EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Public Access

Author manuscript

Environ Sci Technol. Author manuscript; available in PMC 2022 June 15.

About author manuscripts | Submit a manuscript

Published in final edited form as:

Environ Sci Technol. 2021 June 15; 55(12): 7890–7899. doi:10.1021/acs.est.0c07102.

Context is everything: Interacting inputs and landscape characteristics control stream nitrogen

Jiajia Lin[1,2,*], **Jana E. Compton**[1], **Ryan A. Hill**[1], **Alan Herlihy**[1,3], **Robert D. Sabo**[4], **J. Renée Brooks**[1], **Marc Weber**[1], **Brian Pickard**[5], **Steve Paulsen**[1], **John L. Stoddard**[1]

[1]US EPA, Office of Research and Development, Center for Public Health and Environmental Assessment, Pacific Ecological Systems Division, Corvallis, OR 97333

[2]Oak Ridge Institute for Science and Education, Corvallis, OR 97333

[3]Oregon State University, Department of Fisheries and Wildlife, Corvallis, OR 97333

[4]US EPA, Office of Research and Development, Center for Public Health and Environmental Assessment, HEEAD, Washington, DC 20004

[5]Tetra Tech, Inc, Durham, NC 27709

Abstract

To understand the environmental and anthropogenic drivers of stream nitrogen (N) concentrations across the conterminous US, we combined summer low flow data from 4997 streams with watershed information across three survey periods (2000-2014) of the US EPA's National Rivers and Streams Assessment. Watershed N inputs explained 51% of the variation in log transformed stream total N (TN) concentrations. Both N source and input rates influenced stream NO_3/TN ratios and N concentrations. Streams dominated by oxidized N forms (NO_3/TN ratio > 0.50) were more strongly responsive to N input rate compared to streams dominated by other N forms. NO_3 proportional contribution increased with N inputs, supporting N saturation enhanced NO_3 export to aquatic ecosystems. By combining information about N inputs with climatic and landscape factors, random forest models of stream N concentrations explained 70%, 58%, and 60% of the spatial variation in stream concentrations of TN, dissolved inorganic N, and total organic N, respectively. The strength and direction of relationships between watershed drivers and stream N concentrations and forms varied by N input intensity. Model results for high N input watersheds indicated potential contributions from contaminated groundwater to high stream N concentrations, but also the mitigating role of wetlands.

Graphical abstract

*Corresponding author. Address: USEPA, 200 SW 35TH Street, Corvallis, OR 97333. Jlin42@Outlook.com.

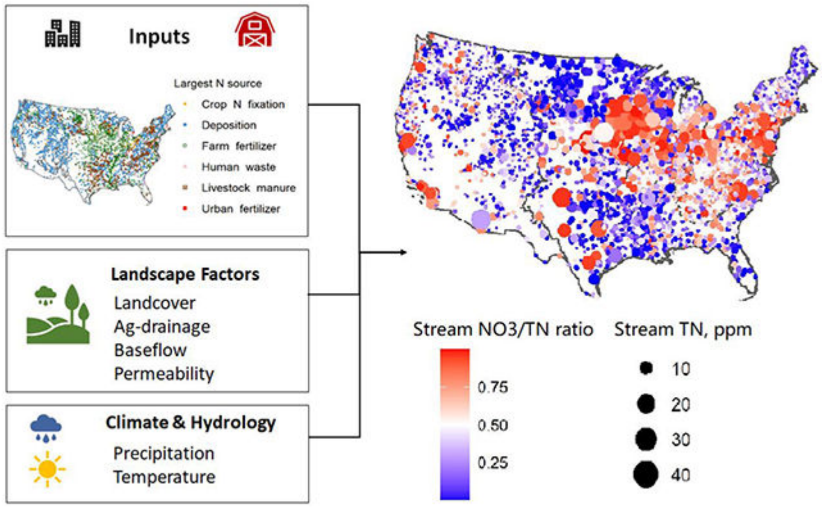
Supporting Information

R code for random forest modeling. Text describing study methods and materials. Tables of predictor variables for random forest models and N inventory. Maps of survey sites and ecoregions, and survey results of stream N. Figures demonstrating relationships of stream TN concentration with N input rate at different input level. Figures of results from random forest modeling.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript



Keywords

N inventory; N concentration; N species; water quality; nutrient; conterminous US; spatial variation; random forest; machine learning

Introduction

Human demand for food, fiber, and energy has reshaped the global nitrogen (N) cycle, leading to an approximate five-fold increase in anthropogenic production of reactive N during the last century, causing a cascade of consequences[1–3]. Excess N input to the landscape can contaminate drinking water supplies[4,5], stimulate the production of greenhouse gas and stratospheric ozone depleting nitrous oxide[6,7] and accelerate eutrophication[8–10].

Understanding the controls on nutrient concentrations is important for managing aquatic ecosystems[17–20]. States often develop criteria for nutrient concentrations to help mitigate and manage the negative outcomes caused by excess nutrients in streams[18,21]. Summer N concentrations are particularly important to track in light of recent evidence that N plays a role in cyanobacterial blooms and toxin formation[22–24] and the documented increase in harmful algal blooms[25]. Despite the importance of nutrient concentration for management, much of the focus of previous work has been on understanding the major driver of river loads for watersheds across the US[26]. To complement the historic efforts predicting N loads, our work focuses on the biologically important summer N concentrations of rivers and streams across the US.

Compilation of spatial datasets of nutrient inputs to watersheds has proven instrumental in modeling stream and river loads[26–29]. Several recent studies examined temporal trends in stream and river N concentrations[30–32], but they did not make direct connections between concentrations and N input rates. Other efforts have focused on a limited set of intensively monitored sites, mainly large streams and rivers, that may not adequately characterize the widespread variability of watershed responses to nutrient inputs[26,33]. More studies are

EPA Author Manuscript

needed to connect variation in concentrations to changes in N inventories at large scales. Newly assembled nutrient input inventories [34,35] allow for more direct connections between the variations in N inputs and aquatic N concentrations. Coupled analysis of nutrient inputs and stream N concentrations across many watersheds can aid in assessing water quality issues in rivers and streams at regional or national scales.

In this work, we studied nearly 5000 US watersheds capturing the wide variability in stream N concentrations and the associated watershed characteristics across the country. National nutrient input inventories[34] (NNI) and watershed landscape and climate variables were paired with stream N concentration data from EPA's National Rivers and Stream Assessment (NRSA) to 1) understand whether and to what extent changes in N inputs are reflected in stream concentrations of total N (TN) and N species in the conterminous US (CONUS); 2) understand the relationships between various landscape and climate drivers and stream nutrient levels; and 3) identify the context dependency of these relationships in response to intensifying amounts of N inputs.

Materials and Methods

National Rivers and Streams Assessment (NRSA)

Water chemistry data (concentrations of TN, NO_3, NH_4, and total organic N) originated from the EPA NRSA surveys that have been conducted across the CONUS since 2000 (Figure S1). The first survey cycle was conducted in 2000-2004[36] (Survey 1, n = 1170). Subsequent survey cycles in 2008-2009 (Survey 2, n = 1877) and 2013-2014 (Survey 3, n = 2046).

Across the three surveys, almost all samples were collected during the May-October index period with the goal of indexing summer baseflow conditions. Less than 1% of samples were sampled outside the index period. We present data on NO_3 (NO_3 plus NO_2), NH_4, and TN determined by persulfate digest. We also calculate total organic N (TON) as TN minus NO_3-N and NH_4-N. Further information on field and laboratory protocols can be found in Supporting Information (SI). Roughly 10% of the sites were sampled a second time within the summer index period within each survey year to assess temporal variability relative to spatial variability. Bellmore et al.[38] showed that 80-90% of the variation in NRSA N samples was associated with spatial variation rather than within year variation. In this study, we only used data from the first site visit in our analyses. We combined the data from all three survey cycles into one sample population and conducted analyses of the combined data, to allow focus on the drivers of spatial patterns in N concentrations. Sites were spatially grouped into three ecoregions[39] for our analyses (Figure S1): Western Mountains (also referred to as the West in the study), Plains and Lowlands (also referred to as the Midwest or plains), and Eastern Highlands (also referred to as the East).

We calculated ratios of different N species to TN across sites to explore the spatial distribution and factors influencing watershed N retention and stream chemistry. We categorized NRSA sites based on stream NO_3/TN ratio: the 'oxidized' stream N group with NO_3/TN ratio > 0.50, and the 'reduced' stream N group with NO_3/TN ratio < 0.50. We also separated watersheds into three groups based on N input levels: low input watersheds (n

= 2219) with N input rate <15 kg N ha^{-1}; medium input watersheds (n = 1565) with N input rate ranges between 15 and 50 kg N ha^{-1}; and high input watersheds (n = 1213) with N input rate > 50 kg N ha^{-1}.

