t) = \begin{cases} CONS_{max}(i,t), & CONS_{max}(i,t) < PAST_{\max}(i,t) \\ PAST_{max}(i,t), & CONS_{max}(i,t) \geq PAST_{\max}(i,t) \end{cases} \quad \text{(Eq. 3.14)}$$

where $PAST(i,t)$ is the pasture uptake (kg-P ha^{-1} yr^{-1}) for county *i* in year *t*.

3.2.1.5 Phosphorus in human and food waste

In the county-scale TREND-Phosphorus dataset, human P inputs were estimated using county-scale estimates of human population and literature estimates of annual human P excretion (0.6 kg-P $capita^{-1}$ y^{-1}) (Hong et al., 2011). Population data at the county scale was obtained at decadal time increments for the period 1930-1990 and at annual time increments for the period 1991-2017 (U.S. Census Bureau, n.d., 2007, 2016). The decadal data was interpolated for the 1930-1990 period to obtain annual population estimates. Here, it should be noted that a large fraction of human waste is treated in wastewater treatment plants or septic systems before being discharged to downstream waters. Quantifying these treatment trajectories is beyond the scope of the current work. Food waste P, FOOD, represents the portion of wasted food destined for the landfill and is quantified as a constant fraction (0.11) of human waste (Parfitt et al., 2010).

3.2.1.6 Phosphorus in detergents

We estimated the phosphorus content in laundry, DET_{laun} *(kg-P* y^{-1}), and dish detergents, DET_{dish} *(kg-P* y^{-1}), by considering changing patterns of detergent use and phosphorus content in North America. For dish detergent, estimates of use and phosphorus content were obtained from Han et al. (2012) and Goyette et al. (2016b). Historically, from 1930 to 2010, dish detergents contained 8.7% phosphates by weight. This percentage dropped to 0.5% in 2011 nationally following legislation in 16 U.S. states that mandated a reduction in phosphate content (McCoy, 2011; Utah State University Extension, 2011). We estimated a temporally varying the dishwashing detergent phosphorus use in each county (DET_{dish}) as follows:

$$DET_{dish}(i,t) = DISH(t) * USE_{dish}(t) * P_{dish}(t) * POP(i,t) \quad \text{(Eq. 3.14)}$$

where $USE_{dish}(t)$ national detergent use per capita in year t (spoons cap^{-1} y^{-1}), $P_{dish}(t)$ is the phosphorus content in detergents in year t (kg-P spoon^{-1}), $DISH(t)$ is the fraction of households with dishwashers in year t, and $POP(i,t)$ is the population for county i in year t.

We estimated time-varying laundry detergent phosphorus use, $DET_{laun}(i,t)$, using an approach developed for the Lake Erie basin (H. Han et al., 2012; Van Meter et al., 2021). Han et al. (2012) use historical state-level estimates of laundry detergent phosphorus content and usage patterns. Phosphorus content in detergents was 15% for decades, but after 1973, phosphorus content gradually decreased on a state-by-state basis following state-level bans on high phosphorus content detergents (Litke, 1999). After 2011, with the introduction of a nationwide ban, the phosphorus content in laundry detergents was mandated to be below 0.5%.

$$DET_{laun}(i,t) = USE_{laun}(t) * P_{laun}(i,t) * POP(i,t) \quad \text{(Eq. 3.15)}$$

where USE_{laun}(t) is detergent use in year t (kg-P cap^{-1} yr^{-1}) and $P_{laun}(i,t)$ is the phosphorus content in the deterrent for county i in year t.

3.2.1.7 Atmospheric deposition and weathering

Atmospheric phosphorus deposition, *ATM (kg-P y^{-1})*, originates from sea salt, dust, biogenic aerosol particles, and fuel combustion emissions (R. Wang et al., 2017). County-scale deposition data originated from global deposition estimates (2.5^{o}x1.27^{o} resolution) from an aerosol chemistry-climate model (R. Wang et al., 2017). The model estimated deposition in 1850, from 1960 to 1990 at ten-year intervals, and from 1997 to 2013 annually. Between 1850 and 1996, county-scale deposition rates were interpolated, and after 2013, it was assumed that deposition rates remained constant.

Gridded weathering phosphorus rates, WTHR *(kg-P ha^{-1} yr^{-1})*, are derived from modeled data compiled by Hartmann et al. (2014). The model calculates weathering rates as a function of runoff, lithologies, temperature, and soil properties, then scales weathering rates by phosphorus contents in different lithological classes. We determined the county-scale phosphorus weathering rate by taking the mean weathering rate in each county.

3.2.2 Creation of the gTREND-Phosphorus gridded dataset

County-scale P fluxes were downscaled to the 250-m grid scale according to land use/land cover and population densities, as described below. For the downscaling, P fluxes were divided into

three different land-use categories: (1) agricultural, (2) developed/urban, and (3) all land (See **Figure 2.3**). Agricultural P fluxes were considered to include four major landscape flux categories: (1) fertilizer, (2) livestock manure P, (3) P uptake by crops and pasture. Developed/urban fluxes were considered to include (1) P in human waste, (2) P in food waste, and (3) P in detergents and (4) non-farm P fertilizer. Finally, atmospheric P deposition was considered to be an all-land flux. While natural land use including wetlands and water areas received atmospheric deposition, they were not assigned any agricultural or urban P fluxes.

Land-use trajectories were established based on two different land use and land cover (LULC) datasets: (1) a modeled historical LULC dataset created at a resolution of 250 m (hereafter referred to as the Sohl LULC dataset)(T. Sohl et al., 2016b; T. L. Sohl et al., 2014), and (2) the widely used National Land Cover Database (NLCD), available at a resolution of 30 m (T. L. Sohl et al., 2018). These two datasets were integrated and harmonized to reconstruct temporally continuous and spatially explicit land-use trajectories from 1930-2017 at the 250-m grid scale for the contiguous United States (CONUS). Specifically, we use here the annual 250-m Sohl LULC data from 1938 to 2005, and then use 30-m NLCD data for the years 2006, 2008, 2011, 2013, 2016. The NLCD data was resampled from 30 m to 250 m using the QGIS-python Nearest Neighbor Function, version 3.28 (*QGIS*, n.d.). Consistent land use was assumed for the years 1930-1937 and in the years between the years of NLCD availability after 2005.

Gridded population-density data was obtained from the 1-km gridded, decadal population maps developed for the contiguous U.S. by Fang & Jawitz (2018, 2019). These maps were created using a parsimonious modeling approach, in which decennial U.S. census data was brought together with a variety of other factors, including socioeconomic desirability, suitability, and inhabitability. While five different population models of varying complexity were created by Fang & Jawitz (2018), we use the results of the M3 model, which was of intermediate complexity but found to be the most accurate based on validation using county subdivision population data. To provide annual estimates of population for the gTREND-Phosphorus downscaling work, we used linear interpolation between the modeled years, and then assumed a constant population for the years 2011-2017. The annual 1-km maps of human population were resampled to a 250-m resolution using the QGIS-python Nearest Neighbor Function, version 3.28 to align with other gridded components of the gTREND-Phosphorus dataset.

3.2.2.1 Agricultural phosphorus fluxes

To prepare for downscaling of the agricultural P fluxes, annual, gridded binary land-use maps were created to classify 250-m grid cells as either agricultural or non-agricultural, g$LU_{ag}(k,t)$ =

0/1, where *k* refers to the individual grid cell and *t* to year. Agricultural land use is defined here as areas used for the production of crops, areas being actively tilled, and areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay. Specifically, note that agricultural land was considered to include both cultivated cropland and hay/pasture in both the Sohl LULC dataset and the NLCD. Land not falling into either the developed or agricultural land classes was considered to be non-agricultural land, $gLU_{ag}(k,t) = 0$.

Total agricultural area in each county *i* in each year *t*, $AREA_{ag}(i,t)$ (ha), was then calculated as the sum of the gridded agricultural land. For each agricultural component of the P mass balance, *j*, county-scale P flux values normalized by agricultural area, $P_{ag,cnty}(i,j,t)$ (kg ha-agland^{-1} y^{-1}) were calculated by dividing the total county P flux by the time-varying agricultural area of each county, i. County-scale vector maps of $P_{ag,cnty}$ values were then rasterized to a 250-m grid scale. To appropriately assign agricultural P fluxes only to grid cells with agricultural land use, the county-scale flux values were multiplied by the 0 or 1 values in the binary land-use grid.

3.2.2.2 Phosphorus fluxes on developed land

As noted above, P fluxes on developed land were considered to include non-farm P fertilizer and human waste. For the non-farm fertilizer, the county-scale estimates of non-farm fertilizer, $FERT_{nf}$ (*kg-P y^{-1}*), were downscaled based on developed land use, assuming that homes and businesses using fertilizer would not occur on either agricultural or "other" land. Using a method similar to that developed for the agricultural P fluxes, we created binary maps to classify 250-m grid cells as either developed or non-developed, $gLU_{dev}(k,t) = 0/1$, where *k* refers to the individual grid cell and *t* to year. Land not falling into either the developed or agricultural land classes was considered to be non-developed, $gLU_{dev}(k,t) = 0$.

Total developed area in each county in each year, $AREA_{dev}(i,t)$ (ha), was then calculated as the sum of the gridded developed land in that county in year *t*. County-scale vector maps of $FERT_{nf,dev}$ values were then rasterized to a 250-m grid scale. Next, to assign the non-farm P fertilizer fluxes only to grid cells with developed land use, the county-scale flux values were multiplied by the 0 or 1 values in the binary land-use grid.

For annual grid-scale estimates of P human waste, *gHUM(k,t) (kg-P ha^{-1} y^{-1})*, food waste *gFOOD(k,t) (kg-P ha^{-1} y^{-1})*, and detergents, *gDET(k,t) (kg-P ha^{-1} y^{-1})*, were developed using gridded estimates of human population (Fang & Jawitz, 2018). We converted the gridded population to the respective components using the same literature-based estimates and approach

outlined in the country-scale analysis above, using gridded population instead of county-scale population.

3.2.2.3 All-land phosphorus fluxes (atmospheric deposition and weathering)

The county-scale atmospheric P deposition magnitudes, *DEP(i,t)* were normalized by county area, *AREA(i,t)*:

$$DEP_{norm}(i,t) = \frac{DEP(i,t)}{AREA_{cnty}(i,t)} \quad \text{(Eq. 3.16)}$$

For the all-land fluxes, all grid cells of our 250-m gridded map of the U.S. were assigned 1 values, $gLU_{all}(k,t) = 1$, where *k* refers to the individual grid cell and *t* to year. The normalized deposition and weathering flux values were then multiplied by the 1 values in the land-use grid.

Finally, for the rock P weathering, we simply rescaled the raster to a 250m grid and assume the weathering rates are constant from 1930-2017.

3.2.2.4 Gridded phosphorus surplus

The annual, grid-scale estimates of the P surplus, gPS, were calculated by raster summation of the gridded components of the P mass balance. Here, we define the P surplus as a sum of P inputs (fertilizer P, atmospheric P deposition, P rock weathering, manure P, human waste P, food waste P, and detergent P) and non-hydrologic P outputs (crop and pasture P uptake) (Byrnes et al., 2020):

$$gDOM(k,t) = gHUM(k,t) + gFOOD\ (k,t) + gDET(k,t) \quad \text{(Eq. 3.17a)}$$

$$gPS(k,t) = gFERT(k,t) + gDEP(k,t) + gMAN(k,t) + gDOM(k,t) - gCROP(k,t) - gPAST(k,t) \quad \text{(Eq. 3.17b)}$$

where gFERT is the sum of farm and non-farm P fertilizer, gDEP is P deposition, gMAN is P in livestock manure, gCROP is P uptake in crops, gPAST is P uptake by pasture grasses consumed by livestock, and gDOM is the P from P in human waste (gHUM), P in food waste (gFOOD), P in laundry and dish detergents (gDET) all in units of kg ha^{-1} y^{-1}.

3.3 Data Validation

3.3.1 Validation of gTREND-Phosphorus with county-scale data

The gTREND-Phosphorus dataset was created from disaggregating county-scale TREND-Phosphorus components to land use. To ensure the conservation of mass between estimates of

annual P fluxes, we re-aggregated the gridded data to the origins county-scale to compare (**Figure 3.2)**. We find a strong correlation between the gridded and county-scale datasets for agricultural components, deposition, and weathering (R2 > 0.99; **Figure 3.2 A-F**). In population-based components, we find a difference because of the different underlying population data used in the county and gridded datasets. However, the two datasets are still highly correlated (R2 > 0.97; **Figure 3.2 G-J**).

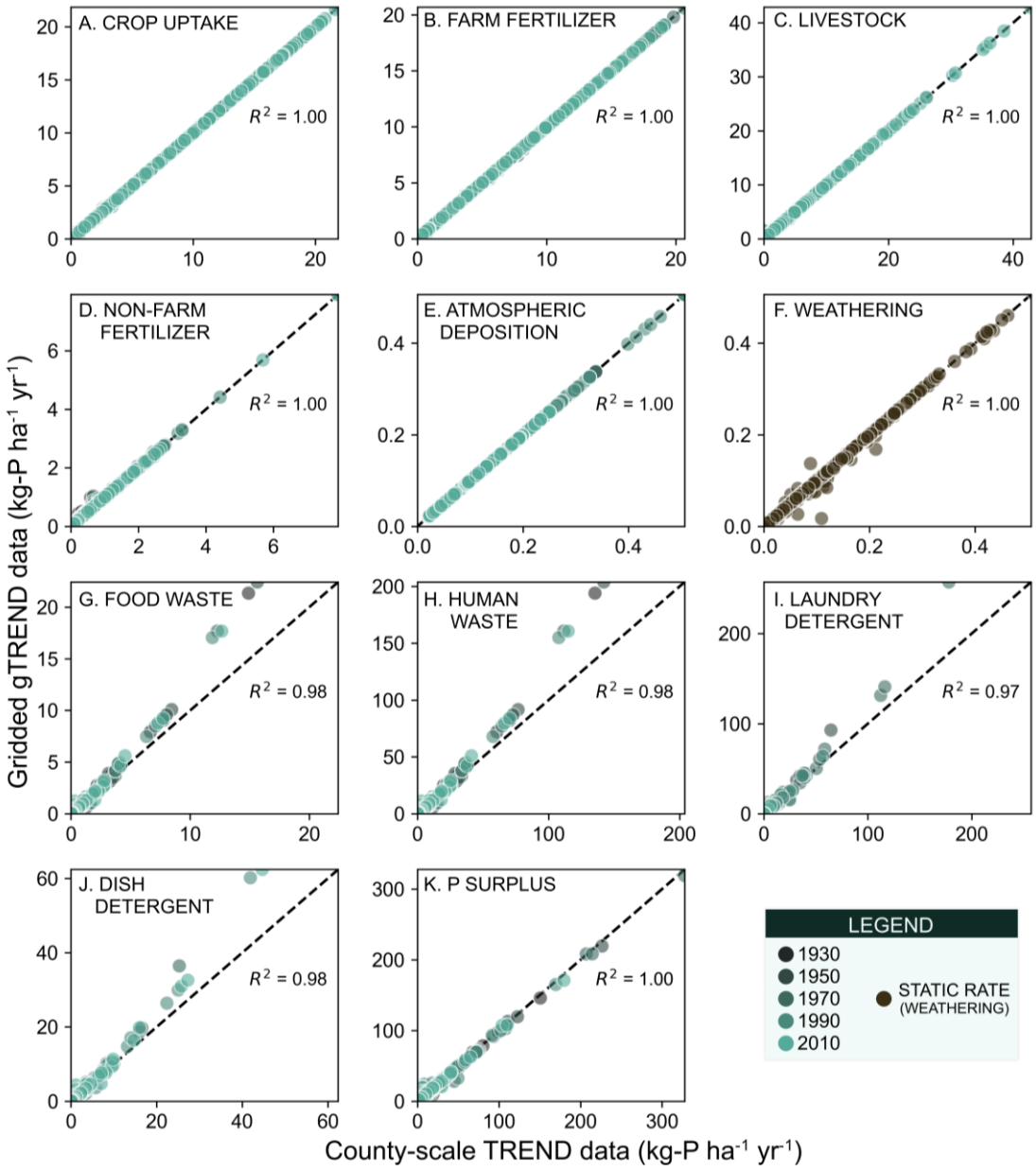


Figure 3.2 Relationship between county-scale TREND-Phosphorus estimates of P mass balance components and the gridded gTREND-Phosphorus estimates, re-aggregated to the county scale. Data points represent estimates for each county every 20-years 1930 to 2010. The dashed lines represent a 1:1 relationship.

3.5 Limitations

The TREND-Phosphorus datasets have limitations regarding the scope of sources considered. While we account for major anthropogenic inputs (synthetic fertilizer, manure, human waste, detergents, weathering, and atmospheric deposition), several other potential phosphorus sources and fates were not included. Specifically, phosphorus from pesticide and herbicide applications, removal through wastewater treatment, and riverine export were excluded from the calculations. While pesticides and herbicides can be regionally important, at the national scale they represent less than 1% of the total inputs considered (Sabo et al., 2019). Furthermore, reporting of pesticide and herbicide use at the county scale is limited before 1992 (Baker and Stone, 2015; Thelin and Stone, 2013). Similarly, riverine phosphorus export at the county or gridded scale is highly uncertain because export is frequently modeled at the watershed level. These data limitations have led us to exclude these sources and sinks from the surplus calculations to be later estimated through modeling exercises.

3.5 Data Records

3.4.1 TREND-Phosphorus dataset

The TREND-Phosphorus dataset covers the contiguous U.S. at the county scale and includes 20 P mass balance components for the years 1930-2017. The current version of the TREND-Phosphorus dataset will be available on Figshare upon publication. Each P mass balance component is in a comma-separated text file (.txt) with columns containing the county IDs, area, and the component for a given year. The GEOIDs are geographical identifiers corresponding to the 2017 county-boundary shapefile, available in the zipped file repository. A detailed description of the file structure is provided in **Table 3.2**. The repository also contains a zipped folder containing the county boundaries associated with the final product.

Table 0.1 Structure and File Names of the TREND-P Dataset in the Data Repository.

Component Name	Text file name
Atmospheric P Deposition	Atmospheric_Deposition.txt
Crop Uptake - Cropland	CropUptake_Cropland.txt
Crop Uptake - Pasture	CropUptake_Pasture.txt
Farm Fertilizer	Fertilizer_Agriculture.txt
Non-farm Fertilizer	Fertilizer_NonAgriculture.txt
Livestock Waste P - Milk Cows	Lvst_DairyCow.txt
Livestock Waste P - Beef Cows	Lvst_BeefCow.txt
Livestock Waste P - Other Cattle	Lvst_OtherCattle.txt
Livestock Waste P - Broilers	Lvst_Broilers.txt
Livestock Waste P - Layers and Pullets	Lvst_LayersPullets.txt
Livestock Waste P - Turkeys	Lvst_Turkeys.txt
Livestock Waste P - Hogs	Lvst_Hogs.txt
Livestock Waste P - Equine	Lvst_Equine.txt
Livestock Waste P - Sheeps and Goats	Lvst_SheepGoat.txt
Human Waste P	Human_Waste.txt
Food Waste P	Food_Waste.txt
Dish Detergent P	Dish_Detergent.txt
Laundry Detergent P	Laundry_Detergent.txt
Rock P Weathering	Weathering.txt
County shapefile	County_Boundaries_2017.zip

Each component text file contains 92 columns, with key identifiers and data fields structured as follows:

- StateFIPS: A 2-digit Federal Information Processing Standard (FIPS) code for unique identification of U.S. states.
- CountyFIPS: A 3-digit county Federal Information Processing Standard (FIPS) code for unique identification at the county level.
- GEOID: A 5-digit concatenation of the StateFIPS and CountyFIPS, forming a geographic identifier allowing consistent spatial referencing across datasets.
- AREA (HA): The county's area reported in hectares.
- yYYYY: The remaining columns are designated as "y" followed by the four-digit year (e.g., y1930, y1940), representing the data year for specific components recorded in the file.

3.4.2 gTREND-Phosphorus dataset

The gTREND-Phosphorus dataset covers the contiguous U.S. and includes 22 P mass balance components for the years 1930-2017. The current version of the gTREND-Phosphorus dataset will be available on Figshare upon publication. The gTREND-Phosphorus data is zipped into subfolders and organized around different components of the P mass balance. Data is available in the Tag Image File Format (TIFF) inside the zipped folders. Each TIFF file represents an annual P mass balance component with the unit of kg-P $ha^{-1}y^{-1}$ at 250-m grid scale. The projection for these TIFF files is EPSG:5070 - NAD83 / Conus Albers.

The dataset is organized into multiple zipped subfolders, with 21 GRID_scale subfolders. The zipped GRID_scale folders include one subfolder per P mass balance component. Each zipped folder then contains 88 TIFF files corresponding to the years 1930-2017. A detailed description of the data structure and files provided in the data repository is provided in **Table 3.3**.

Code for the TREND-Phosphorus county-scale data is available upon request and will be available on a later date at github.com/danykakbyrnes/TREND_Nutrients.git.

Table 0.2 Structure and File Names of the Gridded gTREND-P Dataset in the Data Repository. In the file names, the YYYY values refer to the data year.

Zipped Subfolder Name	File name	Zipped Folder Content	
		File Naming Convention	Number of Files
Atmospheric P Deposition	Atmospheric_Deposition.zip	Atmospheric_YYYY.tif	88
Crop and Pasture P Uptake	CropUptake_Ag.zip	CropUptake_Ag_YYYY.tif	88
Crop P Uptake	CropUptake_Cropland.zip	CropUptake_Cropland_YYYY.tif	88
Pasture P Uptake	CropUptake_Pasture.zip	CropUptake_Pasture_YYYY.tif	88
Farm P Fertilizer	Fertilizer_Ag.zip	Fertilizer_Ag_YYYY.tif	88
Non-Farm P Fertilizer	Fertilizer_Dev.zip	Fertilizer_Dev_YYYY.tif	88
Human Waste P	Human.zip	Human_YYYY.tif	88
Food Waste P	Food.zip	Food_YYYY.tif	88
Dish Detergent P	Dish_Detergent.zip	Dish_Detergent_YYYY.tif	88
Laundry Detergent P	Laundry_Detergent.zip	Laundry_Detergent_YYYY.tif	88
Livestock Waste P - Milk Cows	MilkCow.zip	MilkCow_YYYY.tif	88
Livestock Waste P - Beef Cattle	BeefCow.zip	BeefCow_YYYY.tif	88
Livestock Waste P - Other Cattle	OtherCattle.zip	OtherCattle_YYYY.tif	88
Livestock Waste P - Horses	Equine.zip	Equine_YYYY.tif	88
Livestock Waste P - Hogs	Hogs.zip	Hogs_YYYY.tif	88
Livestock Waste P - Layers & Pullets	LayerPullets.zip	LayerPullets_YYYY.tif	88
Livestock Waste P - Broilers	Broilers.zip	Broilers_YYYY.tif	88
Livestock Waste P - Turkeys	Turkeys.zip	Turkeys_YYYY.tif	88
Livestock Waste P - Sheep & Goats	SheepGoat.zip	SheepGoat_YYYY.tif	88
Livestock Waste P - All	Livestock.zip	Livestock_YYYY.tif	88
P Surplus	Surplus_P.zip	Surplus_P_YYYY.tif	88

Chapter 4 Nitrogen Memoryscapes: Decoupling of Nitrogen Inputs and Riverine Export across the Continental United States

Outline

Nitrogen pollution has severely impacted our inland and coastal waters, contributing to algal blooms, eutrophication, and contamination of drinking water. Despite concerted efforts to address nitrogen pollution, we have not seen widespread improvements in water quality. Most research to date aims to understand what drives stream nitrogen export rather than exploring what drives changes in export. In this work, we investigate the drivers of changes in nitrogen export using a multi-decadal (1990-2017) time series of nitrogen inputs and export for over 400 watersheds across the continental U.S. We find 62 watersheds where both nitrogen surplus and export have been significantly increasing over the last two decades. These input-driven watersheds are characterized by high livestock density, agricultural area, and tile drainage. In contrast, nitrogen surplus and export have decreased in 127 *bright spot* watersheds, characterized by high population density and urban land use. Nitrogen surplus also decreases in the 60 *transitioning* watersheds, but export increases as nitrogen surplus decreases. We argue that these watersheds are Transitioning from agriculture to more urban areas, such that even though fertilizer inputs have decreased, legacy nitrogen stores drive the higher nitrogen export. Finally, 35 watersheds demonstrate a *delayed response*, decreasing nitrogen export despite increasing nitrogen surplus. Climate appears to be the driver of response in these watersheds, with aridity likely driving lower nitrogen export despite increasing inputs. When managing nutrients, accounting for each watershed's unique characteristics and histories is crucial in developing effective strategies to achieve desired environmental outcomes.

4.1 Introduction

Over the last 100 years, the global population has increased fourfold, significantly altering the nitrogen cycle. This transformation has been propelled by the intensification of agricultural practices, an increasing reliance on fossil fuels, and widespread changes in land use to meet the demands of a growing population. Consequently, there has been a s-1e Gulf of Mexico(Diaz et al., 2017; Harding et al., 2019; Malone & Newton, 2020; Rabalais, Turner, & Scavia, 2002; U.S. Environmental Protection Agency (USEPA), 2020). Moreover, approximately 1.4 million people are estimated to depend on groundwater sources in regions with nitrogen concentrations that exceed the safe limits for drinking water(Ransom et al., 2021; Ward et al., 2018). Addressing nitrogen pollution in surface and groundwater systems in the face of a growing population and

agricultural intensification is crucial to preserving water and food security.

To manage nitrogen pollution in surface and groundwater systems, it is important to understand the main climate and landscape drivers of these increased concentrations. This question has been studied for the last two decades, starting from some of the earliest works by Howarth et al.(1996), who found a close linear relationship between the net nitrogen inputs to the landscape and stream nitrogen exports. The slope of this relationship varies across watersheds, from 6%-40% across the U.S. and Canada (Boyer et al., 2002; Goyette et al., 2016a; H. Han & Allan, 2008; Sobota et al., 2009; Swaney et al., 2012), 15% in watersheds draining to the Baltic Sea(Hong et al., 2017), 28% in western Europe (Ehrhardt et al., 2021), and 25%-38% in China (Huishu et al., 2018; A. Wang et al., 2016). In more recent work, machine learning models developed to predict stream nitrogen export also found that spatial patterns of nitrogen export across over 2000 streams in the U.S. are driven primarily by nitrogen inputs, while precipitation and soil type act as secondary controls (Sadayappan et al., 2022).

The recognition that agricultural inputs are largely responsible for nitrogen pollution has contributed to various legislation in different countries to limit fertilizer use and increase nitrogen use efficiency to improve water quality (Basu et al., 2022; R. Howarth et al., 2012a). Concerted efforts have been made to improve nitrogen use efficiency through decreasing fertilizer use, increasing crop yields, and mitigating atmospheric deposition, which has led to widespread decreases in net nitrogen inputs. For example, from 1990 to 2018, the European Union saw a reduction of 11.8 kg-N ha^{-1} yr^{-1}, or 29%, in net nitrogen inputs following the implementation of the Nitrate Directive (Batool et al., 2022; Bouraoui & Grizzetti, 2011). Similarly, net nitrogen inputs have also decreased in some U.S. regions, with the most prominent reductions in the eastern U.S., ranging between 2.5 to 50 kg-N ha^{-1} yr^{-1} over ten years (Byrnes et al., 2020; Sabo et al., 2019). In China, while the country still has increasing net nitrogen inputs (Y. Han et al., 2020), a few areas have had decreasing net nitrogen inputs in recent years (Pei et al., 2022; Wu et al., 2022; Zhao et al., 2022). For example, the net nitrogen inputs in the north central region of the Yellow Basin have plateaued at 80 kg-N ha^{-1} yr^{-1} after a 2004 peak (Pei et al., 2022).

However, in many areas, changes in net nitrogen inputs have not resulted in proportionate improvements in water quality. For instance, in China's Songhuajiang River Basin and Raohe River Basin, nitrogen inputs decreased by 3 kg ha^{-1} yr^{-1} and 7 kg ha^{-1} yr^{-1}, respectively, over the last decade, yet nitrogen exports continued to increase in both watersheds, with the Raohe River Basin having an increase of 17 kg ha^{-1} yr^{-1} (Gao et al., 2023; Zhong et al., 2022). In the

Susquehanna watershed in the U.S., inputs plateaued around 40 kg-N ha^{-1} yr^{-1} from the 1970s to the mid-2000s. Despite the plateau of inputs, loads continued to increase from 5 to nearly 8 kg-N ha^{-1} yr^{-1}(Van Meter et al., 2017).

This nonlinearity is described as a hysteresis effect, meaning that watershed nitrogen loads are not just a function of current-year nitrogen inputs but can also depend on the history of inputs to the watershed (Basu et al., 2022; Chang et al., 2021; Dupas et al., 2018; Van Meter, Byrnes, et al., 2023; Van Meter & Basu, 2017). Van Meter et al.(2023) explored this effect across the U.S., finding that one-third of studied watersheds exhibited clear hysteresis, with loads lagging behind input changes. These findings highlight that while nitrogen inputs drive spatial patterns in nitrogen exports, a decrease in inputs might not contribute to a concomitant decrease in exports. In other words, what drives nitrogen loads may differ from what drives changes in nitrogen loads.

Instead, legacy nutrients built up over decades of excess fertilizer application may play an important role in controlling how a watershed responds to changing nitrogen inputs (Basu et al., 2022; Van Meter, Byrnes, et al., 2023). Legacy nutrients can accumulate in soils as soil organic nitrogen, also referred to as biogeochemical legacy, or dissolved in pore water and groundwater as hydrologic legacy (Van Meter & Basu, 2015). The challenge is that these legacy stores' magnitude and release rates into the stream are largely unknown, complicating predictions of watershed responses to changing input patterns.

The nonlinear relationship between nitrogen inputs and export has been quantified in some studies using the metric lag time – the time lag between the changes in management leading to decreases in the net nitrogen inputs and a measurable decrease in stream export. Such time lags have been quantified across watersheds globally. For instance, in a study of 238 Western European watersheds, lag times varied between 0.2 to 34 years, with shorter lag times associated with wetter catchments (Ehrhardt et al., 2021). A study of 16 catchments in Canada had lag times vary from 12 to 34 years, with lag times controlled by nutrient source and distribution, watershed topography, soil type, climate, tile-drainage densities, and groundwater travel time (Van Meter & Basu, 2017). We also see lag times spanning approximately ten years in seven catchments in France (Dupas et al., 2018), ranging from 7 to 22 years in three nested catchments in Germany (Ehrhardt et al., 2019), and from 1 to 12 years in 77 catchments in New Zealand (McDowell et al., 2021).

Modeling studies have quantified the mismatch between decreased inputs and reduced exports by simulating various input reduction scenarios to estimate the time required to reach load

targets.(Basu et al., 2022; Guillaumot et al., 2021; Ilampooranan et al., 2019; Lee et al., 2016; Nguyen et al., 2022; L. Wang et al., 2012). For instance, using the ELEMeNT model within the Mississippi River basin led to the estimates of over a 30-year time lag to achieve a 60% decrease in load (Van Meter et al., 2018). Similarly, curtailing inputs into the tributary watersheds of Chesapeake Bay in the U.S. and the Grand River watershed in Canada would take up to 17 and 24 years, respectively, to achieve a 25% reduction in nitrogen loads (Chang et al., 2021; Liu et al., 2021).

Of course, all these studies focus primarily on answering the question, when nitrogen inputs decrease, how long does it take for nitrogen exports to decrease? But what about scenarios where nitrogen inputs are increasing, or nitrogen inputs are decreasing while nitrogen exports are increasing? To improve our predictions of watershed responses to changing input patterns and better manage water quality, we need an improved understanding of the different types of watershed nitrogen input and export relationships and their drivers. To address these issues, we analyze nitrogen dynamics across 492 watersheds across the U.S. using nitrogen surplus and riverine load. We use surplus, which is a difference between inputs and non-hydrological export (crop N removal), effectively measuring the nitrogen available to pollute after accounting for crop uptake. By examining both watershed inputs and exports, we investigate the relationship between changes in nitrogen inputs and exports through four key questions: **(1) What drives the spatial patterns in the relationship between landscape net nitrogen input and stream nitrogen export? (2) How have nitrogen inputs and nitrogen exports changed over the last 30 years across the U.S.? (3) What is the relationship between changes in inputs and changes in export? (4) How do climate, land use, and land management drive this relationship?**

4.2 Methods

4.2.1 Data sources and site selection for streamflow and nitrogen concentrations

We analyzed streamflow and nitrate concentration data between 1990 and 2017 across the contiguous U.S. Streamflow data was obtained from the USGS National Water Information System (NWIS), while concentration data was obtained from two different sources for different time periods: (1) 1965-2012: Oelsner et al. (2017), and (2) 2013-2017: National Water Information system and Water Quality Portal. The Oelsner et al. (2017) dataset was compiled as a part of the United States Geological Survey's National Water-Quality Assessment Project, which included colocating streamgages and water quality monitoring stations. However, the dataset ended in 2012, which required us to acquire more recent data for our analysis. We used the watershed IDs compiled by Oelsner et al. (2017) to retrieve data from federal, state, tribal, regional, and local government agencies, and nongovernmental organizations through the dataRetrieval and EGRET package in R (DeCicco et al., 2021; R. Hirsch et al., 2021). We then used a set of criteria based on Oelsner et al (2017) to exclude watersheds with incomplete data: (1) at least 70% of the years had quarterly samples; (2) at least 10% of the samples were taken during high flows; (3) gaps between sampling days were less than 4 years; (4) no more than six consecutive days with missing streamflow per year; (5) no more than 30 days with zero streamflow per year; and (6) at least 10 complete years of data were available, (7) first record began in 2003 or earlier. Some stations that had data discontinuities were included, as long as there was at least a 10-year period that met the above criteria by removing the timeframes with problematic. These criteria led to a selection of 512 stations used for further analysis.

4.2.1.1 Estimation of Daily Loads

We modeled daily watershed loads for 512 stations using Weighted Regressions on Time, Discharge, and Season (WRTDS). WRTDS is a regression model particularly suited to sites with sparse concentration measurements and daily flow data. The model estimates daily concentration from sparse concentration measurements and daily flow data using time, flow, and season as explanatory variables (R. M. Hirsch & De Cicco, 2015; Oelsner, Sprague, Murphy, Zuellig, Johnson, Ryberg, Falcone, Stets, Vecchia, Riskin, De Cicco, et al., 2017). Prior to running WRTDS, the concentration and flow data were cleaned. Specifically, when multiple records were available for a given day, we averaged the multiple records to get a single concentration. We also removed outlier concentrations that were six standard deviations from mean in log space. For stations with discontinuous flow data of 6 days or less, we used linear interpolation between the nearest days to estimate missing flow data. After using WRTDS, the quality of the model's fit was assessed using a flux bias metric. We excluded any stations where the flux bias exceeded ±30%, resulting in a total of 492 watersheds in our dataset. Annual loads were calculated by summing daily loads.

4.2.1.2 Watershed nitrogen inputs and outputs

We used the Gridded TRajEctories Nutrient Dataset for Nitrogen (gTREND-N), which is a 88-year 250m gridded dataset of annual nitrogen inputs and outputs Chang and Byrnes et al. (*in prep*). The dataset includes the following annual fluxes: atmospheric nitrogen deposition (kg $ha^{-1}y^{-1}$), inorganic nitrogen fertilizer (kg $ha^{-1}y^{-1}$), biological nitrogen fixation (kg $ha^{-1}y^{-1}$), nitrogen in livestock manure (kg $ha^{-1}y^{-1}$), crop and pasture nitrogen uptake (kg $ha^{-1}y^{-1}$), and nitrogen in human waste (kg $ha^{-1}y^{-1}$). Annual nitrogen surplus is calculated as the difference between anthropogenic nitrogen inputs and non-hydrological outputs (crop and pastureland nitrogen uptake). The gridded inputs were aggregated to the watershed scale by taking the mean value for each watershed zonal statistics from the the *rasterstats* package in Python (Perry, 2023).

4.2.1.3 Spatial and temporal statistical analysis

To explore the spatial patterns of the relationships between nitrogen surplus and exports in individual watersheds, we used the export ratio (E_r) metric. This ratio represents a fraction of a watershed's nitrogen surplus loss through riverine export. We averaged surplus and load data from the 2003-2017 or up to the end of each time series, whichever was more recent. The aim was to use this ratio along with watershed properties encompassing aspects of land use, climate,

geomorphology, and land management practices to identify the dominant controls on these surplus-export relationships.

We obtained climate, land use and land management data for all of the study watersheds, and used these attributes to identify what drives spatial and temporal trends in surplus and export. Data regarding watershed area, land use, wetland density, soil type were obtained from Falcone (2017). The aridity index (AI = Evapotranspiration/Precipitation) data was obtained from CGIAR Global Aridity Index dataset averaged over 1970-2000 (Zomer & Trabucco, 2022).

To determine the extent of the watersheds' agricultural land that is tile-drained, we performed a two-step process. First, we used a 30-meter resolution dataset of tile-drained areas to calculate the percent of the watershed that is tile-drained (Valayamkunnath et al., 2020). Next, divide the percent of the watershed that is tile-drained by agricultural land to determine the percent of agricultural land that is tile-drained, referred to as "percent tile-drained area".

To explore temporal patterns of the relationships between nitrogen surplus and exports in individual watersheds, we estimated trends in nitrogen surplus and export trajectories between 1990 and 2017, or the available time series between periods. We used the Mann-Kendall non-parametric test for monotonic trend (Helsel et al., 2020). We implemented a variance correction to address the auto-correlation found in 36% of export and 86% of nitrogen surplus time series, to improve the accuracy of the Mann-Kendall test (Millard, 2013; Patakamuri & O'Brien, 2021). The variance correction method is developed by Yue and Wang (2004), first removes any trend from the data, and then it estimates the serial correlation among the detrended data points. Afterward, it employs the concept of effective sample size to adjust for the impact of this estimated serial correlation on statistical tests or analyses. We then used the Sen-Theil slope to quantify the rate of change (Patakamuri & O'Brien, 2021). We normalized the slope by the mean of the time series to express them as annual percentage change. We then categorized each watershed into four primary export response patterns, based on the positive or negative nature of the percent change in surplus and export.

We used watershed characteristics and compared them across the four primary export response patterns to explore the drivers of each pattern. The characteristics we used are population density, land use, percent tile-drained area, aridity index, as well as nitrogen input magnitudes and export. Both 2010 population density and 2011 NLCD land use were compiled by Falcone (2017). We used three land use categories in our temporal analysis, including agricultural land use, urban land use, and newly generated natural land use. Natural land use is an aggregation of

existing land use categories: wetland, open water, forested, shrubland, grassland, and barren land use. Finally, watersheds' percent tile-drained area and aridity index are processed as previously discussed above.

In addition to these attributes, we used the export ratio (E_R), as defined above, as well as average total agricultural inputs, manure input to total fertilizer (inorganic fertilizer and manure) ratio, and nitrogen loads. All these characteristics were averaged from 2003 to 2017 or up to the end of each export time series, whichever was more recent. We also used a 10-year cumulative net agricultural inputs (NAI) in 2017, or at the end of the export time series, to capture the upper bound of legacy nitrogen stored in the landscape. Lastly, we calculated annual percent change of manure inputs, inorganic fertilizer and fixation inputs, and crop uptake using the Mann-Kendall trend test approach described above. Further details on this metric can be found in **Supplemental Table 4.1**. Finally, we used the Wilcoxon rank-sum test to determine whether the distributions of watershed attributes across different quadrants are significantly different (Helsel et al., 2020).

4.3 Results and discussion

4.3.1 Spatial patterns in nitrogen surplus and nitrogen export

Nitrogen surplus and export magnitudes vary across the 492 watersheds, with export ranging from 0.009 to 40.4 kg ha^{-1} y^{-1} and nitrogen surplus ranging from 0.77 to 191 kg ha^{-1} y^{-1}. Spatial variability in nitrogen surplus explained 34% of the variability in export (R^2 = 0.34; **Supplemental Figure 4.1**), and an average of 15% of the nitrogen surplus exported from the watershed through the river network. These findings are consistent with previous studies that have found watersheds export between 6% and 40% of their nitrogen surplus (Boyer et al., 2002; Goyette et al., 2016a; Sobota et al., 2009; Swaney et al., 2012).

The watershed export ratio (E_R = nitrogen export/nitrogen surplus) varied across the CONUS (median = 0.09, IQR 0.04 - 0.21) as a function of climate, landscape and management controls, with the highest export ratios clustered in the Midwestern U.S. (**Figure 4.1a**). Export ratio increased with increasing tile-drained area (R^2 = 0.2; $p < 0.001$), and decreasing wetland coverage (**Figure 4.1b and 1c**). While higher tile density is known to increase nitrogen export through subsurface pathways (David et al., 2010), it is often challenging to disentangle the effects of tile drainage and nitrogen inputs, given tile-drained landscapes are also under intensive agriculture with high nitrogen inputs. Our results highlight that even in watersheds with similar magnitudes of nitrogen inputs, the presence of tile drains increases nitrogen export. The negative

relationship between export ratios and wetland density can be attributed to the nutrient retention functions of wetlands (**Figure 4.1c**). The export ratio is lower in more arid watersheds, and this is likely because the increase in the aridity index decreases streamflow that leads to a decrease in export ratio (**Figure 4.1d**). Finally, soil type has a counterintuitive effect on the ratio, with increase in the percent silt and clay content in soils leading to increase in the export ratio (**Figure 4.1e**). This is likely because watersheds with high clay and silt content are also more likely to have subsurface drainage, reducing the nitrogen removal processes and increasing the export ratio (Van Meter & Basu, 2017).

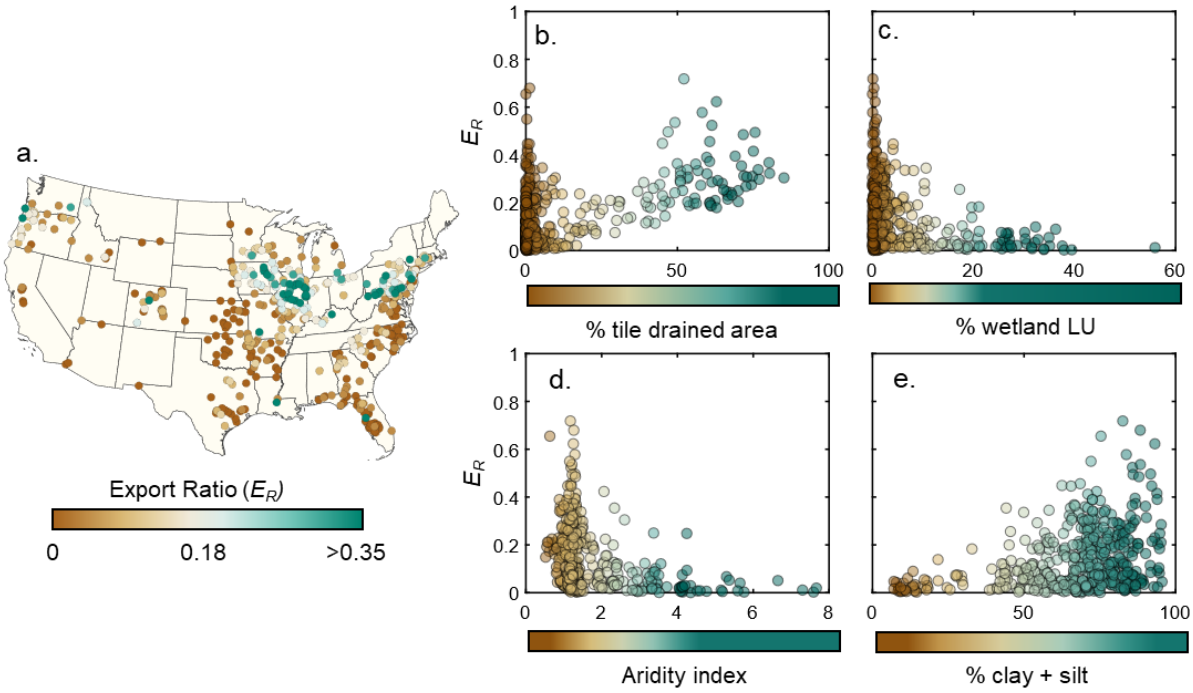


Figure 4.1. Nitrogen export ratio (E_R = Export/Surplus) as a function of climate, land use and management controls (2003-2017). (a) Export ratios vary in space but (b) increase with an increase in the percent tile-drained area, (c) decrease with an increase in the percent wetland area, (d) decrease with an increase in the aridity index, and (e) increase with an increase in percent clay and silt content.

4.3.2 Temporal trends in nitrogen surplus and nitrogen export

Exploring trends over the last 30 years (1990 - 2017), we find nitrogen surplus to be significantly decreasing ($p < 0.1$) in 54% of the watersheds (median rate: -1.29% per year; range: -8.87 to -0.07% per year) and increasing ($p < 0.1$) in 30% of the watersheds (median rate: 1.27% per year; range: 0.08 to 6.81% per year), while there was no significant change in 16% of the watersheds

(**Figure 4.2a-d**). Watersheds with increasing nitrogen surplus predominantly drain the interior U.S., including the Mississippi River Basin and Missouri River Basin, and a few western watersheds, including the Columbia River Basin. On the other hand, watersheds with decreasing nitrogen surplus are concentrated in the Eastern U.S., Florida, and Eastern Mississippi River Basin tributaries.

Nitrogen export trends over the same time frame (1990-2017) follow similar patterns but with fewer decreasing and more insignificant trends. Specifically, we find decreasing ($p < 0.1$) trends in 39% of the watersheds (median rate: -1.86% per year; range: -13.0 to -0.25% per year) and increasing trends ($p < 0.1$) in 29% of watersheds (median rate = 1.89% per year; range: 0.27 to 8.11% per year), while 31% of the watersheds exhibited no significant trend (**Figure 4.2e-h**). Watersheds with increasing export are in Florida, the Mississippi River basin and the Missouri River Basin, while watersheds with decreasing export are concentrated in the Eastern U.S. and watersheds in Oklahoma and Kansas. The greater proportion of decreasing trends in nitrogen surplus compared to decrease in nitrogen export highlights the role of the watershed as a filter that can store past inputs as legacy nitrogen and release them later, thus decoupling the trends in inputs to stream responses (Basu et al., 2011). Trends in nitrogen surplus are a function of land use trajectories that determine nitrogen inputs and thus are driven by broader scale socio-economic and environmental factors, such as increased fertilizer applied to corn for ethanol production from the Energy Policy Act (Donner & Kucharik, 2008; Lark et al., 2022), reduced atmospheric deposition arising from the Clean Air Act (Byrnes et al., 2020; Lloret & Valiela, 2016), increased concentration of livestock operations in smaller areas (Byrnes et al., 2020a; Glibert, 2020), or changing crop varieties and nitrogen use efficiency (Swaney et al., 2018b). Trends in nitrogen export, on the other hand, are a function of both current and past nitrogen surplus trends, and landscape geomorphology and climate attributes that determine how nitrogen is being processed and transported through the landscape. This buffering effect of the landscape contributes to a larger proportion of the watersheds showing no significant trend in nitrogen exports, despite a decreasing or increasing trend in nitrogen surplus.

