11-1-2017

Aquatic Nitrate Retention at River Network Scales Across Flow Conditions Determined Using Nested In Situ Sensors

Wilfred Wollheim
University of New Hampshire, Durham, wil.wollheim@unh.edu

Gopala K. Mulukutla
University of New Hampshire, Durham, gopal.mulukutla@unh.edu

C. Cook
University of New Hampshire, Durham

R. O. Carey
University of New Hampshire, Durham

Recommended Citation
Wollheim, W. M., Mulukutla, G. K., Cook, C., & Carey, R. O. (2017). Aquatic nitrate retention at river network scales across flow conditions determined using nested in situ sensors. Water Resources Research, 53, 9740–9756. https://doi.org/10.1002/2017WR020644

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2017WR020644

Special Section:
Continuous Nutrient Sensing in Research and Management: Applications and Lessons Learned Across Aquatic Environments and Watersheds

Aquatic Nitrate Retention at River Network Scales Across Flow Conditions Determined Using Nested In Situ Sensors

W. M. Wollheim[1,2], G. K. Mulukutla[1,2], C. Cook[1,2], and R. O. Carey[1,2]

[1]Department of Natural Resources and Environment, University of New Hampshire, Durham, NH, USA, [2]Earth Systems Research Center, University of New Hampshire, Durham, NH, USA

Abstract Nonpoint pollution sources are strongly influenced by hydrology and are therefore sensitive to climate variability. Some pollutants entering aquatic ecosystems, e.g., nitrate, can be mitigated by in-stream processes during transport through river networks. Whole river network nitrate retention is difficult to quantify with observations. High frequency, *in situ* nitrate sensors, deployed in nested locations within a single watershed, can improve estimates of both nonpoint inputs and aquatic retention at river network scales. We deployed a nested sensor network and associated sampling in the urbanizing Oyster River watershed in coastal New Hampshire, USA, to quantify storm event-scale loading and retention at network scales. An end member analysis used the relative behavior of reactive nitrate and conservative chloride to infer river network fate of nitrate. In the headwater catchments, nitrate and chloride concentrations are both increasingly diluted with increasing storm size. At the mouth of the watershed, chloride is also diluted, but nitrate tended to increase. The end member analysis suggests that this pattern is the result of high retention during small storms (51–78%) that declines to zero during large storms. Although high frequency nitrate sensors did not alter estimates of fluxes over seasonal time periods compared to less frequent grab sampling, they provide the ability to estimate nitrate flux versus storm size at event scales that is critical for such analyses. Nested sensor networks can improve understanding of the controls of both loading and network scale retention, and therefore also improve management of nonpoint source pollution.

Plain Language Summary High frequency *in situ* nitrate sensors reveal whole river network nitrate retention is high during small storms but declines rapidly to no retention during large storms. Nonpoint nitrate dilutes during storms in headwaters but increases at basin mouth, while chloride dilutes in both headwaters and basin mouth. Wider use of storm even sampling in nested sensor networks increases understanding of catchment biogeochemistry and could improve management of nonpoint pollution.

Key Points:
- Nested network of nitrate sensors reveal whole river network nitrate retention is high during small storms but declines rapidly with increasing storm size
- Nonpoint nitrate dilutes during storms in headwaters but increases at basin mouth, while chloride always dilutes
- Wider use of storm event sampling in nested sensor networks could improve management of nonpoint pollution

Supporting Information:
- Supporting Information S1

Correspondence to:
W. M. Wollheim,
wil.wollheim@unh.edu

Citation:
Wollheim, W. M., Mulukutla, G. K., Cook, C., & Carey, R. O. (2017). Aquatic nitrate retention at river network scales across flow conditions determined using nested in situ sensors. *Water Resources Research*, *53*, 9740–9756. https://doi.org/10.1002/2017WR020644

Received 24 FEB 2017
Accepted 25 OCT 2017
Accepted article online 1 NOV 2017
Published online 27 NOV 2017

1. Introduction

Nonpoint sources of nitrogen (N) are a major factor affecting ecosystem health in receiving water bodies (Howarth & Marino, 2006; Smith et al., 1999). While point source N pollution has steadily declined in recent decades, nonpoint sources continue to be a problem (Alexander et al., 2008; Carpenter et al., 1998). Understanding trends and responses to mitigation requires robust measurements of N inputs from different land covers, including during storms. The network of aquatic ecosystems in a watershed can mitigate N fluxes before they reach downstream water bodies (Alexander et al., 2000; Wollheim et al., 2006). While models have indicated the potential importance of network scale retention, observations of retention at network scales are difficult due to the distributed nature of inputs and sinks. To better understand the role of river network nutrient retention - how it varies spatially or through time - requires robust simultaneous measures of both inputs to the stream network and exports at the basin mouth. Nested monitoring networks of *in situ* sensors are a potential approach that could meet this need.