Watershed and Climate variables

To model NRSA stream TN concentrations, we used a suite of landscape, climate, and nutrient input data as independent covariates. The landscape data for each NRSA site came from the EPA StreamCat dataset, which contains a wide variety of watershed metrics that are derived from coverages that characterize both natural and anthropogenic landscape features (see SI and Table S1)[40]. From StreamCat, we selected a set of watershed features that we hypothesized would influence stream TN concentrations based on our understanding of such factors and previous research[38], such as land use present in the watershed of each NRSA site (see SI for more details and Table S1 for other variables).

In addition to StreamCat data, we also summarized annual and monthly mean climate variables from PRISM and NASA Earth Observatory data for each watershed from the corresponding year and month when each site was sampled (Table S1). These data included annual and monthly means of precipitation, snow cover, numbers of fires, land surface temperature, and vegetation cover within the watershed. More details on data sources and methods can be found in the SI.

National Nutrient Input Inventory (NNI)

The NNI (downscaled from Sabo et al.[34]) compiles data for many nitrogen sources for the entire CONUS including agricultural and residential fertilizer, crop biological N fixation, land applied manure, human waste, and deposition. We connected each NRSA survey with its closest year of record; Survey 1—NNI 2002; Survey 2—NNI 2007; and Survey 3—NNI 2012. Input NNI rasters were summarized to the NRSA watershed scale using zonal statistical procedures in GIS software (see details in the SI). Total watershed N inputs were calculated by summing agricultural N fertilizer, N-fixing crop cultivation, livestock N waste, urban N fertilizer, human N waste, and total N deposition inputs. Watersheds were categorized based on the proportionally largest N sources described above. These categories were then used in subsequent analyses as shown in the next section.

Spatial and statistical comparisons

We mapped NRSA sample sites by their largest N source to explore spatial patterns of N inputs and sources. In addition, we used linear regression to initially test the relationships between N inputs and stream N concentration. To further investigate regional variations in both N inputs and stream N concentrations, we calculated means and standard errors of response variables (N inputs and stream N concentrations and ratios) at both regional and national levels.

ANOVA and t-test were used to examine the response of stream TN concentration to N input rate and its dependence on stream N type (‘oxidized’ where NO_3/TN > 0.5 vs ‘reduced’) and input level, respectively. We also applied ANOVA to compare stream NO_3/TN ratios for different N input sources and levels to further understand how stream dominant N speciation

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

changed with land use and input intensity. These tests were also carried out to examine regional differences in N species.

Random forest models

We used random forest[41] modeling to explore relationships between stream N and watershed features, climate, and N inventories by combining data from the three surveys. Random forest models are robust to violations of the assumptions made by parametric statistical approaches including multicollinearity among predictors[41]. However, to improve model interpretation, we minimized collinearity of predictors by pre-selecting a subset of variables before modeling (see details and code in SI). The final set of 21 predictor variables included 15 variables from StreamCat, PRISM and the Earth Observatory Network, and 6 N input rate variables of six sources from the N inventories (Table S1). With these data, we developed a series of random forest models to explain stream TN and different N species. We combined NO_3 and NH_4 as dissolved inorganic nitrogen (DIN) in our models to avoid potential data and modeling issues caused by too many low values (< detection limit) of NH_4 concentration. There were 4997 data points used in our final analyses after removing erroneous outliers that had TN input rate > 1000 kg N ha^{-1} and/or negative TON values.

The first set of models contained data from the full set of sample sites to explain concentrations of stream TN, DIN, and TON. To examine how the importance and behavior of explanatory variables varied among N input levels, sites were separated into three groups in subsequent N models based on the three N input levels we defined above[42]. Although not required by random forest, stream concentration and N inventories data were natural log transformed to improve interpretations when plotting modeled relationships[43]. We added a ‘1’ to the inventory variables so that zero values could be transformed. The dataset for each model was randomly divided into training (75% of total available data) and testing (25%) datasets to evaluate model performance.

All data analyses were conducted using R programming language[44] (version 3.6.2) and the associated packages (dplyr_0.8.4; randomForest_4.6-14; caret_6.0-86).

Results

Connecting the N inventory to stream N

Sources and rates of N inputs were strongly related to stream N concentrations across the CONUS (Figure 1 & S2). Watersheds where agricultural inputs (farm fertilizer, manure, and crop N fixation) were the major N source, had higher total watershed N input rates and stream TN concentrations. Sites with atmospheric N deposition as the largest watershed N source were associated with lower N input rates and stream TN concentrations. At the national scale, stream TN concentrations were strongly correlated with N input (Figure 1a; r^2 = 0.51, p<=0.001).

Nitrogen source also affected stream N forms. We separated the watersheds into three types based on largest N source: agricultural, deposition and human. On average, sites dominated by agricultural N inputs had significantly higher N concentrations and NO_3/TN ratios than sites where deposition was the largest N source (ANOVA, $p < 0.0001$), but the latter had

EPA Author Manuscript

higher proportional TON contributions than agricultural and human types of watersheds (ANOVA, $p < 0.0001$). Watersheds with human waste and urban fertilizer as the largest N source, categorized as the 'human' group, only accounted for about 3.6% of total sampled sites. These sites on average had the highest NO_3/TN ratio and the lowest TON/TN among three land use types (ANOVA, $p < 0.0001$). Within each watershed group where either agricultural inputs or atmospheric deposition was the largest N source, stream NO_3/TN ratio increased significantly with increasing input level (Figure 2a & Table 1a). Within the 'human' group, NO_3/TN ratio changes were either not significant or not computable due to small number of sites (Table 1a). Combining all input groups, the mean NO_3/TN ratios in the low N input, medium N input, and high N input groups were respectively 0.24, 0.37, and 0.51, and mean TON/TN ratios in these watershed groups were 0.71, 0.59, and 0.45, respectively.

The dominant N forms in streams interacted with input rates. Streams where NO_3 was the largest N form, (the 'oxidized' type), exhibited a stronger and steeper relationship with N inputs as compared to the 'reduced' type of streams (Figure 2b). The difference in regression slopes of the two types of watersheds (equal to 0.27) was statistically significant ($p <$ 0.0001). We further examined such difference in slopes at three input levels, and determined that the contrast between 'oxidized' and 'reduced' slopes were minimum and insignificant among low N input watersheds, but grew greater and significant as input level increased (Figure S3).

The three ecoregions (see regions map in Figure S1) varied in N input and stream N chemistry (Table 1b). Western Mountains region, where a large portion of watersheds had deposition as the largest N source, also had the lowest N input rates on average, stream N concentrations, and NO_3/TN ratios compared to the other two regions ($p <$ 0.0001). In contrast, watersheds in the Plains and Lowlands region had the highest mean inputs and stream N concentrations, the latter was enhanced by both high stream NO_3 and TON concentrations. Streams in Eastern Highlands had the highest regional NO_3/TN ratios compared to the other two regions. The 'reduced' type of watersheds had about equal numbers with the 'oxidized' type in the sampled streams in the East, but were more prevalent in the western and plains regions (Table 1b), with streams in the plains showing the highest mean TON and NH_4 concentrations especially in its 'reduced' type. At the regional level, most 'oxidized' type of sites had significantly greater mean N input rate (ANOVA, $p < 0.0001$ except in the Western Mountains) and stream TN concentration (ANOVA, all $p < 0.01$) than those of the 'reduced' type of sites within the same region. (Table 1b).

Random forest models of stream N concentrations

The models that included all sample sites explained 70%, 58%, and 60% of the variation in stream TN, DIN, and TON concentrations, respectively (Figure 3a–3c). Cross validations using the withheld testing dataset yielded r^2 values of 0.90, 0.86, and 0.88, respectively, for TN, DIN, and TON concentrations. The most important predictors varied by N form, indicating that drivers affected cycling of N species in different ways. Agricultural land use and input rates were among the most important variables for all three models. Proportional

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

baseflow contribution and the annual mean Normalized Vegetation Difference Index (NDVI) – a standard remotely sensed index of plant greenness and health and generally positive correlated with net primary productivity — were also important across models. Human waste and atmospheric deposition input rates were key to predicting DIN concentrations, while wetland coverage and annual mean precipitation were more critical for the TON model. Annual mean temperature was only important in predicting TON. Agricultural input rates ranked higher than land use percentages in their importance in predicting TN and DIN concentrations.

Subsequent random forest models and partial dependance plots conducted on data subsetted by input level further revealed how different factors influenced stream N species under each N input level (Figure S4–S6). The lowest input category primarily represented watersheds where deposition was the largest source. The highest input category represented a large proportion of sites where fertilizer was the largest source, and the intermediate category had a mixture of largest source types.