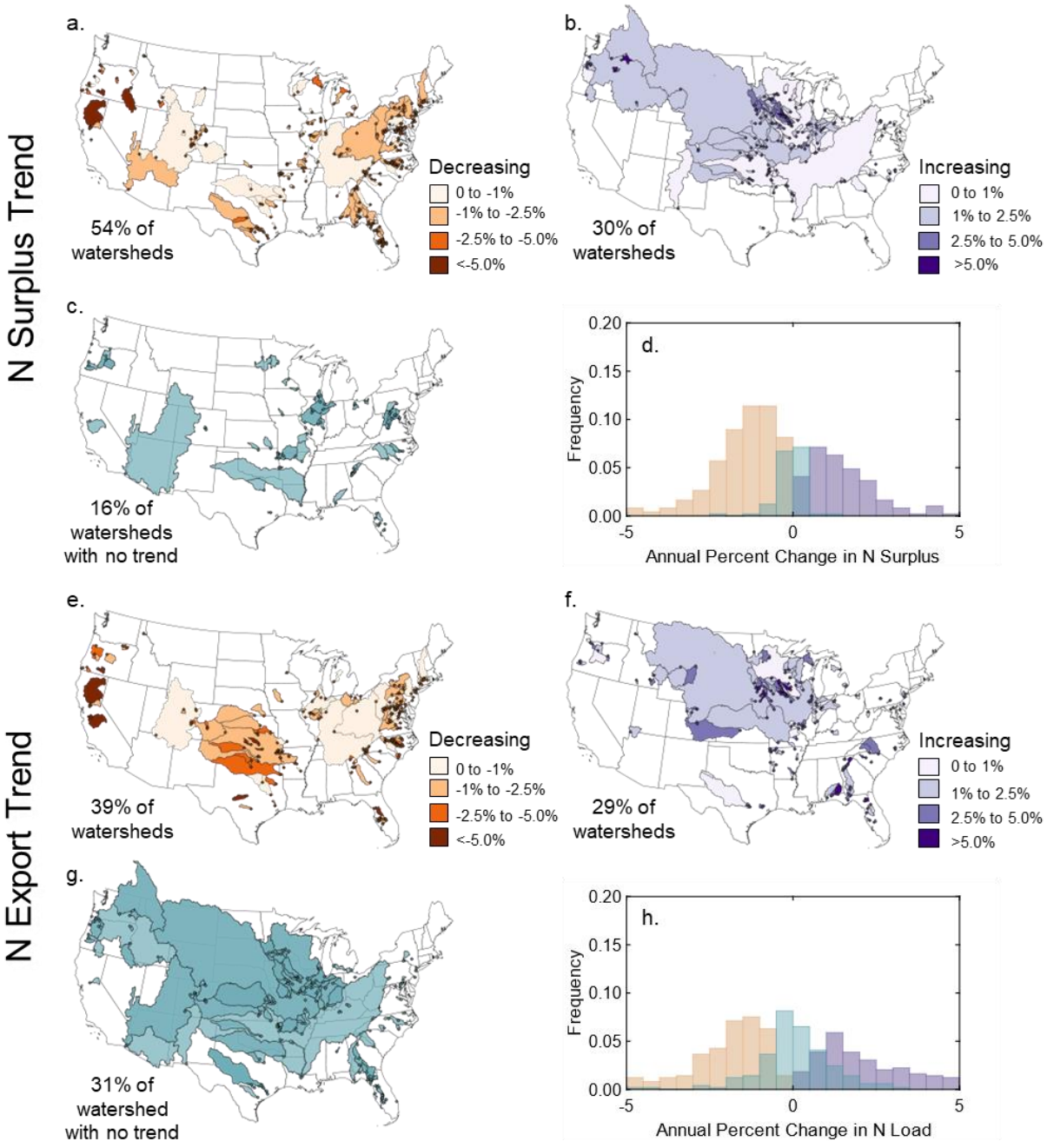


Figure 4.2 Watershed trends in nitrogen surplus (a-d) and export (e-h) over the last three decades (1990 - 2017). (a-c) watersheds with decreasing, increasing ($p < 0.1$), and no significant trend ($p > 0.1$) in nitrogen surplus; (d) histogram of the annual percent change in nitrogen surplus; (e-g) watersheds with decreasing, increasing ($p < 0.1$), and no significant trend ($p > 0.1$) in nitrogen export; (h) histogram of the annual percent change in export. We find that watersheds

exhibit a decoupling of nitrogen surplus and export trends in both direction and magnitude, suggesting the importance of other landscape and climate drivers in changes in nitrogen export.

4.3.3 Typologies of Watershed Responses to Nitrogen Dynamics: Decoupling and Landscape Drivers

We further explored this possible decoupling by grouping watersheds according to their trends in surplus and export. We found four typologies of export response: (1) Input-driven (Q1; 107 watersheds, with 62 having significant trends in both surplus and export), where both nitrogen surplus and export is increasing; (2) Transitioning (Q2; 116 watersheds, 60 significant), where surplus is increasing while export is decreasing; (3) Bright spots (Q3; 189 watersheds, 127 significant) where both nitrogen surplus and export are decreasing; (4) Delayed-response quadrant (Q4; 80 watersheds, 35 significant) where an increase in nitrogen surplus is paired with a decrease in export (**Figure 4.3a**).

Of the 62 Input-driven watersheds with significantly increasing trends in surplus and export (Q1), 45 are found in the Midwestern states (**Figure 4.3c**). Input-driven watersheds are characterized by the highest agricultural land use (median: 68.7%), agricultural inputs (127.9 kg-N ha^{-1} yr^{-1}), tile-drained area (30% of agricultural land), nitrogen export and stream nitrogen export (9.8 kg-N ha^{-1} yr^{-1}) (**Figure 4.4a and 4e**). Thus, the primary drivers of high and increasing nitrogen export in these Input-driven watersheds are high agricultural inputs as well as increased hydrologic connectivity in the landscape from tile drainage. Despite the significant environmental challenges posed by these watersheds, we consider them "Opportunity Spots" where focusing on conservation practices can yield substantial water quality benefits. Landscape-scale nutrient management, such as reducing fertilizer application or improving livestock management practices, will allow them to transition from increasing nitrogen surplus systems to decreasing nitrogen surplus systems, while the greater connectivity can potentially contribute to shorter lag times and faster water quality improvement.

In contrast to the Input-Driven watersheds, nitrogen surplus has been decreasing in the 60 Transitioning watersheds (Q2); however export still has an increasing trend. We find these watersheds to be spread across the U.S., with some clustering in the southeast (**Figure 4.3b-c**). These watersheds are characterized by the highest natural land use (71%) compared to any other quadrant. Agricultural land use (13.6%; $p < 0.001$), inputs (21.1 kg-N ha^{-1} yr^{-1}; $p < 0.001$), and exports (1.14 kg-N ha^{-1} yr^{-1}; $p < 0.001$) in these watersheds are significantly lower than those in the Q1 quadrant (**Figure 4.4b and 4.4e; Supplemental Table 4.2**). A lower intensity agricultural land use and corresponding inputs in these watersheds likely contribute to them

being in Q2 with nitrogen surplus decreasing. However, the decrease in nitrogen surplus in these watersheds has not yet manifested in measurable water quality improvement. We hypothesize that these watersheds are likely driven by legacy nitrogen that is contributing to increased nitrogen export, and thus water quality might improve in the future as they *transition* into Q3.

The 127 Bright Spot watersheds (Q3) are characterized by decreasing surplus and export and comprise our dataset's largest group. This quadrant also has the highest proportion of watersheds with significant nitrogen surplus and export trends, highlighting that statistically significant water quality improvements are being achieved in these regions (Figure 4.3b). A large proportion of these watersheds occur in the Northeastern U.S. and Illinois (Figure 4.3c) and are characterized by high natural land use (66.2%), population density (48.1 pp km-2), and urban land use (8.2%) (Figure 4.4c and 4e). Watershed attributes are not significantly different ($p > 0.1$) between Q2 and Q3 quadrants, except for population density which is significantly greater in the Q3 watersheds compared to Q2 ($p < 0.1$). The observed decreases in watershed inputs stem from both decreasing agricultural inputs and atmospheric nitrogen deposition. The Clean Air Act has been effective in reducing atmospheric deposition, with the most pronounced effects observed in the eastern United States (Byrnes et al., 2020; Lloret & Valiela, 2016). Decreasing loads in these Bright Spot watersheds from decreasing input are in many cases accompanied by advanced wastewater treatment processes, which remove nitrogen before river discharge (Stets et al., 2020). These Bright Spot watersheds thus demonstrate the combined effectiveness of the Clean Air Act and Clean Water Act in improving water quality (Chang et al., 2021; Harding et al., 2019; Lefcheck et al., 2018; Stets et al., 2020).

Finally, the 35 Delayed Response watersheds (Q4), where nitrogen export is decreasing despite increase in nitrogen surplus, are clustered in Oklahoma, Kansas, and the Southeastern U.S. (**Figure 4.3c**). These watersheds are more arid, with the aridity index (2.0; $p < 0.001$) significantly greater than those in the other three quadrants (**Figure 4.4d and 4e**). The higher aridity index might also be contributing to the significantly lower export ratio in these watersheds, given higher evapotranspiration and smaller lateral flows exporting nitrogen from the field to the stream. Agricultural inputs in these watersheds are moderately high (58.5 kg-N ha^{-1} yr^{-1}), significantly higher than Q2 and Q3 ($p < 0.001$) (**Figure 4.4a**). Tile drainage fraction in Q4 is also lower than that in all other quadrants (0.4%), possibly also contributing to the lower export. It is difficult to predict the possible water quality futures of these watersheds: delayed nitrogen transport might mean greater denitrification; however it can also mean that the increasing nitrogen surplus is still traveling through the landscape, and a possible future can be

an increase in nitrogen export over time.

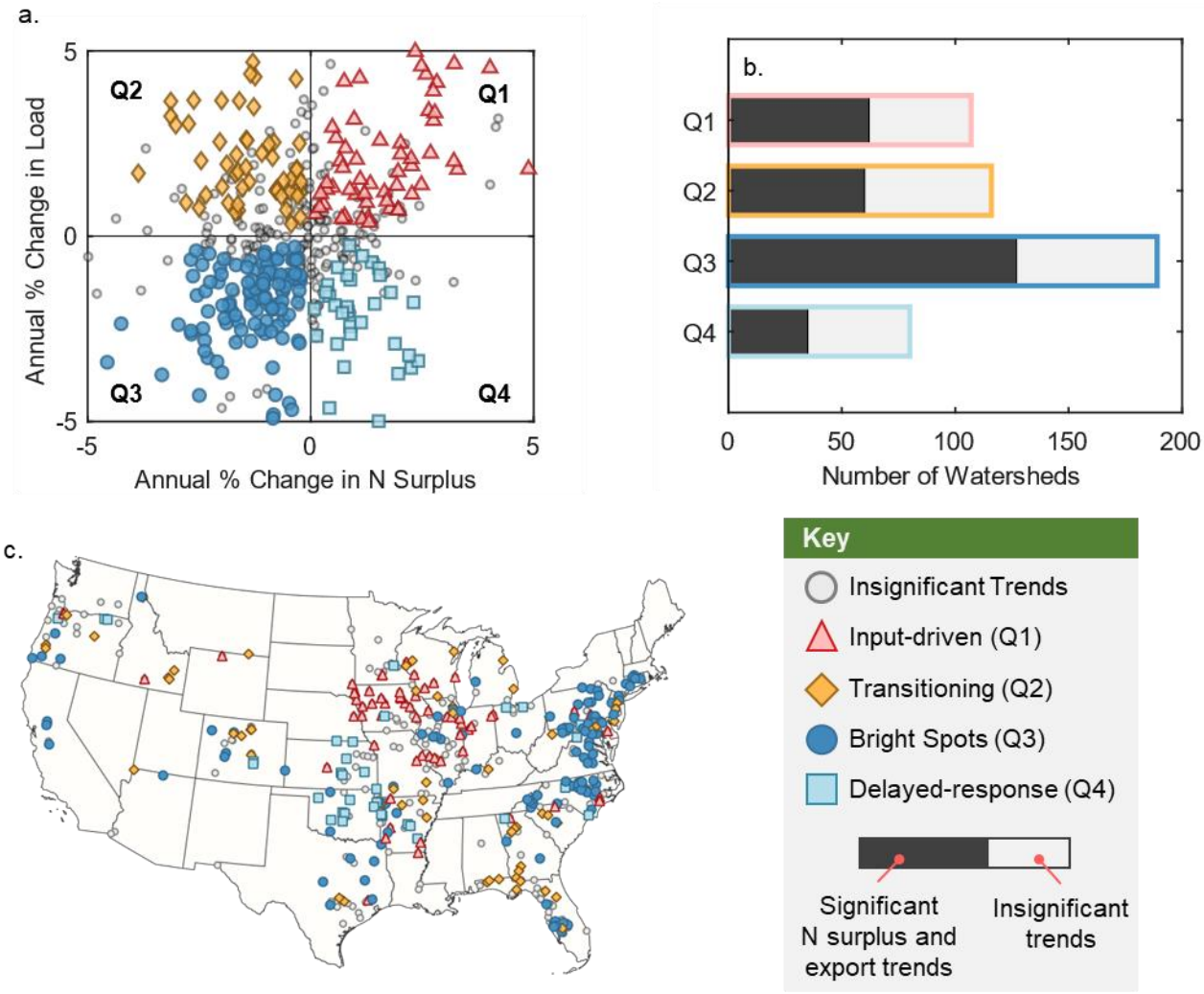


Figure 4.3 Relationship between annual percent change of nitrogen surplus and export. (a) Quadrant plot demonstrating four typologies of watershed response: Input-driven quadrant (Q1) with increases in both nitrogen surplus and export; Transitioning quadrant (Q2) with decreases in nitrogen surplus and increases in export; Bright spots quadrant (Q3) with decreases in both nitrogen surplus and export; Delayed-response quadrant (Q4) with increases in nitrogen surplus and decreases in export. Watersheds with insignificant trends in grey markers. (b) The distribution of significant and insignificant watersheds in each quadrant ($p < 0.1$) in both nitrogen surplus and export. (c) Spatial distribution of the four quadrants, highlighting the geographical variation in watershed response.

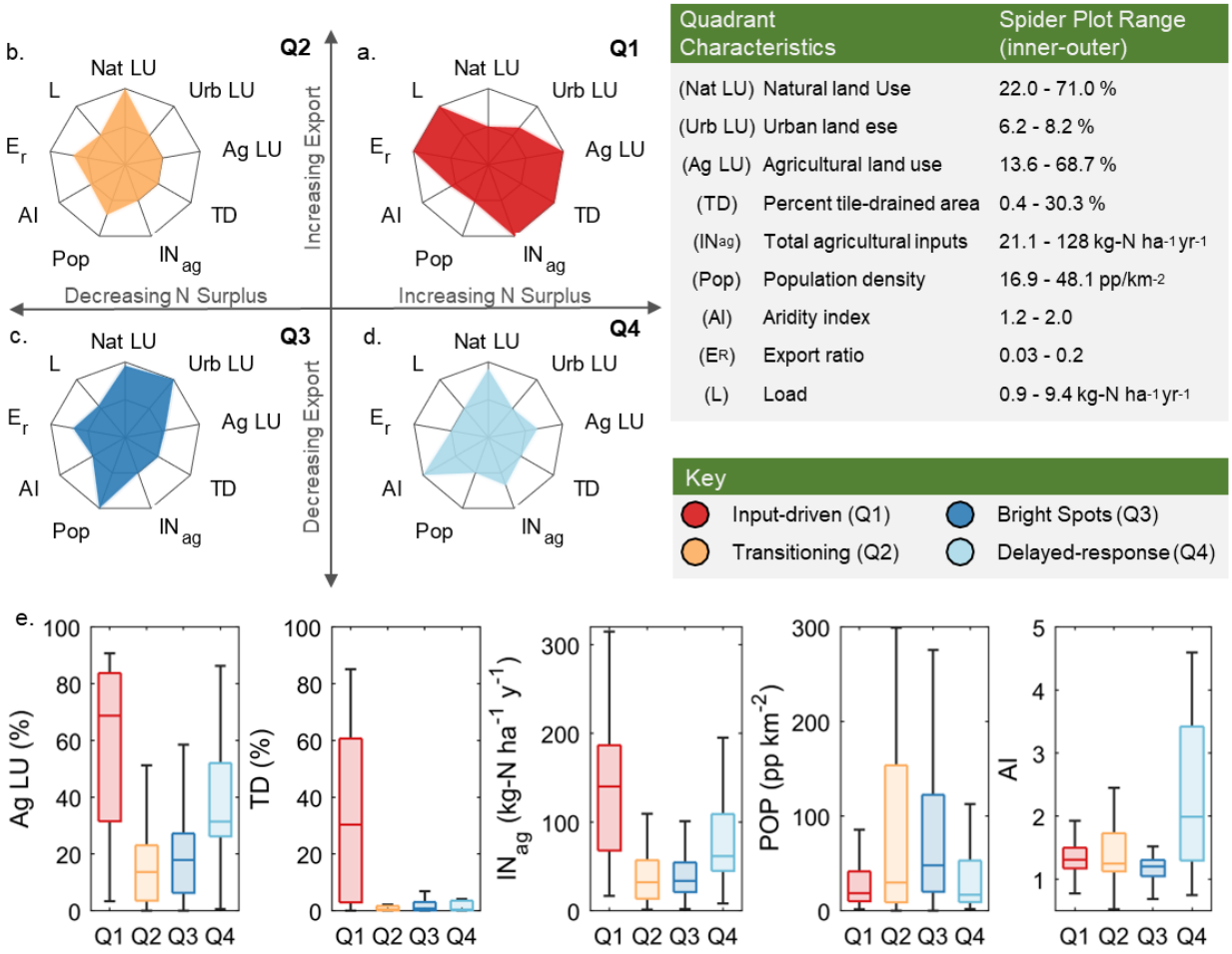


Figure 4.4 (a-d) Spider plot capturing the median characteristics of watersheds in each quadrant and (e) the distribution of select quadrant characteristics. The spider plots display ranges from the inner (minimum) to the outer (maximum) median magnitudes for specific characteristics across the quadrants. We find that Input-driven watersheds (Q1) have high agricultural inputs (IN_{ag}) and increased connectivity from tile drainage (TD). Transitioning watersheds (Q2) have high natural land use (Nat LU) with lower agricultural inputs and land use (IN_{ag} and Ag LU, respectively). Bright spot watersheds (Q3) have high urban land use (Urb LU) and population density (POP). Delayed-response watersheds (Q4) have high aridity index (AI) with moderate agricultural land use and agricultural inputs (Ag LU and IN_{ag}, respectively). All box plots associated with the quadrant characteristics can be found in Supplemental Figure 4.2.

4.3.4 Digging deeper: heterogeneity within quadrants provide insights into nutrient sources

The above analysis highlights some broadscale patterns: nitrogen inputs are decreasing, and water quality is improving in predominantly urban watersheds, while more intensive agricultural watersheds show an increase in both nitrogen inputs and export. However, there is significant heterogeneity in land use and nitrogen inputs within each quadrant (**Figure 4.4b**), and watersheds with very similar attributes can appear in different quadrants.

To explore this further, we focused on 58, predominantly agricultural (> 25% agricultural land use) Midwestern U.S. watersheds (**Figure 4.5a)**. We found that despite having similar nitrogen surplus magnitudes (**Figure 4.5b**), 45 of them belong to the Input-driven quadrant (Q1), four to Transitioning (Q2) and nine to the Bright Spots quadrant (Q3). Given the larger number of watersheds in the Q1 quadrant, it is difficult to attribute drivers effectively. However, some key and interesting trends emerge. The increasing inputs in the Q1 watersheds appear to be driven primarily by increase in manure inputs (1.7% per year; **Figure 4.5c**), while manure inputs are decreasing in both Q2 and Q3 watersheds. Indeed, Iowa, where many of the Input-driven watersheds are located, has the highest number of concentrated animal feeding operations in the country, which have been linked to increased nutrient pollution and degraded water quality (Glibert, 2020; Miralha et al., 2022). The increase in manure inputs in the Input-driven quadrant is especially concerning, given that this creates hotspots of nitrogen sources without availability of crop land for productive recycling of manure as fertilizer (**Figure 4.5c**). Thus the key to management in these regions might lie in managing livestock populations effectively to decrease inputs.

Decreasing inputs, however, is only one part of the puzzle. In Bright Spot watersheds, decreasing inputs have led to decreasing export, while in Transitioning watersheds, decreasing inputs lead to increasing export. We hypothesize that this difference can be attributed to a potential difference in landscape legacy sources and connectivity in these two groups. Specifically, we find that Transitioning watersheds have a higher manure to total fertilizer ratio (M:F), higher net agricultural inputs (NAI_{CUMU}), and lower percent tile-drained area (TD), compared to Bright Spots watersheds (**Figure 4.5b**). Higher manure to fertilizer ratio indicates a potential of build-up of legacy nutrient stores in the landscape, given productive recycling of manure in cropland is not possible, and this is further corroborated by the higher net agricultural inputs. Higher tile density in Q3 watersheds likely increased hydrologic connectivity and contributed to a faster response time – thus, decreasing inputs led to a decrease in export in these watersheds.

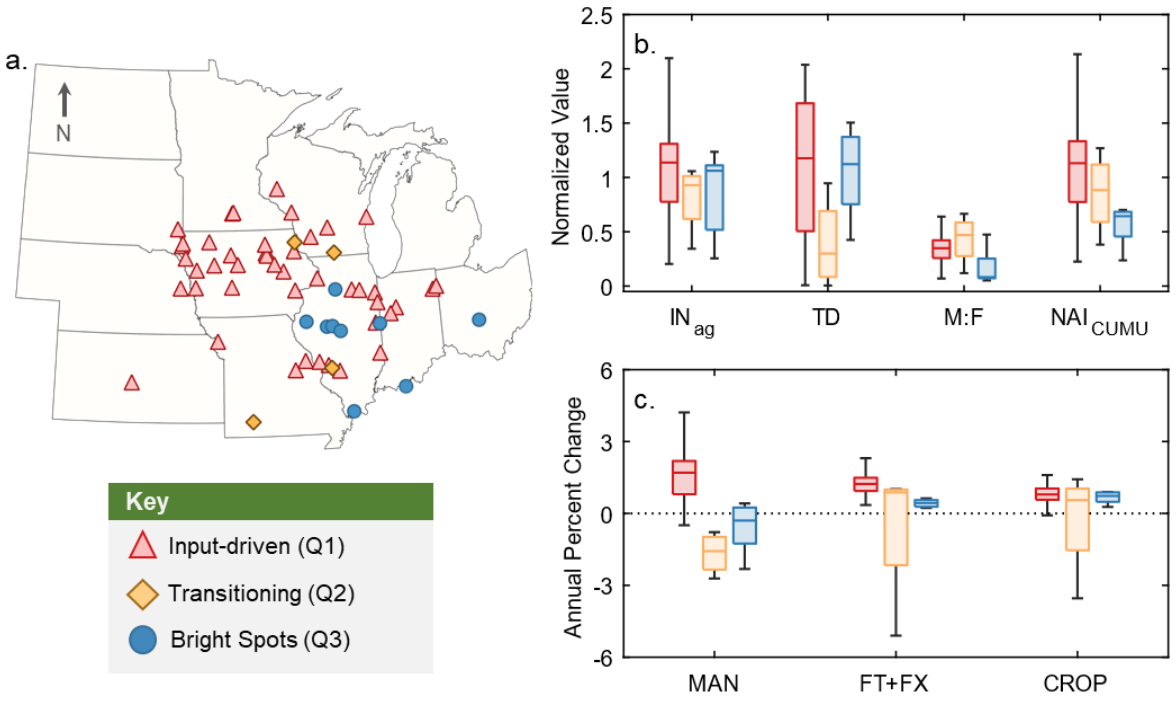


Figure 4.5 Comparison of the distribution of agricultural midwestern watershed attributes in Q1, Q2, and Q3 (agricultural land use > 25%). (a) All watersheds have pour points in Illinois, Iowa, Indiana, Kansas, Wisconsin, Minnesota, Missouri, Nebraska, Michigan, Ohio, or South Dakota. (b) The distribution of watershed characteristics shows that input magnitudes (IN_{ag}) are similar magnitudes between all watersheds, while agricultural land with tile drainage (TD), manure to total fertilizer ratio (M:F), and 10-year net agricultural inputs (NAI_{CUMU}) differ. Characteristics are normalized using the mean magnitude of all midwestern watersheds. (c) The annual percent change of manure inputs (MAN) magnitudes differ across quadrants in comparison to inorganic fertilizer and fixation (FT+FX) and crop uptake (CROP).

4.4 Conclusion

We analyzed nitrogen input and export trajectories over the last 30 years across 492 watersheds in the continental U.S. We find that while nitrogen surplus has been decreasing in 54% of the watersheds, exports have been decreasing in only 39% of the watersheds. The mismatch between the number of watersheds with decreasing surplus versus export role of the watershed on the way nitrogen is processed in the landscape and the relationship between inputs to stream responses.

We identify four typologies of export response patterns: Input-driven, Transitioning, Bright Spots, and Delayed-response. Broadscale patterns show that Input-driven watersheds' high exports are coupled to high watershed nitrogen inputs. These watersheds are highly responsive due to tile drainage that increases hydrologic connectivity, quickly moving dissolved nitrogen

from the landscape to rivers. Delayed-response watersheds' exports are decreasing despite increasing inputs. These watersheds have a high aridity index, which may indicate slow groundwater movement in arid regions. Transitioning and Bright Spot watersheds, characterized by their mixed land use with relatively high urbanized characteristics, show different dynamics; bright spots watersheds' export are likely driven in large part by effective management of human waste, whereas Transitioning watersheds, like Delayed-response watersheds, show a decoupling of nitrogen surplus and nitrogen export trends, indicating the influence of legacy nitrogen.

As a whole, our results provide insights into the complex relationship dynamics of nitrogen surplus and export, as well as how the landscape, climate, and legacy nitrogen can influence these relationships. The decoupling observed between nitrogen surplus and export in many watersheds calls for a more tailored approach to nitrogen management, one that accounts for the unique characteristics and histories of each watershed. Only by understanding these nuances can we develop more effective strategies for managing watersheds and achieving desired environmental outcomes.

Chapter 5 Nitrogen Retention across Continental US

Outline

Increasing nitrogen fertilizer application, livestock densities, and population have led to widespread nitrogen contamination and surface water eutrophication. While the role of legacy nitrogen stores in hindering water quality improvements is well-established, less is known about the extent of legacies in the subsurface and the drivers of nitrogen retention across diverse landscapes. In this work, we aim to quantify nitrogen retention, the nitrogen that is stored in legacy pools or lost to denitrification, across the contiguous U.S. We take a mass balance approach using nitrogen inputs and machine learning modeled watershed runoff to estimate the mass of nitrogen retained in HUC8 across the U.S. Agricultural watersheds have the largest mass of retained nitrogen, particularly those with tile-drained agricultural land, however they retain a smaller fraction of their total inputs because of high export and crop uptake. We find that rates of nitrogen retention are highest in watersheds with large manure inputs and high population densities and watersheds with low inputs. But, it is important to note that not all retention indicates legacy stores. Regions with high population densities likely have a large part of their retention lost through wastewater treatment removal, and watersheds with low inputs can likely assimilate the excess nitrogen in their landscape. However, watersheds with large manure inputs are more at risk of legacy build up. We further parse the proportion of retained nitrogen into the legacy pools in the Upper Mississippi region, estimating that a median of 42% of the nitrogen in the landscape is stored as biogeochemical legacies in soil organic matter, and 16.5% is stored as hydrological legacies in groundwater, with the remainder lost through denitrification. Our results characterize nitrogen retention and the spatial dynamics at the watershed scale across U.S. watersheds. It also offers an improved understanding of legacy dynamics and estimates that can be used to better constrain watershed management models.

5.1 Introduction

Increasing nitrogen use stemming from a growing populations, increasing energy demands, and a shift towards more high-protein foods, has had an adverse effect on water quality and ecosystem health worldwide (Byrnes et al., 2020a; Glibert, 2016; Godfray et al., 2018). The U.S., for instance, is experiencing the consequences of nutrient overuse, triggering persistent eutrophication and hypoxia in important coastal ecosystems, including the Chesapeake Bay and the Gulf of Mexico (Basu et al., 2022; Glibert, 2020; Van Meter et al., 2018). Despite decades of concerted efforts to address excess nitrogen loads on these drainage basins, we have mostly only

seen marked decreasing nitrogen exports in urban watersheds attributed to successful policies targeting point source loading. Agricultural watersheds, on the other hand, have had much less success in export reductions (Stets et al., 2020). In the agricultural watersheds seeing improvements, significant improvements in water quality often remain elusive, with observable responses lagging by decades (Chang et al., 2021; Murphy et al., 2022; Tango & Batiuk, 2016; Van Meter et al., 2017; Q. Zhang et al., 2018).

Research over the past two decades has focused on quantifying the relationship between net nitrogen inputs in watersheds and riverine nitrogen loads to address the growing threat of nitrogen pollution. For instance, Metson et al.(2020) used a mass balance approach to understand the decoupling of watershed anthropogenic inputs on watershed loads. They found that 70% of the nitrogen inputs in Oregon's Willamette River watershed were unaccounted for after considering crop uptake and riverine export. This unaccounted fraction of nitrogen is common in watersheds worldwide, with studies showing that of the nitrogen remaining after crop removal, riverine export accounts can only account between 6% in Mediterranean climates up to 38% in temperate watersheds (Boyer et al., 2002; Ehrhardt et al., 2021; Goyette et al., 2016a; H. Han & Allan, 2008; Hong et al., 2017; Huishu et al., 2018; Metson et al., 2020; Sobota et al., 2009; Swaney et al., 2012; A. Wang et al., 2016).

It is now increasingly understood that delayed improvements are partly due to legacy nitrogen—the accumulated mass of nitrogen within the landscape from decades of nitrogen overuse(Ascott et al., 2017; Basu et al., 2022; Chang et al., 2021; Van Meter et al., 2017, 2018; Q. Zhang et al., 2022). Legacy nitrogen can be stored in the organic nitrogen in the root zone of agricultural soils (biogeochemical legacy) or dissolved nitrate in groundwater and soil water (hydrologic legacy). To estimate the magnitude of biogeochemical and hydrological legacies and their effects on future water quality, Van Meter et al. (2017) developed a process-based model to estimate the magnitude of legacy storage pools in the Susquehanna and Mississippi River basins and their varied dynamics across space and time. Their work captured the broad age range of nitrogen at the drainage basin outlet, spanning from one to over 50 years old, with the Susquehanna having a nitrogen age distribution skewing younger because of the density of point source nitrogen loading. Since then, process-based models have been used in a few watersheds to predict the time lags to meet nitrogen load targets. For example, it would take 20 years to meet nitrogen load targets in the Chesapeake Bay tributaries under a business-as-usual scenario (Chang et al., 2021), 24 years in the Grand River, ON tributaries (Liu et al., 2021), 18 years in South Fork River watershed, IA if nitrogen fertilizer application ceased (Ilampooranan et al., 2019), and over 30 years in the Gulf of Mexico if excess nitrogen inputs in the watershed ceased (Van Meter et al.,

2018). While the scientific community is increasingly aware of the existence of nitrogen legacies, reliance on highly parameterized process-based models underscores the limited scope of current approaches. By and large, the magnitudes of legacy nitrogen and its impact on water quality at a continental scale remain largely unexplored, especially in non-temperate watersheds.

Watersheds act as filters, where microorganisms, vegetation, and hydrological processes work to transform, transport, and remove nitrogen. In these processes, nitrogen can be temporarily retained in the root zone, unsaturated zone, or groundwater pool as legacy nitrogen or permanently removed through denitrification (Dupas et al., 2020; Grizzetti et al., 2015). Characterizing nitrogen retention and its drivers in the landscape are key challenges in managing legacy nitrogen in watersheds. Attempts have been made to estimate legacy nitrogen in different landscape compartments. A recent study by Van Meter et al. (2023) is the first to provide estimates of the mass of accumulated nitrogen stored in the groundwater pool within a large agroecosystem. They used a data-driven machine learning approach to estimate spatial- and depth-varying mass of groundwater nitrogen in the Upper Mississippi River basin (UMRB), finding that 328 kg-N ha^{-1} had been stored in the groundwater recharged over the last 50 years. Similarly, Ascott et al. (2017) used an empirical modeling approach to estimate the mass of nitrogen stored in the vadose zone globally. Despite these advances, many uncertainties remain about the scale of nitrogen retention and legacy stores.

This work introduces a novel data-driven approach to quantify nitrogen retention. Using the newly developed gridded nitrogen mass balance dataset (gTREND-Nitrogen) and watershed properties, we train a machine learning model to estimate watershed nitrogen runoff for over 2000 HUC8 watersheds across the contiguous U.S. We use our modeled results to characterize nitrogen retention and its drivers in watersheds to understand: **(1) How much nitrogen is retained across the contiguous U.S? (2) How does that change regionally? (3) What are the drivers of different retention patterns across the U.S.? and (4) Can we parse the distribution of nitrogen retention stored in groundwater, SON pools, and lost through denitrification within the Upper Mississippi River basin**?

5.2 Methods

We used disparate data sources and long-term water quality records to predict the nitrogen retained in watersheds across the contiguous U.S. (**Figure 5.1**). The following sections describe this process.

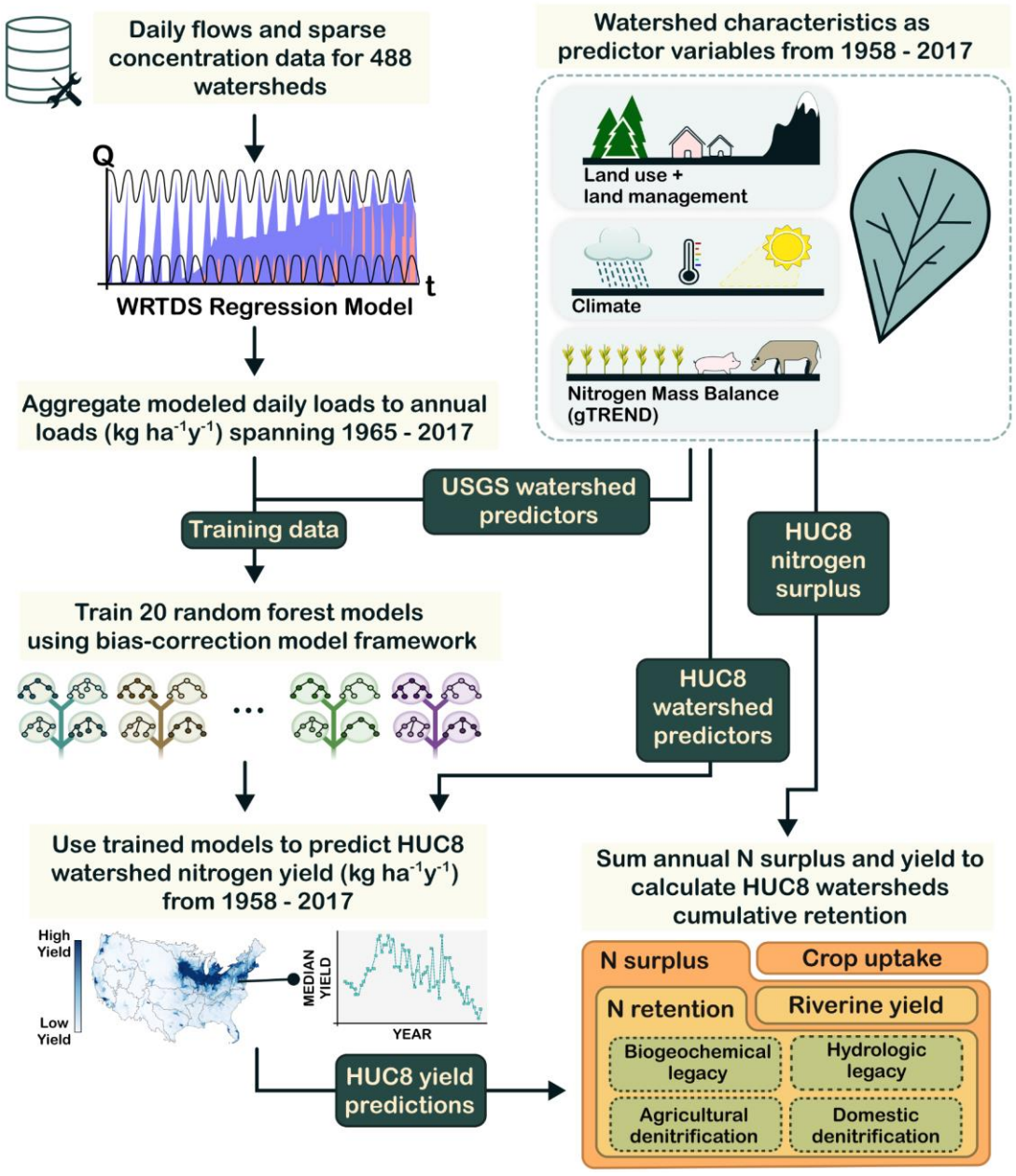


Figure 5.1 Conceptual figure of our methods for estimating the nitrogen retained in HUC8 watersheds across the U.S.

5.3.1 Model training data

5.3.1.1 Site selection and data sources for training data

We used the same water quality data as in Chapter 3. The following summarizes the data compilation methods outlined in the prior chapter. To compile a dataset of annual nitrogen loads between 1968 and 2017, we first compiled concentration data from Oelsner et al. (2017) and NWIS and Water Quality Portal using the dataRetrieval package in R (DeCicco et al., 2024). Daily discharge data was retrieved from NWIS. We first selected a set of watersheds to model based on the following criteria: (1) at least 70% of the years had quarterly samples; (2) at least 10% of the samples were taken during high flows; (3) gaps between sampling days were less than 4 years; (4) no more than six consecutive days with missing streamflow per year; (5) no more than 30 days with zero streamflow per year; and (6) at least 10 complete years of data were available, (7) first record began in 2003 or earlier. Some stations with data discontinuities were included, as long as there was at least a 10 year period that met the above criteria by removing the timeframes with problematic. We cleaned the data by removing concentrations that were six standard deviations from the mean in log space and averaging multiple daily concentration measurements to get a single concentration. If flow data had discontinuities of 6 days or less, we used linear interpolation between the nearest days to estimate missing flow data.

5.3.1.2 Estimation of annual watershed nitrogen export

We modeled daily concentrations and loads for the set of watersheds using Weighted Regressions on Time, Discharge, and Season (WRTDS). WRTDS is a regression model that estimates daily concentration from sparse concentration measurements and daily flow data using time, flow, and season as explanatory variables (R. M. Hirsch & De Cicco, 2015; Oelsner, Sprague, Murphy, Zuellig, Johnson, Ryberg, Falcone, Stets, Vecchia, Riskin, De Cicco, et al., 2017). The quality of the model's fit was assessed using a flux bias metric, and we excluded any stations where the flux bias exceeded ±30%. This resulted in a total of 488 watersheds in our dataset. Annual loads were calculated by summing daily loads.

5.3.2 Hydrological, climatic, and land management predictor dataset

We used many different static and temporally varying datasets to compile our machine-learning model's initial set of 30 predictors. For the training dataset we used watershed metrics compiled by Falcone (2017) including watershed’s NWALT land use, soil type, housing density, dam density, wastewater treatment plant density, watershed slope, base flow index, and stream order.

NWALT land use started in 1978, and thus 1978 was the first year of record in the training data. We supplemented these data with annual climate metrics from TerraClimate (Abatzoglou et al., 2018), including precipitation, potential evapotranspiration, temperature, snow water equivalent. From these metrics, we derived the aridity index, defined as the ratio of potential evapotranspiration to precipitation. We also compiled the fraction of the watershed's agricultural land that is tile drained using a dataset of tile-drained croplands by Valayamkunnath et al. (2020) and NLCD agricultural land use in 2011 compiled by Falcone (2017). Lastly, the gridded components of the N mass balance are derived from the Gridded TRajEctories Nutrient Dataset for Nitrogen (gTREND-N). We calculated the net agricultural nitrogen inputs from the components, which is the difference between nitrogen inputs (fertilizer, livestock manure, and biological nitrogen fixation) and crop removal. We also created a cumulative net agricultural input, the cumulative total within a 10-year window. Further details on the predictors can be found in **Supplemental Table 5.1**.

We processed geospatial datasets for 2110 HUC8 watersheds across the contiguous U.S. using the same datasets and methodology as the USGS watershed predictors. For the HUC8 watersheds, we compiled watershed characteristics using the same or comparable datasets to those used by Falcone.

5.3.3 Training model to estimate nitrogen loading in ungauged watersheds

5.3.3.1 Model development

We developed a set of random forest machine learning models to predict watershed nitrogen export (kg-N ha^{-1} yr^{-1}) from watersheds based on climate, land use, nitrogen inputs, and landscape management (Scikit-learn *v.1.3.0*) (Pedregosa et al., 2011). Random forest models build hundreds of small regression trees from bootstrapped sampled training data. Each tree is built using a random subset of watersheds and a subset of predictors, and the final model predictions are the product of all the trees. They are widely used in large part because of their ability to handle noisy, nonlinear, and intercorrelated data typical of environmental and hydrological data (Hastie et al., 2009; Tyralis et al., 2019). Traditional random forest models have an inherent bias, resulting in underestimating large values and overestimating small ones (Belitz & Stackelberg, 2021; Messager et al., 2021). We applied a bias-correction method that uses the sum of the results of two random forest models to enhance prediction accuracy: the first random forest model predicts watershed export while the second model estimates the bias in the prediction (Belitz & Stackelberg, 2021; P. Feng et al., 2019; G. Zhang & Lu, 2012). The bias-correcting random forest is trained on the residuals of the predicted loads from the first model

with load and original predictor variables serving as the predictor set. This approach leverages the strength of random forest models in reducing variance while correcting for some inherent systematic bias.

For our models, we divided the dataset into training, tuning, and testing sets with proportions of 60%, 10%, and 30%, respectively. We used the training data to train the models, the tuning set for hyperparameter model calibration, and the testing set to evaluate the models. We calibrated the following hyperparameter in the random forest models: (1) number of decision trees in the forest; (2) number of features to consider when looking for the best split; (3) minimum number of samples required to be at a leaf node; and (4) maximum depth of the tree. We employed an exhaustive grid search to select the best parameter set, evaluating all possible permutations within the hyperparameter range using 5-fold cross-validation and selecting the set that produced the lowest mean squared error.

5.3.3.2 Predictor variable selection and interpretation

We initially trained the model with 30 predictors and then recursively eliminated predictors to balance parsimony with model performance. We used R-squared, RMSE, and the median KGE of individual watersheds to evaluate model performance after each eliminated predictor. To prioritize which predictors to remove, we used permutation importance, which ranks the impact of each predictor by permuting each one 100 times and observing the median change in model performance scores. After identifying correlated predictors, we utilized partial dependence plots to visualize the marginal effects of individual predictors on the predicted loads. This recursive process involved assessing the importance rankings of the predictors, investigating the correlations of lower-ranked predictors, and removing them one by one. After each iteration, we reassessed the model's performance to understand which predictors had the largest influence and how predictors related to one another, allowing us to select those that performed intuitively. Ultimately, this approach led to selecting 14 predictors for use in the random forest models.

To enhance the interpretability of the final model, we used SHAP (SHapley Additive exPlanations) values to understand the contribution of each predictor to the predictions (Lundberg et al., 2020). SHAP values offer insights into the operational mechanics of the model and how each selected predictor influenced the prediction at the watershed level (Lundberg et al., 2020). The magnitude of a SHAP value reflects the strength of a predictor's effect on the model: larger SHAP values indicate a greater influence on the model's predictions, with positive values indicating an increase in prediction magnitude and negative values indicating a decrease.

5.3.4 Predicting nitrogen export, fate and legacy pools

5.3.4.1 Predicting annual watershed yield

We used the set of RF models to predict watershed N runoff (kg-N ha^{-1} y^{-1}) for 2110 HUC8 watersheds from 1958 to 2017 across the contiguous U.S. To account for model uncertainty in the export predictions and the predictor importance, we built and trained 20 unique model sets. For each of the 20 training runs of the models carried out, we stored the permutation importance scores of each predictor and their relative ranking. We used the trained set of random forest models to make predictions of HUC8 watershed export from 1958-2017 across the contiguous U.S. We used the median of the 20 models for our load estimate (**Supplemental Figure 5.1**).

5.3.4.2 Estimating nitrogen retention across the U.S.

To understand the relationship between nitrogen inputs, losses, and fate within HUC8 watersheds across the U.S., we calculate the total mass of inputs, losses, and storage over 60 years (1958-2017). We focused on cumulative metrics of nitrogen inputs (including livestock manure, fertilizer, biological nitrogen fixation, human waste, and atmospheric deposition), exports (crop uptake and watershed N runoff), surplus (the difference between nitrogen inputs and crop uptake), and retention. 'Retained nitrogen' has three fates: stored as legacy nitrogen in the soil organic pool, stored as legacy nitrogen in the groundwater pool, and lastly, denitrified and lost from the system. There remains a large uncertainty of denitrification rates because of challenges in estimating the rates across time and large spatial scales. For this reason, while denitrified nitrogen is a loss and not truly retained in the system, it is commonly grouped with the other pools of nitrogen retained within the landscape.

We calculated the cumulative input and export mass (kg-N ha^{-1}) by summing each annual estimate from 1958–2017. Similarly, we estimated the cumulative retention to represent the total mass in each watershed stored in the soil or groundwater pool, or lost to denitrification. By summing the annual retention, we can calculate the cumulative nitrogen retention, R_{cumu} (kg ha^{-1}), for the watershed for the period 1958–2017:

$$R_{CUMU} = \sum FERT(t) + BNF(t) + MANU(t) + ATM(t) + HUM(t) - CROP(t) - Y(t) \quad \text{(Eq. 5.1)}$$

where FERT is fertilizer input for year *t*, BNFis biological nitrogen fixation for year *t*, MANU is livestock manure input for year *t,* ATM is atmospheric deposition for year *t*, HUM is the human waste input for year *t*,. Nitrogen is removed through crop uptake (CROP) and watershed nitrogen runoff (Y) predicted from the random forest model set. Retention can also be conceptualized as

the mass of the nitrogen surplus remaining after watershed nitrogen runoff is removed. All units in kg-N ha^{-1} yr^{-1}.

To understand the drivers of nitrogen retention across HUC8 watersheds, we calculate the percentage of cumulative inputs exported through crop uptake (crop use efficiency) (Eq. 5.2a), the percentage of cumulative inputs exported through runoff (input runoff) (Eq. 5.2b), and the percentage of cumulative inputs retained in the watersheds (input retention) (Eq. 5.2c), as follows:

$$Crop\ use\ efficiency\ (\%) = \frac{CROP_{CUMU}}{INPUTS_{CUMU}} \quad \text{(Eq. 5.2a)}$$

$$Input\ runoff\ (\%) = \frac{Y_{CUMU}}{INPUTS_{CUMU}} \quad \text{(Eq. 5.2b)}$$

$$Input\ retention\ (\%) = \frac{R_{CUMU}}{INPUTS_{CUMU}} \quad \text{(Eq. 5.2c)}$$

5.3.4.3 Partitioning nitrogen retention into legacy pools and losses in the Upper Mississippi Region

We estimated the mass stored in the soil organic nitrogen pool, groundwater pool, and loss to denitrification in HUC8 watersheds within the Upper Mississippi region (Region 07). To achieve this, we calculated the cumulative mass of landscape nitrogen retention (L_{CUMU}) from 1972–2017. Cumulative landscape nitrogen retention (kg-N ha^{-1}) differs from cumulative retention (Eq. 5.1) because it also accounts for the mass of nitrogen lost through wastewater treatment removal, as follows:

$$L_{CUMU} = \sum FERT(t) + BNF(t) + MANU(t) + ATM(t) + (HUM(t) - HUM(t) * \lambda_{WWTP}) - CROP(t) - Y(t) \quad \text{(Eq. 5.3)}$$

where λ_{WWTP} is a static removal rate of 0.83, applied to human inputs to estimate the nitrogen removed annually through wastewater treatment (Van Meter et al., 2017). Nitrogen is also removed through crop uptake (CROP), watershed nitrogen runoff (Y). All units in kg-N ha^{-1} yr^{-1}. A diagram of this calculation is available in **Supplemental Figure 5.2**.

The static wastewater removal value is within the same magnitude of N removal rates for tertiary treatment used in other studies (Van Drescht et al., 2009). Although not all wastewater treatment plants in the Upper Mississippi River region employ tertiary treatment for N removal (Ivahnenko, 2008), the largest urban centers which contribute the majority of wastewater N do utilize advanced treatment systems. Thus, applying a tertiary-level treatment rate effectively

captures nitrogen removal in high-input urban areas while having negligible impact on estimates in nonurban areas where human waste inputs are minimal.

Cumulative landscape nitrogen retention is either stored as biogeochemical legacies in the soil organic nitrogen pool, as hydrological legacies in the groundwater pool, or removed via landscape denitrification, as follows:

$$L_{CUMU} = SON_{CUMU} + GW_{CUMU} + DET_{CUMU} \quad \text{(Eq. 5.4)}$$

where L_{CUMU} is the cumulative landscape retention, SON_{CUMU} and GW_{CUMU} are the cumulative mass stored in the SON and GW legacy pools, respectively. DET_{CUMU} is the mass of nitrogen lost through denitrification and is estimated to be the remaining mass not accounted for in cumulative SON and groundwater storage.

To estimate the mass of legacy nitrogen stored in the Upper Mississippi region (HUC2 region 7), we partitioned the landscape retention mass of each HUC8 based on existing literature estimates of regional legacy mass storage and storage rates. For nitrogen in the groundwater, we used data-driven estimates of nitrogen mass at the 250m scale in the Upper Mississippi basin. Van Meter et al .(2023) used a random forest model to estimate the mass of nitrogen in 'modern groundwater', defining modern groundwater as the groundwater recharged within the past 50 years (from 1972 to 2021) (Gleeson et al., 2016). This data allowed us to sum the mass of nitrogen stored in groundwater beneath each HUC8. For the SON pool, we used the lower bound estimates of soil organic nitrogen accumulation rates of 25 kg-N $ha^{-1}y^{-1}$ on cropland across the Mississippi River Basin from Van Meter et al. (2016). We used the upper bound of the SON accretion value from Van Meter et al. (2016) which ranged from 25 to 70 kg-N ha^{-1} yr^{-1} because the upper bound led SON storage exceeding retention magnitudes in many watersheds. Field studies have estimated SON accumulation rates to be within this bound, including a field study Weitzman et al. (2022), estimating retention in the vadose zone to be 58 kg-N ha^{-1} yr^{-1}, while assuming denitrification was negligible. Since the cumulative landscape retention only covered 1972 to 2017, a shorter period than the SON and groundwater estimates, we extrapolated each HUC8's landscape retention mass from 2017 to 2021 using the average annual retention rate from 1972 to 2017.

We determined the spatially varying SON accumulation rate for each HUC8 from 1972 to 2021 using the estimate of SON mass accretion in cropland 25 kg-N ha^{-1} y^{-1}. We converted this cropland SON accretion rate to HUC8-specific rates using annual agricultural land use fractions. Specifically, we use 250-m annual cropland and pastureland data from 1972 to 2005 by Sohl et al.(2016a; 2014) and then use 30-m NLCD data for the years 2006, 2008, 2011, 2013, 2016,

resampled to 250-m (Dewitz, J. , and U.S. Geological Survey, 2022). We used the land use from the nearest year with available data for missing years. The cumulative SON for each HUC8 watershed was then calculated by summing the annual values.

It is important to note that we did not account for the mass of dissolved nitrogen stored in the pore space in the vadose zone. While Ascott et al. (2017) identified the vadose zone as a significant hydrological legacy nitrogen pool in regions with thick vadose zones and intensive agricultural history, their estimates for the Upper Mississippi River Basin region were low, suggesting that this pool is not a significant source of legacy nitrogen in our study area.

5.3 Results and Discussion

5.3.1 USGS watershed model training data and results

The final nitrogen load training dataset consists of 13,298 samples across the contiguous U.S. from 1978 to 2017, with at least 134 samples per year (**Figure 5.2A**). Load magnitudes range from 0.002 to 69.85 kg-N ha^{-1} y^{-1}, with half of the load below 2.37 kg-N ha^{-1} y^{-1}. The density of monitored USGS watersheds varied across the U.S. (**Figure 5.2B**). We find a higher density of samples in eastern regions, specifically in the South Atlantic-Gulf Region (region 03), Upper Mississippi Region (region 07), Mid Atlantic Region (region 02), Arkansas-White-Red Region (region 11) (**Figure 5.2 C-D**). These regions, which drain to the Gulf of Mexico and the Chesapeake Bay, are more densely monitored for a more extended period because of the persistent issues of excess nutrients and hypoxia. Data is sparse for regions that are small, arid, or have less cropland area, including the region 09, 13, and 16 (Region 16 is not shown in the maps because there are no training watersheds with load data in this region).

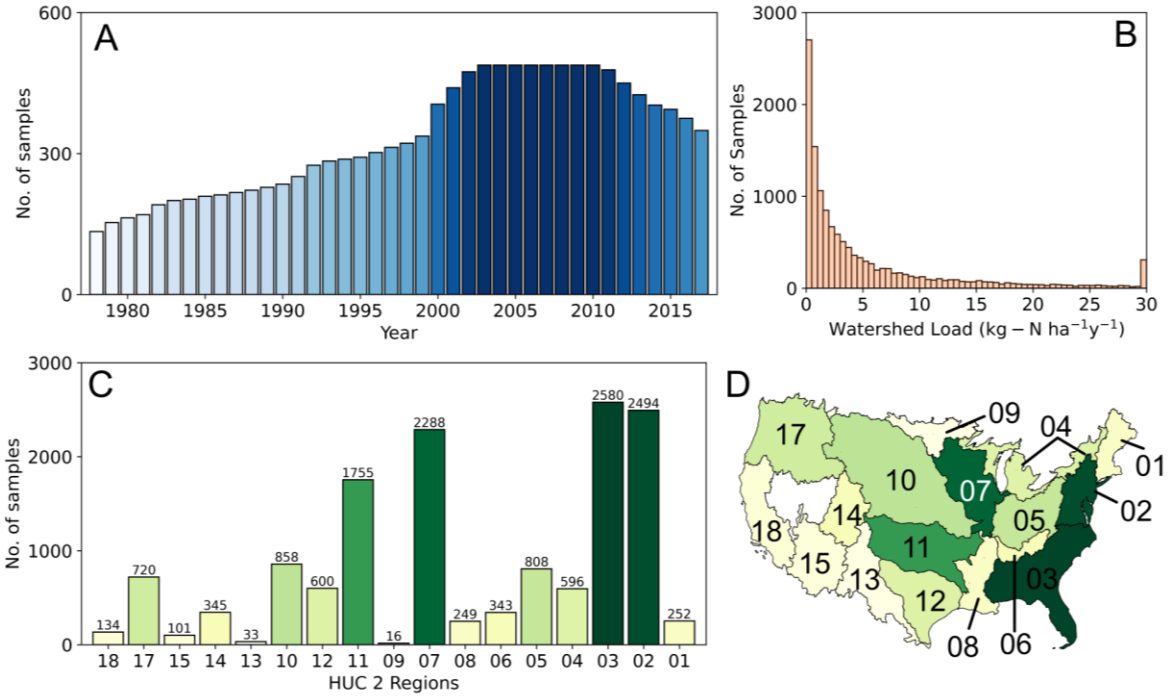


Figure 5.2 Summary of USGS riverine load training data. (a) Number of training data points per year. (b) Histogram of load magnitudes. (c-d) Number of training data points per HUC2 region. USGS watersheds can overlap multiple HUC2 regions, in which cases their samples are counted in multiple HUC2 regions. HUC2 regions are ordered from west to east, based on the region's centroid.

We calculated the mean of the evaluation metrics across all 20 model iterations to assess overall performance. The average R-squared values achieved were 0.99 for the training data and 0.89 for the test data. The RMSE values were 0.603 for the training set and 2.54 for the test set, while the PBIAS values were -0.01% and -1.02% for the training and test sets, respectively (**Supplemental Figure 5.3**). The model framework exhibits strong predictive capability, though performance does decrease at the tails of the distribution (**Figure 5.2 B**), where there is a tendency to regress to the mean and underpredict high loads. This underprediction aligns with the known limitations of ensemble tree models, which are prone to systematic bias. We addressed some of this bias through a bias-correction method, which improved model performance.

We used permutation importance and SHAP (SHapley Additive exPlanations) values to assess the influence of each predictor on the final model predictions. We find that the four most important model predictors of riverine load are percent of agricultural area that is tile-drained (TD), precipitation (PPT), human inputs (HUM), and fertilizer input (FERT) (**Figure 5.3A**). We find that higher magnitudes of TD, HUM, and FERT have higher SHAP values, indicating that

regions with higher tile drainage, fertilizer input, and human inputs have large positive impact N loads magnitudes (**Figure 5.3B and Supplemental Figure 5.4**). These correlations are expected as higher inputs means that there is more N available for export (Basu et al., 2022; Goyette et al., 2016b; R. Howarth et al., 2012b; Swaney et al., 2012). Furthermore, heavily tile-drained watersheds and watersheds with higher precipitation will have decreased water residence times and decreased denitrification (David et al., 2010; Dupas et al., 2020; Ritter & Rao Chitikela, 2020; Van Meter & Basu, 2017). Higher magnitudes of precipitation, PPT, SHAP values, however, do not monotonically increase with increasing precipitation, likely capturing the importance of other confounding climate predictors in this correlation (**Figure 5.4)**. Percent slope (SLOPE), average temperature (T), cumulative net agricultural inputs (cNAI), potential evapotranspiration (PET), and livestock inputs (LVSTK) are between 5th to 9th most important predictors (**Figure 5.3A)**. Percent slope shows lower values having higher impact on the model prediction, likely an effect of land with flatter slopes being used for agricultural and urban land. With temperature and PET, higher magnitudes are associated with larger decreases in model predictions. This can be physically explained as increasing temperature and PET lead to more water being lost to evapotranspiration and less water to flush nitrogen from the watershed. Both cumulative net agricultural inputs (cNAI) and livestock inputs (LVSTK) have increasing SHAP values with increasing predictor magnitudes, likely for similar reasons as fertilizer, higher inputs are correlated with higher riverine loads. While ranked in the middle of the predictors, livestock inputs have relatively high SHAP values, possibly indicating that livestock manure inputs are regionally important drivers (**Figure 5.3B and Supplemental Figure 5.4**). Lastly, watershed area (AREA), snow water equivalent (SWE), atmospheric deposition inputs (DEP), baseflow index (BFI), and net agricultural inputs (NAI) are ranked as the 10th to 14th predictors in the model. Their relatively low ranks show they only impact some model predictions. However, we can see that there are a few predictions with a SHAP value of ~2, which indicates that these predictors are important in specific temporal, geographical, or environmental conditions, which might not be prevalent across all data but are critical where they occur (**Figure 5.3B**). For example, in the case of NAI, while it is highly correlated with fertilizer and livestock manure inputs, it accounts for changes in the annual magnitudes of crop uptake and can capture the effects of surplus from a year of poor yield.

However, it should be noted that RF models are empirical in nature, relying on correlations between predictors and the target variable rather than establishing causality. Although these methods are valuable for understanding model behavior, the significance of predictors within a specific model is highly context-dependent and different relative degrees of importance may

emerge with different input data or when modeling similar data across different spatial or temporal domains. As such, the results should primarily be used to interrogate the model itself, with caution exercised when attempting to infer universal real-world phenomena (Grömping, 2009; Van Meter, Schultz, et al., 2023; Vowels, 2022). These considerations are crucial in any discussion of the relative importance of predictors.

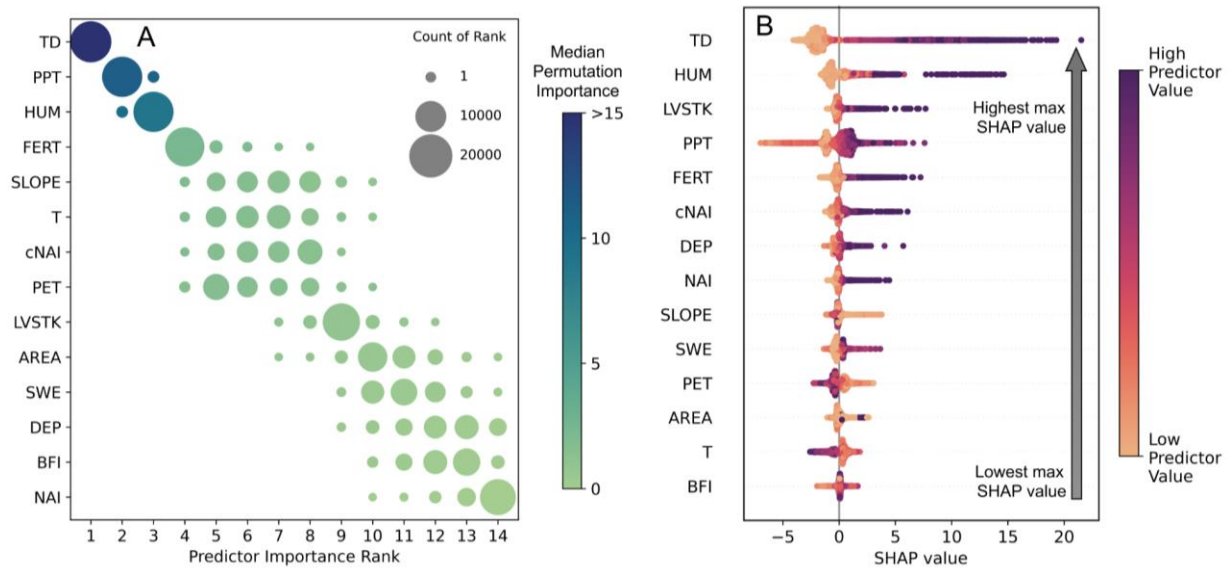


Figure 5.3. Model predictor importance quantified through predictor importance values and SHapley Additive exPlanations (SHAP) values. (A) We rank the order of importance based on the permutation importance value (decrease in model performance) for each of the 100 iterations for each 20 model runs. Predictors are sorted based on median predictor importance value. (B) We use a beeswarm plot to display SHAP values from the best performing model iteration ranked by R-squared, percent bias, and root mean squared error. Each dot represents a datapoint, depicting the impact of individual predictors on the model outcome, allowing to visualize the distribution of each predictor's influence.

5.3.2 Regional variability in nitrogen retention and the role of input sources.

The median cumulative retention (1958 to 2017) across HUC8 watersheds in the U.S. is 683.1 kg-N ha^{-1} with the 5-95th percentile ranging from 76.0-2618 kg-N ha^{-1} (**Figure 5.4A**). The total mass retained is highest in the eastern and midwestern regions (regions 07, 02, and 05) at 1960, 1598, and 1371 kg-N ha^{-1}, respectively. Retention is lowest in western regions 13–17, where regional magnitudes are all under 400 kg-N ha^{-1} (**Figure 5.4B**). HUC8 hotspots largely consist of watersheds with high fertilizer input, though some also show elevated retention due to high manure or human waste inputs (**Figure 5.4A-C**).

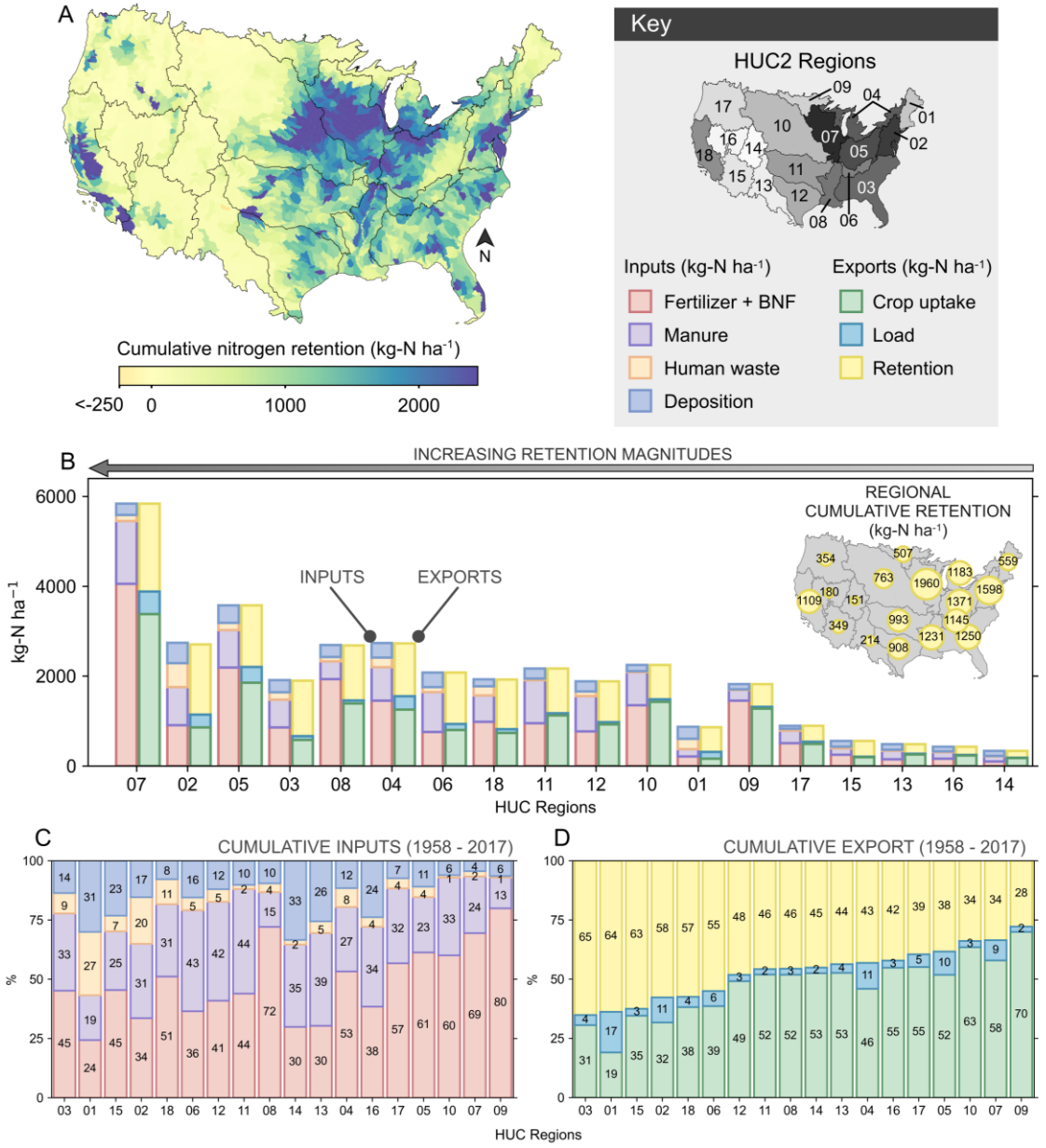


Figure 5.4 HUC2 regions input and fate of cumulative nitrogen (1958–2017). Panel (A) is the map of retention at the HUC8 watershed scale. Panel (B) shows the HUC2 aggregated cumulative mass (kg-N ha^{-1}) of inputs and exports, broken down by input and export components and ordered by magnitude of nitrogen surplus. Panel (C) and (D) display the percentage contribution of each component to the total nitrogen inputs or export for each HUC2 region.

In the previous section, we look at important model predictors to understand the drivers of load. We can use this as a starting point to investigate the drivers of retention. We find that the drivers of retention differ across the country. Regions 07, 02, and 05 have similar retention magnitudes.

However, their nitrogen inputs differ significantly, with regions 02 (2744 kg-N ha^{-1}) and 05 (3577 kg-N ha^{-1}) having nearly half the inputs of region 07 (5841 kg-N ha^{-1}) due to its efficient nitrogen removal via crop uptake (**Figure 5.4A-B**). While regions with high fertilizer usage show greater retention magnitudes, areas dominated by manure inputs retain a larger proportion of their nitrogen inputs. We find that regions with larger proportions of the cumulative inputs from manure, such as regions 03, 06, and 18, retain a large proportion of their cumulative nitrogen inputs (greater than 57%; **Figure 5.4C-D**). For example, region 03 has retention magnitudes similar to regions with much higher cumulative inputs, such as region 08 (1250 versus 1231 kg-N ha^{-1}, respectively). Region 03 has relatively low proportions of inputs that are removed through crop uptake (582 kg-N ha^{-1} or 31% of inputs removed by crops), all of which is met with fertilizer and BNF inputs (858 kg-N ha^{-1} or 45% of inputs; **Figure 5.4B-D**). Therefore, 33% of total inputs come from manure, which only creates a larger surplus of nutrients in the region (**Figure 5.4A**). This demonstrates that on the macro scale, manure is typically seen as a waste product and not used as a substitute for fertilizer to meet crop demands (Flynn et al., 2023; Spiegal et al., 2020).

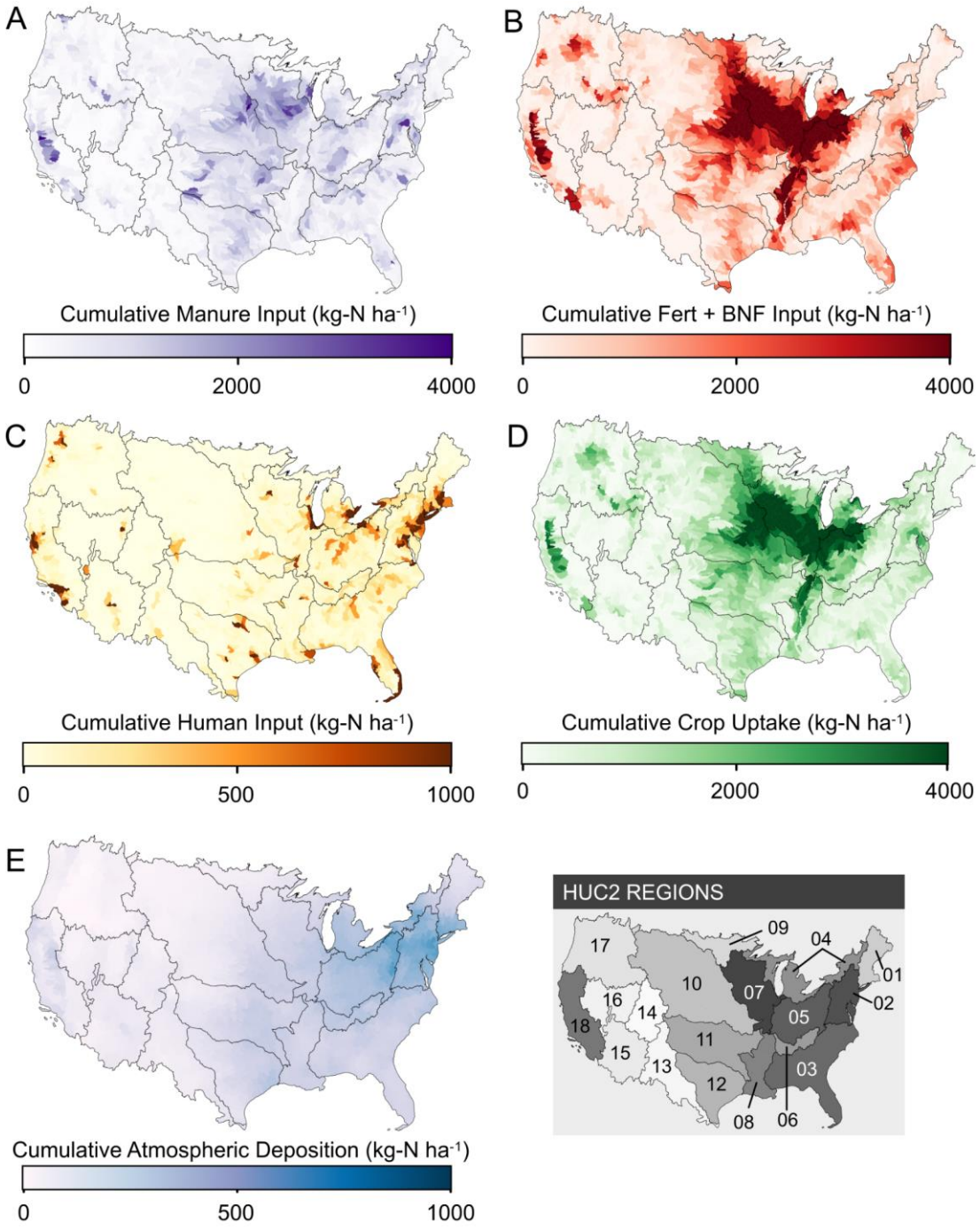


Figure 5.5. Spatial distribution of cumulative nitrogen inputs and exports across HUC8 watersheds (1958–2017) measured in kg-N ha^{-1}. The maps include (A) N inputs from livestock manure, (B) N from fertilizer use and crop biological nitrogen fixation, (C) N from human waste, (D) N removed from crop uptake, and (E) atmospheric N deposition

In New England (region 1) and the Mid-Atlantic (region 2), nitrogen retention rates reach 64% and 58%, respectively, with human inputs contributing more to the total than in other regions (**Figure 5.4A-B**). Cumulative human inputs are low across much of the country (**Figure 5.5C**). However, there is a broad range of magnitudes across space (median: 30.6 kg-N ha^{-1}, 5th-95th percentile: 1.1-526.8 kg-N ha^{-1}). Dense urban centers are found mainly along coastal areas, a pattern shaped by urban expansion and demographic shifts over recent decades (Byrnes et al., 2020a; Rieniets, 2009). High retention rates in these areas are due to a higher proportion of inputs from urban sources because, since the passing of the Clean Water Act in the 1970s and the Clean Air Act in 1990s. The Clean Water Act galvanized upgrades to wastewater treatment plants processes particularly for phosphorus removal. However, improvements have been to upgrade processes to also remove nitrogen through denitrification (Bernhardt et al., 2008; Schlesinger & Bernhardt, 2020). While this is counterintuitive, nitrogen removed through wastewater denitrification is considered 'retained' and is rarely separated from this term because of the challenges in estimating denitrification at large spatial scales. The Clean Air Act led to improvements in NOx pollutants in mobile and stationary sources, which in turn reduced oxidized nitrogen atmospheric deposition (Lloret & Valiela, 2016).

The analysis of the HUC2 regions above highlights some broadscale regional patterns of the relationship between cumulative inputs and exports. However, there is significant heterogeneity within each region. We used the cumulative percentages of each export fate–crop uptake, runoff, and retention–relative to the total inputs for each HUC8 watershed to understand retention patterns across space and the drivers within each region. The 1:1 line on **Figure 5.6** indicates no cumulative retention because all inputs are lost through crop removal or runoff. Watersheds further from this line retain a higher proportion of inputs. Across regions, three dominant relationships emerge: steep negative linear, shallow negative linear, or no relationship.

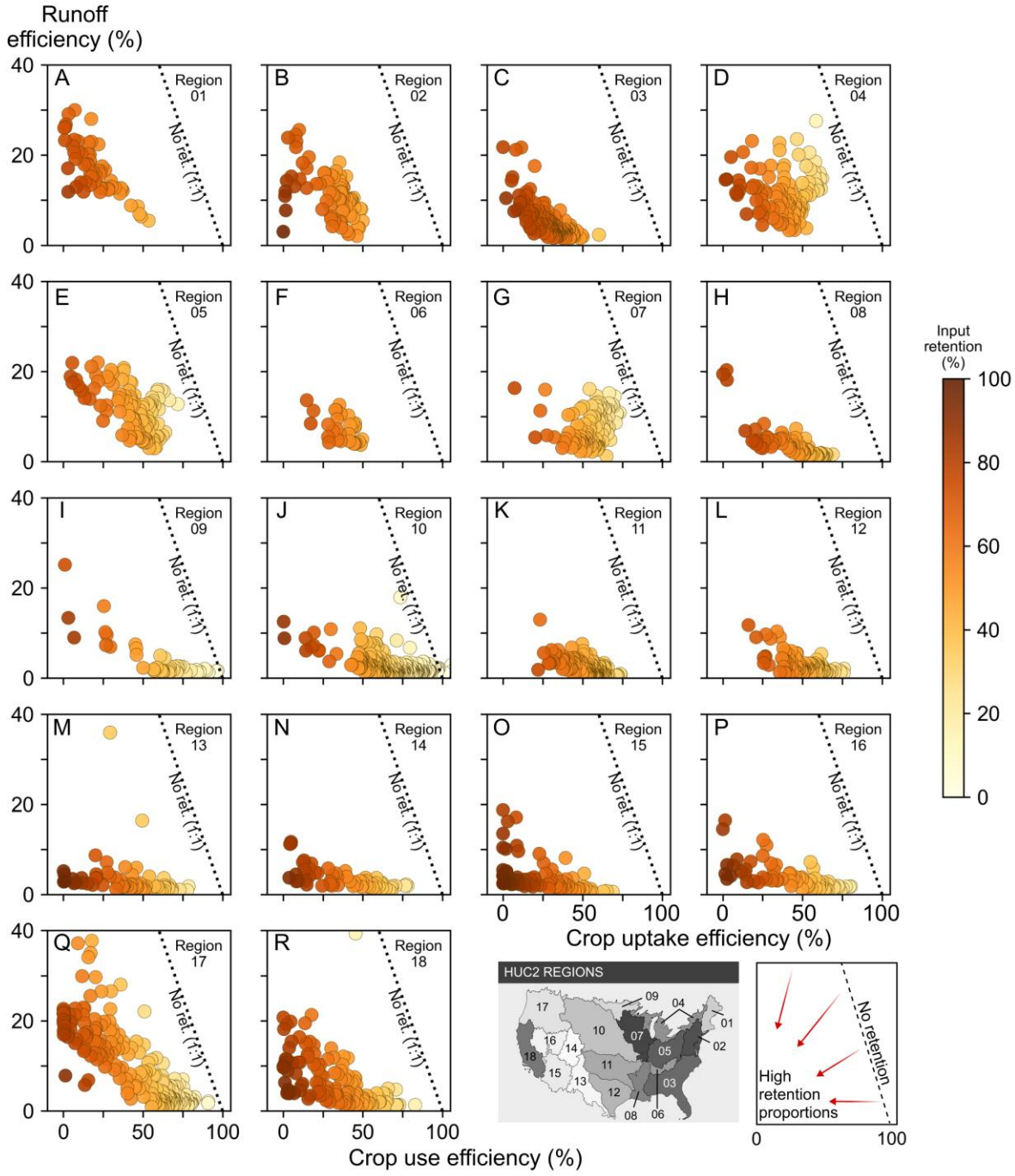


Figure 5.6. Spatially varying relationship between crop use efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by the retention efficiency (%).

Regions 01-03, 08, and 17, which tend to have humid climates, show steep negative relationships

(**Figure 5.6A-C, H, Q)**. In these HUC8 watersheds, low crop uptake efficiency leads to excess nitrogen lost through runoff, resulting in lower proportions of cumulative inputs retained in the watersheds. In contrast, areas such as 09-16 and 18 show (**Figure 5.6I-P, R**) a shallower relationship, where cumulative runoff proportions increase with decreasing crop uptake efficiency but less drastically. These regions tend to have lower nitrogen inputs overall, but are more prone to these inputs being retained. Meanwhile, regions like 04-07 (**Figure 5.6D-G**) do not have a relationship between crop use efficiency and input runoff. We find that many watersheds with high crop use efficiency and input runoff in these regions have very little capacity to retain nitrogen.

To further explore the drivers of retention across these regions, we used agricultural land use, the percentage of agricultural land that is tile drained, the proportion of total inputs from atmospheric deposition, and the proportion of total inputs from human waste (**Figure 5.7 and Supplemental Figure 5.5-5.9)**.

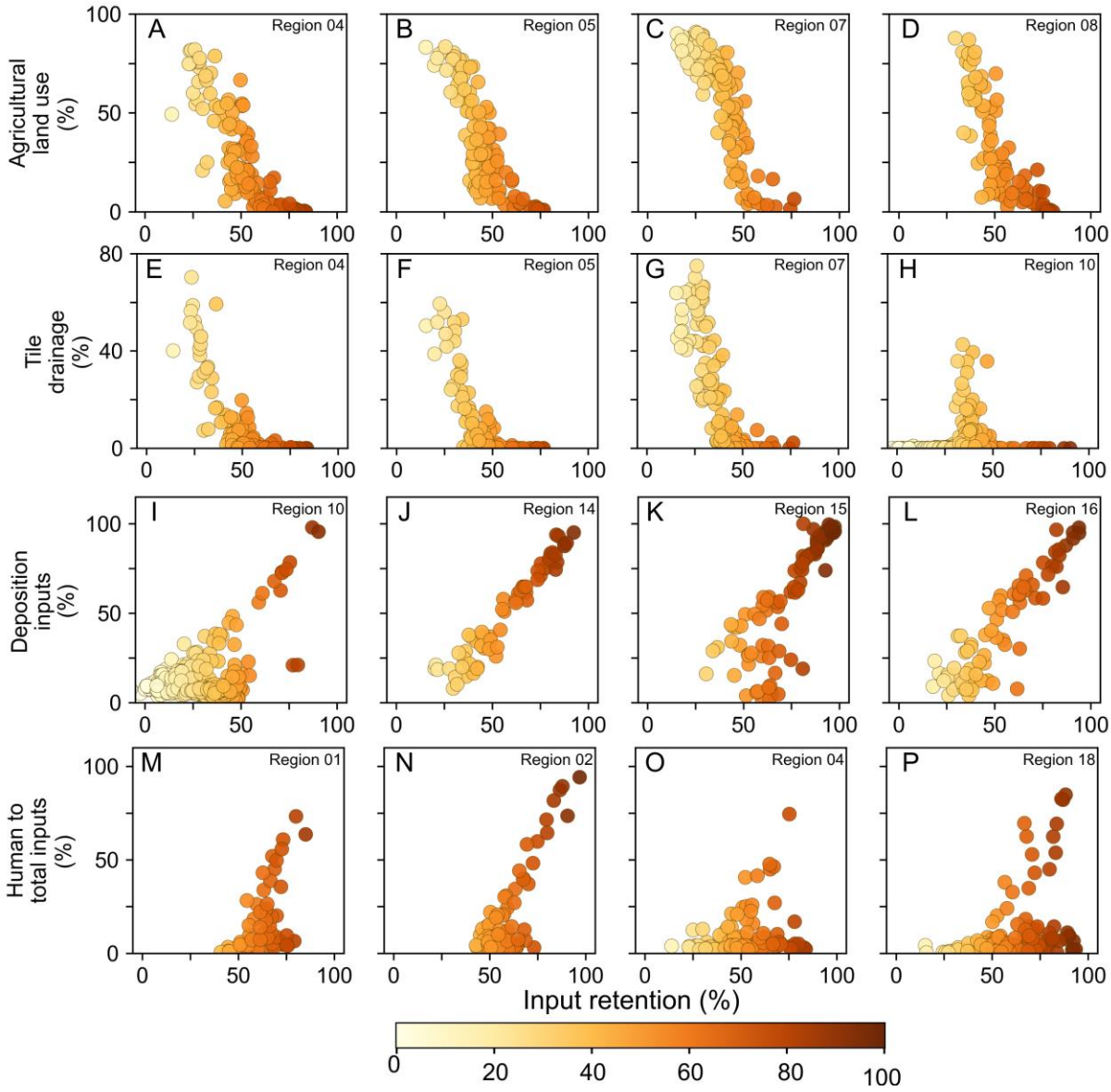


Figure 5.7 Relationships between input retention (%) and watershed properties across different HUC02 regions. The watershed properties include (A-D) percent agricultural land use, (E-H) percent of agricultural land that is tile drained, (I-L) percent of total inputs from atmospheric deposition, and (M-P) percent of total inputs from human waste. Markers are colored based on retention (%) for each region.

In regions 04, 05, 07, and 08, which have been intensively cultivated for many decades, we can see that as agricultural land use increases, the proportion of inputs retained decreases by more than 50% (75% to under 25% of inputs are retained) (**Figure 5.7 A-D**). We also find that the driver of the low retention rates in Regions 04-07 is the large proportions of tile drainage (**Figure**

5.7 E-G). Prolonged high nitrogen inputs into these agricultural watersheds have saturated their nitrogen storage capacity, leading to increasing riverine runoff export and reducing their ability to retain additional inputs (Newcomer et al., 2021). Region 08 has higher input retention proportions than the other three regions (**Figure 5.7 D**) because regions 04-07 have more tile drains. Tile-drains decrease nitrogen's transit time in the subsurface, bypassing retention mechanisms such as soil organic nitrogen storage or denitrification which will compound the low retention due to saturation.

Next, we find that watersheds within regions 10, 14-16 with inputs dominated by atmospheric deposition also have higher proportions of retained inputs (**Figure 5.7 I-L**). This is likely because areas with most of their inputs from deposition will likely be areas of low input (**Supplemental Figure 5.10**). Research by Newcomer et al. (2021) suggests that watersheds with low inputs can retain a higher proportion of inputs, as they have yet to reach nitrogen saturation and can assimilate excess nitrogen through soil organic pools or denitrification.

Lastly, regions 01, 02, 04, and 18 all contain urban centers with high magnitudes of cumulative human inputs (**Figure 5.5C**). In these regions, watersheds with higher proportions of human waste inputs retain more nitrogen due partly to upgrades in wastewater treatment plants that remove nitrogen through denitrification (Bernhardt et al., 2008). For example, region 02 has four watersheds with human waste inputs accounting for over 80% of all inputs and retaining over 80% of total inputs, which aligns with modeled wastewater removal rates in this region(Van Meter et al., 2017) (**Figure 5.7N**). We observe some variation in retention rates, especially in regions 04 and 18, where watersheds with similar high human waste inputs show different retention levels. This could be due to losses during treatment, either from varying wastewater treatment standards across states or untreated sewage leaks (Bernhardt et al., 2008; Flint et al., 2022).

5.3.3 Legacy pools in the Upper Mississippi region

In the previous analysis, we used cumulative retention, which is the difference between inputs and exports from crop uptake and watershed nitrogen runoff. The ability to further distinguish between temporary (legacy) and permanent (denitrification) removal of nitrogen is a major challenge in modeling nitrogen retention at the river basin scale owing to uncertainty in retention dynamics introduced from the heterogeneity in the subsurface and across watersheds(Basu et al., 2022; Grizzetti et al., 2015). In many cases, it is not possible to further parse the fates of retained nitrogen, but existing literature estimates of legacy storage existing in region 07 allow us to attempt to provide estimates of nitrogen stored in the SON, groundwater pool or lost to

denitrification. We attempt to use our model and literature legacy pool estimates to parse the mass stored in legacy pools and lost to denitrification from 1972 to 2021 in region 07, which has the largest cumulative nitrogen retention mass.

The cumulative retention used in the previous analysis includes the mass of nitrogen removed through wastewater treatment. Therefore, to achieve a more accurate estimate of the mass of nitrogen applied to the landscape, we calculated cumulative *landscape* retention (L_{CUMU}). Cumulative landscape retention differs from cumulative retention as it includes not only the nitrogen losses from crop removal and runoff, but also the mass lost in wastewater treatment (**Eq. 5.3**). See **Supplemental Figure 5.2** for a diagram of landscape retention calculation methodology. To calculate cumulative landscape retention, we used a wastewater treatment static removal rate estimated at the Mississippi River basin in a previous modeling study (Van Meter et al., 2017). The cumulative landscape retention can be further divided into the mass in the SON pool, the groundwater pool, and the mass of nitrogen lost through denitrification (**Eq 5.4**). We parse cumulative landscape retention using literature estimates of annual SON accumulation rates in the Mississippi River basin and the accumulated groundwater nitrogen mass in the Upper Mississippi region from 1972 to 2021 (HUC2 region 07).

The median landscape nitrogen retention for HUC8 landscapes from 1972 to 2021 is 1813 kg-N ha^{-1}, with an interquartile range (IQR) of 1402 to 2196 kg-N ha^{-1} (**Figure 5.8A)**. To estimate the legacy nitrogen stored in the SON pool across HUC8 watersheds, we used cropland SON accretion estimates of 25 kg-N $ha_{cropland}^{-1}$ y^{-1} and annual agricultural land area. A median of 42% of the retention mass is stored in the SON pool (5th - 95th percentile: 16 - 71%) (**Figure 5.8B**), with the northwest and southeastern watersheds having the highest magnitudes of SON and proportion of retained nitrogen in SON pool (**Figure 5.8F**). In line with our previous analysis, landscape retention is highest in watersheds with more manure inputs and less tile-drained agricultural area (**Supplemental Figure 5.11A-B, E**).

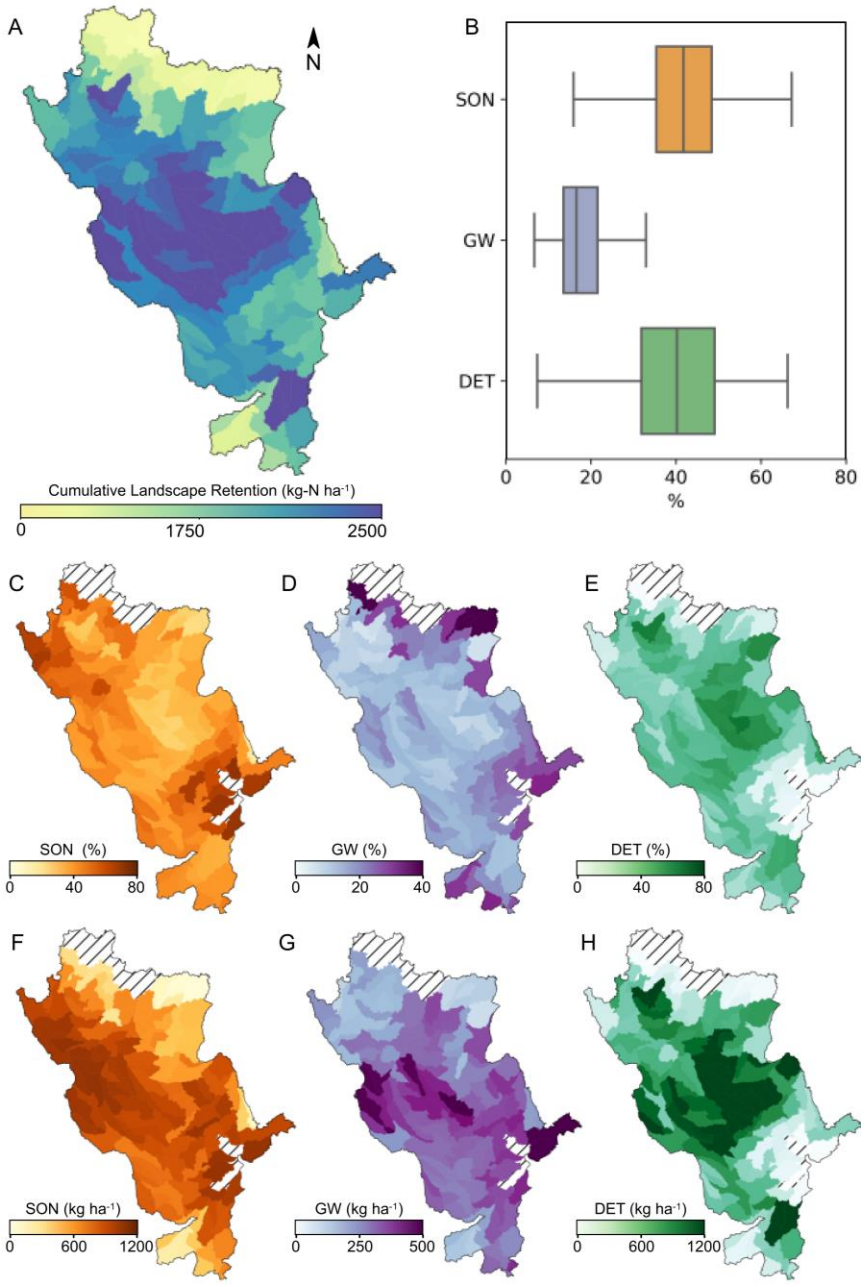


Figure 5.8 Predicted mass of landscape retention, nitrogen stored in the soil organic and groundwater pool, and the nitrogen loss through denitrification. Panel (A) is the landscape nitrogen retention at the HUC8 scale across the Upper Mississippi region. Panel (B)-(E) shows the distribution of HUC8 landscape nitrogen retention stored in the soil organic nitrogen pool (SON), groundwater pool (GW), or loss through denitrification (DET). Panels (F)-(H) are spatially varying estimates of HUC8 landscape retention mass stored in the SON, GW, and DET, respectively. Hatched areas represent areas where soil organic nitrogen pool (SON) and groundwater pool (GW) is larger than the estimated landscape retention.