For low N input watersheds, vegetation index and baseflow contribution were important predictors and exhibited negative relationships with TON and DIN concentrations (Figure S4). Other than agriculture related variables, wetland coverage also had a positive relationship with stream TON concentration, which was its second most important predictor. N deposition rate and monthly mean precipitation exhibited positive relationships with DIN concentration. Monthly mean areal fire counts (per 1000 km^2 per day) were positively associated with TON, DIN, and TN concentrations when count was > 1. Variable importance and correlations with TN concentration mostly followed those in the TON sub-model (Figure 3d & S4), as a result of TON being a more important contributor to stream N in low N input watersheds.

For watersheds with intermediate N input level, predictors associated with direct human activities became important in DIN and TN sub-models, such as urban land use (%), NPDES site densities, and human waste N input rate (Figure S5). However, human waste input rate showed a negative relationship with TON concentration but a positive one with DIN concentration. Baseflow also played an opposing role in affecting TON and DIN concentrations. It appeared to dilute stream TON but increased DIN, leading to a non-monotonic relationship with TN concentration. Annual mean of vegetation greenness and precipitation were negatively associated with TON and TN concentrations, but their impact on DIN was varying. Monthly mean precipitation was positively associated with TON concentration but negatively with TN concentration.

For high input watersheds, one of the most important predictors of DIN and TN concentrations was baseflow contribution, which was positively associated with DIN and TN concentrations (Figure 3e & S6). In contrast, baseflow contribution had a negative relationship with TON concentration. Similarly, wetlands also affected N forms in different ways: increasing wetland coverage was associated with decreasing stream DIN and TN concentrations but increasing TON concentration (Figure 3e & S6). Predictors related to human activities were also crucial in the high N input sub-models. Crop N fixation rate was the most important predictor for TN concentration probably due to its positive association

EPA Author Manuscript

with agricultural practices, rather than because it was the ubiquitously most important source of high N inputs in these watersheds. Monthly mean fire numbers had a positive relationship with TON but a negative relationship with DIN concentration.

Environmental Implications

N inputs establish spatial patterns of stream N concentrations

Watershed N inputs alone explained approximately half of the spatial variation in stream TN concentrations, like a previous finding of 47%[38]. Spatially explicit models[45] of seasonal nutrient concentrations across the US also uncovered spatial patterns similar to our findings, but these models did not include nutrient inputs, and yielded correlation coefficients ranging between 0.61 and 0.78, thus explained 37-61% of the variation in TN concentrations. Our model results based on all sample sites explained 58-70% of the variation in stream N concentrations, affirming that model predictive power may be improved by including watershed-specific and time-specific nutrient inventory variables.

Recent studies[26,46] demonstrated that temporal variations in hydrological nutrient export were largely controlled by hydrological factors, but the spatial patterns were better predicted by nutrient inputs. Our results support the finding that inputs drive the spatial patterns. Spatial variations and relationships between N inputs and stream concentrations are thought to be somewhat stable over time[47], which in part could be related to time lags and nutrient legacies[48,49], and relatively static land cover and inputs over the recent decade[31]. The long-term imprint of high N inputs, largely from agricultural sources, has proven challenging to erase over time despite many years of efforts to reduce the release of N.

Impact of N inputs on stream N speciation and watershed retention

The speciation or forms of stream N varied with N input level. Stream TN was dominated by reduced N, especially TON in areas with low N inputs. This was probably due to wetland contributions of organic N in these watersheds. But NO_3/TN ratio increased significantly with enhanced N inputs (Table 1a & Figure 2a). The elevated level and proportional contribution of in-stream NO_3 as N inputs increase supports N saturation theory, where organic N dominates in less disturbed watersheds, but stream NO_3 export increases with added inputs as terrestrial and aquatic ecosystems lose their ability to retain added N[50–54].

Our work provides new insights about the factors leading to watershed N saturation and changing stream N forms. The stronger and steeper relationship between stream TN concentration and N input rate in NO_3-dominated streams (Figure 2b & S3) suggests that NO_3-dominated streams respond more strongly to further N inputs and become more saturated at a faster rate than the ‘reduced’ group. Therefore, increasing nutrient inputs not only affects immediate aquatic nutrient level and speciation, but can also alter long-term nutrient cycling by limiting N retention capacity in some watersheds and exacerbating nutrient loss and aquatic eutrophication.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Impact of climate and landscape drivers

While spatial variation of stream N concentrations corresponds closely to patterns of N inputs and sources[26], there remains substantial scatter around this relationship[38,55,56]. Many factors can affect watershed nutrient dynamics and stream nitrogen concentrations, including hydrology, soils, and geology[46,57–59]. Random forest modeling validated that a complex variety of climate and physical watershed features can interact to drive instream N concentrations across the CONUS.

The relationship between hydrology and nutrient concentration can help interpret the interaction between the role of transport (precipitation) and supply (nutrient source)[60]. In low N input watersheds, stream DIN concentration increased with both monthly mean precipitation and deposition input rate, demonstrating that wet deposition was an important source of DIN in these systems. On the other hand, in watersheds where monthly mean precipitation was > 100 mm, stream TON decreased with increasing monthly mean precipitation, indicating dilution of the organic N source by precipitation and/or the importance of baseflow input to the summer stream flow (Figure S4).

Separating the influence of summer mean precipitation from deposition and total N inputs is confounding since all three increase from west to east. This spatial gradient could be driving the apparent negative association between stream TN and monthly mean precipitation in the medium N input group: watersheds in the Plains and Lowlands region on average had lower monthly mean precipitation but higher N input rate compared to the eastern watersheds. For the high N input group, the majority (> 65%) of sites were located in the Plains and Lowlands region, and concentrations of TON, DIN and TN all increased with monthly mean precipitation, suggesting these systems were transport limited rather than (N) supply limited[60]. Positive relations between hydrology and stream N concentration are typical for agricultural watersheds, which may result from increased soil leaching and surface runoff during periods of high rainfall and associated elevated flows from agricultural tile drainage[59].

Wildfires can have large impacts on water quality and the aquatic ecosystem by altering stream nutrient species and concentrations[61–63]. The positive relationship between monthly mean fire counts and N concentrations (when > 1 fire within 1000km^2 per day) suggested elevated post-fire N transport to streams, either via transport of N from the soil and forest floor or via increased ash and atmospheric deposition[64]. The impact of wildfires on reorganizing watershed N cycling is a complex process. For example, severe wildfires could change forms of organic N by converting it to the more soluble form or volatilize N[61,65]. How fast forest systems can recover from fires also affects stream N concentrations[62,63,66]. More careful separation of the drivers is required to conduct a more definitive analysis on wildfire impact.

Two critical landscape factors also interacted with N inputs to shape stream TN: wetland cover and the ratio of baseflow to total flow (i.e., the baseflow index)[67]. By dividing and modeling N concentrations at three levels of N inputs, we revealed that the context of N input can influence these relationships.

EPA Author Manuscript

Wetlands play important but varying roles—Wetlands can behave as a potential source of dissolved organic N[68,69]. In the previous examination of the NRSA 2008-2009 survey period, TON was positively correlated with the proportion of wetland in the watershed, and our findings reinforce this finding[38]. The net effect on TN concentrations depends on wetland conditions[70,71].

Covarying spatial gradients across the CONUS in wetland distribution and N input contributed to the positive relationship between wetland coverage and stream TN concentration in low input watersheds. Watersheds in the plains, especially the coastal plains and upper Midwest regions, had the highest wetland coverages and the greatest input rates. In general, watersheds in the Western Mountains had the lowest wetland coverage and N input level. Smith et al.[72] also identified this spatial gradient in natural background stream TN concentration across the CONUS. More analyses are needed to separate the impact of nutrient gradients confounding with gradients in land uses. However, in the Eastern Highlands region, especially the Appalachians, stream TN concentration in low input watersheds had a positive relationship with wetland coverage, whereas N input rates decreased with increasing wetland coverage, indicating the importance of wetland to stream nutrient level (Figure S7).

The negative association of wetland cover with stream DIN and TN concentration in high N input watersheds (Figure 3e) supports previous work on the strong N removal potential of wetlands if they are present and hydrologically connected[73–75]. Wetlands were also an important source of TON in high N input sites (Figure S6). Increasing wetland coverage was accompanied by decreasing agricultural land coverage in many regions, hence often coincided with lower watershed N input rate. But wetlands can serve as a sink of stream TN through denitrification even though wetland N retention rate varies by region and site[76]. We found that some NRSA watersheds with the highest N input rate (>100 kg N ha^{-1} yr^{-1}) and high wetland percentages (> 10%) had TN concentration < 2 mg N L^{-1} or even < 1 mg N L^{-1}, indicating the effectiveness of wetland N retention and potential conservation efforts in these areas. For those watersheds, the location of wetlands in flow paths and in relation to the N input source regulate their ability to serve as removal barriers to excess nitrogen[74].