We estimate the hydrological legacy stores in each HUC8 using gridded nitrogen mass stored in the modern groundwater across the Upper Mississippi basin derived by Van Meter et al.(2023). The groundwater pool makes up 16.5 % of the retention (ranging from 10% to 32%), with the most substantial accumulations observed in the central watersheds (**Figure 5.8D and 5.8G)**. The remaining mass attributed to denitrification suggests that denitrification results in a median loss of 40% of retained nitrogen, though the extent of loss varies dramatically from 2% to 59%, with the highest rates in central regional watersheds (**Figure 5.8E and 5.8H)**. The estimates of SON, groundwater, and denitrification losses align with other studies in temperate watersheds, in which it was estimated that nitrogen storage in the SON pool ranges from 28% to 88% of nitrogen retention (Ilampooranan et al., 2019; Liu et al., 2021; Schlesinger & Bernhardt, 2020; Van Meter, Basu, Veenstra, et al., 2016; Weitzman et al., 2022; Worrall et al., 2015), and in the groundwater pool from 3% to 19% (R. W. Howarth et al., 1996; Ilampooranan et al., 2019; Liu et al., 2021; Worrall et al., 2015). The exact magnitudes of these processes, particularly denitrification—which can account for 8.4% to 77% of nitrogen retention—remain largely unconstrained (Bouwman et al., 2009; Bouwman et al. 2013; David et al. 2009; Hofstra and Bouwman. 2005; R. W. Howarth et al., 1996; Ilampooranan et al., 2019; Seitzinger et al., 2006; Worrall et al., 2015). This variability is due to the challenges in direct measurement at regional scales, often resulting in estimates calculated based on differences between nitrogen inputs and all other storage and loss, or unconstrained model parameter calibration.

There are HUC8 marked by the hatched area, where estimated SON and groundwater mass exceeded the magnitude of landscape retention. Our literature values could have poorly captured the nitrogen retention partitioning in these regions for a few reasons. First, we assume that for each HUC8, the mass of nitrogen in the groundwater was nitrogen applied to the agricultural land in that HUC8. However, the groundwater nitrogen may originate from an upgradient HUC8. Furthermore, the northern area has minimal agricultural land (**Supplemental Figure 5.11E**), and the estimated accumulation from the Mississippi River Basin study poorly represents likely SON retention dynamics. Lastly, the eastern watersheds have low retention mass despite the high agricultural land use (**Supplemental Figure 5.11 A, E)**. These regions have low manure inputs to total input ratio (**Supplemental Figure 5.11 B-C)** and a high percentage of tile-drained agricultural land (**Supplemental Figure 5.11 D, F**). Our previous analysis shows that these regions have relatively low retention. Therefore, the SON accumulation rates in these HUC8 watersheds are perhaps below those for the entire Mississippi River Basin.

5.4 Conclusion

Excess nitrogen from a growing population, increasing production of corn ethanol biofuels, and shifting diets to meat-heavy diets have had repercussions on water quality across the U.S. The changes to the nitrogen cycle have caused increasing incidences of algal blooms and degrading water quality (Glbert, 2020). Over 1 trillion dollars spent to address water pollution since the passing of the Clean Water Act in 1972 (Keiser & Shapiro, 2017), we still see lagged water quality improvements and persistent hypoxic zones in some of the most important coastal zones in the U.S. This is in part due to the Clean Water Act does not directly regulate nonpoint source, despite it being the largest source of nitrogen to many watersheds (Porter et al. 2015).

Previous modeling studies in the Mississippi River Basin (Van Meter et al., 2017), Chesapeake Bay and Susquehanna River Basins (Chang et al., 2021; Van Meter et al., 2017), and the Grand River (Liu et al., 2021) have shown that legacy stores in agroecosystems can delay water quality improvements for up to decades. However, large uncertainties remain concerning the magnitudes of legacy stores across North America. This current work is the first attempt to use data-driven methodologies to estimate nitrogen retention across the U.S. and quantify the mass of legacy nitrogen stored in large agroecosystem landscapes.

In broad strokes, agricultural watersheds tend to have a larger pool of nitrogen retention by virtue of having higher inputs compared to other watersheds. However, these watersheds are also more efficient at removing nitrogen and tend to retain less of their nitrogen inputs than other watersheds. We find retention rates are highest in watersheds with large manure inputs, high population densities, and low total inputs. But, it is important to note that not all retention contributes to legacy stores in the landscape. Regions with high population densities likely have a large part of their retention lost through wastewater treatment removal, and watersheds with low inputs can likely assimilate the excess nitrogen in their landscape. However, watersheds with large manure inputs are likely regions where we will find these legacy build up.

For the first time, we also use data-driven approaches to partition nitrogen retained in the landscape to legacy stores in the soil organic nitrogen pool, groundwater pool, or the fraction lost to denitrification. In the median, 42% of the mass is stored in the SON pool, 16.5 % of the retention is stored in the groundwater pool, and 40% is lost to denitrification.

Chapter 6 Phosphorus Sustainability in U.S. Agriculture: A Quadrant Framework Analysis

Abstract

Phosphorus bolsters today's global food system as a crucial input for agricultural intensification to achieve food security. However, excess phosphorus usage has led to phosphorus loss rates surpassing the safe planetary boundary for phosphorus. The problem of phosphorus overabundance coincides with concerns for scarcity stemming from unmitigated soil erosion and the complex geopolitics around phosphorus rock reserves. To move toward more sustainable phosphorus management, we need a better understanding of the temporal and spatial patterns of phosphorus use. We employ annual and cumulative phosphorus surplus and use efficiency indicators to assess these socio-ecological challenges and explore temporal and spatial variability in phosphorus dynamics. We find that the evolution of agricultural phosphorus management is shaped by changing fertilizer management, the proliferation of concentrated animal operations, climate, and the accumulation of phosphorus legacies within the landscape. To quantify phosphorus sustainability in intensively managed landscapes, we developed an integrated framework that uses both cumulative phosphorus surplus and use efficiency. This framework allows us to synthesize indicators of phosphorus dynamics over time and across spaces, providing a comprehensive perspective of the state and progress of phosphorus management.

6.1 Introduction

Phosphorus is a critical nutrient for food production. However, the phosphorus runoff from agricultural fields and livestock feeding operations has caused increased incidences of eutrophication in freshwater lakes and reservoirs (Glibert, 2020). Eutrophication has become a pervasive problem in lakes and reservoirs globally, from Lake Winnipeg in Canada, to Lake Erie in North America, to Lake Taihu in China (Schindler et al., 2012; Sharpley & Wang, 2014). The overgrowth of algae poses a risk to human health from microcystin toxicity and has detrimental effects on aquatic systems, causing habitat degradation, loss of biodiversity, and fish kills from a lack of dissolved oxygen (Glibert, 2016; Le Moal et al., 2019; Schindler et al., 2016; W. Steffen et al., 2015; Vitousek et al., 1997).

The increased incidence of agriculture-related eutrophication events in the past few decades is the outcome of changes to agricultural systems. Over the last century, technological development and policies promoting food security fueled unprecedented increases in crop production and led to a system to optimize for efficiency. Mineral fertilizer increased to meet crop demands,

replacing the need for manure for fertilization (Gardner, 2002, p. 22; Glibert, 2020). The shift from many small farms to fewer, larger livestock operations has also led to a shortage of nearby cropland that can accommodate the amount of manure produced, leading to these operations spreading manure at rates exceeding crop needs(Dimitri et al., 2005; Long et al., 2018; Swaney & Howarth, 2019b). Simultaneously, to guarantee crop nutrient needs, farmers apply excess inorganic phosphorus fertilizer as a form of 'insurance' to guarantee yields(Jarvie et al., 2015; Roberts & Johnston, 2015). Decades of these practices have led to a build-up of legacy phosphorus stores in soils. These stores act as long-term, persistent sources of phosphorus pollution, which can delay progress in achieving decreases in stream phosphorus loads for decades (Haygarth et al., 2014; Powers et al., 2016; Sharpley et al., 2013; Van Meter et al., 2021).

While we face an overabundance of phosphorus in the environment, there are also risks to food security due to threat of phosphorus scarcity, including the loss from unmitigated soil erosion (Alewell et al., 2020), rising price of fertilizer, and uneven global distribution of phosphorus rock reserves which are subject to geopolitical complexities(Brownlie et al., 2021; Cordell et al., 2009; G. K. MacDonald et al., 2011). These problems of both overabundance and scarcity capture the necessity of sustainable phosphorus management and improved efficiency(Schröder et al., 2011).

To establish guidelines for sustainable phosphorus management, it is crucial that we better understand the temporal and spatial patterns of phosphorus use. However, there has yet to be a comprehensive quantification of regional phosphorus stores and fluxes across long-term timescales. To understand the history of phosphorus use and management, we need to identify the regions where improvements are being made and those where progress is lacking and the reasons for these differences.

In the current work, our objective is to develop a quantitative framework for evaluating the use and management of phosphorus in agricultural landscapes in the context of both food security and environmental outcomes. Working at the scale of the contiguous U.S., we leverage the new gTREND-Phosphorus mass balance dataset, which includes long-term (1930–2017), 250m x 250m phosphorus mass balance data, to explore both spatial variability and multi-decadal trends in agricultural phosphorus use.

In this work, we attempt to answer the following questions: **(1) How have agricultural phosphorus surplus magnitudes and phosphorus use varied temporally and spatially across the U.S.? (2) Are phosphorus surplus magnitudes and phosphorus use efficiency effective**

metrics for evaluating the sustainability of current and future phosphorus use? What are the weaknesses and strengths of the PUE metric and phosphorus surplus estimates? (3) Can we develop an integrated framework to better understand the state and progress of phosphorus management across space and time?

6.2 Methods

6.2.1 Data sources

We analyzed phosphorus fluxes and accumulation from 1930 to 2017 across the contiguous U.S. To do this, we used the agricultural components from the gTREND-Phosphorus dataset discussed in Chapter 3. Since our analysis is concerned with the phosphorus in agricultural land, we only considered fertilizer and manure as phosphorus inputs into each grid cell. We did not include atmospheric deposition on agricultural land because it accounted for only a small amount of total P inputs in these systems (Zou et al., 2022).

6.2.2 Phosphorus sustainability metrics

Three different metrics were evaluated in the current work: (1) the annual agricultural P surplus, (2) the cumulative agricultural P surplus, and (3) phosphorus use efficiency. The methods for calculating these metrics are described below.

6.2.2.1 Agricultural surplus

The agricultural phosphorus surplus was calculated at the grid resolution as the downscaled phosphorus mass balance data. The agricultural phosphorus surplus, $SURP_{ag}$ (kg-P ha^{-1} y^{-1}), is defined as a simple difference between P inputs (fertilizer, livestock manure) and P outputs (crop and pasture P uptake), as follows:

$$SURP_{ag}(i,t) = FERT_f(i,t) + MAN(i,t) - gCROP(i,t) \qquad \text{(Eq. 6.1)}$$

where $FERT_f$ is farm inorganic phosphorus fertilizer for grid cell *i* at time *t*, *MAN* is phosphorus from livestock excretions for grid cell *i* at time *t*, and *CROP* is the phosphorus removed through crop harvest and consumption on pastureland by grazing livestock for grid cell *i* at time *t*. All reported as kg-P ha^{-1} yr^{-1}.

6.2.2.2 Cumulative agricultural phosphorus surplus

The cumulative phosphorus surplus, $SURP_{CUMU}$ *(kg ha^{-1} y^{-1})*, was calculated as the sum of agricultural surplus from 1980 to 2017:

$$SURP_{cumu}(i) = \sum SURP_{ag}(i,t) \quad \text{(Eq. 6.2)}$$

Cumulative phosphorus surplus is an upper bound estimate of the total legacy phosphorus stores and does not account for the phosphorus loss through erosion and riverine export. Estimating the temporal riverine export at the national scale across long timescales is complex and has yet to be accomplished. For a single year, we found a median ratio of export to total surplus for 400 watersheds of 14% (Oelsner, Sprague, Murphy, Zuellig, Johnson, Ryberg, Falcone, Stets, Vecchia, Riskin, Cicco, et al., 2017) (**Supplemental Figure 6.1**). However, because we cannot easily parse the proportion of riverine export due to domestic inputs (wastewater, food waste, and detergent), estimating how much agricultural phosphorus is lost in these systems remains challenging.

6.2.2.3 Phosphorus use efficiency

Annual phosphorus use efficiency (PUE) was calculated to assess the effectiveness of phosphorus use in agricultural landscapes over time. It should be noted that PUE can be calculated in various ways depending on the context (Biswas Chowdhury & Zhang, 2021; Cassman et al., 2002; Chang et al., 2021). In our study, we adopt an approach commonly used at larger scales, where agricultural PUE is defined as the ratio of phosphorus yield to inputs on agricultural land (Lassaletta et al., 2014; Swaney & Howarth, 2019b; Van Meter et al., 2021), which is the following approach:

$$PUE(i,t) = \frac{gCROP(i,t)}{gFERT(i,t) + gMAN(i,t)} \quad \text{(Eq. 6.3)}$$

where *PUE(i,t)* is the phosphorus use efficiency for a given grid cell *i* in year *t*, gCROP(i, t) is gridded phosphorus uptake from crops, and *gFERT(i, t)* and *gMAN(i, t)* represent the phosphorus input from gridded fertilizers and manure inputs, respectively. PUE values range from 0 to infinity. A PUE lower than 1 indicates a system where not all applied phosphorus is used by crops, a PUE of 1 signifies a balanced system where crop uptake equals the applied phosphorus, and a PUE greater than 1 implies a system where crops are removing more phosphorus than is being applied.

Gridded PUE values across the U.S. were aggregated to the HUC 2 scale using the median. Subsequently, the 18 HUC2 watersheds were grouped into 9 regions based on similarities in PUE and component trajectories. Gridded PUE values across the U.S. were aggregated to the HUC 2 scale using the median. We then aggregated the 18 HUC2 watersheds into 9 regions based on similarities between PUE and component trajectories.

6.3 Results and Discussion

6.3.1 Agricultural phosphorus usage across the contiguous United States

Agricultural phosphorus fluxes across the U.S. have changed over the last century, largely due to the industrialization of agriculture **(Figure 6.1).** The spatiotemporal trends in phosphorus use across the contiguous U.S. are characterized by three important features: (1) the initial rise and subsequent plateauing of phosphorus fertilizer use, (2) the steady increase of crop phosphorus uptake attributed to advances in agricultural technology, and (3) the development of regional manure hotspots due to concentrated agricultural practices.

Over the decades, changes in fertilizer use and crop uptake have been brought about by the industrialization of agricultural systems. From 1930 to 1980, the amount of fertilizer used increased from 0.68 to 8.7 kg-P ha^{-1} y^{-1} in the Corn Belt, Mississippi Alluvial Plain, and the California Central Valley (**Figure 6.1a-c, j**). The spatial and temporal patterns of crop uptake and fertilizer are tightly coupled, meaning regions with high fertilizer use typically exhibit high levels of crop uptake (**Figure 6.1g-i)**. Although fertilizer inputs have plateaued since the 1980s, crop phosphorus uptake has continued to rise. Nationally, median crop uptake nearly tripled crop phosphorus uptake since the 1930s from 4.8 to 12.8 kg-P ha^{-1} y^{-1} in 2017 (**Figure 6.1r**), driven in large part by technological advancements, plant genetics, and the removal of resource limitations through the use of irrigation and agrochemicals (Kucharik et al., 2020; Kucharik & Ramankutty, 2005).

Meanwhile, median manure phosphorus input has steadily increased from 4.1 kg-P ha^{-1} y^{-1} in 1930 to 6.5 kg-P ha^{-1} y^{-1} in 2017 **(Figure 6.1k)**. We see pronounced regional hotspots with manure inputs, driven by the transition towards concentrated animal feeding operations, where a large livestock population is concentrated in small areas (**Figure 6.1d-f**). This is evidenced by the doubling of the 95th percentile magnitude (15.9 to 31.0 kg-P ha^{-1} y^{-1} from 1980 to 2017), suggesting that most of the mass accretion occurs in concentrated patches of agricultural land (Byrnes et al., 2020a). Often, manure generated within these concentrated operations is applied in excess to crop demands to fields nearby, which can lead to the nutrients greatly exceeding

crop demands, contributing to the growing legacy stores (Kleinman et al., 2019; Long et al., 2018; Mullins, 2009).

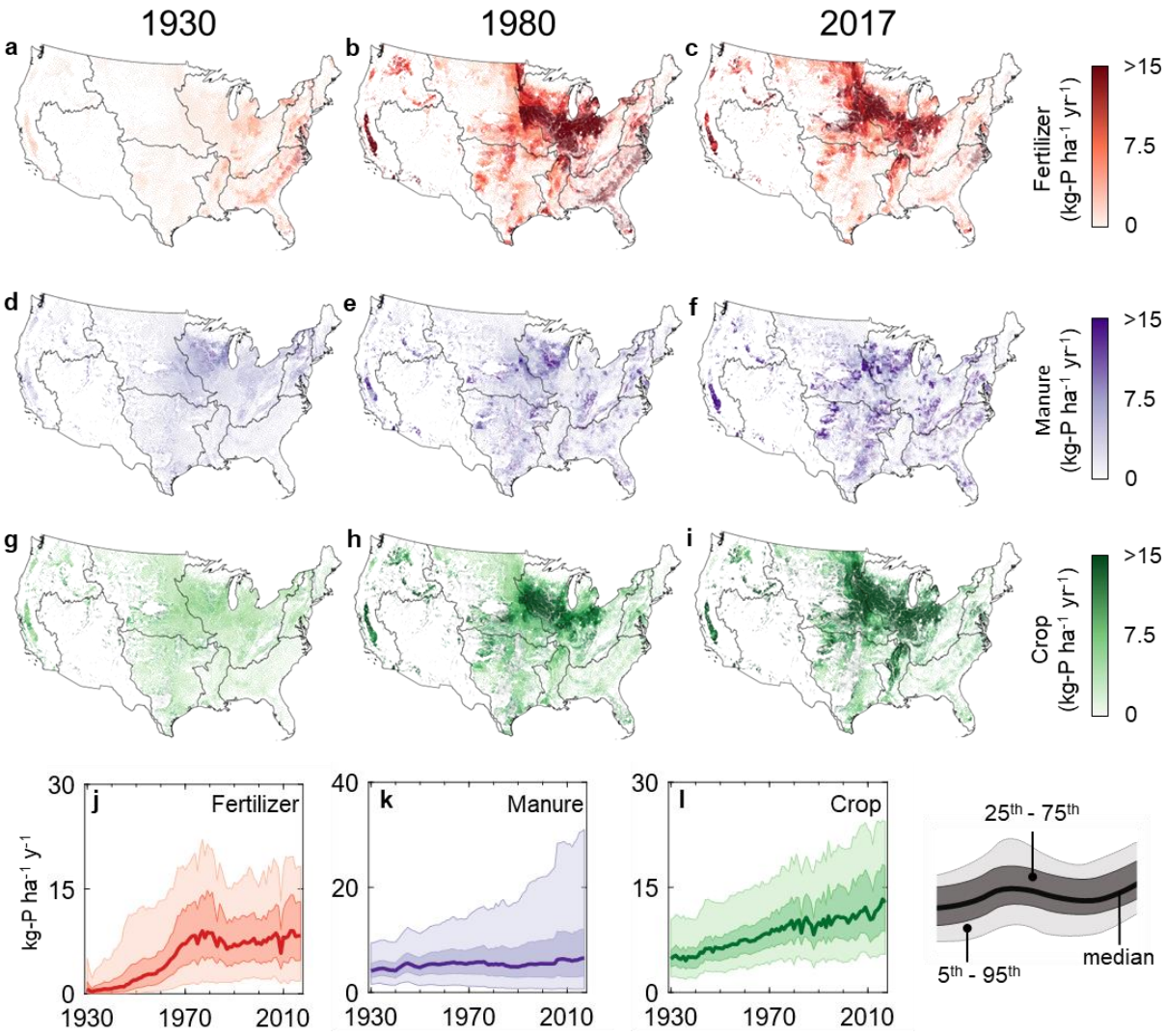


Figure 6.1 Maps of the major components for the agricultural phosphorus mass balance: (a-c) fertilizer P, (d-f) manure P, and (g-i) crop phosphorus uptake, all in kg ha^{-1} y^{-1}. Each map shows three time periods to capture the start and the end of the time series, as well as the peak around 1980, corresponding with the highest national median of fertilizer use. Time series panels (s-w) show the different percentiles, with the darkest color as the median value, the medium shade shows the interquartile range (25th-75th percentiles), and the lightest shade outlines the broader distribution from the 5th to the 95th percentiles.

Indicators are used as tools to monitor and communicate the status or sustainability of environmental systems over time (Goyette et al., 2016b; Lun et al., 2017; G. K. MacDonald et

al., 2011; Swaney & Howarth, 2019b; J. Wang et al., 2022; Zou et al., 2022). We use two indicators to explore the past and current state of phosphorus use across the U.S.: (1) agricultural phosphorus surplus and (2) phosphorus use efficiency. Annual agricultural phosphorus surplus, defined as the difference between manure and fertilizer inputs and crop phosphorus removal, is an effective indicator to track these spatiotemporal changes to nutrient mass inflow and outflows in agricultural systems. Phosphorus use efficiency (PUE), defined here as crop phosphorus removed per unit of applied phosphorus, is another widely used agronomic indicator (G. K. MacDonald et al., 2011; Swaney & Howarth, 2019a; Zou et al., 2022), providing insights into the efficiency of crop system and soil health (Schröder et al., 2011; Zou et al., 2022). We complement these indicators with the spatiotemporal trajectories of the three features – fertilizer, manure, and crop uptake – to understand the drivers of the indicators.

Agricultural phosphorus surplus was 0.25 kg-P ha^{-1} y^{-1} in 1930 and peaked at 5.8 kg-P $ha^{-1}y^{-1}$ in 1976, mirroring the increases in agricultural fertilizer use (**Figure 6.2g)**. Post-1980, however, the median phosphorus surplus decreased due to plateauing fertilizer and increasing crop uptake, settling at 2.6 kg-P $ha^{-1}y^{-1}$ in 2017. In regions like the Midwest and the Mississippi Alluvial Plain, the phosphorus surplus has dropped below zero, indicating a transition to extracting phosphorus from existing soil stores (**Figure 6.2f**). Negative surplus values suggest that crops are using soil phosphorus, which may adversely affect soil health. However, the background phosphorus content in the parent material of U.S. soils is estimated to be substantial, ranging from 1000 to 3000 kg-P ha^{-1}, providing sufficient background phosphorus to maintain crop yields in cases of negative phosphorus surplus (X. Yang et al., 2013). While much of this background phosphorus is not immediately available for plant uptake, field studies by Zhang et al. (2004) found that unfertilized plots were able to maintain crop yields for 10 years without depleting labile phosphorus (M-3 extractable soil P), as crop residues and soil organic phosphorus were mineralized into plant-available forms. PUE trajectory has mirrored the shifting inputs trajectories, increasing crop uptake and changing surplus. At a national scale, the PUE decreased from 0.96 in 1930 to 0.64 in 1980, coinciding with the increase in inorganic fertilizer inputs (**Figure 6.1j)**. However, with improvements in fertilizer management and increasing crop yields, PUE has seen an upturn, increasing to 0.83 by 2017.

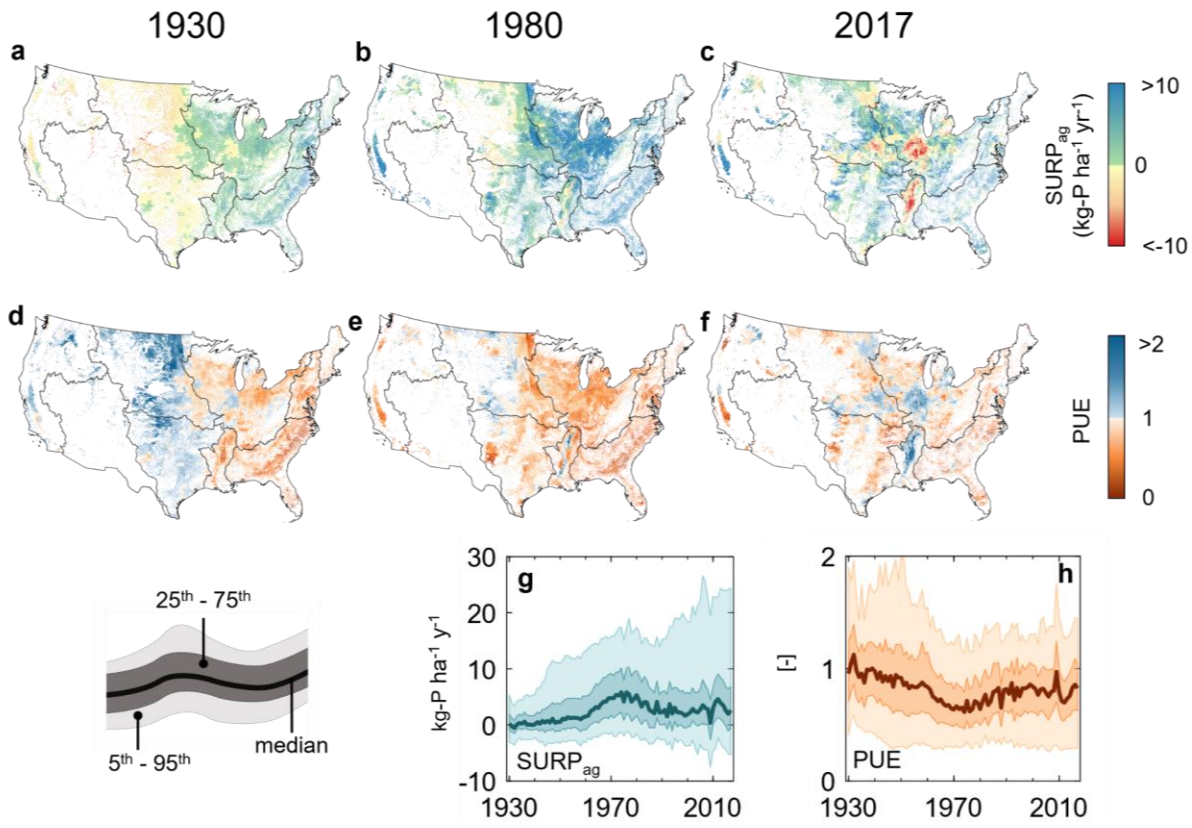


Figure 6.2 Maps of (a-c) agricultural phosphorus surplus, and (d-f) phosphorus use efficiency. Each map is divided showing three time periods to capture the start and the end of the timeseries, as well as the peak around 1980 corresponding with the highest national median of fertilizer use. Time series panels (g-h) show the different percentiles, with the darkest color as the median value, the medium shade shows the interquartile range (25th-75th percentiles), and the lightest shade outlines the broader distribution from the 5th to the 95th percentiles.

The trajectories of phosphorus surplus and PUE across the U.S. suggest a path toward recovery from the peak of excess phosphorus use. However, these indicators at this scale are not without limitations and are restricting the insights they can provide. First, agricultural surplus and PUE as indicators are scale-dependent. National trends show recovering trajectories and signs of improving nutrient management, however, we see very different PUE trajectories across regions (**Figure 6.2a-c**). The differences are largely driven by the complex interplay of regional agricultural practices, land use, and climate (Dimitri et al., 2005; Glibert, 2020; Spiegal et al., 2020). For instance, when comparing regions, stark differences in inorganic fertilizer management emerge: the eastern U.S. has reduced its fertilizer use, while the western region has not. In Region 6 we see a V-shaped PUE pattern mirroring the national trend (**Figure 6.3b**) because fertilizer use increased rapidly until the 1980s and then plateaued despite increasing crop

yields (**Figure 6.3e**). Whereas, the arid climate in Region 4 has traditionally restricted the land's suitability for row crop production, leading to predominantly low-input pasture and livestock grazing, which resulted in high PUE (Borchers & Nickerson, 2012). After the 1980s however, with the increasing use of irrigation, the region could support row crop agriculture resulting in increasing fertilizer use and decreasing PUE (**Figure 6.3a,d**) (Glibert, 2020).

The influence of manure on PUE reveals more nuanced interactions at the regional scale. In Region 7, manure has steadily risen since 1930 and has increased drastically since the 1990s from the rise of intensive livestock operations (Glibert, 2020). Therefore, despite decreasing fertilizer inputs, the region's low PUE is sustained by an increase in manure production without parallel increases in crop production (**Figure 6.3c,f**).

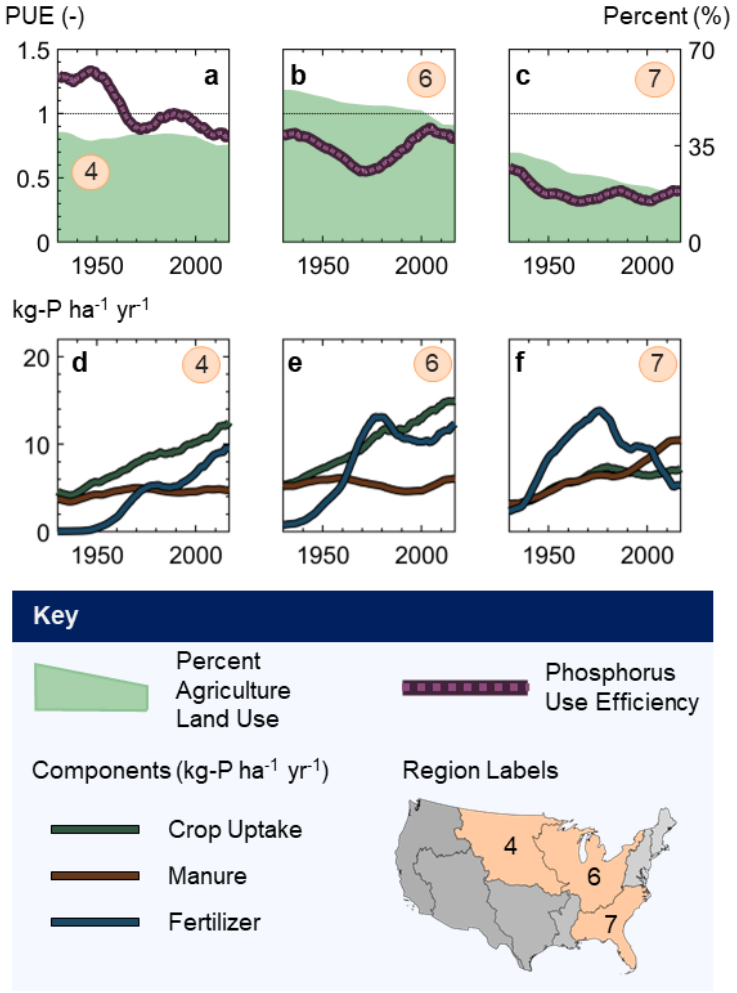


Figure 6.3 Phosphorus use efficiency and input for three distinct hydrological regions across the U.S. Panels (a-c) show the percent agricultural land and the median phosphorus use efficiency for each region. Panels (d-f) show the median trajectory of fertilizer and manure inputs and crop uptake for each region. Each region is denoted by its respective panel letter corresponding to the map legend. Figure with all regions is available as Supplemental Figure 6.2 and 6.3.

Given the long-established knowledge that phosphorus accumulates in soils over decades of intensive agriculture (Dolph et al., 2023; Pavinato et al., 2020; Sharpley et al., 2013; Van Meter et al., 2021), focusing solely on annual fluctuations of phosphorus surplus is inadequate, as it fails to capture the long-term *memory* of the landscape. Going beyond agricultural surplus and using cumulative surplus provides this essential historical context. We see the value of cumulative surplus in examining regions with negative surpluses. In 2017, 29% of U.S. agricultural land reported negative annual phosphorus surpluses, yet 61% of these areas have positive cumulative surpluses due to historical phosphorus overuse. Although a negative surplus indicates soil phosphorus mining, which can threaten long-term soil health, the cumulative surplus highlights that these areas likely draw on previously accumulated stores to sustain crop yields. Evidence of mining happening offers a pathway to potentially improve environmental outcomes, a promising perspective for future research and action.

Moreover, there are regional disparities in the magnitude of cumulative phosphorus stores. For example, in 1980, Region 4 had virtually no cumulative surplus (~ -1 kg ha^{-1}), while Regions 6 and 7 had substantial and increasing accumulated stores, reaching 286 and 578 kg-P ha^{-1} by 2017, respectively (**Figure 6.4a**). We can see that much of Region 4 remains with a negative surplus, but cumulative surplus in the eastern part of the region is intensifying and has similar magnitudes to Region 6.

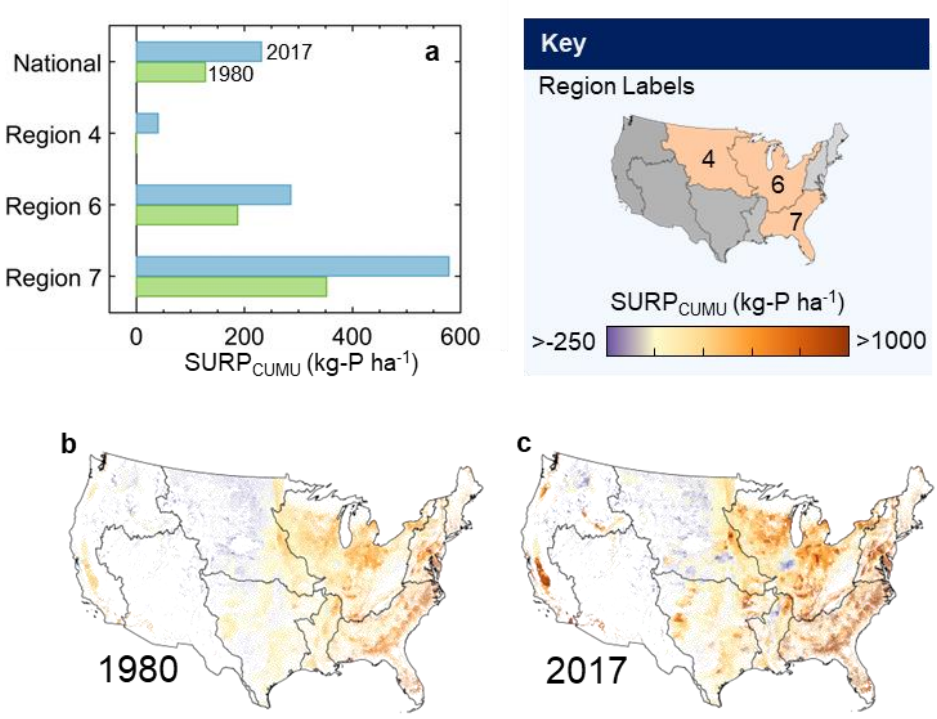


Figure 6.4 National and regional cumulative phosphorus surplus from 1930 - 1980 and 1930 - 2017. Cumulative phosphorus surplus can be used as an integrative metric for decades of phosphorus surplus as it provides historical context in assessing the current state of an agricultural system. Regions are aggregated HUC2 hydrological regions. Panel (a) captures the total agricultural area weighted mass of cumulative surplus at the national and regional scale, and panel (b-c) shows the spatial distribution of positive and negative cumulative surplus stores in 1980 and 2017. Bar charts for all regions are available as Supplemental Figure 6.4.

Lastly, relying solely on either surplus or PUE for assessments can oversimplify systems that, in reality, differ vastly in their efficiency or environmental impact. For example, PUE as a production indicator can quantify how effectively phosphorus inputs are used in crop production. However, as a relative metric, PUE does not quantify the residual surplus nutrients available to erode and contribute to water quality degradation (Schröder et al., 2011). A system with a PUE of 0.8 can be associated with vastly different scales of phosphorus use, from minor to significant surpluses dictated by the magnitude of inputs into the system. Therefore, while PUE provides essential context, leveraging multiple indicators can provide a more comprehensive state of the sustainability of the landscape.

6.3.2 Toward a holistic approach for landscape socio-environmental evaluation

Interpreting the state and progress of diverse landscapes through a single indicator poses a challenge in drawing broad conclusions about agricultural sustainability and environmental

health. To address the indicator limitations, we propose an integrated framework that takes into account both cumulative phosphorus surplus (1930—present) and phosphorus use efficiency. Integrating these two metrics allows us to consider long-term phosphorus-use magnitudes as well as the efficiency of use and likely future trajectories (**Figure 6.5**).

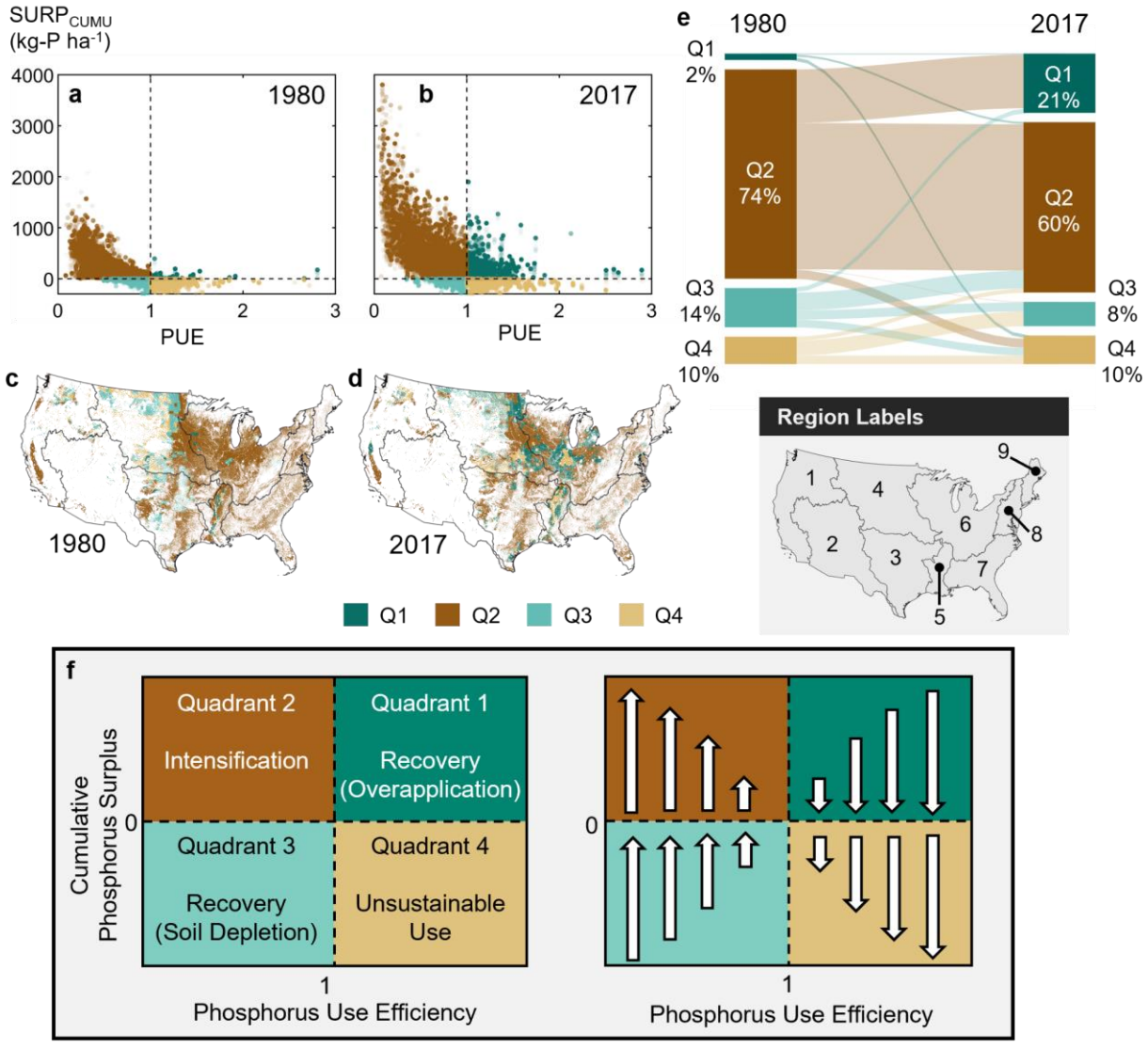


Figure 6.5 A composite diagram of the classification of agricultural land in the multi-indicator framework and how the quadrant classification changes over time. Panels a and c show the cumulative agricultural phosphorus surplus and phosphorus use efficiency capturing grid of agricultural land in 1980 and b, d in 2017. Panel e shows a Sankey diagram capturing the distribution and change of quadrant classification of agricultural land. Intensification remains the dominant quadrant, but has decreased from 1980 to 2017. F. Grid cells are categorized into four modes: Q1 Recovery (overapplication), Q2 Intensification, Q3 Recovery (soil depletion), and Q4

Unsustainable Use. The quadrant plot of cumulative phosphorus surplus and PUE demonstrates that most areas in 1980 are undergoing intensification with a PUE < 1 and a positive cumulation phosphorus surplus.

Within this framework, each indicator serves a specific purpose: cumulative surplus measures the total phosphorus stock in agricultural lands, representing the system's current state, while PUE acts as a directional vector, indicating whether the cumulative surplus is increasing or decreasing. A PUE further from 1 implies a faster rate of change in the cumulative stock (**Figure 6.5f)**. As shown in **Figure 6.5a-b**, using the cumulative phosphorus surplus and PUE as two different axes of analysis, we can identify four different states of landscape phosphorus history and management (**Figure 6.5f)**.

Quadrant 1 (Q1) we identify as a "Recovery" class, meaning that while there is a positive cumulative phosphorus surplus, PUE is greater than 1, indicating that accumulated phosphorus is being used to support crop growth and that no additional excess phosphorus is being added to the landscape. Quadrant 2 (Q2) is the "Intensification" class, characterized by a decreasing positive cumulative phosphorus surplus with a PUE of less than 1. This indicates that the system is accumulating phosphorus. Quadrant 3 (Q3), identified as "Recovery (Soil Depletion)," is marked by an increasing negative cumulative surplus driven by a PUE below 1 and suggests that the depleted systems are recovering from soil phosphorus mining. Lastly, we identify Quadrant 4 (Q4) as "Unsustainable Use," with a decreasing negative cumulative surplus and a PUE above 1, indicating that the area is mining phosphorus from the soil, leading to long-term soil health issues.

Our analysis spans two periods: 1980, coinciding with the peak of inorganic fertilizer usage, and 2017, reflecting current conditions. In 1980, the majority (74%) of U.S. agricultural land was situated in the Intensification quadrant (Q2), a result of phosphorus inputs exceeding crop phosphorus removal (**Figure 6.5c**). While the majority of agricultural land (60%) remains in the Intensification quadrant, 21% of the land has improved phosphorus management and shifted to the Recovery (Overapplication) quadrant (Q1), with a large portion of which originated in the Intensification quadrant (**Figure 6.5d**). Areas in the Recovery (Overapplication) quadrant (Q1) are effectively drawing down cumulative phosphorus stores. However, croplands with prolonged PUE above 1 risk depleting soil phosphorus cumulative stores and moving to Unsustainable Use (Q4) eventually which may result in mining background stores of phosphorus, threatening soil health and risking crop yield reductions (Brownlie et al., 2021; G. K. MacDonald et al., 2011). We see this occurring in Illinois and the Mississippi Alluvial Plain, where decades of steadily increasing soybean and corn yields and plateauing fertilizer have gradually depleted the

cumulative surplus stores (Sabo et al., 2021). Phosphorus in parent material in these regions can exceed 1000 kg-P ha^{-1} (X. Yang et al., 2013), thus, it is possible that phosphorus soil stores are unsustainably being used to support current crop production. Indeed, estimates of contemporary phosphorus content in the top 30 cm of soil in the Q4 quadrant Illinois remain elevated, ranging from 4,000 to 6,000 kg-P ha^{-1} (Dolph et al., 2023). We, however, do not account for the transport of manure from outside regions of high manure production (Spiegal et al., 2020). While these transportation routes are not ubiquitous, they may influence the surplus in some regions.

Lastly, we see a portion of the Recovery (Soil Depletion) moving to the Intensification quadrant due to expanding row crop agriculture in what was once low-productivity pasture land. While national fertilizer use decreased, the western regions are seeing increased use. These regions are intensifying as Recovery (Soil Depletion) shrinks between 1980 and 2017, with many of these areas in 1980 shifting to the Intensification quadrant.

We find that areas with proportionately more fertilizer input have shifted from the Intensification quadrant to the Recovery (Overapplication) quadrant. In 1980, Q1-3 had manure and inorganic fertilizer contributions in roughly equal proportions (percent manure of total inputs = 36-51%) (**Figure 6.6a**). With the decrease in inorganic fertilizer starting in the 1980s and the sustained or increase in manure inputs, we find that the areas in the *Intensification* quadrant (Q2) have manure as the predominant input (54% IQR 33-75%; Wilcoxon rank sum test p-value < 0.001). This suggests that phosphorus intensification is largely by livestock manure generation (**Figure 6.6b**). Thus, manure management remains underemphasized, particularly for areas in the *Intensification* quadrant (Q2), despite known links between livestock manure degrading water quality (Chadwick et al., 2011; Jones et al., 2019; Miralha et al., 2022; Secchi & Mcdonald, 2019). Phosphorus management must account for regional phosphorus dynamics and address inefficient phosphorus use by improving phosphorus recycling (Spiegal et al., 2020; Van Meter et al., 2021).

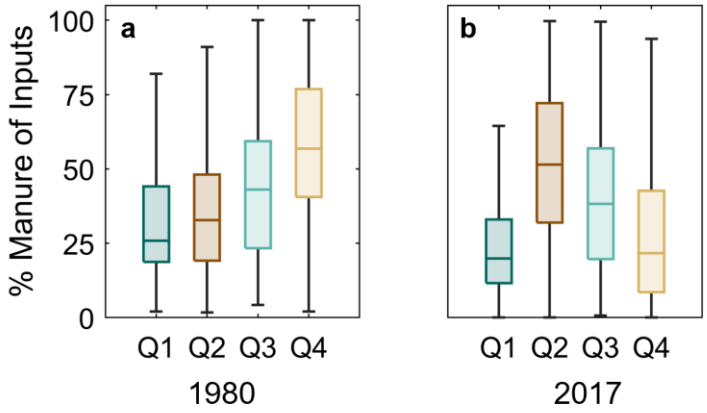


Figure 6.6 Box plot of the percent of the total inputs are from manure for (a) 1980 and (2) 2017 for each quadrant. The intensification is predominantly driven by inputs manure P, with the boxplot showing that the median grid cell in Q4 has manure accounting for over 54% of inputs. Assume that an ideal system has no cumulative phosphorus surplus and has a PUE of 1, as in this scenario, there are no cumulative stores and crops consume all the applied phosphorus.

6.4 Conclusion

Phosphorus inputs are crucial for food production to meet the demands of a growing population. Nevertheless, its current use threatens the ecological integrity of the fresh and coastal water systems we depend on. Quantifying the state of the sustainable use of phosphorus in agricultural systems frequently relies on indicators, including phosphorus surplus and use efficiency. Our analysis using these indicators reveals an evolution in agricultural phosphorus management across the U.S., characterized by the changing fertilizer management, the proliferation of concentrated animal operations, and the landscape's memory of past phosphorus use. With varied regional trajectories influenced by disparate drivers, including climate, fertilizer and manure inputs, and low agricultural land use.

Management relies on measurement; however, using metrics in isolation can oversimplify complex systems. To facilitate the assessment of phosphorus management in agricultural systems, we developed a framework using PUE and cumulative surplus. This framework enables a comprehensive understanding of the historical state and the current progress of phosphorus sustainability, which is essential for addressing long-term phosphorus use issues.

Agricultural land is divided into four specific states reflecting the rate of change and the state of accumulated phosphorus stores in the soils: Recovery (Overapplication), Intensification, Recovery (Soil depletion), and Unsustainable Use. We applied the framework across decades, showing the transition of 29.5% of U.S. agricultural land from the Intensification quadrant to the

Recovery (Overapplication) quadrant, indicating successful measures to mitigate excess fertilizer phosphorus. However, 70% of the agricultural area in the U.S. remains in the *Intensification* quadrant, where we find that these areas have disproportionately more of their inputs from manure, pointing to the need to treat manure as a resource instead of the current approach of treating it as a waste product. The findings underscore the importance of nuanced phosphorus management strategies to address the diverse challenges of different agricultural systems and prevent unintended environmental consequences.

Chapter 7 Conclusions and Perspectives

7.1 Summary of major findings

In my thesis, I aim to contribute to our understanding of the relationship between nutrient inputs and outputs and how different processes and watershed properties modulate these relationships.

In order to achieve this, Chapters 2 and 3 showcased the nitrogen and phosphorus mass balance datasets we developed, TREND and gTREND for nitrogen and phosphorus. These datasets quantify long-term anthropogenic nutrient fluxes, capturing the landscapes' nutrient sources and sinks. These datasets are novel because of their spatial and temporal resolution, covering the contiguous U.S. annually from 1930 to 2017 at the county and 250m gridded scales. Ultimately, these data allow for the investigation of the effects of the long-term source and sink fluxes in the landscape on various environmental outcomes.

Chapter 4 uses the gTREND-Nitrogen dataset to understand the relationship between watershed nitrogen inputs and exports. Decades of research have supported that nitrogen inputs into the watersheds drive load magnitudes, but see that changing nitrogen input does not always result in changes in watershed nitrogen export. Ultimately, we see that what drives nitrogen loads is different than what drives changes in nitrogen loads. We use over 400 watershed nitrogen inputs and load time series to investigate the relationship between net nitrogen inputs, or surplus, and loads. We placed watersheds into four typologies based on their input-load relationship. In general, we found that watersheds act as filters where different properties will change the way in which nitrogen is transported and processed. For instance, tile-drained agricultural watersheds tend to have shorter residence times and are more responsive to changes in net nitrogen inputs. Indeed, a few of these watersheds' are the only agricultural watersheds where we find decreasing loads. However, watersheds with properties that lead to longer subsurface residence times and loads remain less responsive to changing inputs, indicating legacy stores may be modulating the watershed response to changing net nitrogen inputs.

In Chapter 5, we aim to further characterize legacy nitrogen across the contiguous U.S. To achieve this, we use nitrogen retention--defined as the nitrogen stored in the soil organic nitrogen pool, the groundwater pool, or lost through denitrification--to better understand how much nitrogen is being temporarily (legacy) or permanently (denitrification) stored in the landscape. We built a machine learning model, trained on watershed inputs from gTREND-Nitrogen, climate, and land use characteristics to predict nitrogen load for over 2000 HUC8 watersheds,

covering the entire contiguous U.S. Together with inputs and predicted loads, we calculated the cumulative mass of retained nitrogen (1958-2017) and used this to investigate what drives retention. We find that agricultural regions in the U.S. have high magnitudes of cumulative retention but retain a lower proportion of their inputs. Meanwhile, watersheds with more input from manure or those with low inputs tend to retain a higher proportion of their inputs.

Typically, further parsing retention into temporary (legacy) and permanent (denitrification) removal of nitrogen is often impossible because of the uncertainty around legacy dynamics and denitrification rates. However, due to existing literature estimates of legacy storage in the Upper Mississippi region, we attempt to estimate the proportion of nitrogen entering each HUC8 watershed stored in the SON, groundwater pool, or lost to denitrification. We estimate that across all HUC8 watersheds, a median of 42% of nitrogen retained in the soil is stored as soil organic nitrogen, 16.5% in groundwater, and 40% is lost to denitrification. These estimates can help constrain process-based models used to understand legacies in the Upper Mississippi region.

In my final chapter, I shifted from nitrogen to phosphorus. I explored current and historical agricultural phosphorus use and management in landscapes within the context of both food security and environmental health. We use the agricultural components from the newly developed gTREND-Phosphorus to characterize the extent of phosphorus depletion and excess. We use annual and cumulative phosphorus surplus and phosphorus use efficiency (PUE) to achieve this. We find that regional phosphorus inputs and crop export trajectories are influenced by changing fertilizer management, densification of livestock operations, and the landscape's memory of past phosphorus use. We use cumulative phosphorus surplus and PUE for each agricultural land grid-cell in a multi-indicator framework, which characterizes how phosphorus is being managed and where are areas of concern for phosphorus excess or limitations. Between 1980 and 2017, we found that nearly 30% of the agricultural land has moved from a state of overuse of phosphorus to a state of recovery by depleting cumulative stores. However, 70% of the agricultural land is still accumulating phosphorus. These areas also correspond with areas with more of their inputs originating from manure inputs, indicating that manure is being treated as a waste product instead of a resource used to meet crop nutrient needs.

By leveraging novel datasets, we can use a data-driven approach to explore nutrient use across space and time and its impact on environmental outcomes. I have made significant contributions towards understanding the magnitude and distribution of cumulative net nutrient input stores in the landscape, as well as the ways in which intrinsic watershed properties, climate, land management, and historical nutrient use can modulate the relationship between nutrient inputs

and export. Our findings underscore the importance of nuanced, place-based, and context-dependent nutrient management strategies, focusing on manure management, to address the diverse challenges of different agricultural systems and prevent unintended environmental consequences.

7.2 Recommendations for future work

My thesis findings aim to improve our understanding of legacy dynamics by contributing to the overall characterization of nutrient accumulation in the subsurface. It advances our understanding of the role of land use, nutrient management, and climate on accumulation and retention dynamics and how these factors may ultimately impact nitrogen export. The applied nature of my work highlights two critical pathways for future research aimed at refining our understanding and using the existing knowledge to implement targeted approaches to improve our water quality future.

Improve publicly-available datasets

Data-driven research is empirical and relies on available datasets to investigate correlations. At the national scale, a large disparity remains between available data and the data required to validate findings. In this study, we encountered instances of negative nitrogen or phosphorus surpluses. While negative surpluses are not necessarily indicative of estimation errors—particularly in areas with soils that have high background nutrient concentrations—the lack of reliable data on these background concentrations limits our ability to validate some of these surpluses.

Additionally, there is limited national-scale data on various watershed properties. There are known drivers of nutrient pollution and legacy effects investigated in previous studies, but we cannot fully assess their influence on water quality patterns due to data scarcity. For example, the size of concentrated animal operations is known to significantly impact water quality (Miralha et al., 2022). However, the national-scale location data for these operations is not publicly available, preventing us from considering their impact in this context. Thus, greater nuance and accuracy could be achieved in future analyses with improved access to open-source data.

Balancing complexity to accommodate large spatial and temporal scales

My research spans long-time scales and large spatial scales. As models move to cover larger scales, it is common for simplifications to be used to balance reliability and complexity. Creating

datasets at the contiguous U.S. scale requires simplification across many aspects, including model parameters and processes needed to build a flexible and reliable mass-balance model. While increasing model complexity may seem desirable to improve accuracy, we may see decreases in precision with the addition of new processes due to errors associated with the heterogeneity in processes across spatial and temporal scales and a lack of adequate data to model these complex processes. For example, future dataset iterations could benefit from improved characterization or reduced uncertainty in:

- the spatial and temporal differences in crop nutrient concentrations;
- the spatial and temporal differences in livestock excretion parameters due to feed type and use of supplements;
- the locations and rate of nutrient removal from pasture grazing;
- the agricultural fields that receive manure from nearby livestock operations (Spiegal et al., 2020);
- the rate of biological nitrogen fixation for legumes;
- volatilization and denitrification gas losses;
- soil nitrogen and phosphorus retention mechanisms, rates, and stores;
- inputs into groundwater;
- natural biological nitrogen fixation

Furthermore, we also take a parsimonious approach to understanding nutrient dynamics and drivers. While my research focuses on nitrogen and phosphorus in silos, nitrogen, phosphorus, and carbon cycles are tightly coupled, and the presence of each other modulates their dynamics. However, such simplifications are needed to understand nutrient dynamics at large spatial and long temporal scales. Working to integrate the biogeochemical cycles would likely offer further insights into nutrient and legacy dynamics.

System-level planning for local solutions

Certain watersheds in the U.S., such as the Chesapeake Bay (ex. Chang et al., 2022; Van Meter et al., 2017; Zhang et al., 2023) and the Mississippi River Basin (ex. Stackpoole et al. 2021; Van Meter et al. 2017), are better characterized because of their regional importance. Not all watersheds have been as well studied, and we need improved characterization of the dominant sources of nutrient legacies. Our work was able to identify some of these drivers. For instance, at the national scale, we found that crop nutrient needs are often met by inorganic fertilizer, while nutrients from manure are often in excess of crop nutrient needs. To address this challenge, we

must continue and prioritize stakeholder-driven research to develop feasible, system-level strategies for optimizing manure use to meet crop nutrient demands. The concept of "manuresheds" offers a promising approach to redistributing manure more effectively between areas of surplus and deficit.

Improving water quality requires long-term thinking and sustained effort. However, it is challenging to determine whether implemented conservation interventions are ineffective or need more time to see improvements (Basu et al., 2022; Jarvie et al., 2013). A system-level approach that delivers both short- and long-term benefits can minimize skepticism and disillusionment among stakeholders about the efficacy of interventions due to long timelines. For example, nutrient management and cover crops effectively minimize excess nutrients in the root zone. However, their full benefits may take decades to materialize due to the persistent legacy of nutrients below the root zone. To accelerate progress, targeted interventions can be placed in watersheds with shorter travel times to achieve faster reductions in nutrient loads. We must also implement interventions that can intercept legacy fluxes or use legacy nutrient pools as a resource. This might include remediating reactive interfaces such as wetlands or floodplains to capture legacy-laden groundwater or using nitrogen-rich groundwater for irrigation, which can reduce lag times by preventing legacy nutrients from reaching rivers and surface water bodies.

References

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, *5*(1), 1–12.

Aegerter, B., Smith, R., Gaskell, M., & Rilla, E. (2013). *Pumpkin Production in California*. https://anrcatalog.ucanr.edu/pdf/7222.pdf

Alewell, C., Ringeval, B., Ballabio, C., Robinson, D. A., Panagos, P., & Borrelli, P. (2020). Global phosphorus shortage will be aggravated by soil erosion. *Nature Communications*, *11*(1), 4546.

Alexander, R. B., & Smith, R. A. (1990). *County-level estimates of nitrogen and phosphorus fertilizer use in the United States, 1945 to 1985* (- ed.). https://doi.org/10.3133/ofr90130

Ascott, M. J., Gooddy, D. C., Fenton, O., Vero, S., Ward, R. S., Basu, N. B., Worrall, F., Van Meter, K. J., & Surridge, B. W. J. (2021). The need to integrate legacy nitrogen storage dynamics and time lags into policy and practice. *The Science of the Total Environment*, 146698.

Ascott, M. J., Gooddy, D. C., Wang, L., Stuart, M. E., Lewis, M. A., Ward, R. S., & Binley, A. M. (2017). Global patterns of nitrate storage in the vadose zone. In *Nature Communications* (Vol. 8, Issue 1). https://doi.org/10.1038/s41467-017-01321-w

Ashley, K., Cordell, D., & Mavinic, D. (2011). A brief history of phosphorus: from the philosopher's stone to nutrient recovery and reuse. *Chemosphere*, *84*(6), 737–746.

Backer, H., Leppänen, J.-M., Brusendorff, A. C., Forsius, K., Stankiewicz, M., Mehtonen, J., Pyhälä, M., Laamanen, M., Paulomäki, H., Vlasov, N., & Haaranen, T. (2010). HELCOM Baltic Sea Action Plan--a regional programme of measures for the marine environment based on the Ecosystem Approach. *Marine Pollution Bulletin*, *60*(5), 642–649.

Basu, N. B., Thompson, S. E., & Rao, P. S. C. (2011). Hydrologic and biogeochemical functioning of intensively managed catchments: A synthesis of top-down analyses. *Water Resources Research*, *47*(10). https://doi.org/10.1029/2011wr010800

Basu, N. B., Van Meter, K. J., Byrnes, D. K., Van Cappellen, P., Brouwer, R., Jacobsen, B. H., Jarsjö, J., Rudolph, D. L., Cunha, M. C., Nelson, N., Bhattacharya, R., Destouni, G., & Olsen, S. B. (2022). Managing nitrogen legacies to accelerate water quality improvement. *Nature Geoscience*, *15*(2), 97–105.

Batool, M., Sarrazin, F. J., Attinger, S., Basu, N. B., Van Meter, K., & Kumar, R. (2022). Long-term annual soil nitrogen surplus across Europe (1850–2019). *Scientific Data*, *9*(1), 1–22.

Beaulieu, J. J., DelSontro, T., & Downing, J. A. (2019). Eutrophication will increase methane emissions from lakes and impoundments during the 21st century. *Nature Communications*, *10*(1), 1375.

Belguerri, H., Villar, J. M., Pascual, M., Fatmi, A., Amadeu, A., & Rufat, J. (2016). A Proposal of Nitrogen Balance in a Very High Density Olive Orchard. *Journal of Fundamental and Applied Sciences*, *8*(2), 639–654.

Belitz, K., & Stackelberg, P. E. (2021). Evaluation of six methods for correcting bias in estimates from ensemble tree machine learning regression models. *Environmental Modelling & Software*, *139*, 105006.

Bernhardt, E. S., Band, L. E., Walsh, C. J., & Berke, P. E. (2008). Understanding, managing, and minimizing urban impacts on surface water nitrogen loading. *Annals of the New York Academy of Sciences*, *1134*, 61–96.

Biswas Chowdhury, R., & Zhang, X. (2021). Phosphorus use efficiency in agricultural systems: A comprehensive assessment through the review of national scale substance flow analyses. *Ecological Indicators*, *121*, 107172.

Borchers, A., & Nickerson, C. J. (2012). *How is land in the United States used? A focus on agricultural land*. https://www.ers.usda.gov/amber-waves/2012/march/data-feature-how-is-land-used/

Bottoms, T. G., Hartz, T. K., Cahn, M. D., & Farrara, B. F. (2013). Crop and soil nitrogen dynamics in annual strawberry production in California. *HortScience: A Publication of the American Society for Horticultural Science*, *48*(8), 1034–1039.

Bouraoui, F., & Grizzetti, B. (2011). Long term change of nutrient concentrations of rivers discharging in European seas. *The Science of the Total Environment*, *409*(23), 4899–4916.

Bouwman, A. F., Beusen, A. H. W., & Billen, G. (2009). Human alteration of the global nitrogen and phosphorus soil balances for the period 1970-2050. *Global Biogeochemical Cycles*, *23*(4). https://doi.org/10.1029/2009GB003576

Bouwman, A. F., Van Drecht, G., & Van der Hoek, K. W. (2005). Global and regional surface nitrogen balances in intensive agricultural production systems for the period 1 1/4—21/431/4. *Pedosphere*, *15*(2), 137–155.

Bouwman, A. F., Van Drescht, G., & Van der Hoek, K. W. (2005). Global and Regional Surface Nitrogen Balances in Intensive Agricultural Production Systems for the Period 1970-2030. *Pedosphere*, *15*(2), 135–155.

Boyer, E. W., Goodale, C. L., Jaworski, N. A., & Howarth, R. W. (2002). Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A. In *Biogeochemistry* (Vol. 57, pp. 137–169). Springer. https://doi.org/10.2307/1469687

Brakebill, J. W., & Gronberg, J. M. (2017). *County-Level Estimates of Nitrogen and Phosphorus from Commercial Fertilizer for the Conterminous United States, 1987-2012*. U.S. Geological Survey. https://doi.org/10.5066/F7H41PKX

Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., … Zhang, J. (2018). Declining oxygen in the global ocean and coastal waters. *Science*, *359*(6371). https://doi.org/10.1126/science.aam7240

Brownlie, W. J., Sutton, M. A., Reay, D. S., Heal, K. V., Hermann, L., Kabbe, C., & Spears, B. M. (2021). Global actions for a sustainable phosphorus future. *Nature Food*, *2*(2), 71–74.

Byrnes, D. K., Van Meter, K. J., & Basu, N. B. (2020a). Long-term shifts in U.S. nitrogen sources and sinks revealed by the new TREND-nitrogen data set (1930–2017). *Global Biogeochemical Cycles*, *34*(9), e2020GB006626.

Byrnes, D. K., Van Meter, K. J., & Basu, N. B. (2020b). Long-term shifts in U.S. nitrogen sources and sinks revealed by the new TREND-nitrogen dataset (1930-2017). *Global Biogeochemical Cycles*. https://doi.org/10.1029/2020GB006626

Byrnes, D. K., Van Meter, K. J., & Basu, N. B. (2020c). Long-term shifts in U.S. nitrogen sources and sinks revealed by the new TREND-nitrogen data set (1930–2017). *Global Biogeochemical Cycles*, *34*(9). https://doi.org/10.1029/2020gb006626

Byrnes, D. K., Van Meter, K. J., & Basu, N. B. (2020d). Trajectories nutrient dataset for nitrogen (TREND-nitrogen). *PANGAEA Https://doi. org/10.1594/PANGAEA*. https://scholar.google.ca/scholar?cluster=14185913947556576 63&hl=en&as_sdt=0,5&sciodt=0,5

Byrnes, D. K., Van Meter, K. J., & Basu, N. B. (2022). *Trajectories Nutrient Dataset for Nitrogen (TREND-Nitrogen) v2* [Dataset]. https://doi.org/10.6084/m9.figshare.20915989.v1

California State Water Resources Control Board. (2017). *Nitrate Project*. https://www.waterboards.ca.gov/water_issues/programs/nitrate_project/

Cao, P., Lu, C., & Yu, Z. (2018). Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: application rate, timing, and fertilizer types. *Earth System Science Data*, *10*(2), 969–984.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Applications*, *8*(3), 559–568.

Cassman, K. G., Dobermann, A., & Walters, D. T. (2002). Agroecosystems, nitrogen-use efficiency, and nitrogen management. *Ambio*, *31*(2), 132–140.

Chadwick, D., Sommer, S., Thorman, R., Fangueiro, D., Cardenas, L., Amon, B., & Misselbrook, T. (2011). Manure management: Implications for greenhouse gas emissions. *Animal Feed Science and Technology*, *166-167*, 514–531.

Chang, S., Byrnes, D., Van Meter, K., & Basu, N. (n.d.). *Gridded TRajEctories Nutrient Dataset for Nitrogen (gTREND-N): Long-Term Nitrogen Mass Balance Data for the Contiguous United States (1930-2017)*.

Chang, S. Y., Zhang, Q., Byrnes, D. K., Basu, N. B., & Van Meter, K. J. (2021). Chesapeake legacies: the importance of legacy nitrogen to improving Chesapeake Bay water quality. *Environmental Research Letters: ERL [Web Site]*, *16*(8), 085002.

Chen, D., Huang, H., Hu, M., & Dahlgren, R. A. (2014). Influence of Lag Effect, Soil Release, And Climate Change on Watershed Anthropogenic Nitrogen Inputs and Riverine Export Dynamics. *Environmental Science & Technology*, *48*(10), 5683–5690.

Chen, D., Hu, M., & Dahlgren, R. A. (2014). A dynamic watershed model for determining the effects of transient storage on nitrogen export to rivers. *Water Resources Research*, *50*(10), 7714–7730.

Chen, D., Shen, H., Hu, M., Wang, J., Zhang, Y., & Dahlgren, R. A. (2018). Chapter Five - Legacy Nutrient Dynamics at the Watershed Scale: Principles, Modeling, and Implications. In D. L. Sparks (Ed.), *Advances in Agronomy* (Vol. 149, pp. 237–313). Academic Press.

Chen, Y., Shen, H., Shih, J.-S., Russell, A. G., Shao, S., Hu, Y., Odman, M. T., Nenes, A., Pavur, G. K., Zou, Y., Chen, Z., Smith, R. A., Burtraw, D., & Driscoll, C. T. (2020). Greater contribution from agricultural sources to future reactive nitrogen deposition in the United States. *Earth's Future*, *8*(11). https://doi.org/10.1029/2019ef001453

Chesapeake Bay Program. (1988). *Chesapeake Bay Agreement*. https://www.chesapeakebay.net/content/publications/cbp_12510.pdf

Chesapeake Bay Program. (2020). *WIP Water Quality Standards Attainment & Monitoring Outcomes Management Strategy*. https://www.chesapeakebay.net/what/goals/water-quality

Ciampitti, I. A., & Vyn, T. J. (2012). Physiological perspectives of changes over time in maize yield dependency on nitrogen uptake and associated nitrogen efficiencies: A review. *Field Crops Research*, *133*, 48–67.

Conant, R. T., Berdanier, A. B., & Grace, P. R. (2013). Patterns and trends in nitrogen use and nitrogen recovery efficiency in world agriculture. *Global Biogeochemical Cycles*, *27*(2), 558–566.

Conley, D. J., Carstensen, J., Aigars, J., Axe, P., Bonsdorff, E., Eremina, T., Haahti, B.-M., Humborg, C., Jonsson, P., Kotta, J., Lännegren, C., Larsson, U., Maximov, A., Medina, M. R., Lysiak-Pastuszak, E., Remeikaité-Nikiené, N., Walve, J., Wilhelms, S., & Zillén, L. (2011). Hypoxia is increasing in the coastal zone of the Baltic Sea. *Environmental Science & Technology*, *45*(16), 6777–6783.

Cordell, D., Drangert, J.-O., & White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change: Human and Policy Dimensions*, *19*(2), 292–305.

Dagan, S. (1987). Nitrogen fertilization in citrus orchards. *Plant and Soil*, *9*(1770), 1–9.

David, M. B., Drinkwater, L. E., & McIsaac, G. F. (2010). Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, *39*(5), 1657–1667.

Davidson, E. A., David, M. B., Galloway, J. N., Goodale, C. L., Haeuber, R., Harrison, J. A., Howarth, R. W., Jaynes, D. B., Lowrance, R. R., Thomas, N. B., & Others. (2011). Excess Nitrogen in the U.S. Environment: Trends, Risks, and Solutions. *Issues in Ecology*, *15*. https://pubs.er.usgs.gov/publication/70032270

DeCicco, L., Hirsch, R., Lorenz, D., Watkins, D., & Johnson, M. (2021). *dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services* (Version v.2.7.7). U.S. Geological Survey. https://doi.org/10.5066/P9X4L3GE

DeCicco, L., Hirsch, R., Lorenz, D., Watkins, D., & Johnson, M. (2024). *dataRetrieval: R packages for discovering and retrieving water data available from U.S. federal hydrologic web services* (Version 2.7.15). U.S. Geological Survey. https://doi.org/10.5066/P9X4L3GE

De Moura R., G. J. C. J. C. W. A. K. K. K. K. &. (n.d.). *Sample costs to establish an orchard and produce sweet cherries*. University of California Cooperative Extension. Retrieved November 14, 2023, from https://coststudyfiles.ucdavis.edu/uploads/cs_public/6f/9b/6f9b0a93-163b-4060-ba94-8a3fc689e97d/cherryvn2011.pdf

Destouni, G., Fischer, I., & Prieto, C. (2017). Water quality and ecosystem management: Data-driven reality check of effects in streams and lakes. In *Water Resources Research* (Vol. 53, Issue 8, pp. 6395–6406). https://doi.org/10.1002/2016wr019954

Dewitz, J. , and U.S. Geological Survey. (2022). *National Land Cover Database (NLCD) 2019 products (ver. 2.0, June 2021)*. https://www.sciencebase.gov/catalog/item/5f21cef582cef313ed940043

Diaz, Selman, & Chique-Canache. (2017). *Global Eutrophic and Hypoxic Coastal Systems: Eutrophication and Hypoxia--Nutrient Pollution in Coastal Waters*. World Resources Institute. https://wedocs.unep.org/bitstream/handle/20.500.11822/19229/eutrophication_and_hypoxia_in_coastal_areas.pdf?sequence=1

Dimitri, C., Effland, A., & Conklin, N. (2005). *The 20th Century Transformation of U.S. Agriculture and Farm Policy The 20th Century and Farm Policy: Economic Information Bulletin Number 3* (p. 14).

http://www.ers.usda.gov/publications/EIB3/eib3.pdf

Doerge, T., Roth, R. L., & Gardener, B. R. (1991). *Nitrogen Fertilizer Management in Arizona*. College of Agriculture, The University of Arizona.

Dolph, C. L., Cho, S. J., Finlay, J. C., Hansen, A. T., & Dalzell, B. (2023). Predicting high resolution total phosphorus concentrations for soils of the Upper Mississippi River Basin using machine learning. *Biogeochemistry*, *163*(3), 289–310.

Donner, S. D., & Kucharik, C. J. (2008). Corn-based ethanol production compromises goal of reducing nitrogen export by the Mississippi River. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4513–4518.

Drew J. Lyon, Paul A Burgener, Karen L. DeBoer, Robert M Harveson, Gary L Hein, Gary W. Hergert, Lenis A. Nelson. (2008). *Producing and Marketing Proso Millet in the Great Plains*. https://extensionpublications.unl.edu/assets/pdf/ec137.pdf

Dubrovsky, N. M., Burow, K. R., Clark, G. M., Gronberg, J. M., Hamilton, P. A., Hitt, K. J., Mueller, D. K., Munn, Nolan, B. T., Puckett, L. J., Rupert, M. G., Short, T. M., Spahr, N. E., Sprague, L. A., & Wilber, W. G. (2010). *The quality of our Nation's waters: Nutrients in the Nation's streams and groundwater, 1992-2004*. U.S. Geological Survey Circular 1350. https://doi.org/10.3133/cir1350

Dupas, R., Ehrhardt, S., Musolff, A., Fovet, O., & Durand, P. (2020). Long-term nitrogen retention and transit time distribution in agricultural catchments in western France. *Environmental Research Letters: ERL [Web Site]*, *15*(11), 115011.

Dupas, R., Minaudo, C., Gruau, G., Ruiz, L., & Gascuel-Odoux, C. (2018). Multidecadal Trajectory of Riverine Nitrogen and Phosphorus Dynamics in Rural Catchments. *Water Resources Research*, *54*(8), 5327–5340.

Durack, :J, & Taylor, K. E. (2019). *input4MIPs: Boundary Condition and Forcing Datasets for CMIP6* [Dataset]. https://esgf-node.llnl.gov/search/input4mips/

Ehrhardt, S., Ebeling, P., Dupas, R., Kumar, R., Fleckenstein, J. H., & Musolff, A. (2021). Nitrate transport and retention in western European catchments are shaped by hydroclimate and subsurface properties. *Water Resources Research*, *57*(10). https://doi.org/10.1029/2020wr029469

Ehrhardt, S., Kumar, R., Fleckenstein, J. H., Attinger, S., & Musolff, A. (2019). Trajectories of nitrate input and output in three nested catchments along a land use gradient. *Hydrology and Earth System Sciences*, *23*(9), 3503–3524.

Falcone, J. A. (2017). Watershed characteristics for study sites of the Surface Water Trends project [Dataset]. In *National Water Quality Program: U.S. Geological Survey data release*. https://doi.org/10.5066/F7TX3CKP

Falcone, J. A. (2021). *Estimates of county-level nitrogen and phosphorus from fertilizer and manure from 1950 through 2017 in the conterminous United States*. US Geological Survey. https://doi.org/10.3133/ofr20201153

Fang, Y., & Jawitz, J. W. (2018). High-resolution reconstruction of the United States human population distribution, 1790 to 2010. *Scientific Data*, *5*, 180067.

Fang, Y., & Jawitz, J. W. (2019). The evolution of human population distance to water in the USA from 1790 to 2010. *Nature Communications*, *10*(1), 430.

Feng, L., Wang, Y., Hou, X., Qin, B., Kuster, T., Qu, F., Chen, N., Paerl, H. W., & Zheng, C. (2024). Harmful algal blooms in inland waters. *Nature Reviews. Earth & Environment*, 1–14.

Feng, P., Wang, B., Liu, D. L., & Yu, Q. (2019). Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agricultural Systems*, *173*, 303–316.

Flint, E. M., Ascott, M. J., Gooddy, D. C., Stahl, M. O., & Surridge, B. W. J. (2022). Water supply processes are responsible for significant nitrogen fluxes across the United States. *Global Biogeochemical Cycles*, *36*(9). https://doi.org/10.1029/2022gb007340

Flynn, K. C., Spiegal, S., Kleinman, P. J. A., Meinen, R. J., & Smith, D. R. (2023). Manureshed management to overcome longstanding nutrient imbalances in US agriculture. *Resources, Conservation and Recycling*, *188*, 106632.

Food and Agriculture Organization of the United Nations (FAO). (2022). *Fertilizers by Nutrient* [Dataset]. https://www.fao.org/faostat/en/#data/RFN

Galloway, J. N., & Cowling, E. B. (2002). Reactive nitrogen and the world: 200 years of change. *Ambio*, *31*(2), 64–71.

Galloway, J. N., & Cowling, E. B. (2021). Reflections on 200 years of Nitrogen, 20 years later. *Ambio*, *50*(4), 745–749.

Gao, Y., & Cabrera Serrenho, A. (2023). Greenhouse gas emissions from nitrogen fertilizers could be reduced by up to one-fifth of current levels by 2050 with combined interventions. *Nature Food*, *4*(2), 170–178.

Gao, Y., Tian, Y., Zhan, W., Li, L., Sun, H., Zhao, T., Zhang, H., Meng, Y., Li, Y., Liu, T., & Ding, J. (2023). Characterizing legacy nitrogen-induced time lags in riverine nitrogen reduction for the Songhuajiang River Basin: Source analysis, spatio-seasonal patterns, and impacts on future water quality improvement. *Water Research*, *242*(120292), 120292.

Gardner, B. L. (2002). *American Agriculture in the Twentieth Century: How it Flourished and What it Cost*. Harvard University Press.

Geisseler, D. (2016). *Nitrogen concentrations in harvested plant parts - A literature overview* (p. 157).

Geisseler, D., & Horwath, W. R. (n.d.). *Production of processing tomatoes in California*. Retrieved November 15, 2023, from https://www.cdfa.ca.gov/is/ffldrs/frep/FertilizationGuidelines/pdf/Tomato_Production_CA.pdf

Geisseler, D., & Horwath, W. R. (2016). *Lettuce Production in California*. Department of Land, Air and Water Resources, University of California, Davis. https://apps1.cdfa.ca.gov/FertilizerResearch/docs/Lettuce_Production_CA.pdf

Geisseler, D., Ortiz, R. S., & Diaz, J. (2022). Nitrogen nutrition and fertilization of onions (Allium cepa L.)–A literature review. *Scientia Horticulturae*, *291*, 110591.

Gleeson, T., Befus, K. M., Jasechko, S., Luijendijk, E., & Bayani Cardenas, M. (2016). The global volume and distribution of modern groundwater. In *Nature Geoscience* (Vol. 9, Issue 2, pp. 161–167). https://doi.org/10.1038/ngeo2590

Glibert, P. M. (2016). Eutrophication, harmful algae and biodiversity - Challenging paradigms in a world of complex nutrient changes. *Marine Pollution Bulletin*, *April*.

https://doi.org/10.1016/j.marpolbul.2017.04.027

Glibert, P. M. (2020). From hogs to HABs: impacts of industrial farming in the US on nitrogen and phosphorus and greenhouse gas pollution. *Biogeochemistry*, *150*(2), 139–180.

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

Goyette, J.-O., Bennett, E. M., Howarth, R. W., & Maranger, R. (2016a). Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. *Global Biogeochemical Cycles*, *30*(7), 1000–1014.

Goyette, J.-O., Bennett, E. M., Howarth, R. W., & Maranger, R. (2016b). Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. *Global Biogeochemical Cycles*, *30*(7), 1000–1014.

Grain Conversion Table. (2013). Agriculture Financial Services Corporation. https://afsc.ca/news/resource/grain-conversion-table/

Great Lakes Water Quality Agreement Nutrient Annex Subcommittee. (2015). *Recommended Binational Phosphorus Targets to Combat Lake Erie Algal Blooms*. https://www.epa.gov/sites/default/files/2015-06/documents/recommended-binational-phosphorus-targets-20150625-8pp.pdf

Gren, I.-M., & Destouni, G. (2012). Does divergence of nutrient load measurements matter for successful mitigation of marine eutrophication? *Ambio*, *41*(2), 151–160.

Grizzetti, B., Passy, P., Billen, G., Bouraoui, F., Garnier, J., & Lassaletta, L. (2015). The role of water nitrogen retention in integrated nutrient management: assessment in a large basin using different modelling approaches. *Environmental Research Letters*, *10*(6), 065008.

Grömping, U. (2009). Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, *63*(4), 308–319.

Gu, B., Ge, Y., Chang, S. X., Luo, W., & Chang, J. (2013). Nitrate in groundwater of China: Sources and driving forces. *Global Environmental Change: Human and Policy Dimensions*, *23*(5), 1112–1121.

Guillaumot, L., Marçais, J., Vautier, C., Guillou, A., Vergnaud, V., Bouchez, C., Dupas, R., Durand, P., de Dreuzy, J.-R., & Aquilina, L. (2021). A hillslope-scale aquifer-model to determine past agricultural legacy and future nitrate concentrations in rivers. *The Science of the Total Environment*, *800*, 149216.

Hagy, J. D., Boynton, W. R., Keefe, C. W., & Wood, K. V. (2004). Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow. *Estuaries*, *27*(4), 634–658.

Haines, M., Fishback, P., & Rhode, P. (2018). *United States Agriculture Data, 1840 - 2012*. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR35206.v4

Hamilton, S. K. (2012). Biogeochemical time lags may delay responses of streams to ecological restoration. *Freshwater Biology*, *57*(s1), 43–57.

Hamlin, Q. F., Kendall, A. D., Martin, S. L., Whitenack, H. D., Roush, J. A., Hannah, B. A., & Hyndman, D. W. (2020). Quantifying landscape nutrient inputs with spatially explicit nutrient source estimate

maps. *Journal of Geophysical Research. Biogeosciences*, *125*(2). https://doi.org/10.1029/2019jg005134

Han, H., & Allan, J. D. (2008). Estimation of nitrogen inputs to catchments: comparison of methods and consequences for riverine export prediction. *Biogeochemistry*, *91*(2), 177–199.

Han, H., Allan, J. D., & Bosch, N. S. (2012). Historical pattern of phosphorus loading to Lake Erie watersheds. *Journal of Great Lakes Research*, *38*(2), 289–298.

Han, H., Bosch, N., & Allan, J. D. (2011). Spatial and temporal variation in phosphorus budgets for 24 watersheds in the Lake Erie and Lake Michigan basins. *Biogeochemistry*, *102*(1-3), 45–58.

Hansen, B., Thorling, L., Kim, H., & Blicher-Mathiesen, G. (2019). Long-term nitrate response in shallow groundwater to agricultural N regulations in Denmark. In *Journal of Environmental Management* (Vol. 240, pp. 66–74). https://doi.org/10.1016/j.jenvman.2019.03.075

Hansen, B., Thorling, L., Schullehner, J., Termansen, M., & Dalgaard, T. (2017). Groundwater nitrate response to sustainable nitrogen management. *Scientific Reports*, *7*(1), 8566.

Hanson, E. J., & Howell, G. S. (1995). Nitrogen Accumulation and Fertilizer Use Efficiency by Grapevines in Short-season Growing Areas. *HortScience: A Publication of the American Society for Horticultural Science*, *30*(3), 504–507.

Han, Y., Feng, G., Swaney, D. P., Dentener, F., Koeble, R., Ouyang, Y., & Gao, W. (2020). Global and regional estimation of net anthropogenic nitrogen inputs (NANI). *Geoderma*, *361*, 114066.

Harding, L. W., Jr, Mallonee, M. E., Perry, E. S., Miller, W. D., Adolf, J. E., Gallegos, C. L., & Paerl, H. W. (2019). Long-term trends, current status, and transitions of water quality in Chesapeake Bay. *Scientific Reports*, *9*(1), 6709.

Harter, T., Lund, J. R., Darby, J., Fogg, G. E., Howitt, R., Jessoe, K. K., Stuart Pettygrove, G., Quinn, J. F., Viers, J. H., Boyle, D. B., Canada, H. E., DeLaMora, N., Dzurella, K. N., FryjoffHung, A., Hollander, A. D., Honeycutt, K. L., Jenkins, M. W., Jensen, V. B., King, A. M., … Rosenstock, T. S. (2012). *Addressing Nitrate in California's Drinking Water: With a Focus on Tulare Lake Basin and Salinas Valley Groundwater*. State Water Resources Control Board Report. http://groundwaternitrate.ucdavis.edu/files/138956.pdf

Hartmann, J., Moosdorf, N., Lauerwald, R., Hinderer, M., & Joshua West, A. (2014). Global chemical weathering and associated P-release — The role of lithology, temperature and soil properties. In *Chemical Geology* (Vol. 363, pp. 145–163). https://doi.org/10.1016/j.chemgeo.2013.10.025

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Science & Business Media.

Haygarth, P. M., Jarvie, H. P., Powers, S. M., Sharpley, A. N., Elser, J. J., Shen, J., Peterson, H. M., Chan, N.-I., Howden, N. J. K., Burt, T., Worrall, F., Zhang, F., & Liu, X. (2014). Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. *Environmental Science & Technology*, *48*(15), 8417–8419.

Healy, J. (2018, November 3). Rural America's Own Private Flint: Polluted Water Too Dangerous to Drink. *The New York Times*. https://www.nytimes.com/2018/11/03/us/water-contaminated-rural-america.html

Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). *Statistical methods in water resources* (Nos. 4-A3). U.S. Geological Survey. https://doi.org/10.3133/tm4A3

Hirsch, R., DeCicco, L., & Murphy, J. (2021). *Exploration and Graphics for RivEr Trends (EGRET)* (Version 3.0.5). U.S. Geological Survey. https://pubs.usgs.gov/tm/04/a10/

Hirsch, R. M., & De Cicco, L. A. (2015). User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data. In *Techniques and Methods*. US Geological Survey. https://doi.org/10.3133/tm4a10

Holman, I. P., Whelan, M. J., Howden, N. J. K., Bellamy, P. H., Willby, N. J., Rivas-Casado, M., & McConvey, P. (2008). Phosphorus in groundwater—an overlooked contributor to eutrophication? *Hydrological Processes*, *22*(26), 5121–5127.

Hong, B., Swaney, D. P., & Howarth, R. W. (2011). A toolbox for calculating net anthropogenic nitrogen inputs (NANI). *Environmental Modelling and Software*, *26*(5), 623–633.

Hong, B., Swaney, D. P., & Howarth, R. W. (2013). Estimating net anthropogenic nitrogen inputs to U.S. watersheds: Comparison of methodologies. *Environmental Science and Technology*, *47*(10), 5199–5207.

Hong, B., Swaney, D. P., McCrackin, M., Svanbäck, A., Humborg, C., Gustafsson, B., Yershova, A., & Pakhomau, A. (2017). Advances in NANI and NAPI accounting for the Baltic drainage basin: spatial and temporal trends and relationships to watershed TN and TP fluxes. *Biogeochemistry*, *133*(3), 245–261.

Hong, B., Swaney, D. P., Mörth, C. M., Smedberg, E., Eriksson Hägg, H., Humborg, C., Howarth, R. W., & Bouraoui, F. (2012). Evaluating regional variation of net anthropogenic nitrogen and phosphorus inputs (NANI/NAPI), major drivers, nutrient retention pattern and management implications in the multinational areas of Baltic Sea basin. *Ecological Modelling*, *227*, 117–135.

Hossain, Z., Wang, X., Hamel, C., Knight, J. D., Morrison, M. J., & Gan, Y. (2016). Biological nitrogen fixation by pulse crops on the semiarid Canadian Prairie. *Canadian Journal of Plant Science. Revue Canadienne de Phytotechnie*, *CJPS-2016-0185*. https://doi.org/10.1139/cjps-2016-0185

Houlton, B. Z., Boyer, E. W., Finzi, A., Galloway, J., Leach, A., Liptzin, D., Melillo, J., Rosenstock, T. S., Sobota, D., & Townsend, A. R. (2013). Intentional versus unintentional nitrogen use in the United States: Trends, efficiency and implications. *Biogeochemistry*, *114*(1-3), 11–23.

Howarth, R., Swaney, D., Billen, G., Garnier, J., Hong, B., Humborg, C., Johnes, P., Mörth, C.-M., & Marino, R. (2012a). *Nitrogen fluxes from the landscape are controlled by net anthropogenic nitrogen inputs and by climate* [Dataset]. https://doi.org/10.1890/100178

Howarth, R., Swaney, D., Billen, G., Garnier, J., Hong, B., Humborg, C., Johnes, P., Mörth, C.-M., & Marino, R. (2012b). Nitrogen fluxes from the landscape are controlled by net anthropogenic nitrogen inputs and by climate. *Frontiers in Ecology and the Environment*, *10*(1), 37–43.

Howarth, R. W., Billen, G., Swaney, D., Townsend, A., Jaworski, N., Lajtha, K., Downing, J. A., Elmgren, R., Caraco, N., Jordan, T., Berendse, F., Freney, J., Kudeyarov, V., Murdoch, P., & Zhao-Liang, Z. (1996). Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. In *Nitrogen Cycling in the North Atlantic Ocean and its Watersheds* (pp. 75–139). https://doi.org/10.1007/978-94-009-1776-7_3

Howarth, R. W., Chan, F., Conley, D. J., Garnier, J., Doney, S. C., Marino, R., & Billen, G. (2011). Coupled biogeochemical cycles: Eutrophication and hypoxia in temperate estuaries and coastal marine ecosystems. *Frontiers in Ecology and the Environment*, *9*(1), 18–26.

Huishu, L., Qiuliang, L., Xinyu, Z., Haw, Y., Hongyuan, W., Limei, Z., Hongbin, L., Huang, J.-C., Tianzhi, R., Jiaogen, Z., & Weiwen, Q. (2018). Effects of anthropogenic activities on long-term changes of nitrogen budget in a plain river network region: A case study in the Taihu Basin. *The Science of the Total Environment*, *645*, 1212–1220.

Hyytiäinen, K., Blyh, K., Hasler, B., Ahlvik, L., Ahtiainen, H., Artell, J., & Ericsdotter, S. (2014). *Environmental economic research as a tool in the protection of the Baltic Sea: – costs and benefits of reducing eutrophication*. Nordic Council of Ministers.

Ilampooranan, I., Van Meter, K. J., & Basu, N. B. (2019). A Race Against Time: Modeling Time Lags in Watershed Response. *Water Resources Research*, *475*, 65.

IPNI Estimates of Nutrient Uptake and Removal. (n.d.). Retrieved October 17, 2022, from http://www.ipni.net/article/IPNI-3296

Jarvie, H. P., Sharpley, A. N., Flaten, D., Kleinman, P. J. A., Jenkins, A., & Simmons, T. (2015). The pivotal role of phosphorus in a resilient water-energy-food security nexus. *Journal of Environmental Quality*, *44*(4), 1049–1062.

Jarvie, H. P., Sharpley, A. N., Spears, B., Buda, A. R., May, L., & Kleinman, P. J. A. (2013). Water quality remediation faces unprecedented challenges from "legacy phosphorus." *Environmental Science & Technology*, *47*(16), 8997–8998.

Jenkinson, D. S., Poulton, P. R., Johnston, A. E., & Powlson, D. S. (2004). Turnover of nitrogen-15-labeled fertilizer in old grassland. *Soil Science Society of America Journal. Soil Science Society of America*, *68*(3), 865–875.

Jones, C. S., Drake, C. W., Hruby, C. E., Schilling, K. E., & Wolter, C. F. (2019). Livestock manure driving stream nitrate. *Ambio*, *48*(10), 1143–1153.

Keiser, D. A., & Shapiro, J. S. (2017). *Consequences of the Clean Water Act and the Demand for Water Quality*. National Bureau of Economic Research.

Kellogg, R. L., Lander, C. H., Moffitt, D. C., & Gollehon, N. (2000a). Manure nutrients relative to the capacity of cropland and pastureland to assimilate nutrients: Spatial and temporal trends for the United States. *Proceedings of the Water Environment Federation*, *2000*(16), 18–157.

Kellogg, R. L., Lander, C. H., Moffitt, D. C., & Gollehon, N. (2000b). Manure Nutrients Relative to The Capacity Of Cropland And Pastureland To Assimilate Nutrients: Spatial and Temporal Trends for the United States. In *Proceedings of the Water Environment Federation* (Vol. 2000, Issue 16, pp. 18–157). https://doi.org/10.2175/193864700784994812

Kim, H., Surdyk, N., Møller, I., Graversgaard, M., Blicher-Mathiesen, G., Henriot, A., Dalgaard, T., & Hansen, B. (2020). Lag Time as an Indicator of the Link between Agricultural Pressure and Drinking Water Quality State. *WATER*, *12*(9), 2385.