Proportion of baseflow: An indicator of legacy groundwater contamination in high N input areas?—The impact of baseflow contribution[77] on stream N concentrations also diverged at different input levels (Figure 3d–e). The negative relationship between baseflow and stream N concentrations in low N input watersheds may indicate dilution by subsurface contributions, especially for stream TON. Denitrification along subsurface flow paths could further reduce DIN concentration in streams with high baseflow contribution[78]. In contrast, we found a positive relationship between baseflow contribution and stream DIN in high N input watersheds, suggesting that potential baseflow inputs of NO_3-contaminated groundwater contributed to stream N in these watersheds. We also found a negative relationship between baseflow contribution and TON concentration in these high N input watersheds, indicating that baseflow might also serve as a dilution of TON, likely as a result of mineralization of TON along the flow path and/or low efficiency of subsurface transport of dissolved organic matter controlled by sorption processes[79–82]. NO_3/TN ratios were > 0.95 at some high TN concentration sites (>10 mg N L^{-1}), confirming the subsurface source

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

of N. Watersheds with long agricultural histories and high nutrient input can experience nutrient legacies, which potentially accumulate in soil and groundwater over decades of excess fertilizer application, acting as a long-term source that slowly releases N into surface waters and can hinder nutrient reduction efforts[48,83]. Our findings support the hypothesis that groundwater sources critical to summer low flows can be contaminated with excess N in high N input areas, where streams are more vulnerable to legacy N contamination[84].

Model implications

By careful quantification of N inputs combined with the random forest approach to tease out relationships by N input level, we were able to identify context-dependent connections. Critical variables may have diverging impacts on stream chemistry depending on climate, hydrology, and long-term land use history, and can lead to either improving or degrading water quality.

The long-term trajectory of N inputs is important to consider – decades of high N inputs in some areas may lead to a very complex N input-output relationship. Where groundwater is contaminated with NO_3, and baseflow is an important component supplying streamflow, thus N export may be driven by NO_3 inputs from baseflow that reflect legacy N inputs and preferential leaching of mobile NO_3. This interaction between legacy N inputs and groundwater contamination provides continued challenges to reducing stream N loads and concentrations[83,85–87]. By coupling information about baseflow proportion, current and legacy N inputs, and stream N concentrations and fluxes, it may be possible to identify these streams with shorter and longer timeframes for recovery and include this information in environmental decision making.

Limited resources are available for watersheds restoration and assessment. Spatially explicit datasets on watershed characteristics and water quality models can provide critical support to help evaluate the success of N reduction efforts. A recent study, by applying spatial records on nutrient inventory, identified key areas to boost nutrient use efficiency and to achieve biggest gains to water quality restoration[88]. Our high-performing random forest models can be applied to efficiently make predictions of stream N concentrations in all NHD streams in the CONUS, and interpolate water quality information for areas lacking in data. This next-step application of our models will provide managers with valuable insights to prioritize monitoring efforts (such as groundwater/baseflow monitoring), identify key spots for effective actions in the absence of additional in-situ measurements, and design N reduction strategies within or across different regions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors thank Karen Blocksom, Tom Kincaid, Tony Olsen, and David Peck for valuable input regarding NRSA datasets and statistical analyses. We thank Heather Golden for comments on the manuscript. The USEPA and the Oak Ridge Institute for Science and Education Postdoctoral Research Fellowship Program and the National Academies of Science Research Associateship Program funded through USEPA supported this project. This manuscript has been subjected to Agency review and has been approved for publication. The views expressed

EPA Author Manuscript

in this paper are those of the authors and do not necessarily reflect the views or policies of the USEPA. Dataset for this study has been submitted to the USEPA data repository and will be publicly available upon approval at: doi:10.23719/1519428. Code for random forest modeling has been submitted as part of the Supporting Information.

References

(1). Fowler D; Coyle M; Skiba U; Sutton MA; Cape JN; Reis S; Sheppard LJ; Jenkins A; Grizzetti B; Galloway JN; Vitousek P; Leach A; Bouwman AF; Butterbach-Bahl K; Dentener F; Stevenson D; Amann M; Voss M The Global Nitrogen Cycle in the Twenty-First Century. Philosophical Transactions of the Royal Society B: Biological Sciences 2013, 368 (1621), 20130164. 10.1098/rstb.2013.0164.

(2). Battye W; Aneja VP; Schlesinger WH Is Nitrogen the next Carbon? Earth's future 2017, 5 (9), 894–904.

(3). Galloway JN; Aber JD; Erisman JW; Seitzinger SP; Howarth RW; Cowling EB; Cosby BJ The Nitrogen Cascade. BioScience 2003, 53 (4), 341. https://doi.org/10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.

(4). Pennino MJ; Kaushal SS; Murthy SN; Blomquist JD; Cornwell JC; Harris LA Sources and Transformations of Anthropogenic Nitrogen along an Urban River–Estuarine Continuum. Biogeosciences 2016, 13 (22), 6211–6228. 10.5194/bg-13-6211-2016.

(5). Nolan BT; Hitt KJ Vulnerability of Shallow Groundwater and Drinking-Water Wells to Nitrate in the United States. Environmental science & technology 2006, 40 (24), 7834–7840. [PubMed: 17256535]

(6). Beaulieu JJ; Tank JL; Hamilton SK; Wollheim WM; Hall RO; Mulholland PJ; Peterson BJ; Ashkenas LR; Cooper LW; Dahm CN; Dodds WK; Grimm NB; Johnson SL; McDowell WH; Poole GC; Valett HM; Arango CP; Bernot MJ; Burgin AJ; Crenshaw CL; Helton AM; Johnson LT; O'Brien JM; Potter JD; Sheibley RW; Sobota DJ; Thomas SM Nitrous Oxide Emission from Denitrification in Stream and River Networks. Proc Natl Acad Sci USA 2011, 108 (1), 214–219. 10.1073/pnas.1011464108. [PubMed: 21173258]

(7). Seitzinger SP; Kroeze C; Styles RV Global Distribution of N2O Emissions from Aquatic Systems: Natural Emissions and Anthropogenic Effects. Chemosphere-Global Change Science 2000, 2 (3–4), 267–279.

(8). Gilbert PM; Mayorga E; Seitzinger S Prorocentrum Minimum Tracks Anthropogenic Nitrogen and Phosphorus Inputs on a Global Basis: Application of Spatially Explicit Nutrient Export Models. Harmful Algae 2008, 8 (1), 33–38. 10.1016/j.hal.2008.08.023.

(9). Diaz RJ; Rosenberg R Spreading Dead Zones and Consequences for Marine Ecosystems. science 2008, 321 (5891), 926–929. [PubMed: 18703733]

(10). Anderson DM; Burkholder JM; Cochlan WP; Glibert PM; Gobler CJ; Heil CA; Kudela RM; Parsons ML; Rensel JEJ; Townsend DW; Trainer VL; Vargo GA Harmful Algal Blooms and Eutrophication: Examining Linkages from Selected Coastal Regions of the United States. Harmful Algae 2008, 8 (1), 39–53. 10.1016/j.hal.2008.08.017. [PubMed: 19956363]

(11). Lloret J; Valiela I Unprecedented Decrease in Deposition of Nitrogen Oxides over North America: The Relative Effects of Emission Controls and Prevailing Air-Mass Trajectories. Biogeochemistry 2016, 129 (1–2), 165–180.

(12). Zhang Y; Mathur R; Bash JO; Hogrefe C; Xing J; Roselle SJ Long-Term Trends in Total Inorganic Nitrogen and Sulfur Deposition in the US from 1990 to 2010. Atmospheric chemistry and physics 2018, 18 (12), 9091. [PubMed: 30079084]

(13). Howarth RW Coastal Nitrogen Pollution: A Review of Sources and Trends Globally and Regionally. Harmful Algae 2008, 8 (1), 14–20. 10.1016/j.hal.2008.08.015.

(14). Dubrovsky NM; Burow KR; Clark GM; Gronberg JM; Hamilton PA; Hitt KJ; Mueller DK; Munn MD; Nolan BT; Puckett LJ The Quality of Our Nation's Waters—Nutrients in the Nation's Streams and Groundwater, 1992–2004. US geological survey Circular 2010, 1350 (2), 174.

(15). Tomer MD; Locke MA The Challenge of Documenting Water Quality Benefits of Conservation Practices: A Review of USDA-ARS's Conservation Effects Assessment Project Watershed Studies. Water Science and Technology 2011, 64 (1), 300–310. [PubMed: 22053489]

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

(16). Lintern A; McPhillips L; Winfrey B; Duncan J; Grady C Best Management Practices for Diffuse Nutrient Pollution: Wicked Problems Across Urban and Agricultural Watersheds. Environ. Sci. Technol 2020, acs.est.9b07511. 10.1021/acs.est.9b07511.

(17). Herlihy AT; Sifneos JC Developing Nutrient Criteria and Classification Schemes for Wadeable Streams in the Conterminous US. Journal of the North American Benthological Society 2008, 27 (4), 932–948. 10.1899/08-041.1.

(18). Dodds WK; Welch EB Establishing Nutrient Criteria in Streams. Journal of the North American Benthological Society 2000, 19 (1), 186–196.

(19). Evans-White MA; Haggard BE; Scott JT A Review of Stream Nutrient Criteria Development in the United States. Journal of environmental quality 2013, 42 (4), 1002–1014. [PubMed: 24216352]

(20). Bennett MG; Lee SS Measuring Lotic Ecosystem Responses to Nutrients: A Mismatch That Limits the Synthesis and Application of Experimental Studies to Management. Limnology and oceanography bulletin 2019, 28 (1), 26. [PubMed: 31024221]

(21). Smith AJ; Thomas RL; Nolan JK; Velinsky DJ; Klein S; Duffy BT Regional Nutrient Thresholds in Wadeable Streams of New York State Protective of Aquatic Life. Ecological Indicators 2013, 29, 455–467.

(22). Yuan LL; Pollard AI; Pather S; Oliver JL; D'Anglada L Managing Microcystin: Identifying National-Scale Thresholds for Total Nitrogen and Chlorophyll a. Freshwater biology 2014, 59 (9), 1970–1981.

(23). Kramer BJ; Davis TW; Meyer KA; Rosen BH; Goleski JA; Dick GJ; Oh G; Gobler CJ Nitrogen Limitation, Toxin Synthesis Potential, and Toxicity of Cyanobacterial Populations in Lake Okeechobee and the St. Lucie River Estuary, Florida, during the 2016 State of Emergency Event. PLoS One 2018, 13 (5), e0196278. [PubMed: 29791446]

(24). Newell SE; Davis TW; Johengen TH; Gossiaux D; Burtner A; Palladino D; McCarthy MJ Reduced Forms of Nitrogen Are a Driver of Non-Nitrogen-Fixing Harmful Cyanobacterial Blooms and Toxicity in Lake Erie. Harmful Algae 2019, 81, 86–93. [PubMed: 30638502]

(25). Ho JC; Michalak AM; Pahlevan N Widespread Global Increase in Intense Lake Phytoplankton Blooms since the 1980s. Nature 2019, 574 (7780), 667–670. [PubMed: 31610543]

(26). Sinha E; Michalak AM Precipitation Dominates Interannual Variability of Riverine Nitrogen Loading across the Continental United States. Environmental science & technology 2016, 50 (23), 12874–12884. [PubMed: 27771946]

(27). Howarth R; Swaney D; Billen G; Garnier J; Hong B; Humborg C; Johnes P; Mörth C-M; Marino R Nitrogen Fluxes from the Landscape Are Controlled by Net Anthropogenic Nitrogen Inputs and by Climate. Frontiers in Ecology and the Environment 2012, 10 (1), 37–43. 10.1890/100178.

(28). McCrackin ML; Harrison JA; Compton JE A Comparison of NEWS and SPARROW Models to Understand Sources of Nitrogen Delivered to US Coastal Areas. Biogeochemistry 2013, 114 (1–3), 281–297.

(29). Preston SD; Alexander RB; Schwarz GE; Crawford CG Factors Affecting Stream Nutrient Loads: A Synthesis of Regional SPARROW Model Results for the Continental United States 1. JAWRA Journal of the American Water Resources Association 2011, 47 (5), 891–915. [PubMed: 22457574]

(30). Shoda ME; Sprague LA; Murphy JC; Riskin ML Water-Quality Trends in U.S. Rivers, 2002 to 2012: Relations to Levels of Concern. Science of The Total Environment 2019, 650, 2314–2324. 10.1016/j.scitotenv.2018.09.377.

(31). Stets EG; Sprague LA; Oelsner GP; Johnson HM; Murphy JC; Ryberg K; Vecchia AV; Zuellig RE; Falcone JA; Riskin ML Landscape Drivers of Dynamic Change in Water Quality of U.S. Rivers. Environ. Sci. Technol 2020, 54 (7), 4336–4343. 10.1021/acs.est.9b05344. [PubMed: 32216285]

(32). Argerich A; Johnson SL; Sebestyen SD; Rhoades CC; Greathouse E; Knoepp JD; Adams MB; Likens GE; Campbell JL; McDowell WH Trends in Stream Nitrogen Concentrations for Forested Reference Catchments across the USA. Environmental Research Letters 2013, 8 (1), 014039.

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

(33). Sinha E; Michalak AM; Balaji V Eutrophication Will Increase during the 21st Century as a Result of Precipitation Changes. Science 2017, 357 (6349), 405–408. 10.1126/science.aan2409. [PubMed: 28751610]

(34). Sabo RD; Clark CM; Bash J; Sobota D; Cooter E; Dobrowolski JP; Houlton BZ; Rea A; Schwede D; Morford SL; Compton JE Decadal Shift in Nitrogen Inputs and Fluxes Across the Contiguous United States: 2002–2012. J. Geophys. Res. Biogeosci 2019, 124 (10), 3104–3124. 10.1029/2019JG005110.

(35). Swaney DP; Howarth RW Phosphorus Use Efficiency and Crop Production: Patterns of Regional Variation in the United States, 1987–2012. Science of The Total Environment 2019, 685, 174–188. 10.1016/j.scitotenv.2019.05.228.

(36). US EPA, O. Wadeable Streams Assessment: A Collaborative Survey of the Nation's Streams. https://www.epa.gov/sites/production/files/2014-10/documents/2007_5_16_streamsurvey_wsa_assessment_may2007.pdf.

(37). Strahler AN Quantitative Analysis of Watershed Geomorphology. Eos, Transactions American Geophysical Union 1957, 38 (6), 913–920.

(38). Bellmore RA; Compton JE; Brooks JR; Fox EW; Hill RA; Sobota DJ; Thornbrugh DJ; Weber MH Nitrogen Inputs Drive Nitrogen Concentrations in U.S. Streams and Rivers during Summer Low Flow Conditions. Science of The Total Environment 2018, 639, 1349–1359. 10.1016/j.scitotenv.2018.05.008.

(39). Omernik JM Ecoregions of the Conterminous United States. Annals of the Association of American geographers 1987, 77 (1), 118–125.

(40). Hill RA; Weber MH; Leibowitz SG; Olsen AR; Thornbrugh DJ The Stream-Catchment (StreamCat) Dataset: A Database of Watershed Metrics for the Conterminous United States. JAWRA Journal of the American Water Resources Association 2016, 52 (1), 120–128.

(41). Breiman L Random Forests. Machine learning 2001, 45 (1), 5–32.

(42). Walsh ES; Kreakie BJ; Cantwell MG; Nacci D A Random Forest Approach to Predict the Spatial Distribution of Sediment Pollution in an Estuarine System. PLOS ONE 2017, 12 (7), e0179473. 10.1371/journal.pone.0179473. [PubMed: 28738089]

(43). De'ath G; Fabricius KE Classification and Regression Trees: A Powerful yet Simple Technique for Ecological Data Analysis. Ecology 2000, 81 (11), 3178–3192. https://doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.

(44). Team, R. C. R: A Language and Environment for Statistical Computing; Vienna, Austria, 2013.

(45). Shen LQ; Amatulli G; Sethi T; Raymond P; Domisch S Estimating Nitrogen and Phosphorus Concentrations in Streams and Rivers, within a Machine Learning Framework. Sci Data 2020, 7 (1), 161. 10.1038/s41597-020-0478-7. [PubMed: 32467642]

(46). Hale RL; Grimm NB; Vörösmarty CJ; Fekete B Nitrogen and Phosphorus Fluxes from Watersheds of the Northeast U.S. from 1930 to 2000 Role of Anthropogenic Nutrient Inputs, Infrastructure, and Runoff. Global Biogeochem. Cycles 2015, 29 (3), 341–356. 10.1002/2014GB004909.

(47). Abbott BW; Gruau G; Zarnetske JP; Moatar F; Barbe L; Thomas Z; Fovet O; Kolbe T; Gu S; Pierson-Wickmann A; Davy P; Pinay G Unexpected Spatial Stability of Water Chemistry in Headwater Stream Networks. Ecol Lett 2018, 21 (2), 296–308. 10.1111/ele.12897. [PubMed: 29282860]

(48). Van Meter KJ; Basu NB Catchment Legacies and Time Lags: A Parsimonious Watershed Model to Predict the Effects of Legacy Storage on Nitrogen Export. PLoS ONE 2015, 10 (5), e0125971. 10.1371/journal.pone.0125971. [PubMed: 25985290]

(49). Burt TP; Howden NJK; Worrall F; Whelan MJ Long-Term Monitoring of River Water Nitrate: How Much Data Do We Need? Journal of Environmental Monitoring 2010, 12 (1), 71–79. [PubMed: 20082001]

(50). Stoddard JL Long-Term Changes in Watershed Retention of Nitrogen: Its Causes and Aquatic Consequences. Advances in Chemistry Series 1994, 237, 223–284.