Kleinman, P. J. A., Fanelli, R. M., Hirsch, R. M., Buda, A. R., Easton, Z. M., Wainger, L. A., Brosch, C., Lowenfish, M., Collick, A. S., Shirmohammadi, A., Boomer, K., Hubbart, J. A., Bryant, R. B., & Shenk, G. W. (2019). Phosphorus and the Chesapeake Bay: Lingering Issues and Emerging Concerns for Agriculture. *Journal of Environmental Quality*, *48*(5), 1191–1203.

Knobeloch, L., Salna, B., Hogan, A., Postle, J., & Anderson, H. (2000). Blue babies and nitrate-contaminated well water. *Environmental Health Perspectives*, *108*(7), 675–678.

Kraimer, R. A., Lindemann, W. C., & Herrera, E. A. (2001). Distribution of15N-labeled fertilizer applied

to pecan: A case study. *HortScience: A Publication of the American Society for Horticultural Science*, *36*(2), 308–312.

Kraimer, R. A., Lindemann, W. C., & Herrera, E. A. (2004). Recovery of late-season15N-labeled fertilizer applied to pecan. *HortScience: A Publication of the American Society for Horticultural Science*, *39*(2), 256–260.

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., & Matias, Y. (2023). Caravan - A global community dataset for large-sample hydrology. *Scientific Data*, *10*(1), 61.

Kucharik, C. J., & Ramankutty, N. (2005). Trends and Variability in U.S. Corn Yields Over the Twentieth Century. *Earth Interactions*, *9*(1), 1–29.

Kucharik, C. J., Ramiadantsoa, T., Zhang, J., & Ives, A. R. (2020). Spatiotemporal trends in crop yields, yield variability, and yield gaps across the USA. *Crop Science*, *60*(4), 2085–2101.

Lander, C. H., Moffitt, D., & Alt, K. (1998a). *Nutrients Available from Livestock Manure Relative to Crop Growth Requirements*. U.S. Department of Agriculture, Natural Resources Conservation Service. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/?cid=nrcs143_014175

Lander, C. H., Moffitt, D., & Alt, K. (1998b). *Nutrients Available from Livestock Manure Relative to Crop Growth Requirements (Appendix 1)*. Resource Assessment and Strategic Planning Working Paper 98-1. U.S. Department of Agriculture, Natural Resources Conservation Service. https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/technical/nra/nri/results/?cid=nrcs143_014150

Lark, T. J., Hendricks, N. P., Smith, A., Pates, N., Spawn-Lee, S. A., Bougie, M., Booth, E. G., Kucharik, C. J., & Gibbs, H. K. (2022). Environmental outcomes of the US Renewable Fuel Standard. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(9). https://doi.org/10.1073/pnas.2101084119

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems: The relationship between yield and nitrogen input to cropland. *Environmental Research Letters: ERL [Web Site]*, *9*(10). https://doi.org/10.1088/1748-9326/9/10/105011

Lee, M., Shevliakova, E., Malyshev, S., Milly, P. C. D., & Jaffé, P. R. (2016). Climate variability and extremes, interacting with nitrogen storage, amplify eutrophication risk. *Geophysical Research Letters*, *43*(14), 7520–7528.

Lefcheck, J. S., Orth, R. J., Dennison, W. C., Wilcox, D. J., Murphy, R. R., Keisman, J., Gurbisz, C., Hannam, M., Landry, J. B., Moore, K. A., Patrick, C. J., Testa, J., Weller, D. E., & Batiuk, R. A. (2018). Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(14), 3658–3662.

Le Moal, M., Gascuel-Odoux, C., Ménesguen, A., Souchon, Y., Étrillard, C., Levain, A., Moatar, F., Pannard, A., Souchu, P., Lefebvre, A., & Pinay, G. (2019). Eutrophication: A new wine in an old bottle? *The Science of the Total Environment*, *651*(Pt 1), 1–11.

Litke, D. W. (1999). *Review of phosphorus control measures in the United States and their effects on water quality*. https://pubs.usgs.gov/wri/wri994007/pdf/wri99-4007.pdf

Liu, J., Van Meter, K. J., McLeod, M. M., & Basu, N. B. (2021). Checkered landscapes: hydrologic and biogeochemical nitrogen legacies along the river continuum. *Environmental Research Letters: ERL*. https://doi.org/10.1088/1748-9326/ac243c

Lloret, J., & Valiela, I. (2016). Unprecedented decrease in deposition of nitrogen oxides over North America: the relative effects of emission controls and prevailing air-mass trajectories. *Biogeochemistry*, *129*(1-2), 165–180.

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.

Lovatt, C. (1996). *Nitrogen allocation within the "Hass" avocado*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3a364a9a91d97e13e6a10f90ed464133cfb4b973

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56–67.

Lun, F., Liu, J., Ciais, P., Nesme, T., Chang, J., Wang, R., Goll, D. S., Sardans, J., Penuelas, J., & Obersteiner, M. (2017). *Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency* [Dataset]. PANGAEA - Data Publisher for Earth & Environmental Science. https://doi.org/10.1594/PANGAEA.875296

MacDonald, G. K., Bennett, E. M., Potter, P. A., & Ramankutty, N. (2011). Agronomic phosphorus imbalances across the world's croplands. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3086–3091.

MacDonald, J. M., & McBride, W. D. (2009). *The Transformation of U.S. Livestock Agriculture Scale, Efficiency, and Risks*. https://doi.org/10.2139/ssrn.1354028

Malone, T. C., & Newton, A. (2020). The Globalization of Cultural Eutrophication in the Coastal Ocean: Causes and Consequences. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00670

McCoy, M. (2011). Goodbye, Phosphates. In *Chemical & Engineering News* (p. 011811165646). https://doi.org/10.1021/cen011811165646

McDowell, R. W., Simpson, Z. P., Ausseil, A. G., Etheridge, Z., & Law, R. (2021). The implications of lag times between nitrate leaching losses and riverine loads for water quality policy. *Scientific Reports*, *11*(1), 16450.

McIsaac, G. (2023). [Personal communication].

McIsaac, G. F., David, M. B., Gertner, G. Z., & Goolsby, D. A. (2001). Nitrate flux in the Mississippi River. *Nature*, *414*(6860), 166–167.

Meals, D. W., Dressing, S. A., Tech, T., & Davenport, T. E. (2010). Lag Time in Water Quality Response to Best Management Practices : A Review. *Journal of Environmental Quality*, (39), 85–96.

Meisinger, J. J., & Randall, G. W. (1991). Estimating Nitrogen Budgets for Soil-Crop Systems. In R. F. Follett, D. R. Keeney, & R. M. Cruse (Eds.), *Managing Nitrogen for Groundwater Quality and Farm Profitability* (pp. 85–124). Soil Science Society of America.

Messager, M. L., Lehner, B., Cockburn, C., Lamouroux, N., Pella, H., Snelder, T., Tockner, K.,

Trautmann, T., Watt, C., & Datry, T. (2021). Global prevalence of non-perennial rivers and streams. *Nature*, *594*(7863), 391–397.

Metson, G. S., Lin, J., & Harrison, J. A. (2020). Where Have All the Nutrients Gone? Long-Term Decoupling of Inputs and Outputs in the Willamette River Watershed, Oregon, United States. *Journal of Geophysical Research*. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JG005792

Millard, S. P. (2013). *EnvStats: An R Package for Environmental Statistics*. Springer Science & Business Media.

Miralha, L., Sidique, S., & Logsdon Muenich, R. (2022). The spatial organization of CAFOs and its relationship to water quality in the United States. *Journal of Hydrology*, *613*, 128301.

Mississippi River/Gulf of Mexico Watershed Nutrient Task Force. (2001). *Action Plan for Reducing, Mitigating, and Controlling Hypoxia in the Northern Gulf of Mexico*. Office of Wetlands, Oceans, and Watersheds, U.S. Environmental Protection Agency.

Mohadjer, P., Strik, B. C., Zebarth, B. J., & Righetti, T. L. (2015). *Nitrogen uptake , partitioning and remobilization in "Kotata" blackberry in alternate-year production*. *0316*. https://doi.org/10.1080/14620316.2001.11511434

Müller-Karulis, B., McCrackin, M. L., Dessirier, B., Gustafsson, B. G., & Humborg, C. (2024). Legacy nutrients in the Baltic Sea drainage basin: How past practices affect eutrophication management. *Journal of Environmental Management*, *370*(122478), 122478.

Mullins, G. L. (2009). *Phosphorus, agriculture & the environment*. https://vtechworks.lib.vt.edu/bitstream/handle/10919/55777/424-029.pdf

Multi-Resolution Land Characteristics Consortium. (2016). *National Land Cover Database (NLCD)* [Dataset]. https://www.mrlc.gov/data?f%5B0%5D=category%3ALand%20Cover

Murphy, R. R., Keisman, J., Harcum, J., Karrh, R. R., Lane, M., Perry, E. S., & Zhang, Q. (2022). Nutrient improvements in Chesapeake Bay: Direct effect of load reductions and implications for coastal management. *Environmental Science & Technology*, *56*(1), 260–270.

NADP. (n.d.). *National Atmospheric Deposition Program*. National Trends Network. Retrieved July 2, 2024, from https://nadp.slh.wisc.edu/networks/national-trends-network/

National Atmospheric Deposition Program. (2018). *NADP Total Deposition Maps*. http://nadp.slh.wisc.edu/committees/tdep/tdepmaps/

National IPM Database - 2002 Crop Profile for Squash in Florida: Document Report. (n.d.). Retrieved November 15, 2023, from https://ipmdata.ipmcenters.org/source_report.cfm?view=yes&sourceid=382

Neilsen, D., Millard, P., Neilsen, G. H., & Hogue, E. J. (2001). Nitrogen Uptake , Efficiency of Use , and Partitioning for Growth in Young Apple Trees. *Journal of the American Society for Horticultural Science. American Society for Horticultural Science*, *126*(May 1992), 144–150.

Neilsen, D., & Neilsen, G. H. (2002). Efficient use of nitrogen and water in high-density apple orchards. *HortTechnology*, *12*(1), 19–25.

Newcomer, M. E., Bouskill, N. J., Wainwright, H., Maavara, T., Arora, B., Siirila-Woodburn, E. R., Dwivedi, D., Williams, K. H., Steefel, C., & Hubbard, S. S. (2021). Hysteresis patterns of watershed

nitrogen retention and loss over the past 50 years in United States hydrological basins. *Global Biogeochemical Cycles*, *35*(4). https://doi.org/10.1029/2020gb006777

New Jersey Department of Agriculture. (2015). *Annual Report & Agricultural Statistics*. https://www.nj.gov/agriculture/pdf/2015annualreport.pdf

Nguyen, T. V., Sarrazin, F. J., Ebeling, P., Musolff, A., Fleckenstein, J. H., & Kumar, R. (2022). Toward understanding of long-term nitrogen transport and retention dynamics across German catchments. *Geophysical Research Letters*, *49*(24). https://doi.org/10.1029/2022gl100278

Oak Ridge National Laboratory. (2017). *Fossil-Fuel CO2 Emissions by Nation*. https://cdiac.ess-dive.lbl.gov/trends/emis/tre_coun.html#

Oelsner, G. P., Sprague, L. A., Murphy, J. C., Zuellig, R. E., Johnson, H. M., Ryberg, K. R., Falcone, J. A., Stets, E. G., Vecchia, A. V., Riskin, M. L., Cicco, L. A. D., Mills, T. J., & Farmer, W. H. (2017). *Water-Quality trends in the nation's rivers and streams , 1972 – 2012 — Data Preparation, Statistical Methods, and Trend Results Scientific Investigations Report 2017 – 5006*. U.S. Geological Survey Scientific Investigations Report 2017–5006. https://doi.org/10.3133/sir20175006

Oelsner, G. P., Sprague, L. A., Murphy, J. C., Zuellig, R. E., Johnson, H. M., Ryberg, K. R., Falcone, J. A., Stets, E. G., Vecchia, A. V., Riskin, M. L., De Cicco, L. A., Mills, T. J., & Farmer, W. H. (2017). Water-quality trends in the nation's rivers and streams, 1972–2012—Data preparation, statistical methods, and trend results. In *Scientific Investigations Report*. https://doi.org/10.3133/sir20175006

Parfitt, J., Barthel, M., & Macnaughton, S. (2010). Food waste within food supply chains: quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1554), 3065–3081.

Pastén-Zapata, E., Ledesma-Ruiz, R., Harter, T., Ramírez, A. I., & Mahlknecht, J. (2014). Assessment of sources and fate of nitrate in shallow groundwater of an agricultural area by using a multi-tracer approach. In *Science of The Total Environment* (Vols. 470–471, pp. 855–864). https://doi.org/10.1016/j.scitotenv.2013.10.043

Patakamuri, S. K., & O'Brien, N. (2021). *modifiedmk: Modified Versions of Mann Kendall and Spearman's Rho Trend Tests*. https://cran.r-project.org/web/packages/modifiedmk/index.html

Pavinato, P. S., Cherubin, M. R., Soltangheisi, A., Rocha, G. C., Chadwick, D. R., & Jones, D. L. (2020). Revealing soil legacy phosphorus to promote sustainable agriculture in Brazil. *Scientific Reports*, *10*(1), 15615.

Peacock, B., Christensen, P., & Hirschfeldt, D. (1991). *Best management practices for nitrogen fertilization of grapevines. Tulare County UCCE, Pub NG4-96. 20*, 1–6.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research: JMLR*, *12*, 2825–2830.

Pei, W., Yan, T., Lei, Q., Zhang, T., Fan, B., Du, X., Luo, J., Lindsey, S., & Liu, H. (2022). Spatio-temporal variation of net anthropogenic nitrogen inputs (NANI) from 1991 to 2019 and its impacts analysis from parameters in Northwest China. *Journal of Environmental Management*, *321*, 115996.

Pennino, M. J., Compton, J. E., & Leibowitz, S. G. (2017). Trends in Drinking Water Nitrate Violations

Across the United States. *Environmental Science & Technology*, *51*(22), 13450–13460.

Perry, M. T. (2023). *rasterstats python package* (Version 0.18.0). https://github.com/perrygeo/python-rasterstats

Phosphate Rock - Historical Statistics (Data Series 140) 2018 update. (2018). https://www.usgs.gov/media/files/phosphate-rock-historical-statistics-data-series-140-2018-update

Picetti, R., Deeney, M., Pastorino, S., Miller, M. R., Shah, A., Leon, D. A., Dangour, A. D., & Green, R. (2022). Nitrate and nitrite contamination in drinking water and cancer risk: A systematic review with meta-analysis. *Environmental Research*, *210*, 112988.

Postel, S. L., & Carpenter, S. R. (2012). *Nature's Services: Societal dependence on natural ecosystems* (G. C. Daily (Ed.)). Island Press.

Powers, S. M., Bruulsema, T. W., Burt, T. P., Chan, N. I., Elser, J. J., Haygarth, P. M., Howden, N. J. K., Jarvie, H. P., Lyu, Y., Peterson, H. M., Sharpley, A. N., Shen, J., Worrall, F., & Zhang, F. (2016). Long-term accumulation and transport of anthropogenic phosphorus in three river basins. In *Nature Geoscience* (Vol. 9, Issue 5, pp. 353–356). https://doi.org/10.1038/ngeo2693

Puckett, L. J., Tesoriero, A. J., & Dubrovsky, N. M. (2011). Nitrogen Contamination of Surficial Aquifers — A Growing Legacy. In *Environmental Science & Technology* (Vol. 45, Issue 3, pp. 839–844). https://doi.org/10.1021/es1038358

QGIS. (n.d.). Retrieved July 3, 2024, from https://www.qgis.org/en/site/

Rabalais, N. N., Díaz, R. J., Levin, L. A., Turner, R. E., Gilbert, D., & Zhang, J. (2009). Dynamics and distribution of natural and human-caused coastal hypoxia. In *Biogeosciences Discussions* (Vol. 6, Issue 5, pp. 9359–9453). https://doi.org/10.5194/bgd-6-9359-2009

Rabalais, N. N., & Turner, E. R. (2019). Gulf of Mexico Hypoxia: Past, Present, and Future. In *Limnology and Oceanography Bulletin* (Vol. 28, Issue 4, pp. 117–124). https://doi.org/10.1002/lob.10351

Rabalais, N. N., Turner, R. E., & Scavia, D. (2002). Beyond Science into Policy: Gulf of Mexico Hypoxia and the Mississippi River: Nutrient policy development for the Mississippi River watershed reflects the …. *Bioscience*. https://academic.oup.com/bioscience/article-abstract/52/2/129/341059

Rabalais, N. N., Turner, R. E., & Wiseman, W. J. (2002). Gulf of Mexico Hypoxia, A.K.A. "The Dead Zone." *Annual Review of Ecology and Systematics*, *33*(1), 235–263.

Raese, J. T. (1998). Response of apple and pear trees to nitrogen, phosphorus, and potassium fertilizers. *Journal of Plant Nutrition*, *21*(12), 2671–2696.

Ramankutty, N., & Foley, J. A. (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, *13*(4), 997–1027.

Ransom, K. M., Nolan, B. T., Stackelberg, P. E., Belitz, K., & Fram, M. S. (2021). Machine learning predictions of nitrate in groundwater used for drinking supply in the conterminous United States. *The Science of the Total Environment*, 151065.

Reckhow, K. H., Norris, P. E., Budell, R. J., Di Toro, D. M., Galloway, J. N., Greening, H., Sharpley, A. N., Shirmohammadi, A., & Stacey, P. E. (2011). *Achieving Nutrient and Sediment Reduction Goals in the Chesapeake Bay*. National Academies Press. https://doi.org/10.17226/13131

Rempel, H. G., Strik, B. C., & Righetti, T. L. (2004). *Uptake , Partitioning , and Storage of Fertilizer*

Nitrogen in Red Raspberry as Affected by Rate and Timing of Application. *129*(3), 439–448.

Rieniets, T. (2009). Shrinking cities: Causes and effects of urban population losses in the twentieth century. *Nature and Culture*, *4*(3), 231–254.

Ritter, W. F., & Rao Chitikela, S. (2020). The Mississippi river basin nitrogen problem: Past history and future challenges to solve it. *Watershed Management 2020*, 109–123.

Roberts, T. L., & Johnston, A. E. (2015). Phosphorus use efficiency and management in agriculture. *Resources, Conservation and Recycling*, *105*, 275–281.

Roper, T., DeMoranville, C., Davenport, J., Service, O. S. U. E., & Hart, J. (2000). *Nitrogen for bearing cranberries in North America*. https://ir.library.oregonstate.edu/concern/open_educational_resources/8g84mm58v

Sabo, R. D., Clark, C. M., Bash, J., Sobota, D., Cooter, E., Dobrowolski, J. P., Houlton, B. Z., Rea, A., Schwede, D., Morford, S. L., & Compton, J. E. (2019). Decadal Shift in Nitrogen Inputs and Fluxes Across the Contiguous United States: 2002–2012. In *Journal of Geophysical Research: Biogeosciences* (Vol. 124, Issue 10, pp. 3104–3124). https://doi.org/10.1029/2019jg005110

Sabo, R. D., Clark, C. M., Gibbs, D. A., Metson, G. S., Jason Todd, M., LeDuc, S. D., Greiner, D., Fry, M. M., Polinsky, R., Yang, Q., Tian, H., & Compton, J. E. (2021). Phosphorus Inventory for the Conterminous United States (2002–2012). In *Journal of Geophysical Research: Biogeosciences* (Vol. 126, Issue 4). https://doi.org/10.1029/2020jg005684

Sadayappan, K., Kerins, D., Shen, C., & Li, L. (2022). Nitrate concentrations predominantly driven by human, climate, and soil properties in US rivers. *Water Research*, *226*, 119295.

Sanchez, E. E. (2016). *Nitrogen Management in Orchards*. *January 1995*.

San-Martino, L., Sozzi, G. O., San-Martino, S., & Lavado, R. S. (2010). Isotopically-labelled nitrogen uptake and partitioning in sweet cherry as influenced by timing of fertilizer application. *Scientia Horticulturae*, *126*(1), 42–49.

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.

Schindler, D. W., Carpenter, S. R., Chapra, S. C., Hecky, R. E., & Orihel, D. M. (2016). Reducing Phosphorus to Curb Lake Eutrophication is a Success. *Environmental Science & Technology*, *50*(17), 8923–8929.

Schindler, D. W., Hecky, R. E., & McCullough, G. K. (2012). The rapid eutrophication of Lake Winnipeg: Greening under global change. *Journal of Great Lakes Research*, *38*, 6–13.

Schlesinger, W. H., & Bernhardt, E. S. (2020). *Biogeochemistry: An Analysis of Global Change* (Vol. 4). Elsevier/Academic Press.

Schröder, J. J., Smit, A. L., Cordell, D., & Rosemarin, A. (2011). Improved phosphorus use efficiency in agriculture: a key requirement for its sustainable use. *Chemosphere*, *84*(6), 822–831.

Scott, M. P., Edwards, J. W., Bell, C. P., Schussler, J. R., & Smith, J. S. (2006). *Grain composition and amino acid content in maize cultivars representing 80 years of commercial maize varieties*.

https://hygeia-analytics.com/wp-content/uploads/2018/02/Scott-et-al.-2006-corn-protein-amino-acids-historical-comparisons.pdf

Sebilo, M., Mayer, B., Nicolardot, B., Pinay, G., & Mariotti, A. (2013). Long-term fate of nitrate fertilizer in agricultural soils. *Proceedings of the National Academy of Sciences*, *110* (45), 18185–18189.

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

Seitzinger, S., Harrison, J. A., Böhlke, J. K., Bouwman, A. F., Lowrance, R., Peterson, B., Tobias, C., & Van Drecht, G. (2006). Denitrification across landscapes and waterscapes: a synthesis. *Ecological Applications: A Publication of the Ecological Society of America*, *16*(6), 2064–2090.

Sharpley, A., Jarvie, H. P., Buda, A., May, L., Spears, B., & Kleinman, P. (2013). Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. *Journal of Environmental Quality*, *42*(5), 1308–1326.

Sharpley, A., & Wang, X. (2014). Managing agricultural phosphorus for water quality: lessons from the USA and China. *Journal of Environmental Sciences* , *26*(9), 1770–1782.

Shomar, B., Osenbrück, K., & Yahya, A. (2008). Elevated nitrate levels in the groundwater of the Gaza Strip: Distribution and sources. In *Science of The Total Environment* (Vol. 398, Issues 1-3, pp. 164–174). https://doi.org/10.1016/j.scitotenv.2008.02.054

Smith, D. R., King, K. W., Johnson, L., Francesconi, W., Richards, P., Baker, D., & Sharpley, A. N. (2015). Surface runoff and tile drainage transport of phosphorus in the midwestern United States. *Journal of Environmental Quality*, *44*(2), 495–502.

Sobota, D. J., Harrison, J. A., & Dahlgren, R. A. (2009). Influences of climate, hydrology, and land use on input and export of nitrogen in California watersheds. *Biogeochemistry*, *94*(1), 43–62.

Sohl, T. L., Sayler, K. L., Bouchard, M. A., Reker, R. R., Freisz, A. M., Bennett, S. L., Sleeter, B. M., Sleeter, R. R., Wilson, T., Soulard, C., Knuppe, M., & Van Hofwegen, T. (2018). *Conterminous United States Land Cover Projections - 1992 to 2100: U.S. Geological Survey data release* [Dataset]. https://doi.org/10.5066/P95AK9HP

Sohl, T. L., Sayler, K. L., Bouchard, M. A., Reker, R. R., Friesz, A. M., Bennett, S. L., Sleeter, B. M., Sleeter, R. R., Wilson, T., Soulard, C., Knuppe, M., & Van Hofwegen, T. (2014). Spatially explicit modeling of 1992-2100 land cover and forest stand age for the conterminous United States. *Ecological Applications: A Publication of the Ecological Society of America*, *24*(5), 1015–1036.

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., & Friesz, A. (2016a). Modeled historical land use and land cover for the conterminous United States. In *Journal of Land Use Science* (Vol. 11, Issue 4, pp. 476–499). https://doi.org/10.1080/1747423x.2016.1147619

Sohl, T., Reker, R., Bouchard, M., Sayler, K., Dornbierer, J., Wika, S., Quenzer, R., & Friesz, A. (2016b). *Modeled Historical Land Use and Land Cover for the Conterminous United States: 1938-1992* [Dataset]. https://www.sciencebase.gov/catalog/item/59d3c73de4b05fe04cc3d1d1

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.

Stackpoole, S., Sabo, R., Falcone, J., & Sprague, L. (2021). Long-Term Mississippi River Trends Expose

Shifts in the River Load Response to Watershed Nutrient Balances Between 1975 and 2017. *Water Resources Research*, *57*(11), e2021WR030318.

Standard Weights per Bushel for Agricultural Commodities. (n.d.). Small Farm Canada. Retrieved November 14, 2023, from https://www.smallfarmcanada.ca/resources/standard-weights-per-bushel-for-agricultural-commodities

Steffen, M. M., Davis, T. W., McKay, R. M. L., Bullerjahn, G. S., Krausfeldt, L. E., Stough, J. M. A., Neitzey, M. L., Gilbert, N. E., Boyer, G. L., Johengen, T. H., Gossiaux, D. C., Burtner, A. M., Palladino, D., Rowe, M. D., Dick, G. J., Meyer, K. A., Levy, S., Boone, B. E., Stumpf, R. P., … Wilhelm, S. W. (2017). Ecophysiological Examination of the Lake Erie Microcystis Bloom in 2014: Linkages between Biology and the Water Supply Shutdown of Toledo, OH. *Environmental Science & Technology*, *51*(12), 6745–6755.

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., De Vries, W., De Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223). https://doi.org/10.1126/science.1259855

Stets, E. G., Sprague, L. A., Oelsner, G. P., Johnson, H. M., Murphy, J. C., Ryberg, K., Vecchia, A. V., Zuellig, R. E., Falcone, J. A., & Riskin, M. L. (2020). Landscape Drivers of Dynamic Change in Water Quality of U.S. Rivers. In *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.9b05344

Stoddard, J. L., Van Sickle, J., Herlihy, A. T., Brahney, J., Paulsen, S., Peck, D. V., Mitchell, R., & Pollard, A. I. (2016). Continental-Scale Increase in Lake and Stream Phosphorus: Are Oligotrophic Systems Disappearing in the United States? *Environmental Science & Technology*, *50*(7), 3409–3415.

Swaney, D. P., Hong, B., Ti, C., Howarth, R. W., & Humborg, C. (2012). Net anthropogenic nitrogen inputs to watersheds and riverine N export to coastal waters: a brief overview. *Current Opinion in Environmental Sustainability*, *4*(2), 203–211.

Swaney, D. P., & Howarth, R. W. (2019a). County, subregional and regional phosphorus data derived from the net anthropogenic nitrogen/phosphorus inputs (NANI/NAPI) toolbox. *Data in Brief*, *25*, 104265.

Swaney, D. P., & Howarth, R. W. (2019b). Phosphorus use efficiency and crop production: Patterns of regional variation in the United States, 1987-2012. *The Science of the Total Environment*, *685*, 174–188.

Swaney, D. P., Howarth, R. W., & Hong, B. (2018a). County, subregional and regional nitrogen data derived from the Net Anthropogenic Nitrogen Inputs (NANI) toolbox. *Data in Brief*, *18*, 1877–1888.

Swaney, D. P., Howarth, R. W., & Hong, B. (2018b). Nitrogen use efficiency and crop production: Patterns of regional variation in the United States, 1987-2012. *The Science of the Total Environment*, *635*, 498–511.

Syakila, A., & Kroeze, C. (2011). The global nitrous oxide budget revisited. *Greenhouse Gas Measurement and Management*, *1*(1), 17–26.

Tango, P. J., & Batiuk, R. A. (2016). Chesapeake Bay recovery and factors affecting trends: Long-term monitoring, indicators, and insights. *Regional Studies in Marine Science*, *4*, 12–20.

Tesoriero, A. J., Duff, J. H., Saad, D. A., Spahr, N. E., & Wolock, D. M. (2013). Vulnerability of streams to legacy nitrate sources. *Environmental Science and Technology*, *47*(8), 3623–3629.

Toccalino, P. L., & Hopple, J. (2010). *The quality of our Nation's waters—Quality of water from public-supply wells in the United States, 1993–2007—Overview of major findings: U.S. Geological Survey Circular 1346*. *58*.

Tyralis, H., Papacharalampous, G., & Langousis, A. (2019). A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. In *Water* (Vol. 11, Issue 5, p. 910). https://doi.org/10.3390/w11050910

UC Davis. (n.d.). *California Fertilization Guidelines*. https://www.cdfa.ca.gov/is/ffldrs/frep/FertilizationGuidelines/

United States Department of Agriculture. (n.d.). *FoodData Central*. Retrieved January 9, 2024, from https://fdc.nal.usda.gov/

University of California Cooperative Extension. (2016). *Sample Costs to Establish an Orchard and Produce Almonds*. University of California. https://coststudyfiles.ucdavis.edu/uploads/cs_public/87/3c/873c1216-f21e-4e3e-8961-8ece2d647329/2016_almondsjv_south_final_10142016.pdf

U.S. Census Bureau. (n.d.). *County Population Totals and Components of Change: 2010-2018*. Retrieved May 4, 2019, from https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html

U.S. Census Bureau. (2007). *Census U.S. Decennial County Population Data, 1900-1990*. http://www.nber.org/data/census-decennial-population.html

U.S. Census Bureau. (2016). *Census U.S. Intercensal County Population Data, 1970-2014*. https://www.nber.org/data/census-intercensal-county-population.html

USDA. (2023). *USDA - National Agricultural Statistics Service - Data and Statistics*. https://www.nass.usda.gov/Data_and_Statistics/

USDA Economic Research Service. (n.d.). *Pumpkin yield, nine States, 2012-17*. Retrieved November 15, 2023, from https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=85525

USDA Economic Research Service. (1992). *Weights, Measures, and Conversion Factors for Agricultural Commodities and Their Products*. https://www.ers.usda.gov/webdocs/publications/41880/33132_ah697_002.pdf

USDA Economics, Statistics and Market Information System. (2008). *U.S. Bell and Chile Pepper Statistics*. https://usda.library.cornell.edu/concern/publications/p5547r38s?locale=en

USDA Economics, Statistics and Market Information System. (2010a). *U.S. Asparagus Statistics*. https://usda.library.cornell.edu/concern/publications/d504rk32w?locale=en

USDA Economics, Statistics and Market Information System. (2010b). *U.S. Tomato Statistics*. https://usda.library.cornell.edu/concern/publications/br86b356q?locale=en

USDA Economics, Statistics and Market Information System. (2011a). *U.S. Broccoli Statistics*. https://usda.library.cornell.edu/concern/publications/jd472w44j?locale=en

USDA Economics, Statistics and Market Information System. (2011b). *U.S. Carrot Statistics*.

https://usda.library.cornell.edu/concern/publications/sx61dm28w?locale=en

USDA Economics, Statistics and Market Information System. (2011c). *U.S. Lettuce Statistics*. https://usda.library.cornell.edu/concern/publications/cc08hf60z?locale=en

USDA Economics, Statistics and Market Information System. (2011d). *U.S. Onion Statistics*. https://usda.library.cornell.edu/concern/publications/k643b116n?locale=en

USDA Economics, Statistics and Market Information System. (2013a). *U.S. Cantaloupe Statistics*. https://usda.library.cornell.edu/concern/publications/bv73c0402?locale=en

USDA Economics, Statistics and Market Information System. (2013b). *U.S. Strawberry Industry*. https://usda.library.cornell.edu/concern/publications/8s45q876k?locale=en

USDA-NASS. (2017). *Census of Agriculture*. https://www.nass.usda.gov/AgCensus/index.php

USDA National Agricultural Statistics Service. (2019). *Quick Stats Tools*. United States Department of Agriculture National Agricultural Statistics Service Quick Stats. https://www.nass.usda.gov/Quick_Stats/

USDA - National Agricultural Statistics Service - State Agricultural Overview. (n.d.). Retrieved January 2019, from https://www.nass.usda.gov/Statistics_by_State/

USDA - national agricultural statistics service - surveys - crops/stocks (agricultural survey). (n.d.). Retrieved September 5, 2023, from https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Crops_Stocks/index.php

USDA NRCS. (2008). *Agricultural Waste Management Field Handbook - Chapter 4 Agricultural Waste Characteristics*. https://directives.sc.egov.usda.gov/landingpage/86fc4bf7-b503-40bf-a112-91cc8ee87804

U.S. Department of Agriculture. (n.d.). *Livestock and Meat Domestic Data*. Retrieved September 25, 2023, from https://www.ers.usda.gov/data-products/livestock-and-meat-domestic-data/

U.S. Environmental Protection Agency (USEPA). (2020). *National Rivers and Streams Assessment 2013–2014: A Collaborative Survey*. US EPA. www.epa.gov/national-aquatic-resource-surveys/national-rivers-and-streams-assessment-2013-2014-report

USEPA. (2023). *Mississippi River/Gulf of Mexico Watersdhed Nutrient Task Force: 2023 Report to Congress*. https://www.epa.gov/system/files/documents/2023-11/10305_2023-htf-report-to-congress_508.pdf

U.S. Sweet Corn Statistics (Table 51 and 52). (2010). USDA Economics, Statistics and Market Information System. https://usda.library.cornell.edu/concern/publications/1r66j112r?locale=en

Utah State University Extension. (2011). *Dingy Dishes Blame Reduced Phosphates*. https://extension.usu.edu/archive/dingy-dishes-blame-reduced-phosphates

Valayamkunnath, P., Barlage, M., Chen, F., Gochis, D. J., & Franz, K. J. (2020). Mapping of 30-meter resolution tile-drained croplands using a geospatial modeling approach. *Scientific Data*, *7*(1), 257.

Van Breemen, N., Boyer, E. W., Goodale, C. L., Jaworski, N. A., Paustian, K., Seitzinger, S. P., Lajtha, K., Mayer, B., Van Dam, D., Howarth, R. W., Nadelhoffer, K. J., Eve, M., & Billen, G. (2002). Where did all the nitrogen go? Fate of nitrogen inputs to large watersheds in the northeastern U.S.A. *Biogeochemistry, 5758*, 267–293.

Van Meter, K. J., & Basu, N. (2017). Time lags in watershed-scale nutrient transport: an exploration of dominant controls. *Environmental Research Letters: ERL [Web Site]*, *12*(8), 084017.

Van Meter, K. J., & Basu, N. B. (2015). Catchment legacies and time lags: a parsimonious watershed model to predict the effects of legacy storage on nitrogen export. *PloS One*, *10*(5), e0125971.

Van Meter, K. J., Basu, N. B., & Van Cappellen, P. (2017). Two centuries of nitrogen dynamics: Legacy sources and sinks in the Mississippi and Susquehanna River Basins. *Global Biogeochemical Cycles*, *31*(1), 2–23.

Van Meter, K. J., Basu, N. B., & Veenstra, J. J. (2016). The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. *The Environmentalist*. http://iopscience.iop.org/article/10.1088/1748-9326/11/3/035014/meta

Van Meter, K. J., Basu, N. B., Veenstra, J. J., & Burras, C. L. (2016). The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters: ERL [Web Site]*, *11*(3), 035014.

Van Meter, K. J., Byrnes, D. K., & Basu, N. B. (2023). Memory and management: Competing controls on long-term nitrate trajectories in U.S. rivers. *Global Biogeochemical Cycles*. https://doi.org/10.1029/2022gb007651

Van Meter, K. J., McLeod, M. M., Liu, J., Tenkouano, G. T., Hall, R. I., Van Cappellen, P., & Basu, N. B. (2021). Beyond the mass balance: Watershed phosphorus legacies and the evolution of the current water quality policy challenge. *Water Resources Research*, *57*(10). https://doi.org/10.1029/2020wr029316

Van Meter, K. J., Schultz, V. O., & Chang, S. Y. (2023). Data-driven approaches demonstrate legacy N accumulation in Upper Mississippi River Basin groundwater. *Environmental Research Letters: ERL [Web Site]*, *18*(9), 094016.

Van Meter, K. J., Van Cappellen, P., & Basu, N. B. (2018). Legacy nitrogen may prevent achievement of water quality goals in the Gulf of Mexico. *Science*, *360*(6387), 427–430.

Vero, S. E., Basu, N. B., Van Meter, K. J., Richards, K. G., Mellander, P.-E., Healy, M. G., & Fenton, O. (2017). The environmental status and implications of the nitrate time lag in Europe and North America. *Hydrogeology Journal*, 1–16.

Villamil, M. B., Nafziger, E. D., & Behnke, G. D. (2019). New grain P and K concentration values for Illinois field crops. *Crop, Forage & Turfgrass Management*, *5*(1), 180090.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human Domination of Earth's Ecosystems. *Science*, *277*(5325), 494–499.

Vowels, M. J. (2022). Trying to outrun causality with machine learning: Limitations of model explainability techniques for identifying predictive variables. In *arXiv [stat.ML]*. arXiv. http://arxiv.org/abs/2202.09875

Wang, A., Tang, L., Yang, D., & Lei, H. (2016). Spatio-temporal variation of net anthropogenic nitrogen inputs in the upper Yangtze River basin from 1990 to 2012. *SCIENCE CHINA Earth Sciences*, *59*(11), 2189–2201.

Wang, J., Qi, Z., & Bennett, E. M. (2022). Changes in Canada's phosphorus cycle 1961-2018: Surplus and deficits. *Global Biogeochemical Cycles*. https://doi.org/10.1029/2022gb007407

Wang, L., Stuart, M. E., Bloomfield, J. P., Butcher, A. S., Gooddy, D. C., McKenzie, A. A., Lewis, M. A., & Williams, A. T. (2012). Prediction of the arrival of peak nitrate concentrations at the water table at the regional scale in Great Britain. *Hydrological Processes*, *26*(2), 226–239.

Wang, R., Goll, D., Balkanski, Y., Hauglustaine, D., Boucher, O., Ciais, P., Janssens, I., Penuelas, J., Guenet, B., Sardans, J., Bopp, L., Vuichard, N., Zhou, F., Li, B., Piao, S., Peng, S., Huang, Y., & Tao, S. (2017). Global forest carbon uptake due to nitrogen and phosphorus deposition from 1850 to 2100. *Global Change Biology*, *23*(11), 4854–4872.

Ward, M. H., Jones, R. R., Brender, J. D., de Kok, T. M., Weyer, P. J., Nolan, B. T., Villanueva, C. M., & van Breda, S. G. (2018). Drinking Water Nitrate and Human Health: An Updated Review. *International Journal of Environmental Research and Public Health*, *15*(7). https://doi.org/10.3390/ijerph15071557

Water quality standards attainment and monitoring - Chesapeake progress. (2021). Chesapeake Progress. https://www.chesapeakeprogress.com/clean-water/water-quality

Weber, C., Wechsler, S. J., & Wakefield, H. (n.d.). *Fruit and Tree Nuts Yearbook Tables*. Retrieved January 2019, from https://www.ers.usda.gov/data-products/fruit-and-tree-nuts-data/fruit-and-tree-nuts-yearbook-tables/

Weinbaum, S. a., Muraoka, T. T., Catlin, P. B., & Kelley, K. (1991). *Utilization of fertilizer N by walnut trees*. 317–334.

Westermann, D. T., Kleinkopf, G. E., Porter, L. K., & Leggett, G. E. (1981). Nitrogen sources for bean seed production[1]. *Agronomy Journal*, *73*(4), 660–664.

White, L., Yang, W., Hart, J. M., & Strik, B. (2006). *Nutrient management for blueberries in Oregon*. https://catalog.extension.oregonstate.edu/sites/catalog/files/project/pdf/em8918.pdf

Worrall, F., Howden, N. J. K., & Burt, T. P. (2015). Evidence for nitrogen accumulation: the total nitrogen budget of the terrestrial biosphere of a lowland agricultural catchment. In *Biogeochemistry* (Vol. 123, Issue 3, pp. 411–428). https://doi.org/10.1007/s10533-015-0074-7

Wu, Z., Jiang, M., Wang, H., Di, D., & Guo, X. (2022). Management implications of spatial-temporal variations of net anthropogenic nitrogen inputs (NANI) in the Yellow River Basin. *Environmental Science and Pollution Research International*, *29*(35), 52317–52335.

Xu, H., Paerl, H. W., Qin, B., Zhu, G., Hall, N. S., & Wu, Y. (2015). Determining critical nutrient thresholds needed to control harmful cyanobacterial blooms in eutrophic Lake Taihu, China. *Environmental Science & Technology*, *49*(2), 1051–1059.

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. *The Science of the Total Environment*, *541*, 1592–1602.

Yang, X., Post, W. M., Thornton, P. E., & Jain, A. (2013). The distribution of soil phosphorus for global biogeochemical modeling. *Biogeosciences* , *10*(4), 2525–2537.

Yue, S., & Wang, C. Y. (2004). The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resources Management*. https://doi.org/10.1023/B:WARM.0000043140.61082.60

Zhang, G., & Lu, Y. (2012). Bias-corrected random forests in regression. *Journal of Applied Statistics*, *39*(1), 151–160.

Zhang, Q., Bostic, J. T., & Sabo, R. D. (2022). Regional patterns and drivers of total nitrogen trends in the Chesapeake Bay watershed: Insights from machine learning approaches and management implications. *Water Research*, *218*(118443), 118443.

Zhang, Q., Bostic, J. T., & Sabo, R. D. (2023). Effects of point and nonpoint source controls on total phosphorus load trends across the Chesapeake Bay watershed, USA. *Environmental Research Letters: ERL [Web Site]*, *19*(1), 014012.

Zhang, Q., Murphy, R. R., Tian, R., Forsyth, M. K., Trentacoste, E. M., Keisman, J., & Tango, P. J. (2018). Chesapeake Bay's water quality condition has been recovering: Insights from a multimetric indicator assessment of thirty years of tidal monitoring data. *The Science of the Total Environment*, *637-638*, 1617–1625.

Zhao, Z., Zhang, L., & Deng, C. (2022). Changes in net anthropogenic nitrogen and phosphorus inputs in the Yangtze River Economic Belt, China (1999–2018). *Ecological Indicators*, *145*, 109674.

Zhong, W., Wang, S., Dong, Y., Ni, Z., Fan, Y., & Wu, D. (2022). Trends of the response-relationship between net anthropogenic nitrogen and phosphorus inputs (NANI/NAPI) and TN/TP export fluxes in Raohe basin, China. *Chemosphere*, *286*(Pt 1), 131662.

Zomer, R., & Trabucco, A. (2022). *Global Aridity Index and potential evapotranspiration (ET0) database: Version 6* [Dataset]. figshare. https://doi.org/10.6084/M9.FIGSHARE.7504448.V6

Zou, T., Zhang, X., & Davidson, E. A. (2022). Global trends of cropland phosphorus use and sustainability challenges. *Nature*, 1–7.

Appendix (Supplemental Figures and Tables)

Supplemental

Chapter 2

Supplemental Table 2.1. Normalized livestock weights (1958-2017). Weights of each livestock type are normalized to 2008 values, as reported by the USDA.

Year	Cattle, beef cows	Cattle, milk cows	Cattle, other	Hogs	Sheep	Broilers	Other Chickens	Turkeys	Goats	Horses
1958	0.79	0.79	0.79	0.89	0.71	0.60	0.85	0.52	0.71	1.00
1959	0.81	0.81	0.81	0.89	0.72	0.60	0.85	0.52	0.72	1.00
1960	0.80	0.80	0.80	0.89	0.72	0.60	0.85	0.52	0.72	1.00
1961	0.81	0.81	0.81	0.90	0.71	0.62	0.85	0.54	0.71	1.00
1962	0.80	0.80	0.80	0.90	0.71	0.61	0.84	0.55	0.71	1.00
1963	0.81	0.81	0.81	0.90	0.71	0.62	0.83	0.56	0.71	1.00
1964	0.81	0.81	0.81	0.91	0.71	0.62	0.83	0.56	0.71	1.00
1965	0.79	0.79	0.79	0.90	0.73	0.63	0.82	0.57	0.73	1.00
1966	0.80	0.80	0.80	0.91	0.74	0.63	0.81	0.57	0.74	1.00
1967	0.81	0.81	0.81	0.90	0.73	0.64	0.80	0.59	0.73	1.00
1968	0.80	0.80	0.80	0.90	0.74	0.64	0.80	0.61	0.74	1.00
1969	0.80	0.80	0.80	0.89	0.75	0.65	0.80	0.62	0.75	1.00
1970	0.82	0.82	0.82	0.90	0.76	0.65	0.80	0.62	0.76	1.00
1971	0.81	0.81	0.81	0.89	0.75	0.66	0.78	0.62	0.75	1.00
1972	0.82	0.82	0.82	0.89	0.76	0.67	0.77	0.63	0.76	1.00
1973	0.82	0.82	0.82	0.90	0.78	0.67	0.76	0.62	0.78	1.00
1974	0.82	0.82	0.82	0.91	0.76	0.68	0.76	0.63	0.76	1.00
1975	0.79	0.79	0.79	0.90	0.76	0.67	0.74	0.61	0.76	1.00
1976	0.80	0.80	0.80	0.89	0.79	0.68	0.76	0.63	0.79	1.00
1977	0.81	0.81	0.81	0.89	0.79	0.69	0.74	0.64	0.79	1.00
1978	0.81	0.81	0.81	0.89	0.80	0.70	0.75	0.65	0.80	1.00

1979	0.82	0.82	0.82	0.90	0.81	0.71	0.77	0.65	0.81	1.00
1980	0.83	0.83	0.83	0.90	0.81	0.71	0.76	0.64	0.81	1.00
1981	0.83	0.83	0.83	0.90	0.82	0.72	0.78	0.66	0.82	1.00
1982	0.84	0.84	0.84	0.91	0.83	0.72	0.77	0.67	0.83	1.00
1983	0.84	0.84	0.84	0.91	0.83	0.73	0.78	0.68	0.83	1.00
1984	0.85	0.85	0.85	0.91	0.84	0.75	0.77	0.68	0.84	1.00
1985	0.85	0.85	0.85	0.91	0.84	0.75	0.79	0.69	0.84	1.00
1986	0.86	0.86	0.86	0.92	0.85	0.76	0.78	0.69	0.85	1.00
1987	0.86	0.86	0.86	0.92	0.86	0.77	0.78	0.70	0.86	1.00
1988	0.88	0.88	0.88	0.93	0.90	0.77	0.79	0.72	0.90	1.00
1989	0.89	0.89	0.89	0.93	0.90	0.78	0.82	0.72	0.90	1.00
1990	0.89	0.89	0.89	0.93	0.91	0.78	0.81	0.73	0.91	1.00
1991	0.91	0.91	0.91	0.94	0.91	0.79	0.82	0.73	0.91	1.00
1992	0.91	0.91	0.91	0.94	0.91	0.81	0.82	0.75	0.91	1.00
1993	0.91	0.91	0.91	0.95	0.93	0.82	0.82	0.77	0.93	1.00
1994	0.93	0.93	0.93	0.95	0.91	0.83	0.86	0.78	0.91	1.00
1995	0.92	0.92	0.92	0.96	0.91	0.84	0.88	0.80	0.91	1.00
1996	0.91	0.91	0.91	0.95	0.93	0.86	0.90	0.82	0.93	1.00
1997	0.92	0.92	0.92	0.95	0.97	0.86	0.87	0.83	0.97	1.00
1998	0.94	0.94	0.94	0.96	0.96	0.87	0.89	0.85	0.96	1.00
1999	0.94	0.94	0.94	0.96	0.97	0.89	0.90	0.88	0.97	1.00
2000	0.95	0.95	0.95	0.98	0.99	0.90	0.92	0.88	0.99	1.00
2001	0.95	0.95	0.95	0.98	1.02	0.90	0.92	0.90	1.02	1.00
2002	0.98	0.98	0.98	0.99	0.98	0.92	0.93	0.92	0.98	1.00
2003	0.96	0.96	0.96	0.99	0.98	0.93	0.97	0.92	0.98	1.00
2004	0.97	0.97	0.97	0.99	1.00	0.94	0.98	0.94	1.00	1.00
2005	0.98	0.98	0.98	1.00	1.01	0.96	0.99	0.97	1.01	1.00

2006	0.99	0.99	0.99	1.00	1.00	0.98	1.06	0.97	1.00	1.00
2007	0.99	0.99	0.99	1.00	1.00	0.99	1.04	0.98	1.00	1.00
2008	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2009	1.01	1.01	1.01	1.01	1.01	1.00	0.99	1.00	1.01	1.00
2010	1.00	1.00	1.00	1.02	0.99	1.02	0.99	1.01	0.99	1.00
2011	1.00	1.00	1.00	1.02	1.02	1.04	0.99	1.02	1.02	1.00
2012	1.02	1.02	1.02	1.02	1.06	1.05	0.96	1.03	1.06	1.00
2013	1.03	1.03	1.03	1.03	0.99	1.06	0.97	1.05	0.99	1.00
2014	1.04	1.04	1.04	1.06	0.99	1.08	0.99	1.05	0.99	1.00
2015	1.06	1.06	1.06	1.06	1.00	1.10	1.04	1.04	1.00	1.00
2016	1.06	1.06	1.06	1.05	0.99	1.10	1.04	1.06	0.99	1.00
2017	1.05	1.05	1.05	1.05	0.98	1.11	1.06	1.07	0.98	1.00

Supplemental Table 2.2. Crop uptake parameters (field crops). Parameter values for all field crops, including harvested crop units, crop-unit conversion values, dry matter percentages, and fraction N in crop dry matter.