(51). Aber JD; Nadelhoffer KJ; Steudler P; Melillo JM Nitrogen Saturation in Northern Forest Ecosystems. BioScience 1989, 39 (6), 378–386. 10.2307/1311067.

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

(52). Lovett GM; Weathers KC; Sobczak WV Nitrogen Saturation and Retention in Forested Watersheds of the Catskill Mountains, New York. Ecological Applications 2000, 10 (1), 73–84.

(53). Bernot MJ; Dodds WK Nitrogen Retention, Removal, and Saturation in Lotic Ecosystems. Ecosystems 2005, 8 (4), 442–453.

(54). Perakis SS; Hedin LO Nitrogen Loss from Unpolluted South American Forests Mainly via Dissolved Organic Compounds. Nature 2002, 415 (6870), 416–419. [PubMed: 11807551]

(55). Oliver SK; Collins SM; Soranno PA; Wagner T; Stanley EH; Jones JR; Stow CA; Lottig NR Unexpected Stasis in a Changing World: Lake Nutrient and Chlorophyll Trends since 1990. Glob Change Biol 2017, 23 (12), 5455–5467. 10.1111/gcb.13810.

(56). Sprague LA; Mitchell RM; Pollard AI; Falcone JA Assessing Water-Quality Changes in US Rivers at Multiple Geographic Scales Using Results from Probabilistic and Targeted Monitoring. Environ Monit Assess 2019, 191 (6), 348. 10.1007/s10661-019-7481-5. [PubMed: 31055662]

(57). Comer-Warner SA; Gooddy DC; Ullah S; Glover L; Kettridge N; Wexler SK; Kaiser J; Krause S Seasonal Variability of Sediment Controls of Nitrogen Cycling in an Agricultural Stream. Biogeochemistry 2020, 148 (1), 31–48. 10.1007/s10533-020-00644-z.

(58). Lin J; Compton JE; Leibowitz SG; Mueller-Warrant G; Matthews W; Schoenholtz SH; Evans DM; Coulombe RA Seasonality of Nitrogen Balances in a Mediterranean Climate Watershed, Oregon, US. Biogeochemistry 2019, 142 (2), 247–264. 10.1007/s10533-018-0532-0.

(59). Lin J; Böhlke JK; Huang S; Gonzalez-Meler M; Sturchio NC Seasonality of Nitrate Sources and Isotopic Composition in the Upper Illinois River. Journal of Hydrology 2019, 568, 849–861. 10.1016/j.jhydrol.2018.11.043.

(60). Crawford JT; Stets EG; Sprague LA Network Controls on Mean and Variance of Nitrate Loads from the Mississippi River to the Gulf of Mexico. Journal of Environmental Quality 2019, 48 (6), 1789–1799. 10.2134/jeq2018.12.0435.

(61). Baird M; Zabowski D; Everett RL Wildfire Effects on Carbon and Nitrogen in Inland Coniferous Forests. Plant and Soil 1999, 209 (2), 233–243.

(62). Rhoades CC; Entwistle D; Butler D The Influence of Wildfire Extent and Severity on Streamwater Chemistry, Sediment and Temperature Following the Hayman Fire, ColoradoA. International Journal of Wildland Fire 2011, 20 (3), 430–442.

(63). Rust AJ; Saxe S; McCray J; Rhoades CC; Hogue TS Evaluating the Factors Responsible for Post-Fire Water Quality Response in Forests of the Western USA. Int. J. Wildland Fire 2019, 28 (10), 769–784. 10.1071/WF18191.

(64). Koplitz SN; Nolte CG; Sabo RD; Clark CM; Horn KJ; Thomas RQ; Newcomer-Johnson TA The Contribution of Wildland Fire Emissions to Deposition in the U S: Implications for Tree Growth and Survival in the Northwest. Environ. Res. Lett 2021, 16 (2), 024028. 10.1088/1748-9326/abd26e.

(65). Knicker H; Almendros G; González-Vila FJ; Martin F; Lüdemann H-D 13C- and 15N-NMR Spectroscopic Examination of the Transformation of Organic Nitrogen in Plant Biomass during Thermal Treatment. Soil Biology and Biochemistry 1996, 28 (8), 1053–1060. 10.1016/0038-0717(96)00078-8.

(66). Hauer FR; Spencer CN Phosphorus and Nitrogen Dynamics in Streams Associated with Wildfire: A Study of Immediate and Longterm Effects. International Journal of Wildland Fire 1998, 8 (4), 183–198.

(67). Wolock DM Base-Flow Index Grid for the Conterminous United States. Open-File Report 03-236. US Geological Survey, Reston, VA 2003.

(68). Pellerin BA; Wollheim WM; Hopkinson CS; McDowell WH; Williams MR; Vörösmarty CJ; Daley ML Role of Wetlands and Developed Land Use on Dissolved Organic Nitrogen Concentrations and DON/TDN in Northeastern U.S. Rivers and Streams. Limnology and Oceanography 2004, 49 (4), 910–918. 10.4319/lo.2004.49.4.0910.

(69). Inamdar SP; Mitchell MJ Storm Event Exports of Dissolved Organic Nitrogen (DON) across Multiple Catchments in a Glaciated Forested Watershed. Journal of Geophysical Research: Biogeosciences 2007, 112 (G2).

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

(70). Hayakawa A; Shimizu M; Woli KP; Kuramochi K; Hatano R Evaluating Stream Water Quality through Land Use Analysis in Two Grassland Catchments. Journal of Environmental Quality 2006, 35 (2), 617–627. 10.2134/jeq2005.0343. [PubMed: 16510707]

(71). Flint SA; McDowell WH Effects of Headwater Wetlands on Dissolved Nitrogen and Dissolved Organic Carbon Concentrations in a Suburban New Hampshire Watershed. Freshwater Science 2015, 34 (2), 456–471.

(72). Smith RA; Alexander RB; Schwarz GE Natural Background Concentrations of Nutrients in Streams and Rivers of the Conterminous United States. Environ. Sci. Technol 2003, 37 (14), 3039–3047. 10.1021/es020663b. [PubMed: 12901648]

(73). Cheng FY; Basu NB Biogeochemical Hotspots: Role of Small Water Bodies in Landscape Nutrient Processing: NUTRIENT REMOVAL IN SMALL WATER BODIES. Water Resour. Res 2017, 53 (6), 5038–5056. 10.1002/2016WR020102.

(74). Mengistu SG; Golden HE; Lane CR; Christensen JR; Wine ML; D'Amico E; Prues A; Leibowitz SG; Compton JE; Weber MH Wetland Flowpaths Mediate Nitrogen and Phosphorus Concentrations across the Upper Mississippi River Basin. JAWRA Journal of the American Water Resources Association 2020.

(75). Cheng FY; Van Meter KJ; Byrnes DK; Basu NB Maximizing US Nitrate Removal through Wetland Protection and Restoration. Nature 2020, 1–6.

(76). Mitsch WJ; Day JW; Zhang L; Lane RR Nitrate-Nitrogen Retention in Wetlands in the Mississippi River Basin. Ecological Engineering 2005, 24 (4), 267–278. 10.1016/j.ecoleng.2005.02.005.

(77). Wolock DM Base-Flow Index Grid for the Conterminous United States; 2003.

(78). Gardner JR; Fisher TR; Jordan TE; Knee KL Balancing Watershed Nitrogen Budgets: Accounting for Biogenic Gases in Streams. Biogeochemistry 2016, 127 (2), 231–253. 10.1007/s10533-015-0177-1.

(79). Kothawala DN; Moore TR; Hendershot WH Soil Properties Controlling the Adsorption of Dissolved Organic Carbon to Mineral Soils. Soil Science Society of America Journal 2009, 73 (6), 1831–1842. 10.2136/sssaj2008.0254.

(80). Osburn CL; Oviedo-Vargas D; Barnett E; Dierick D; Oberbauer SF; Genereux DP Regional Groundwater and Storms Are Hydrologic Controls on the Quality and Export of Dissolved Organic Matter in Two Tropical Rainforest Streams, Costa Rica. Journal of Geophysical Research: Biogeosciences 2018, 123 (3), 850–866. 10.1002/2017JG003960.

(81). Murphy DV; Macdonald AJ; Stockdale EA; Goulding KWT; Fortune S; Gaunt JL; Poulton PR; Wakefield JA; Webster CP; Wilmer WS Soluble Organic Nitrogen in Agricultural Soils. Biology and Fertility of Soils 2000, 30 (5–6), 374–387. 10.1007/s003740050018.

(82). Noe GB; Hupp CR; Rybicki NB Hydrogeomorphology Influences Soil Nitrogen and Phosphorus Mineralization in Floodplain Wetlands. Ecosystems 2013, 16 (1), 75–94. 10.1007/s10021-012-9597-0.