Crop Type	Reporting Unit	Conversion (kg-crop/reporting unit)	Fraction Dry Matter (kg-$crop_{dry}$/kg-crop)	Fraction N in Dry Matter (kg-N /kg-$crop_{dry}$)	Citation
Alfalfa hay	dry tons	907.18	0.904	0.026	(Hong et al., 2013; *IPNI Estimates of Nutrient Uptake and Removal*, n.d.)
Barley	bushels	21.77	0.889	0.021	(Swaney et al., 2012)
Beans	bushels	27.22	0.900	0.046	(Lander et al., 1998a)
Beans	hundredweight	45.36	0.900	0.046	(Lander et al., 1998a)
Canola	pounds	0.45	0.911	0.039	(Lander et al., 1998a)
Corn, grain[A]	bushels	25.40	0.867	0.015	(Hong et al., 2013)
Corn, grain[B]	bushels	25.40	0.850	0.013	(Villamil et al., 2019)
Corn, silage	acres	7967	0.284	0.012	(Hong et al., 2013; *USDA - National Agricultural Statistics Service - Surveys - Crops/stocks*

					(agricultural Survey), n.d.)
Corn, silage	green tons	907.18	0.284	0.012	(Hong et al., 2013)
Corn, total[C]	acres	6264	-	-	(*USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey)*, n.d.)
Sweet corn	acres	4774	0.909	0.020	(Lander et al., 1998a; *U.S. Sweet Corn Statistics (Table 51 and 52)*, 2010)
Cotton	bales	226.80	0.922	0.033	(Hong et al., 2013)
Cropland pasture	acres	907.18	1.000	0.025	(Hong et al., 2013)
Flaxseed	bushels	25.40	0.941	0.037	(Lander et al., 1998a)
Hay, (excl alfalfa)	dry tons	907.18	0.867	0.020	(Hong et al., 2013; *IPNI Estimates of Nutrient Uptake and Removal*, n.d.)
Hops	pounds	0.45	0.700	0.032	(Lander et al., 1998a)
Lentils	pounds	0.45	0.887	0.044	(Lander et al., 1998a)
Lentils	hundredweight	45.36	0.887	0.044	(Lander et al., 1998a)
Millet	bushels	22.67	0.900	0.021	(Lander et al., 1998a; USDA Economic Research Service, 1992)
Millet	acres	527.80	0.900	0.021	(Drew J. Lyon, Paul A Burgener, Karen L. DeBoer, Robert M Harveson, Gary L Hein, Gary W. Hergert, Lenis A. Nelson, 2008; Lander et al., 1998a)
Non-Cropland pasture	acres	453.59	1.000	0.020	(Hong et al., 2013)
Oats	bushels	14.51	0.894	0.020	(Hong et al., 2013)
Peanuts	bushels	9.98	0.940	0.043	(Hong et al., 2013; *Standard Weights per Bushel for Agricultural Commodities*, n.d.)
Peanuts	pounds	0.45	0.940	0.043	(Hong et al., 2013)
Potatoes	hundredweight	45.36	0.223	0.016	(Hong et al., 2013)
Potatoes	acres	18688	0.223	0.016	(Hong et al., 2013; *USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey)*, n.d.)
Potatoes	bushels	27.22	0.223	0.016	(Hong et al., 2013; USDA Economic Research Service, 1992)
Rice	hundredweight	45.36	0.889	0.014	(Hong et al., 2013)

Rice	bushels	20.25	0.889	0.014	(Hong et al., 2013; USDA Economic Research Service, 1992)
Rye	bushels	25.40	0.881	0.022	(Hong et al., 2013)
Safflower	pounds	0.45	0.924	0.027	(Hong et al., 2013)
Safflower	bushels	17.24	0.924	0.027	(*Grain Conversion Table*, 2013; Lander et al., 1998a)
Safflower	acres	859.1	0.924	0.027	(Lander et al., 1998a; USDA-NASS, 2017)
Snap beans	acres	3872.6	0.130	0.030	(Meisinger & Randall, 1991; *USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey)*, n.d.)
Sorghum, grain	bushels	25.40	0.894	0.020	(Hong et al., 2013)
Sorghum, silage	green tons	907.18	0.580	0.013	(Hong et al., 2013)
Soybeans[A]	bushels	27.22	0.906	0.065	(Hong et al., 2013)
Soybeans[B]	bushels	27.22	0.870	0.062	(Villamil et al., 2019)
Sugar beets	tons	907.19	0.191	0.012	(Lander et al., 1998a)
Sunflower seeds	bushels	14.50	0.933	0.031	(Lander et al., 1998a; USDA Economic Research Service, 1992)
Sunflower seeds	pounds	0.45	0.933	0.031	(Lander et al., 1998a)
Sweet potato	bushels	22.68	0.318	0.008	(Lander et al., 1998a)
Sweet potato	acres	10482	0.318	0.008	(Lander et al., 1998a; *USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey)*, n.d.)
Tobacco	pounds	0.45	1.000	0.032	(Hong et al., 2013)
Triticale	bushels	25.40	0.891	0.030	(Lander et al., 1998a)
Wheat[A]	bushels	27.20	0.885	0.022	(Hong et al., 2013)
Wheat[B]	bushels	27.20	0.865	0.017	(Villamil et al., 2019)

[A] Value before 1980.
[B] Value after 2014
[C] Parameter used in intermediary calculation, "silage corn (kg) = total corn (kg) – corn for grain (kg)" for 1944.

Supplemental Table 2.3. Crop uptake parameters (Fruit and vegetables). Parameter values for vegetables, including crop harvest units, crop-unit conversion values, and fraction N in crop mass.

Crop Type	Reporting Unit	Conversion (kg-crop/ reporting unit)	Fraction N in crop (kg-N/kg-crop)	Citation
Asparagus	acres	756	0.005	(Doerge et al., 1991; USDA Economics, Statistics and Market Information System, 2010a)
Broccoli	acres	5130	0.009	(Doerge et al., 1991; USDA Economics, Statistics and Market Information System, 2011a)
Cantaloupes	acres	9550	0.002	(Bouwman, Van Drescht, et al., 2005; UC Davis, n.d.; USDA Economics, Statistics and Market Information System, 2013a)
Carrots	acres	15866	0.003	(Bouwman, Van Drescht, et al., 2005; USDA Economics, Statistics and Market Information System, 2011b)
Lettuce	acres	13467	0.002	(Geisseler & Horwath, 2016; USDA Economics, Statistics and Market Information System, 2011c)
Onions	acres	16795	0.0026	(Geisseler et al., 2022; USDA Economics, Statistics and Market Information System, 2011d)
Peppers	acres	11122	0.001	(Bouwman, Van Drescht, et al., 2005; USDA Economics, Statistics and Market Information System, 2008)
Pumpkin	acres	9072	0.001	(Aegerter et al., 2013; Bouwman, Van Drescht, et al., 2005; USDA Economic Research Service, n.d.)
Squash	acres	6008	0.001	(Bouwman, Van Drescht, et al., 2005; National IPM Database - 2002 Crop Profile for Squash in Florida: Document Report, n.d.; New Jersey Department of Agriculture, 2015)
Strawberry	acres	15236	0.003	(Bottoms et al., 2013; USDA Economics, Statistics and Market Information System, 2013b; Weber et al., n.d.)
Tomatoes	acres	20550	0.0013	(Bouwman, Van Drescht, et al., 2005; Geisseler & Horwath, n.d.; USDA Economics, Statistics and Market Information System, 2010b)

Watermelon	acres	8417	0.002	(Bouwman, Van Drescht, et al., 2005; Weber et al., n.d.)
Sugarcane	acres	33109	0.002	(Bouwman, Van Drescht, et al., 2005; USDA - National Agricultural Statistics Service - State Agricultural Overview, n.d.)
Sugarcane	tons	907.19	0.002	(Bouwman, Van Drescht, et al., 2005)

Supplemental Table 2.4. Crop uptake parameters (Fruits and Orchard Crops). Parameter values for fruits and orchard crops, including crop harvest units and crop-unit N uptake values.

Crop Type	Reporting Unit	N Uptake (kg-N/reporting unit)	Citation
Almonds	trees	0.82	(Geisseler, 2016; University of California Cooperative Extension, 2016)
Apple	trees	0.23	(Neilsen et al., 2001; Neilsen & Neilsen, 2002)
Apricot	trees	0.15	(Sanchez, 2016; UC Davis, n.d.)
Avocado	trees	1.06	(Lovatt, 1996; UC Davis, n.d.)
Cherry	trees	0.23	(De Moura R., n.d.; San-Martino et al., 2010)
Grapefruit	trees	0.31	(Dagan, 1987)
Olive	trees	0.06	(Belguerri et al., 2016)
Orange	trees	0.31	(Dagan, 1987)
Peach	trees	0.15	(Sanchez, 2016; UC Davis, n.d.)
Pear	trees	0.23	(Neilsen et al., 2001; Neilsen & Neilsen, 2002; Raese, 1998)
Pecan	trees	1.50	(Kraimer et al., 2001, 2004)
Pistachio	trees	0.44	(UC Davis, n.d.)
Prune and plums	trees	0.38	(UC Davis, n.d.)
Walnut	trees	1.30	(UC Davis, n.d.; Weinbaum et al., 1991)
Grapes	vines	0.03	(Hanson & Howell, 1995; Peacock et al., 1991)
Cranberry	acres	18.14	(Roper et al., 2000)
Blackberries	acres	13.50	(Mohadjer et al., 2015; Rempel et al., 2004)
Blueberry	acres	10.93	(White et al., 2006)
Raspberries	acres	16.20	(Mohadjer et al., 2015; Rempel et al., 2004)

Supplemental

Chapter 3

Supplemental Table 3.1. Aggregation and Disaggregation of Census Cattle and Poultry Categories, 1930-2017

Year	Data Available	Description of Process
Cattle and cows		
1930	Cow & heifers > 2 yr, beef production	Cow and heifers > 2yr for milk and beef were categorized as milk and beef cattle, respectively. 'Other cattle' was calculated by summing the remaining categories.
	Cow & heifers > 2 yr, milk production	
	Cow & heifers < 2 yr, beef production	
	Cow & heifers < 2 yr, milk production	
	Heifers, yearlings	
	Bulls > 1yr	
	Steers & bulls, 1 yr	
	Steers, 2 yrs	
	Steers > 2 yrs	
	Calves	
1935	Cows & heifers > 2 yr	Cows and heifers > 2yrs was split into milk and beef cattle using county-scale fractions (i.e. milk cattle to cows & heifers > 2 yr ratio) from 1940. 'Other cattle' was calculated by summing the remaining categories.
	Heifers < 2 yr	
	Steers & bull > 1 yr	
	Calves	
1940	All cattle	'Other cattle' was calculated by finding the difference between all cattle and milk and beef cattle.
	Cow & heifers > 2 yr, beef production	
	Cow & heifers > 2 yr, milk production	
1945	All cattle	Cows and heifers > 2yrs was split into milk and beef cattle using county-scale fractions (i.e. milk cattle to cows & heifers > 2 yr ratio) from 1940. 'Other cattle' was calculated by finding the difference between all cattle and milk and beef cattle.
	Cows & heifers > 2 yr	

1950	All cattle	The difference between Cows & heifers that have calved and Cows and heifers > 2yrs, milk production was used to calculate beef cattle. 'Other cattle' was calculated by finding the difference between all cattle and milk and beef cattle.
	Cows & heifers > 2 yr	
	Cow & heifers > 2yr, milk production	
1954 to 1964	Cows & heifers that have calved	The difference between Cows & heifers that have calved and Cows and heifers > 2yrs, milk production was used to calculate beef cattle.' Other cattle' was calculated by summing the remaining categories.
	Cow & heifers > 2 yr, milk production	
	Steers & bulls, incld. calves	
	Heifers & heifer calves, number	
1969 to 1997	Cow & heifers > 2 yr, beef production	'Other cattle' was calculated by summing the remaining categories.
	Cow & heifers > 2 yr, milk production	
	Steers & bulls, incld. calves	
	Heifers & heifer calves, number	
2002 to 2017	Cow & heifers > 2 yr, beef production	No calculations required.
	Cow & heifers > 2 yr, milk production	
	Other cattle	

Supplemental Table 3.1 (cont'd)

Chicken and Broiler Chickens		
1930 to 1964	Chickens > 3 mths	Chickens > 3 months categorized as other chickens. Broilers were scaled from 1969 to 1930 using rate of change of broiler consumption (kg consumed) and broiler live weight (kg meat/chicken) to establish the percent change of national broiler inventory for each census year.
1969 to 1974	Chickens > 4 mths	Chickens > 4 months categorized as other chickens.
	Broilers & other meat chickens	
1978 to 1992	Chickens > 3 mths	Chickens > 3 mth and pullets < 3 mths are categorized as other chickens.
	Pullets < 3 mths	
	Broilers & other meat chickens	
1997	Layers & pullets > 3 mths	Layers & pullets > 13 wks and pullet chicks & pullets < 13 wks are categorized as other chickens.
	Pullet chicks & pullets < 3 mths	
	Broilers & other meat chickens	
2002 to 2017	Layers > 20 wks	Layers > 20 wks and pullet checks for laying flock replacement are categorized as other chickens.
	Pullet chickens for laying flock replacement	
	Broilers & other meat chickens	

Supplemental Table 3.2. Crop phosphorus uptake parameters (field crops). Parameter values for all field crops, including harvested crop units, crop-unit conversion values, fraction of P in crops, dry matter percentages, and fraction P in crop dry matter.

Crop Type	Reporting Unit	Conversion (kg-crop/reporting unit)	Fraction P in Crop (kg-P /kg-$crop_{dry}$)	Fraction Dry Matter (kg-$crop_{dry}$/ kg-crop)	Fraction P in Dry Matter (kg-P /kg-$crop_{dry}$)	Citation
Alfalfa hay	dry tons	907.18	-	0.904	0.0026	(IPNI Estimates of Nutrient Uptake and Removal, n.d.; Swaney & Howarth, 2019a)
Barley	bushels	21.77	-	0.889	0.0042	(Swaney & Howarth, 2019a)
Beans	bushels	27.22	-	0.900	0.0054	(Lander et al., 1998a)
Beans	hundredweight	45.36	-	0.900	0.0054	(Lander et al., 1998a)
Canola	pounds	0.45	-	0.911	0.0062	(Lander et al., 1998a)
Corn, grain[A]	bushels	25.40	-	0.867	0.00317	(Swaney & Howarth, 2019a)
Corn, grain[B]	bushels	25.40	-	0.850	0.0027	(Villamil et al., 2019)
Corn, silage	acres	7967	-	0.284	0.0019	(Swaney & Howarth, 2019a; USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey), n.d.)
Corn, silage	green tons	907.18	-	0.284	0.0019	(Swaney & Howarth, 2019a)
Corn, total[C]	acres	6264	-	-	-	(USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey), n.d.)
Cotton	bales	226.80	-	0.922	0.0041	(Swaney & Howarth, 2019a)
Cropland pasture	acres	907.18	-	1.000	0.0055	(Swaney & Howarth, 2019a)
Flaxseed	bushels	25.40	-	0.941	0.0057	(Lander et al., 1998a)
Hay, (excl alfalfa)	dry tons	907.18	-	0.867	0.0026	(IPNI Estimates of Nutrient Uptake and Removal, n.d.)
Hops	pounds	0.45	-	0.700	0.005	(Lander et al., 1998a)
Lentils	pounds	0.45	-	0.887	0.0043	(Lander et al., 1998a)
Lentils	hundredweight	45.36	-	0.887	0.0043	(Lander et al., 1998a)
Millet	bushels	22.67	-	0.900	0.0034	(Lander et al., 1998a; USDA Economic Research Service, 1992)
Millet	acres	527.80	-	0.900	0.0034	(Drew J. Lyon, Paul A Burgener, Karen L. DeBoer, Robert M Harveson, Gary L Hein, Gary W.

						Hergert, Lenis A. Nelson, 2008; Lander et al., 1998a)
Non-Cropland pasture	acres	453.59	-	1.000	0.0044	(Swaney & Howarth, 2019a)
Oats	bushels	14.51	-	0.894	0.0037	(Swaney & Howarth, 2019a)
Peanuts	bushels	9.98	-	0.940	0.0035	(Standard Weights per Bushel for Agricultural Commodities, n.d.; Swaney & Howarth, 2019a)
Peanuts	pounds	0.45	-	0.940	0.0035	(Swaney & Howarth, 2019a)
Potatoes	hundredweight	45.36	-	0.223	0.0025	(Swaney & Howarth, 2019a)
Potatoes	acres	18688	-	0.223	0.0025	(Swaney & Howarth, 2019a; USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey), n.d.)
Potatoes	bushels	27.22	-	0.223	0.0025	(Swaney & Howarth, 2019a; USDA Economic Research Service, 1992)
Rice	hundredweight	45.36	-	0.889	0.0033	(Swaney & Howarth, 2019a)
Rice	bushels	20.25	-	0.889	0.0033	(Swaney & Howarth, 2019a; USDA Economic Research Service, 1992)
Rye	bushels	25.40	-	0.881	0.0037	(Swaney & Howarth, 2019a)
Safflower	pounds	0.45	-	0.924	0.006	(Swaney & Howarth, 2019a)
Safflower	bushels	17.24	-	0.924	0.006	(Grain Conversion Table, 2013; Lander et al., 1998a)
Safflower	acres	859.1	-	0.924	0.006	(Lander et al., 1998a; USDA-NASS, 2017)
Sorghum, grain	bushels	25.40	-	0.894	0.0035	(Swaney & Howarth, 2019a)
Sorghum, silage	green tons	907.18	-	0.580	0.0021	(Swaney & Howarth, 2019a)
Soybeans[A]	bushels	27.22	-	0.906	0.0066	(Swaney & Howarth, 2019a)
Soybeans[B]	bushels	27.22	-	0.870	0.0052	(Villamil et al., 2019)
Sugar beets	tons	907.19	-	0.191	0.0025	(Lander et al., 1998a)
Sugarcane	acres	33109	0.00015	-	-	(Bouwman, Van Drescht, et al., 2005; United States Department of Agriculture, n.d.; USDA - National Agricultural Statistics Service - State Agricultural Overview, n.d.)
Sugarcane	tons	907.19	0.00015	-	-	(Bouwman, Van Drescht, et al., 2005;

						United States Department of Agriculture, n.d.)
Sunflower seeds	bushels	14.50	-	0.933	0.0063	(Lander et al., 1998a; USDA Economic Research Service, 1992)
Sunflower seeds	pounds	0.45	-	0.933	0.0063	(Lander et al., 1998a)
Sweet potato	bushels	22.68	-	0.318	0.0015	(Lander et al., 1998a)
Sweet potato	acres	10482	-	0.318	0.0015	(Lander et al., 1998a; USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey), n.d.)
Tobacco	pounds	0.45	-	1.000	0.0021	(Swaney & Howarth, 2019a)
Triticale	bushels	25.40	-	0.891	0.0034	(Lander et al., 1998a)
Wheat[A]	bushels	27.20	-	0.885	0.0040	(Swaney & Howarth, 2019a)
Wheat[B]	bushels	27.20	-	0.865	0.0027	(Villamil et al., 2019)

[A] Value before 1980
[B] Value after 2014
[C] Parameter used in intermediary calculation, "silage corn (kg) = total corn (kg) – corn for grain (kg)" for 1944.

Supplemental Table 3.3. Crop uptake parameters (vegetables and melons). Parameter values for fruits and vegetables, including harvested crop units, crop-unit conversion values, fraction P in crop mass, fraction of dry matter in crops, and fraction of P in dry crop matter.

Crop Type	Reporting Unit	Conversion (kg-crop/reporting unit)	Fraction P in crop (g-P/kg-crop)	Fraction Dry Matter ($kg\text{-}crop_{dry}/kg\text{-}crop$)	Fraction P in Dry Matter ($kg\text{-}P/kg\text{-}crop_{dry}$)	Citation
Asparagus	acres	756	0.703	-	-	(Doerge et al., 1991; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2010a)
Broccoli	acres	5130	1.318	-	-	(Doerge et al., 1991; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2011a)
Cantaloupes	acres	9550	0.224	-	-	(Bouwman, Van Drescht, et al., 2005; UC Davis, n.d.; United States Department of

						Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2013a)
Carrots	acres	15866	0.705	-	-	(Bouwman, Van Drescht, et al., 2005; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2011b)
Lettuce	acres	13467	0.304	-	-	(Geisseler & Horwath, 2016; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2011c)
Onions	acres	16795	0.528	-	-	(Geisseler et al., 2022; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2011d)
Peppers	acres	11122	0.146	-	-	(Bouwman, Van Drescht, et al., 2005; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2008)
Pumpkin	acres	9072	0.275	-	-	(Aegerter et al., 2013; Bouwman, Van Drescht, et al., 2005; United States Department of Agriculture, n.d.; USDA Economic Research Service, n.d.)
Snap beans	acres	3872.6	-	0.13	0.004	(Meisinger & Randall, 1991; United States Department of Agriculture, n.d.; USDA - National Agricultural Statistics Service - Surveys - Crops/stocks (agricultural Survey), n.d.)
Squash	acres	6008	0.206	-	-	(Bouwman, Van Drescht, et al., 2005; National IPM Database - 2002 Crop Profile for Squash in Florida: Document Report, n.d.; New Jersey Department of

						Agriculture, 2015; United States Department of Agriculture, n.d.)
Sweet corn	acres	4774.0	-	0.909	0.004	(Lander et al., 1998a; United States Department of Agriculture, n.d.; U.S. Sweet Corn Statistics (Table 51 and 52), 2010)
Tomatoes	acres	20550	0.169	-	-	(Bouwman, Van Drescht, et al., 2005; Geisseler & Horwath, n.d.; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2010b)
Watermelon	acres	8417	0.225	-	-	(Bouwman, Van Drescht, et al., 2005; United States Department of Agriculture, n.d.; Weber et al., n.d.)

Supplemental Table 3.4. Crop uptake parameters (fruit and orchard crops). Parameter values for fruits and orchard crops, including harvested crop units, crop-unit conversion values, fraction of P in crops, and crop-unit P uptake values.

Crop Type	Reporting Unit	Conversion (kg-crop/ reporting unit)	Fraction P in crop (g-P/kg-crop)	P Uptake per reporting unit (kg-P/reporting unit)	Citation
Almonds	trees	-	-	0.112	(Geisseler, 2016; United States Department of Agriculture, n.d.; University of California Cooperative Extension, 2016)
Apple	trees	-	-	0.061	(Neilsen et al., 2001; Neilsen & Neilsen, 2002)
Apricot	trees	-	-	0.015	7,8(United States Department of Agriculture, n.d.)
Avocado	trees	-	-	0.172	(UC Davis, n.d.; United States Department of Agriculture, n.d.)
Blackberries	acres	-	-	2.15	(Mohadjer et al., 2015; Rempel et al., 2004; United States Department of Agriculture, n.d.)
Blueberry	acres	-	-	1.11	(United States Department of Agriculture, n.d.; White et al., 2006)

Cherry	trees	-	-	0.013	(De Moura R., n.d.; San-Martino et al., 2010; United States Department of Agriculture, n.d.)
Cranberry	acres	-	-	3.78	(Roper et al., 2000; United States Department of Agriculture, n.d.)
Grapefruit	trees	-	-	0.045	(Dagan, 1987; United States Department of Agriculture, n.d.)
Grapes	vines	-	-	0.0052	(Hanson & Howell, 1995; Peacock et al., 1991; United States Department of Agriculture, n.d.)
Olive	trees	-	-	0.0012	(Belguerri et al., 2016; United States Department of Agriculture, n.d.)
Orange	trees	-	-	0.018	(Dagan, 1987; United States Department of Agriculture, n.d.)
Peach	trees	-	-	0.021	(UC Davis, n.d.; United States Department of Agriculture, n.d.)
Pear	trees	-	-	0.041	(Neilsen et al., 2001; Neilsen & Neilsen, 2002; Raese, 1998; United States Department of Agriculture, n.d.)
Pecan	trees	-	-	0.283	(Kraimer et al., 2001, 2004; United States Department of Agriculture, n.d.)
Pistachio	trees	-	-	0.067	(UC Davis, n.d.; United States Department of Agriculture, n.d.)
Prune and plums	trees	-	-	0.055	(UC Davis, n.d.; United States Department of Agriculture, n.d.)
Raspberries	acres	-	-	2.43	(Mohadjer et al., 2015; Rempel et al., 2004; United States Department of Agriculture, n.d.)
Strawberry	acres	15236	0.675	-	(Bottoms et al., 2013; United States Department of Agriculture, n.d.; USDA Economics, Statistics and Market Information System, 2013b; Weber et al., n.d.)
Walnut	trees	-	-	0.184	(UC Davis, n.d.; United States Department of Agriculture, n.d.; Weinbaum et al., 1991)

Supplemental

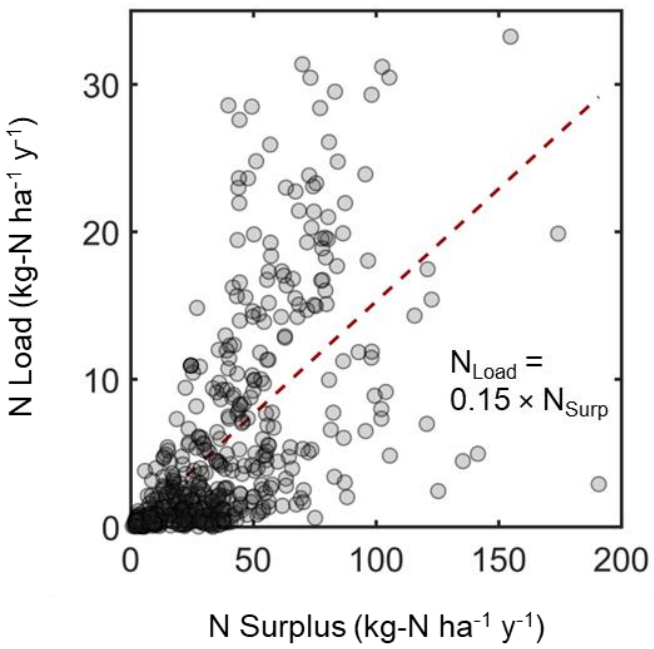
Chapter 4

Supplemental Table 4.1. Watershed properties used in spatial and temporal analyses.

Category	Name (abbreviation in text, if applicable)	Description	Citation
Land Use	Percent wetlands land use	Percent watershed as wetland in 2011.	(Falcone, 2017)
	Percent agricultural land use (Ag LU)	Percent watershed as agricultural land in 2011	(Falcone, 2017)
	Percent urban land use (Urb LU)	Percent watershed as developed land in 2011.	(Falcone, 2017)
	Percent natural land use (Nat LU)	Percent watershed as low impacted land use which includes wetland, open water, forested, shrubland, grassland, and barren.	(Falcone, 2017)
	Population density (POP)	Population density (pop km^{-2}) in 2010 aggregated from 60-m resolution to the watershed-scale.	(Falcone, 2017)
Landscape Management	Percent tile-drained area (TD)	The percentage of agricultural land in the watershed that is currently tile drained.	(Falcone, 2017; Valayamkunnath et al., 2020)
Geomorphology	Percent clay + silt	The silt and clay content as a percentage of total soil (sand, clay, silt).	(Falcone, 2017)
Inputs	Agricultural inputs (IN_{ag})	Total agricultural inputs (livestock manure, inorganic fertilizer, and biological nitrogen fixation) normalized by total watershed area (kg-N ha^{-1} yr^{-1}). Averages were taken from 2003 to the end of the timeseries.	This study

	Cumulative net agricultural inputs (NAI_{CUMU})	The cumulative net agricultural inputs (livestock manure, inorganic fertilizer, biological nitrogen fixation minus crop and pasture uptake) in 2017 or the last year of the export timeseries. Values are normalized by total watershed area (kg-N ha^{-1}).	This study
	Livestock manure to fertilizer ratio (M:F)	The ratio of livestock manure N inputs to the total fertilizer (inorganic fertilizer and manure) normalized by total watershed area (kg-N ha^{-1} yr^{-1}). Averages were taken from 2003 to the end of the timeseries.	This study

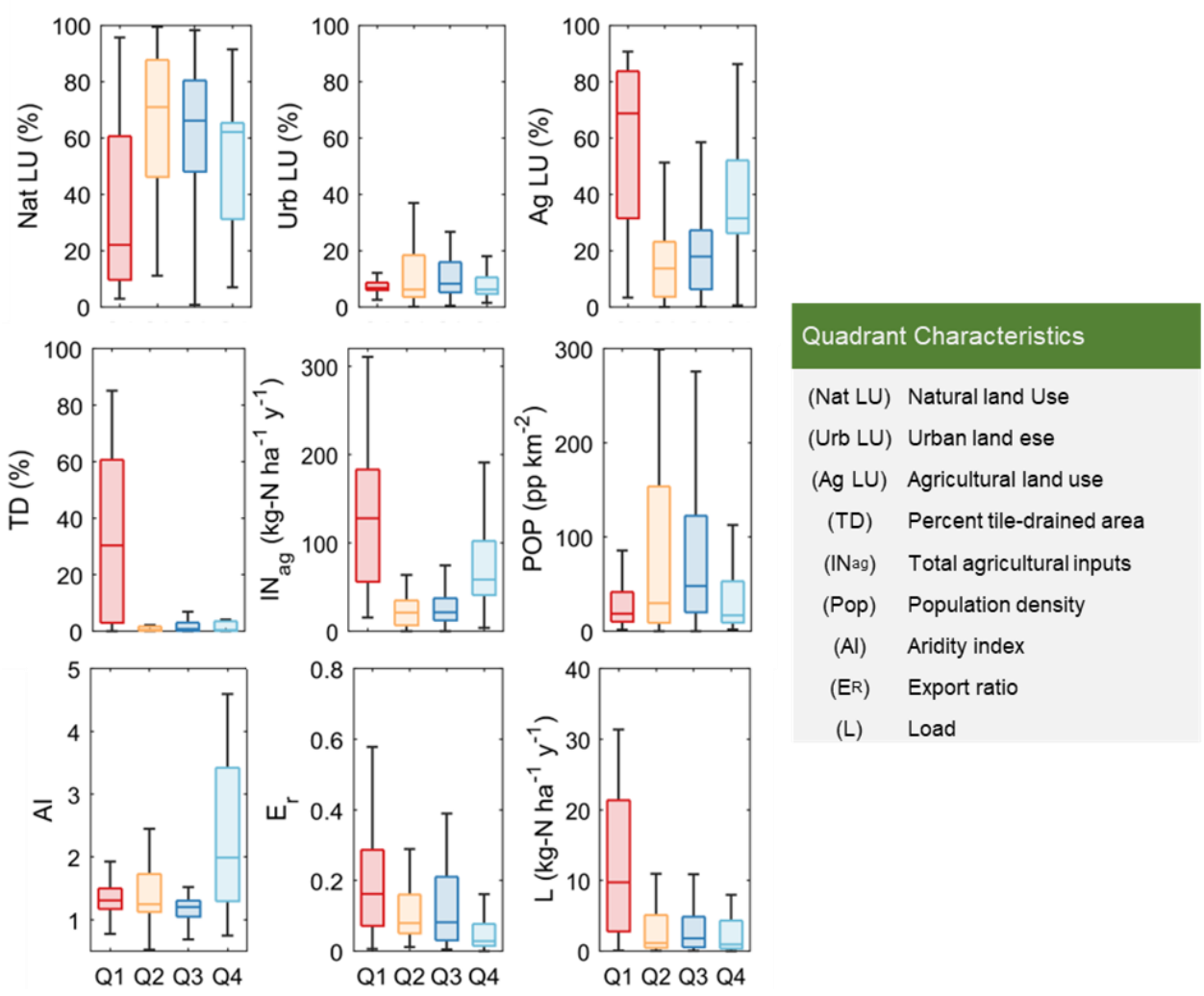
Inputs	Annual percent change in manure (MAN)	The annual percent change of manure inputs in each watershed between 1990 and 2017, or the available time series between periods. Slopes were calculated using Sen slope and normalized by the mean of the time series.	This study
	Annual percent change in fertilizer and fixation (FT+FX)	The annual percent change of inorganic fertilizer and biological nitrogen fixation inputs in each watershed between 1990 and 2017, or the available time series between periods. Slopes were calculated using Sen slope and normalized by the mean of the time series.	This study
	Annual percent change in crop uptake (CROP)	The annual percent change of crop and pasture uptake in each watershed between 1990 and 2017, or the available time series between periods. Slopes were calculated using Sen slope and normalized by the mean of the time series.	This study
Climate	Aridity Index (AI)	Watershed average aridity index, which measures the dryness of a watershed, is calculated by taking the ratio of potential evapotranspiration to precipitation.	[61]
Hydrology	Export Ratio (E_R)	The ratio of the average load to the mean average. Averages were taken from 2003 to the end of the timeseries.	This study
	Load (L)	The average area normalized load from each watershed. Averages were taken from 2003 to the end of the timeseries.	This study



Supplemental Figure 4.1. Nitrogen loads are correlated with surplus (net N inputs) from the landscape across 492 watersheds with loads constituting approximately 15% of N surplus (R^2 = 0.37). The scatter however suggests the influence of other factors in driving N loads.

Supplemental Table 4.2. P value from Wilcoxon rank-sum test for each quadrant between watershed characteristics. Highlighted green boxes are p <0.05, orange boxes are p <0.1, and white boxes are p >0.01.

Characteristics	Q1 vs.			Q2 vs.			Q3 vs.			Q4 vs.		
	Q2	Q3	Q4	Q1	Q3	Q4	Q1	Q2	Q4	Q1	Q2	Q3
Nat LU	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.28	0.02	0.00	0.00	0.02
Urb LU	0.40	0.10	0.31	0.40	0.11	0.99	0.10	0.11	0.07	0.31	0.99	0.07
Ag LU	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.22	0.00	0.00	0.00	0.00
TD	0.00	0.00	0.00	0.00	0.61	0.84	0.00	0.61	0.76	0.00	0.84	0.76
INag	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.42	0.00	0.00	0.00	0.00
LVSTK	0.00	0.00	0.11	0.00	0.23	0.00	0.00	0.23	0.00	0.11	0.00	0.00
POP	0.22	0.00	0.87	0.22	0.08	0.35	0.00	0.08	0.01	0.87	0.35	0.01
AI	0.41	0.00	0.00	0.41	0.04	0.00	0.00	0.04	0.00	0.00	0.00	0.00
ER	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.99	0.00	0.00	0.00	0.00
LOAD	0.00	0.00	0.00	0.00	0.35	0.49	0.00	0.35	0.15	0.00	0.49	0.15



Supplemental Figure 4.2. Distribution of quadrant characteristics.

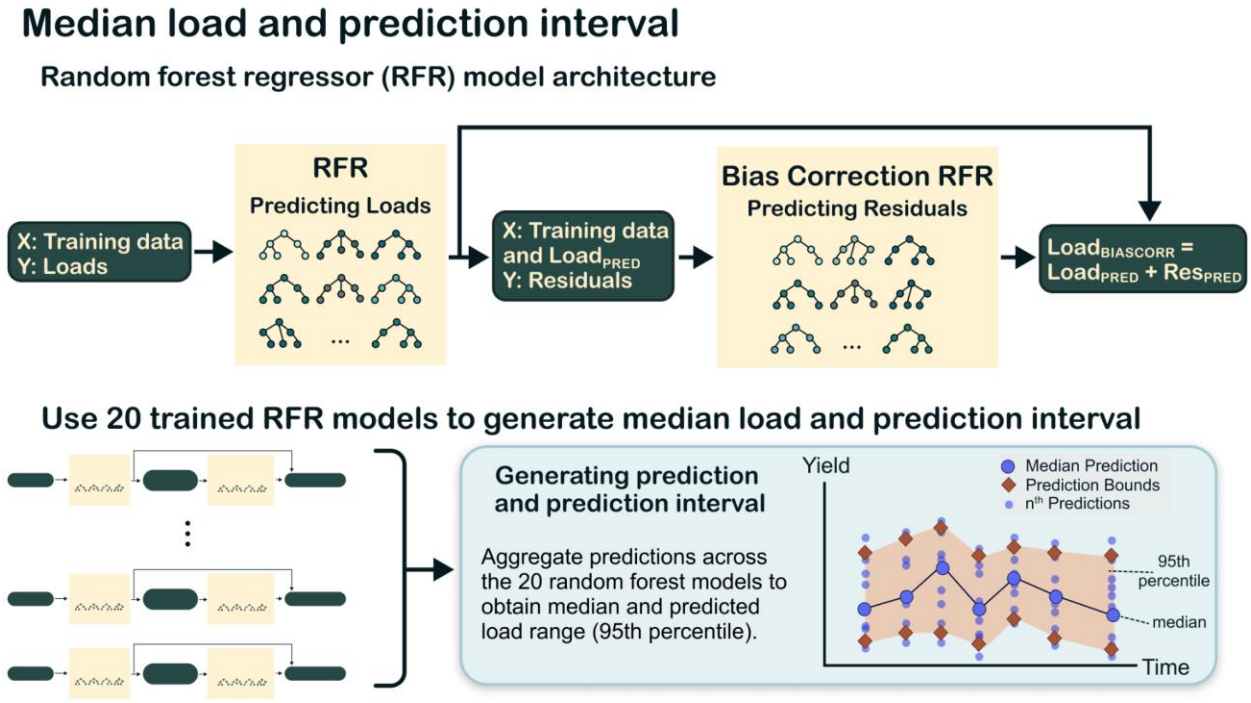
Supplemental

Chapter 5

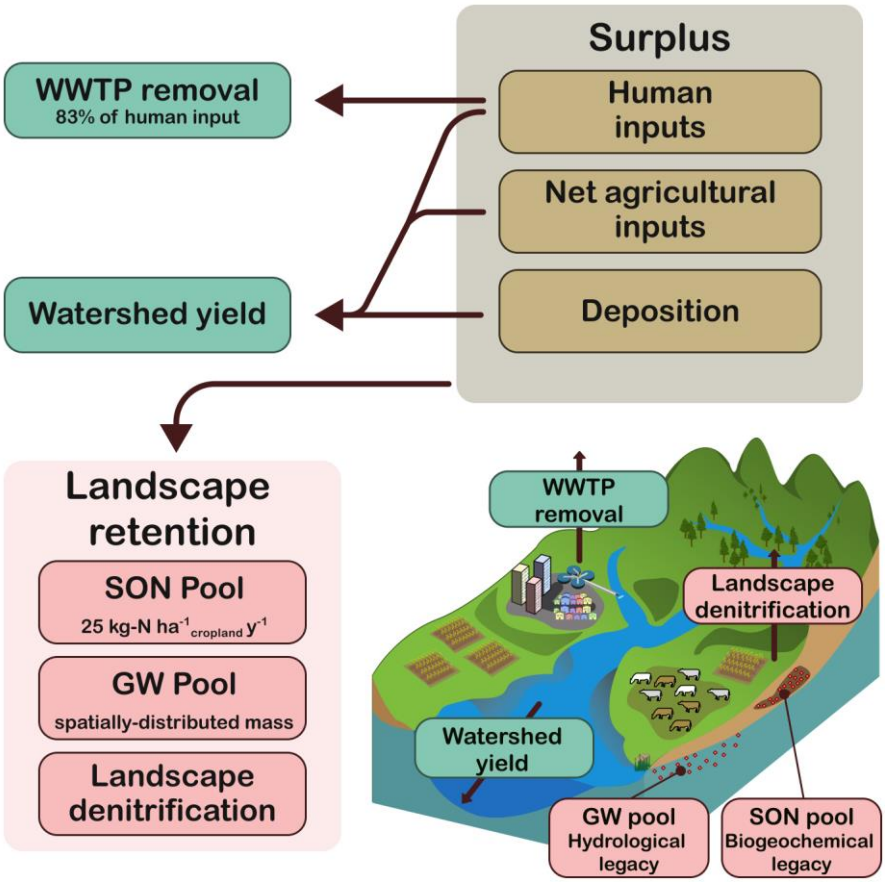
Supplemental Table 5.1. Hyperparameter grid and best parameter for each model iteration. Hyperparameters were selected using Scikit learn GridSearchCV which selects the combination of hyperparameters by maximizing the root mean squared error of the model.