(83). Van Meter KJ; Van Cappellen P; Basu NB Legacy Nitrogen May Prevent Achievement of Water Quality Goals in the Gulf of Mexico. Science 2018, 360 (6387), 427–430. 10.1126/science.aar4462. [PubMed: 29567808]

(84). Tesoriero AJ; Duff JH; Saad DA; Spahr NE; Wolock DM Vulnerability of Streams to Legacy Nitrate Sources. Environmental science & technology 2013, 47 (8), 3623–3629. [PubMed: 23530900]

(85). Van Meter KJ; Van Cappellen P; Basu NB Legacy Nitrogen May Prevent Achievement of Water Quality Goals in the Gulf of Mexico. Science 2018, 360 (6387), 427–430. 10.1126/science.aar4462. [PubMed: 29567808]

(86). Stackpoole SM; Stets EG; Sprague LA Variable Impacts of Contemporary versus Legacy Agricultural Phosphorus on US River Water Quality. Proc Natl Acad Sci USA 2019, 116 (41), 20562–20567. 10.1073/pnas.1903226116. [PubMed: 31548416]

(87). McCrackin ML; Muller-Karulis B; Gustafsson BG; Howarth RW; Humborg C; Svanbäck A; Swaney DP A Century of Legacy Phosphorus Dynamics in a Large Drainage Basin. Global Biogeochem. Cycles 2018, 32 (7), 1107–1122. 10.1029/2018GB005914.

(88). Sabo RD; Clark CM; Compton JE Considerations When Using Nutrient Inventories to Prioritize Water Quality Improvement Efforts across the US. Environ. Res. Commun 2021, 3 (4), 045005. 10.1088/2515-7620/abf296.

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

Synopsis

Watershed nitrogen inputs drive stream nitrogen concentrations and forms, while climate, hydrology, and land cover are also influential

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

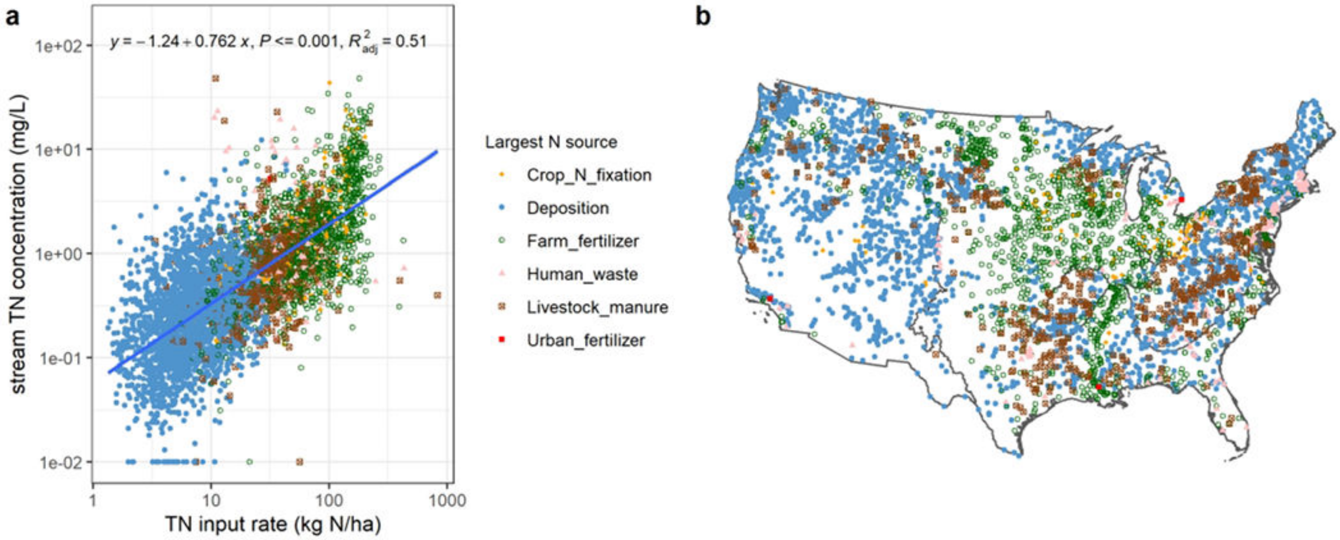


Figure 1.

a. Stream TN concentrations and largest N sources across the CONUS; colors and shapes represent different largest landscape TN sources. b. Stream TN concentrations as a function of N inputs; colors and shapes represent different largest landscape TN sources. Lines are the three regions' boundaries (Figure S1).

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

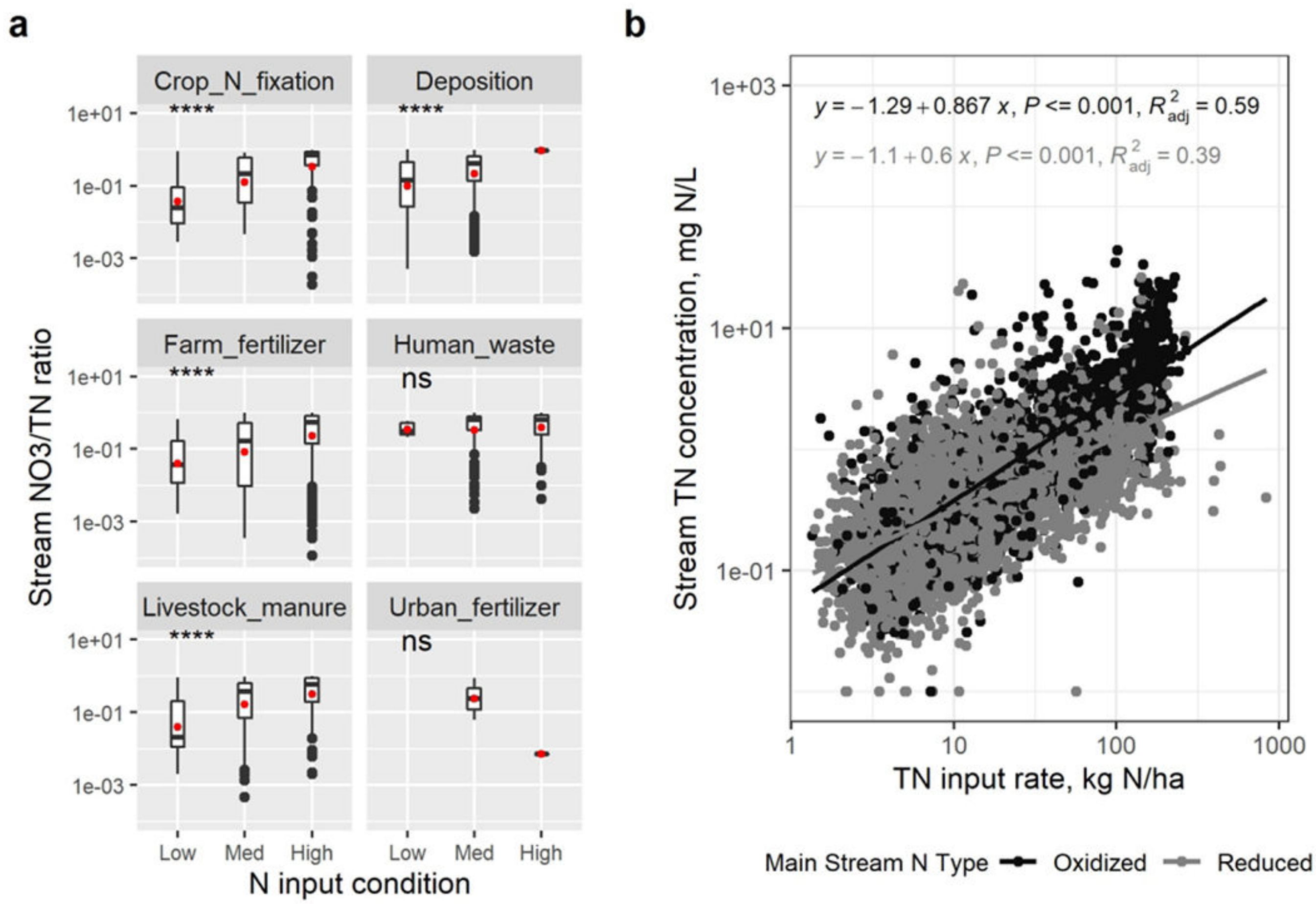


Figure 2.

a. Comparing stream NO3 contribution at three input levels and in watersheds with different larges N sources. Red points indicate group means. ****: p <= 0.0001; ns: not significant.

b. Stream TN concentrations as a function of N inputs; colors represent the main N species (‘oxidized’-black vs ‘reduced’-grey).

EPA Author Manuscript
EPA Author Manuscript
EPA Author Manuscript

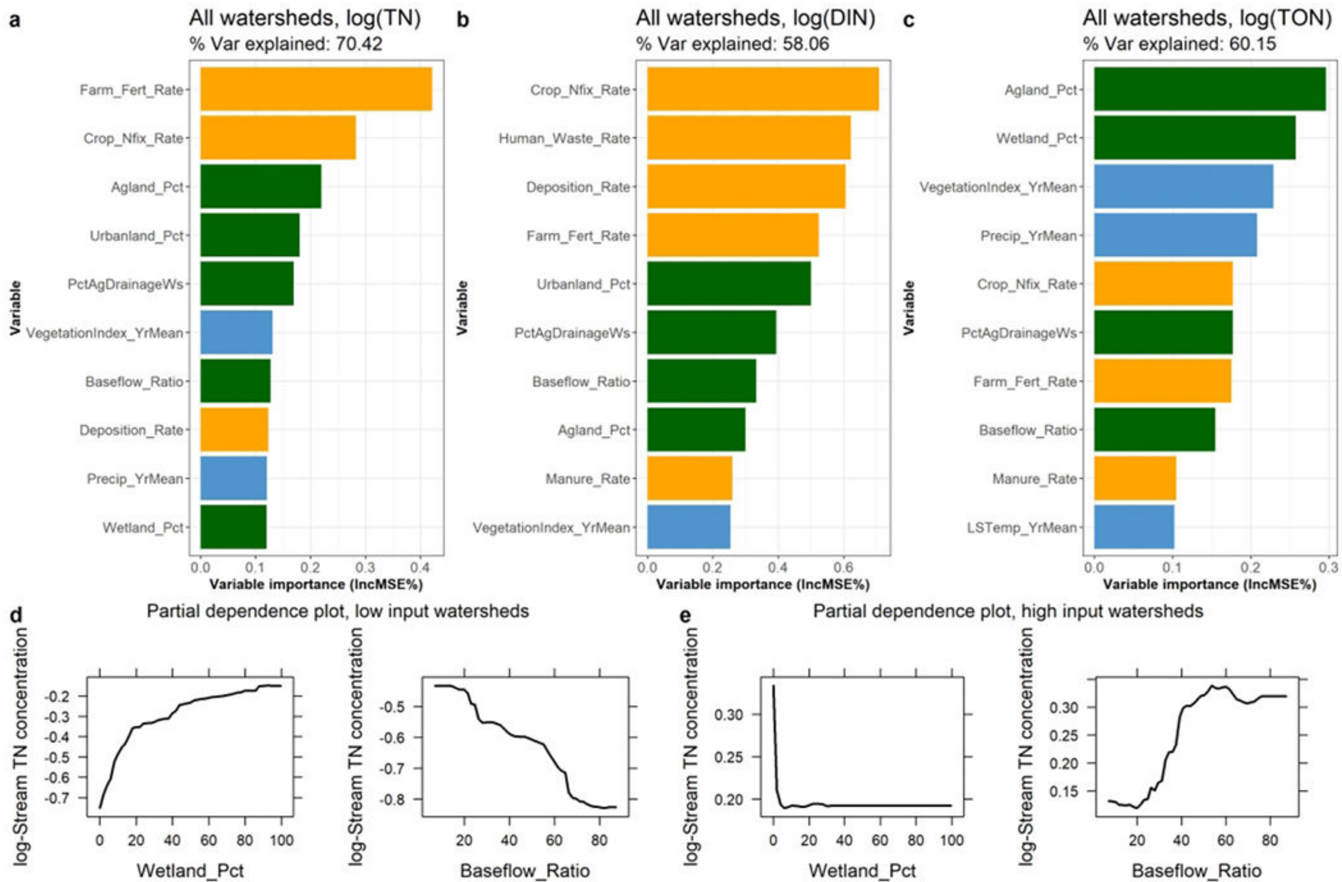


Figure 3.

Variable importance plots of random forest model results for a) stream TN concentration; b) DIN concentration, and c) TON concentration. Each model combined data from three surveys (n = 4997). Only the top 10 variables were plotted. Bar colors represent different predictor categories: Green—StreamCat landscape and land use (%) variables; Blue—climatic variables; Orange— N input rates of six sources. The partial dependence relationships of stream TN concentration vs. wetland coverage and baseflow contribution in low input watersheds (d) exhibited opposite trends with those in high input watersheds (e).

EPA Author Manuscript EPA Author Manuscript EPA Author Manuscript

Table 1.

a. Means and standard errors (SE) of watershed N input rate, stream N concentrations, NO3/TN ratio, and TON/TN ratio of a. watersheds of different land use types and input levels, and b. different regions and stream N types. n = number of streams.

a

Watershed type	Input level	N input rate (kg N ha^{-1})	Stream TN (mg N L^{-1})	Stream NO3 (mg NL^{-1})	Stream TON (mg NL^{-1})	Stream NO3/TN ratio	Stream TON/TN ratio	n
		Mean (SE)						
Agriculture	Low	10.50 (0.25)	0.60 (0.14)	0.21 (0.13)	0.38 (0.03)	0.13 (0.02)	0.83 (0.02)	134
	Medium	32.76 (0.31)	1.10 (0.04)	0.44 (0.04)	0.62 (0.02)	0.33 (0.01)	0.63 (0.01)	895
	High	104.85 (1.54)	3.16 (0.13)	2.21 (0.11)	0.85 (0.04)	0.51 (0.01)	0.45 (0.01)	1132
	All	69.14 (1.15)	2.15 (0.07)	1.35 (0.07)	0.72 (0.02)	0.41 (0.01)	0.55 (0.01)	2161
Deposition	Low	6.93 (0.08)	0.39 (0.01)	0.11 (0.01)	0.27 (0.01)	0.25 (0.01)	0.71 (0.01)	2080
	Medium	22.52 (0.26)	0.71 (0.04)	0.33 (0.03)	0.34 (0.02)	0.40 (0.01)	0.55 (0.01)	577
	High	60.69 (3.48)	2.31 (0.16)	2.14 (0.21)	0.16 (0.05)	0.92 (0.03)	0.07 (0.02)	2
	All	10.36 (0.15)	0.46 (0.01)	0.16 (0.01)	0.28 (0.01)	0.28 (0.01)	0.67 (0.01)	2659
Human	Low	12.15 (0.69)	12.78 (4.07)	3.84 (1.12)	1.40 (0.66)	0.37 (0.07)	0.16 (0.06)	5
	Medium	33.77 (0.86)	1.93 (0.32)	1.51 (0.31)	0.37 (0.02)	0.55 (0.03)	0.41 (0.03)	93
	High	82.69 (5.43)	1.99 (0.27)	1.33 (0.24)	0.50 (0.04)	0.54 (0.03)	0.40 (0.03)	79
	All	54.99 (3.10)	2.26 (0.27)	1.50 (0.20)	0.46 (0.03)	0.54 (0.02)	0.40 (0.02)	177

b

Regions	Stream N type	N input rate (kg N ha^{-1})	Stream TN (mg N L^{-1})	Stream NO3 (mg NL^{-1})	Stream TON (mg NL^{-1})	Stream NO3/TN ratio	Stream TON/TN ratio	n
				Mean (SE)				
Eastern Highlands	Oxidized	36.02 (1.02)	1.24 (0.08)	1.00 (0.08)	0.22 (0.01)	0.71 (0.01)	0.26 (0.00)	642
	Reduced	24.14 (1.50)	0.47 (0.02)	0.11 (0.00)	0.32 (0.02)	0.25 (0.01)	0.69 (0.01)	667
	All	29.97 (0.93)	0.85 (0.04)	0.55 (0.04)	0.27 (0.01)	0.47 (0.01)	0.48 (0.01)	1309
Plains and Lowlands	Oxidized	91.89 (2.20)	3.95 (0.18)	3.33 (0.16)	0.57 (0.03)	0.76 (0.01)	0.22 (0.00)	726

EPA Author Manuscript

EPA Author Manuscript

EPA Author Manuscript

a

Watershed type	Input level	N input rate (kg N ha^{-1})	Stream TN (mg N L^{-1})	Stream NO3 (mg NL^{-1})	Stream TON (mg NL^{-1})	Stream NO3/TN ratio	Stream TON/TN ratio	n
		Mean (SE)						
	Reduced	43.71 (1.10)	1.08 (0.04)	0.15 (0.01)	0.84 (0.03)	0.13 (0.00)	0.82 (0.00)	1593
	All	58.80 (1.12)	1.98 (0.07)	1.14 (0.06)	0.76 (0.02)	0.33 (0.01)	0.63 (0.01)	2319
Western Mountains	Oxidized	11.55 (1.09)	0.78 (0.08)	0.61 (0.07)	0.16 (0.02)	0.69 (0.01)	0.27 (0.01)	312
	Reduced	7.10 (0.42)	0.31 (0.03)	0.05 (0.01)	0.22 (0.01)	0.13 (0.00)	0.82 (0.00)	1057
	All	8.11 (0.41)	0.42 (0.03)	0.18 (0.02)	0.21 (0.01)	0.26 (0.01)	0.69 (0.01)	1369