Hyperparameter	Hyperparameter Values Tested
Max tree depth	10 8 5
Max no. features	0.33 0.5 1
No. of trees	500 1000 4000
Minimum leaf sample	1 3 5 8 10

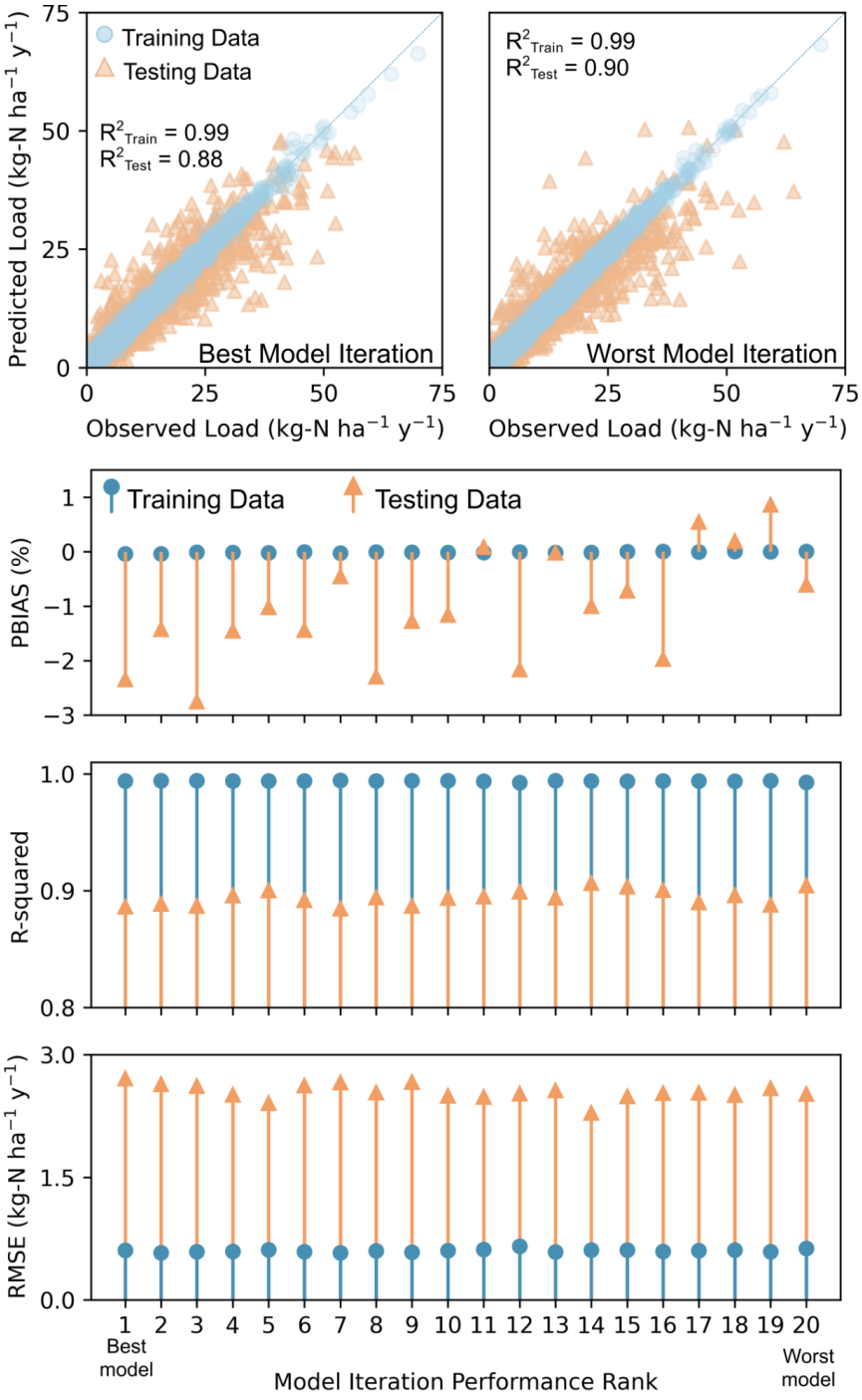
	Random Forest Model				Bias Correction Model			
Itr	Max tree depth	Max no. features	No. of trees	Minimum leaf sample	Max tree depth	Max no. features	No. of trees	Minimum leaf sample
0	10	0.5	4000	1	10	0.5	4000	1
1	10	0.5	500	1	10	0.5	4000	1
2	10	0.5	4000	1	10	0.5	1000	1
3	10	0.5	4000	1	10	0.5	4000	1
4	10	0.5	4000	1	10	0.5	1000	1
5	10	0.5	1000	1	10	0.5	4000	1
6	10	0.5	4000	1	10	0.5	500	1
7	10	0.5	4000	1	10	0.5	4000	1
8	10	0.5	4000	1	10	0.5	1000	1
9	10	0.5	4000	1	10	0.5	500	1
10	10	0.5	4000	1	10	0.5	4000	3
11	10	0.5	500	1	10	0.5	500	1
12	10	0.5	1000	1	10	0.5	500	3
13	10	0.5	500	1	10	0.5	4000	1
14	10	0.5	1000	1	10	0.5	1000	1
15	10	0.5	500	1	10	0.5	4000	1
16	10	0.5	4000	1	10	0.5	4000	1
17	10	0.5	4000	1	10	0.5	1000	1
18	10	0.5	1000	1	10	0.5	1000	1
19	10	0.5	4000	1	10	0.5	4000	1



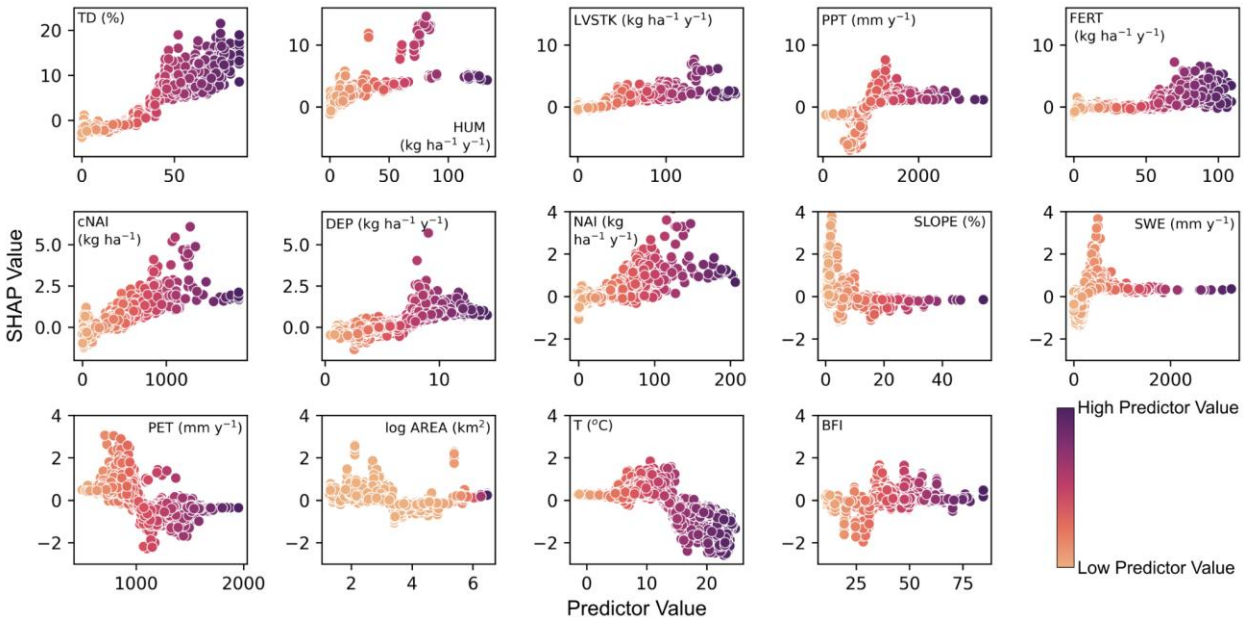
Supplemental Figure 5.1. The process flow of the random forest model framework and applying the set of trained models to predict HUC8 yields and prediction intervals.



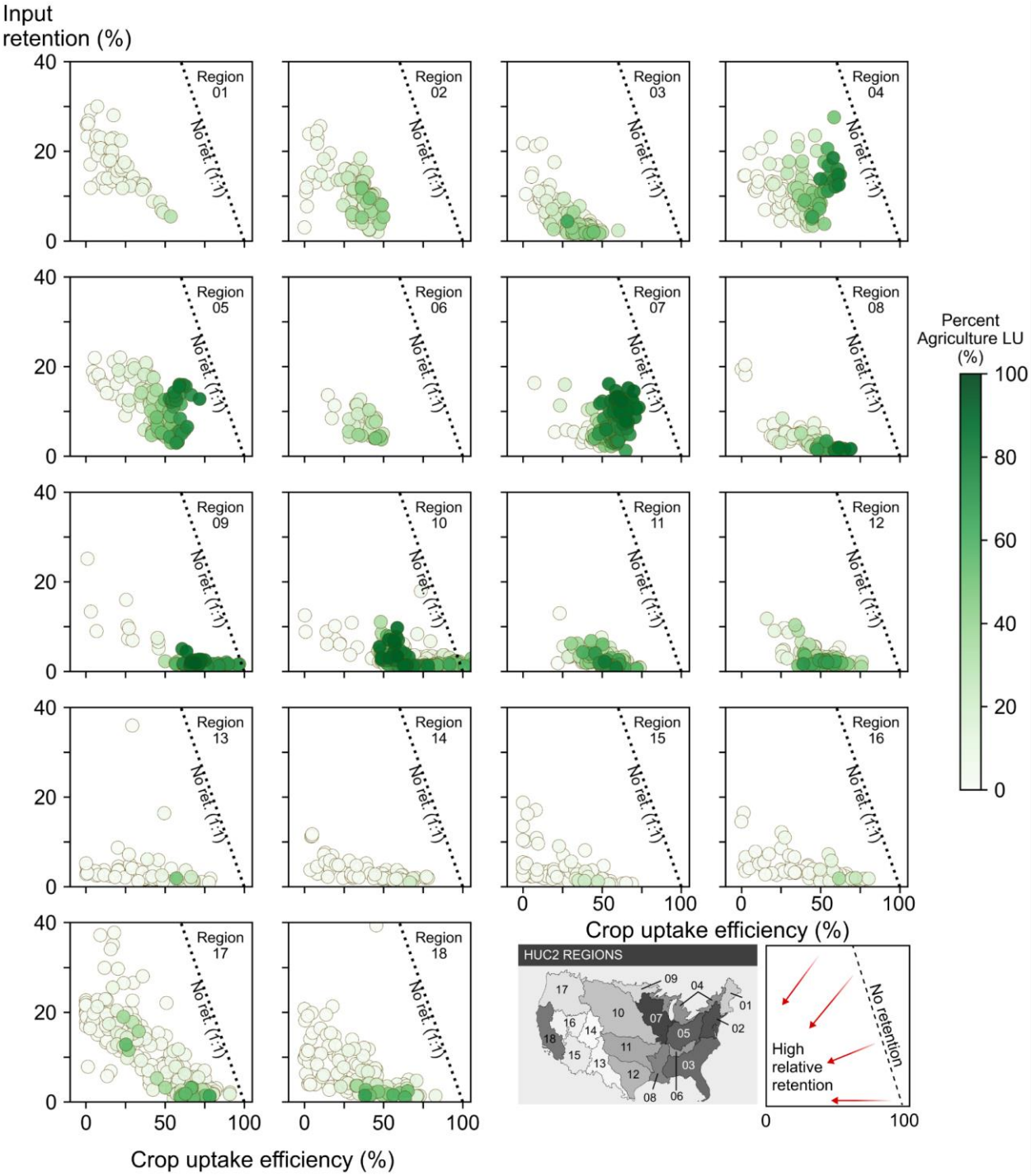
Supplemental Figure 5.2. A schematic depicting the calculation of landscape retention and the pool that encompasses landscape retention.



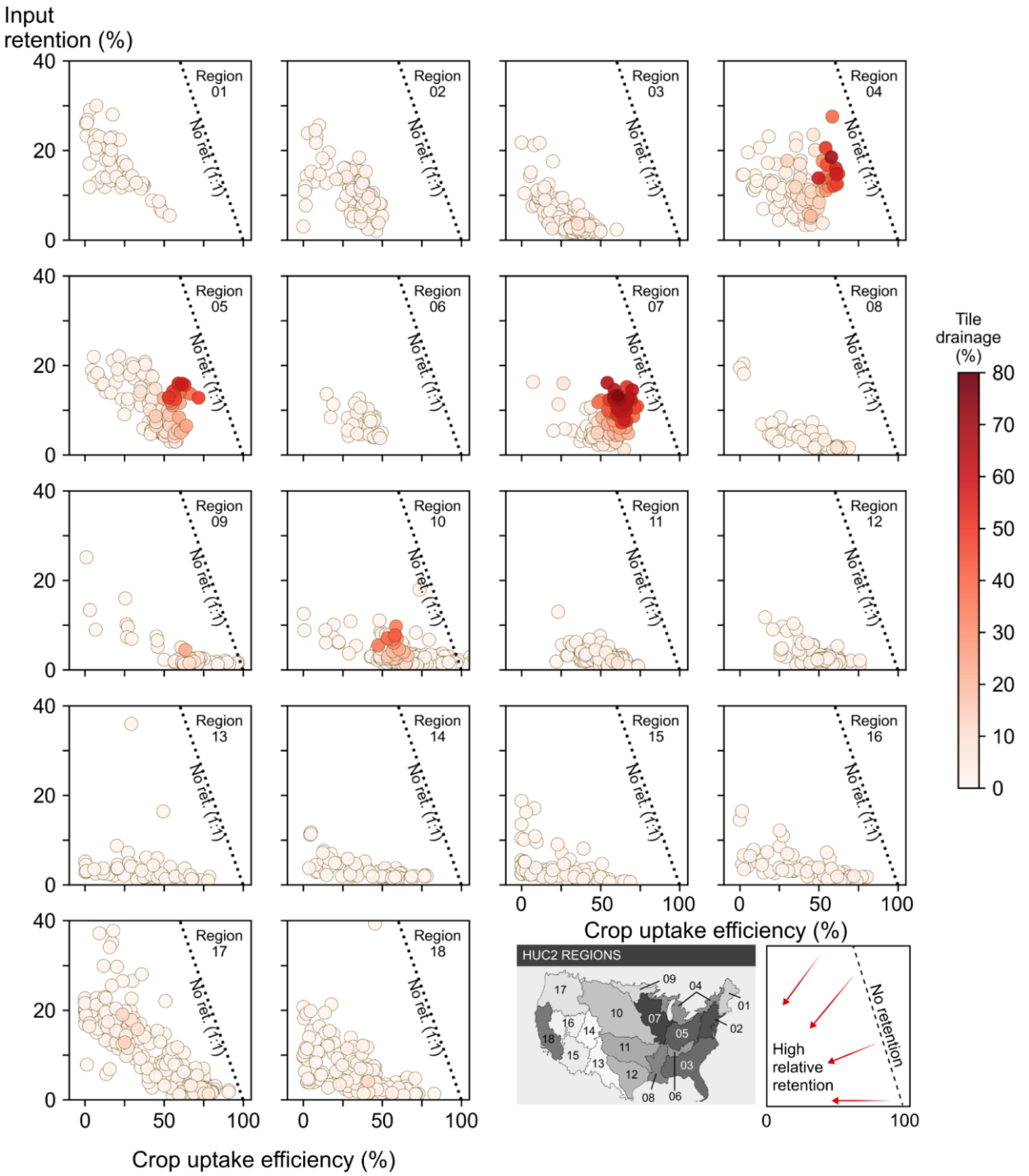
Supplemental Figure 5.3. Summary of model evaluation metrics for the set of 20 trained models. We used all 20 model iterations to generate prediction intervals.



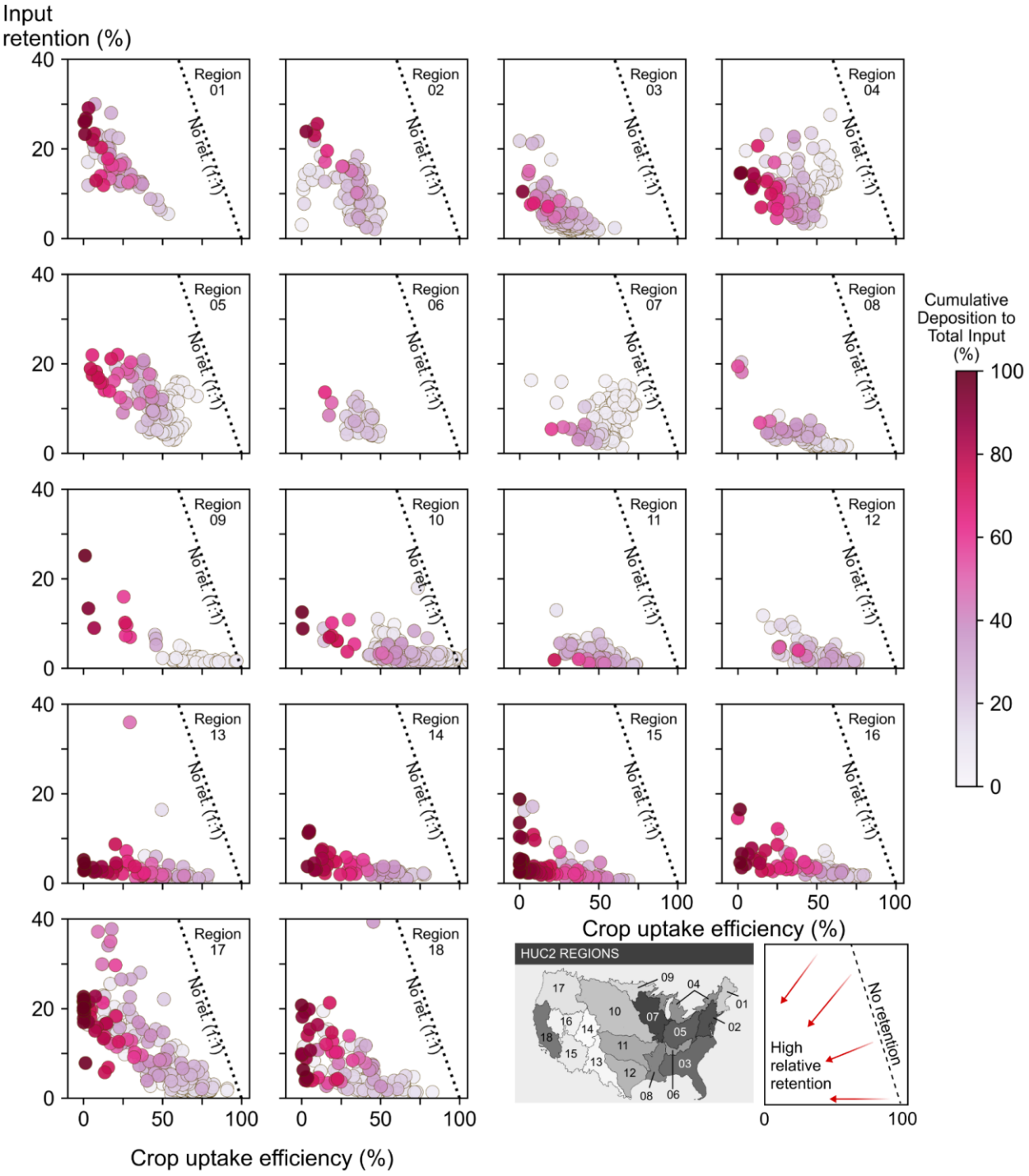
Supplemental Figure 5.4. Partial dependence plot using SHAP values for the best performing model. SHAP values (y-axis) are a measure of the relative importance of the predictor labeled on the top left corner (x-axis) on watershed load prediction. Markers are coloured by the magnitude of the predictor Each panel is a different predictor, and the magnitude of the predictor. Plots ordered by the maximum SHAP value from left to right, indicating the impact of each predictor on the model output. The drainage area predictor' is represented on a logarithmic scale as 'log AREA(km^2)', although the actual values used in the model are not transformed.



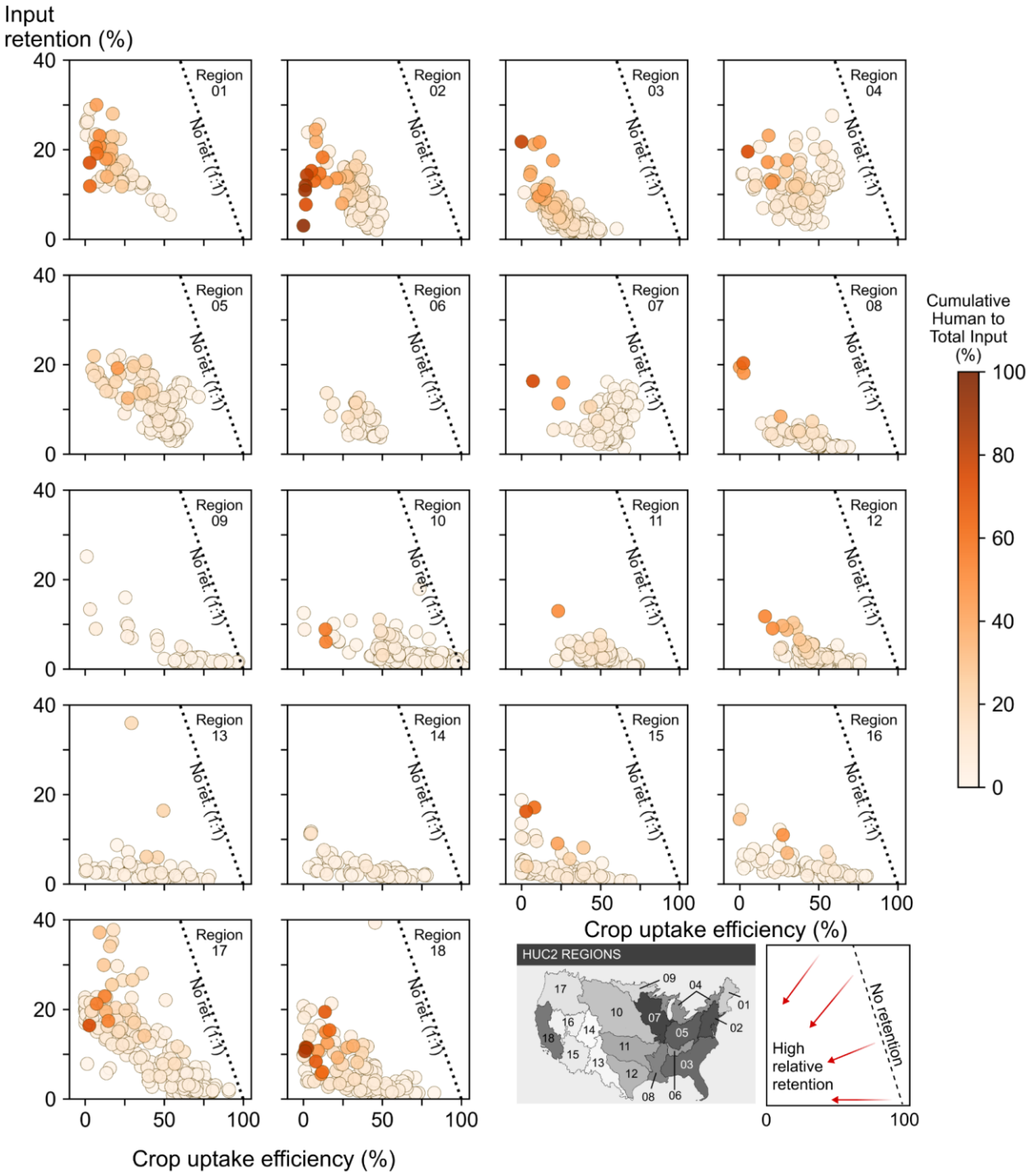
Supplemental Figure 5.5. Spatially-varying relationship between crop N uptake efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by percent agricultural land use.



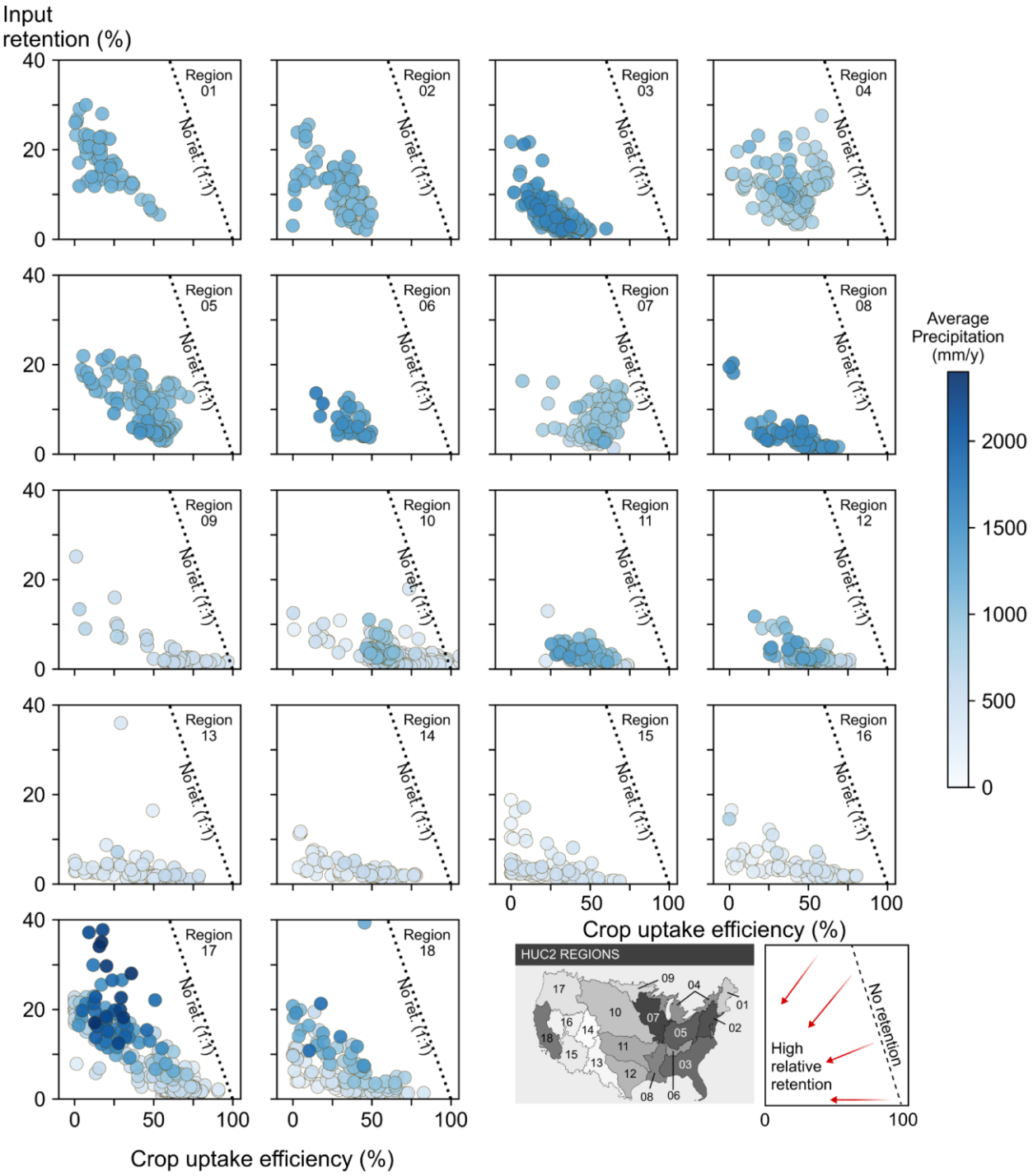
Supplemental Figure 5.6. Spatially-varying relationship between crop N uptake efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by percent tile drainage in agricultural land.


Input retention (%)
Region 01
Region 02
Region 03
Region 04
Region 05
Region 06
Region 07
Region 08
Region 09
Region 10
Region 11
Region 12
Region 13
Region 14
Region 15
Region 16
Region 17
Region 18
No ret. (1:1)
0
20
40
0
50
100
Crop uptake efficiency (%)
Cumulative Deposition to Total Input (%)
100
80
60
40
20
0
HUC2 REGIONS
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
No retention
High relative retention

Supplemental Figure 5.7. Spatially-varying relationship between crop N uptake efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by percent of total inputs from cumulative atmospheric deposition.



Supplemental Figure 5.8. Spatially-varying relationship between crop N uptake efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by percent of total inputs from human waste.



Supplemental Figure 5.9. Spatially-varying relationship between crop N uptake efficiency (%) and N runoff efficiency (%) for HUC8 watersheds within 18 HUC2 regions. Markers are colored by average annual precipitation (1958 to 2017; mm/y).

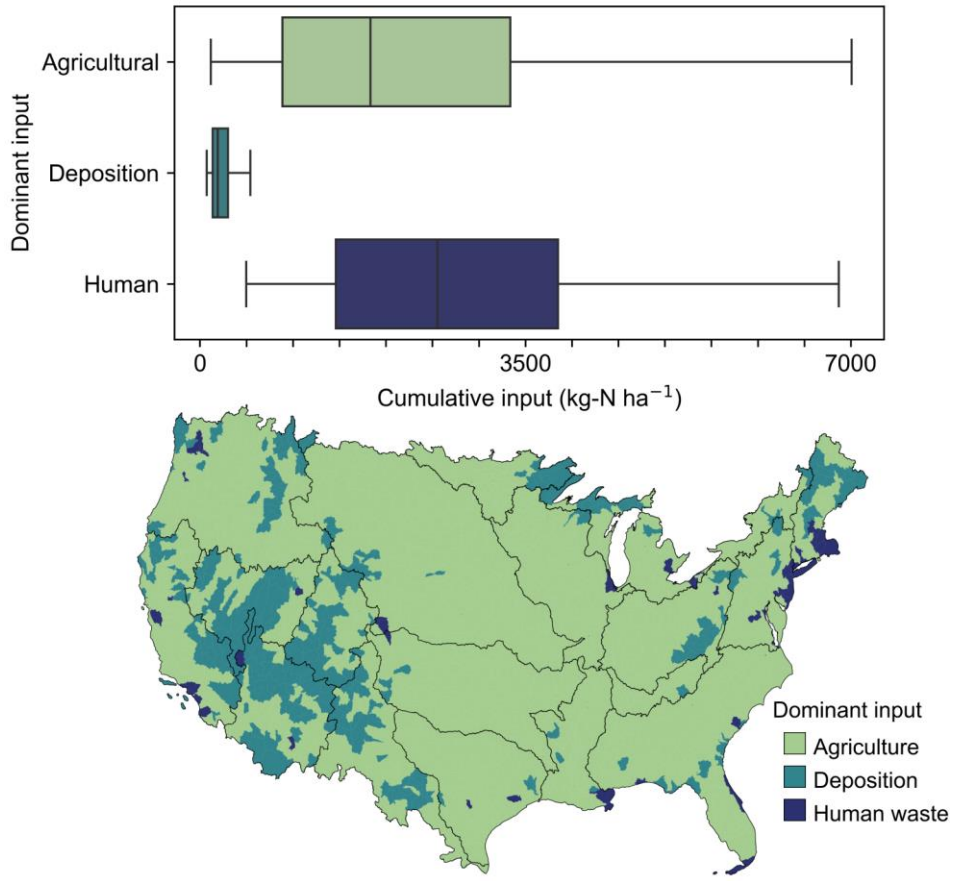
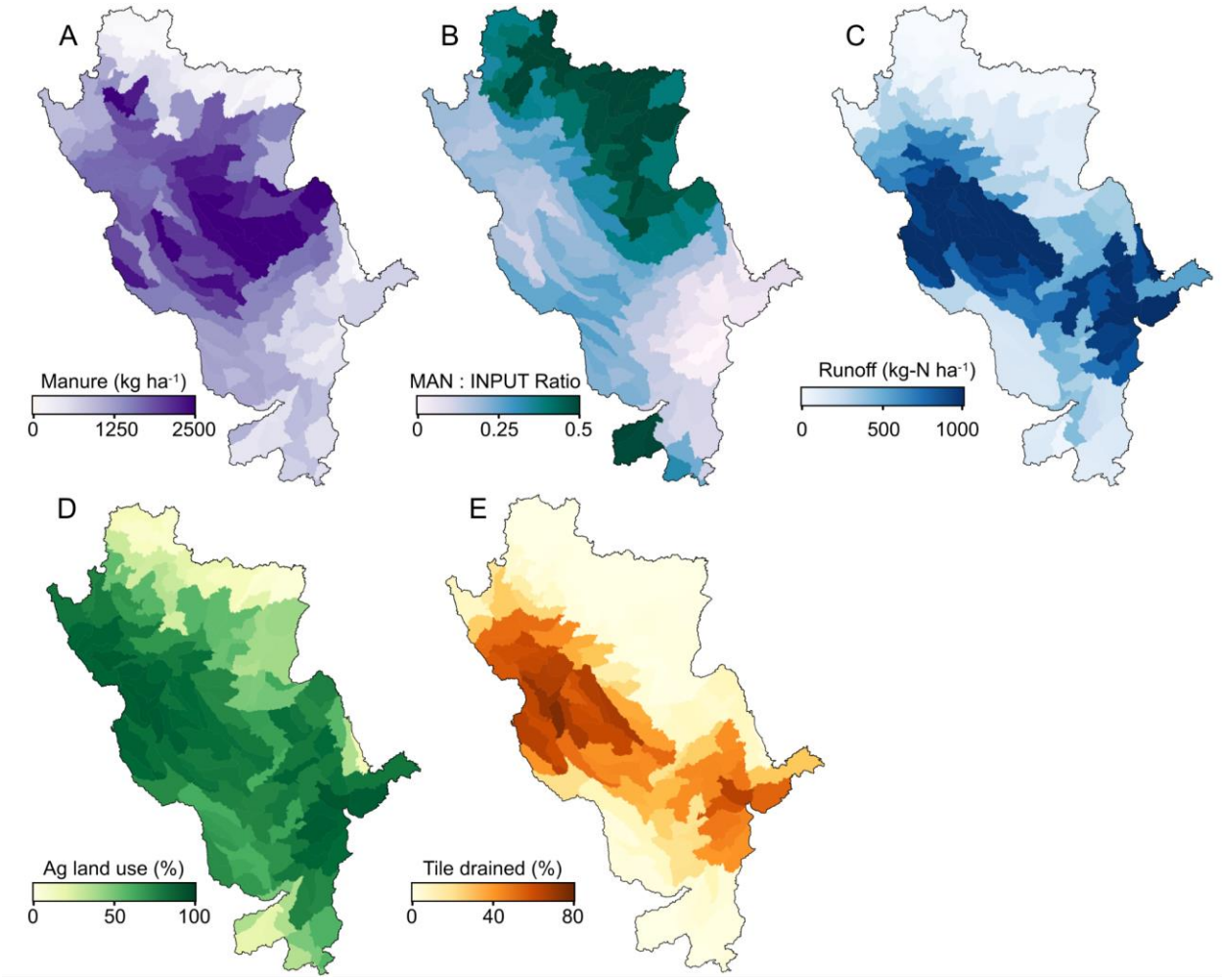


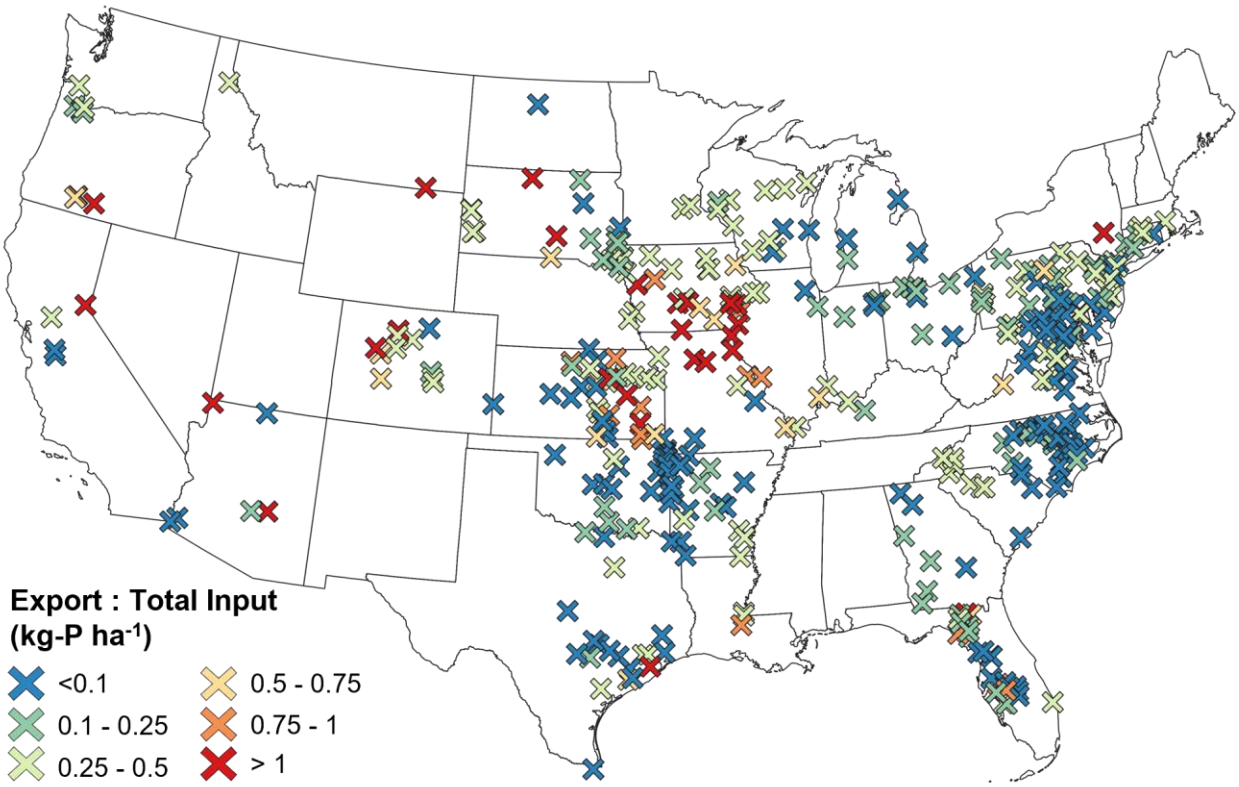
Figure 5.10. The dominant input for each HUC8 watershed and the magnitude of total cumulative inputs. Atmospheric deposition magnitudes are much lower than all other inputs, except in low impact areas.



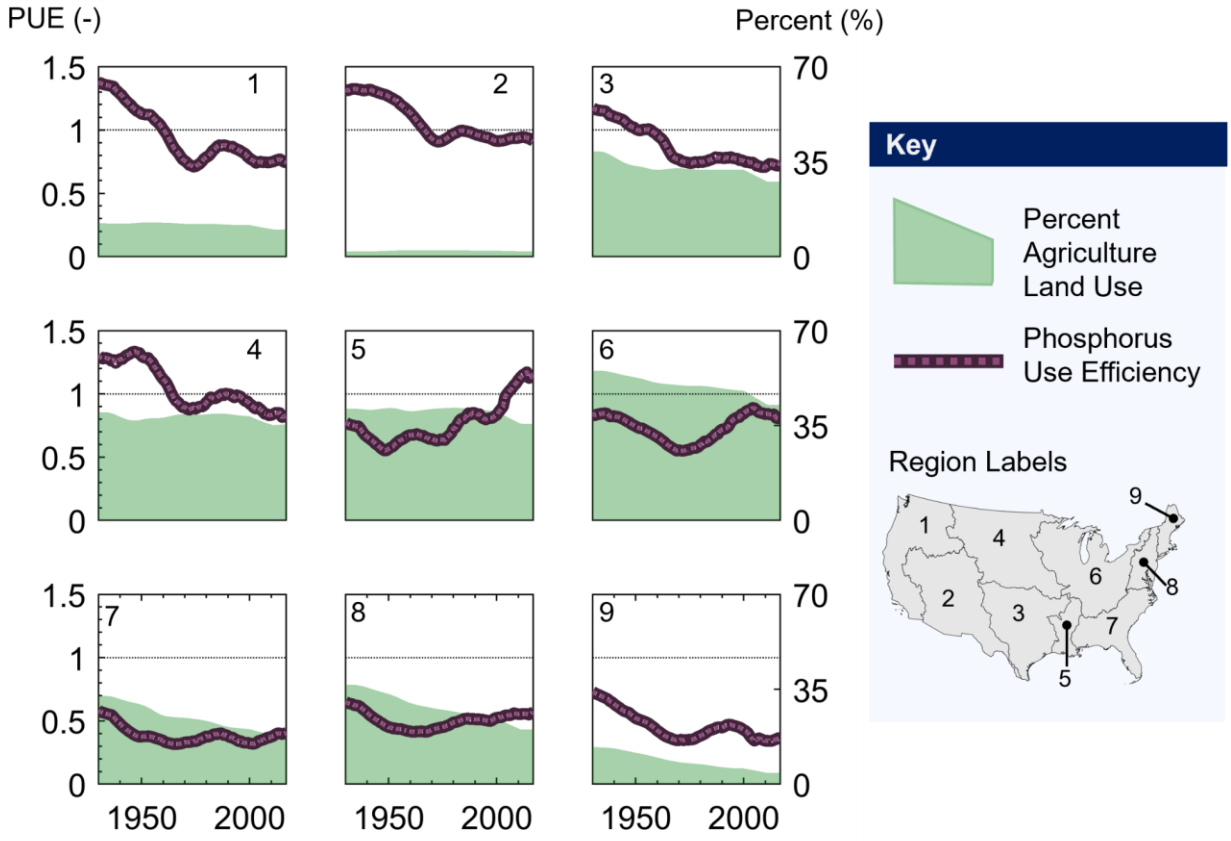
Supplemental Figure 5.11. Maps of the Upper Mississippi region (region 7) including (A) cumulative manure input, (B) ratio of manure to total nitrogen input, (C) cumulative nitrogen yield, (D) percent agricultural land use, and (E) the percent of agricultural land use that is tile-drained.

Supplemental

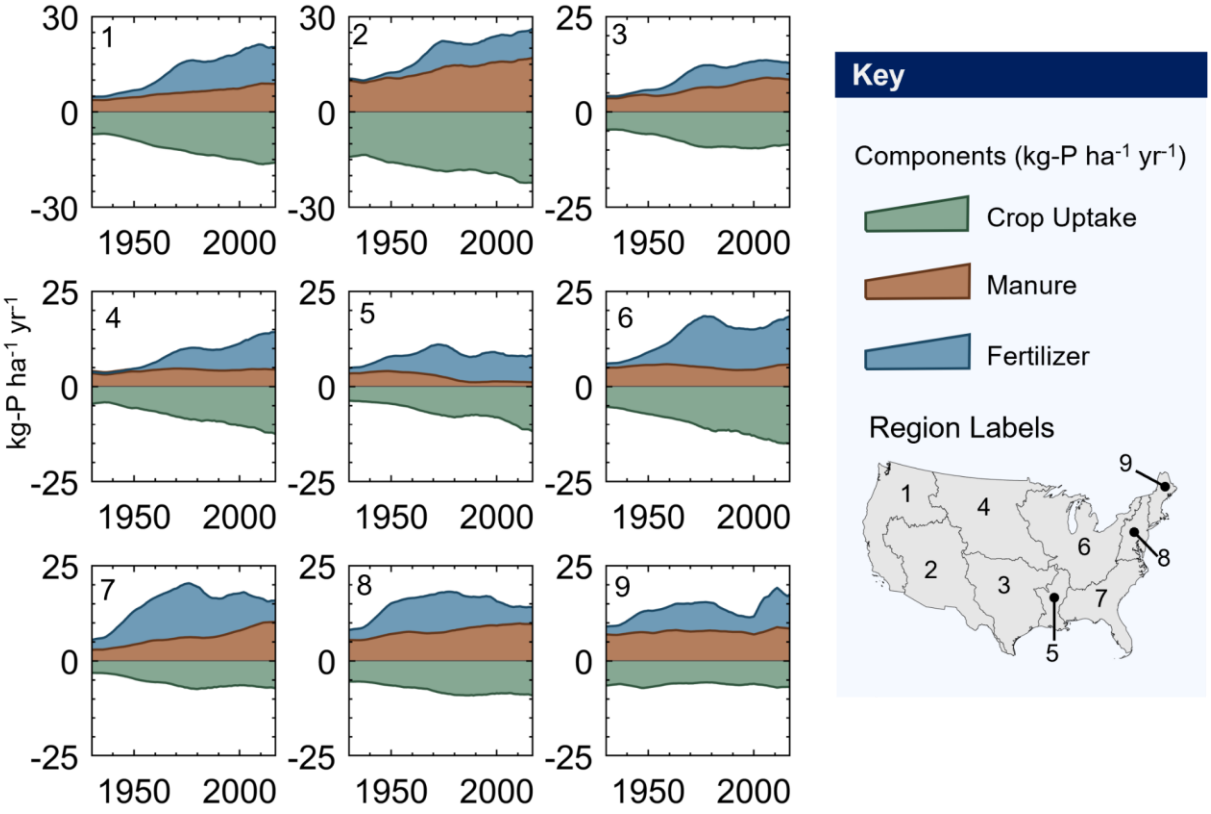
Chapter 6



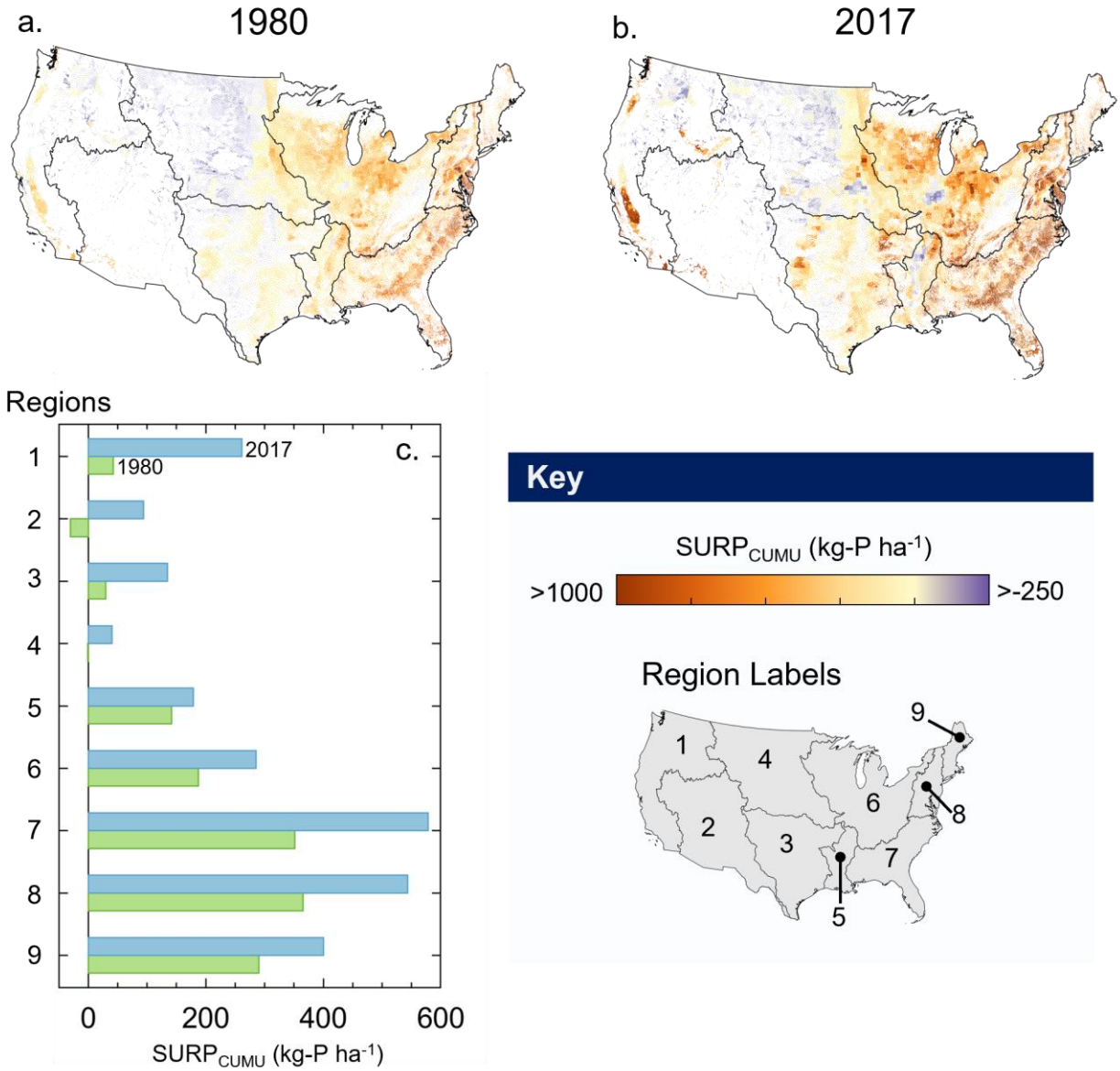
Supplemental Figure 6.1. The ratio of total phosphorus riverine export to the phosphorus surplus (P inputs from manure, fertilizer, atmospheric deposition, and domestic inputs minus crop and pasture uptake) in 2010 for 400 watersheds across the contiguous U.S. Data use is collected for the National Water-Quality Assessment (NAWQA) Project.[62]



Supplemental Figure 6.2. Phosphorus use efficiency and percent agricultural land use for hydrological regions across the U.S. from 1930-2017.



Supplemental Figure 6.3. Manure input, fertilizer inputs, and crop uptake for hydrological regions across the U.S. from 1930-2017.



Supplemental Figure 6.4. National and regional cumulative phosphorus surplus from 1930 - 1980 and 1930 - 2017. Spatial distribution of cumulative P surplus in 1980 (a) and 2017 (b). Panel c captures the cumulative surplus for each region during those same time periods.