Unlike point sources, nonpoint sources are difficult to quantify because their inputs are diffuse and variable in the landscape. Samples from headwater streams are often used to infer quantities of nonpoint nutrient sources from different land covers (e.g., Wollheim et al., 2008a). However, this is problematic for two reasons. First, it is difficult to get an adequate number of samples throughout the hydrograph, particularly in flashy urban streams with abundant impervious surfaces. Measurement of flow itself is essential, yet gaging

stations in small catchments of different land cover are relatively rare. Further, in impacted catchments, nutrient concentrations are likely to be variable through time due to temporal heterogeneity in anthropogenic activity that contribute sources and storm characteristics (size, intensity) that mobilize and transport nutrients (Bende-Michl et al., 2013). Second, in-stream processes are likely to modify terrestrial signals (Bernhardt et al., 2005a), especially for extremely reactive nutrients like ammonium and phosphate (Peterson et al., 2001), though less reactive forms like nitrate and dissolved organic nitrogen are likely to be less affected over relatively short headwater reaches (Wollheim, 2016; Wollheim et al., 2006). High frequency sampling using *in situ* sensors is likely to greatly improve estimates of nonpoint source loading from land to water, particularly during storm events when most loads can occur (Doyle, 2005).

While in-stream processes are less likely to have altered the terrestrial signal within headwater streams, especially during storms when residence times are very short, in-stream processes may accumulate throughout river networks (Helton et al., 2011; Wollheim, 2016). Nutrient measurements at downstream sites in larger watersheds are likely to be a signal of both terrestrial loading and in-stream processes (Miller et al., 2016). Miller et al. (2016) used time varying patterns at downstream locations to infer network scale retention seasonally. However, this approach assumed that concentrations of sources to river networks are constant through time, which may not hold due to dependence of loading concentration on discharge, or seasonal variability in terrestrial retention that controls transfers to aquatic systems. Further, the representativeness of any intensively monitored headwaters must also be assessed against a wider array of catchments that are sampled less frequently.

Here we assess the ability of an entire river network to retain nitrate across flow conditions using a nested monitoring network within a single river network. The monitoring network integrates high frequency measurement of conservative and reactive solutes, less frequent grab sampling, and spatially extensive sampling, all distributed from headwaters to basin mouth. We have two goals: 1) assess whether high frequency *in situ* sensors improve loading estimates of nitrate to impacted headwater streams over more traditional grab sampling, 2) evaluate whether the river network of the watershed as a whole alters the signature of nutrient export patterns at time scales ranging from storm event to annual. We hypothesized that *in situ* sensors would improve estimates of loading in flashy headwaters because they better capture short-term variability that is difficult using grab samples alone. We also hypothesized that the river network alters the signature of nitrate export during low to moderate storms, but not large storms.

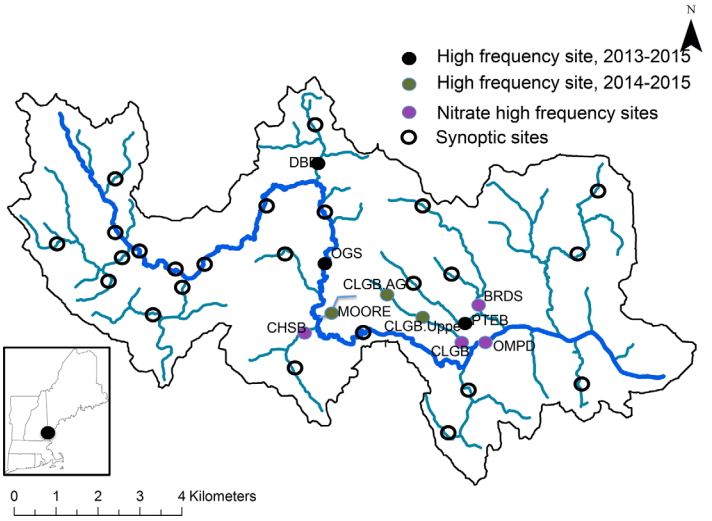

Figure 1. Map of the Oyster River watershed, NH, USA, showing sampling location of different frequency and time frames. The main stem below OMPD is tidal and not part of this study. High frequency sites were also included in the synoptic survey.

2. Methods

2.1. Study Area

The Oyster River watershed is a 50.6 km^2 coastal watershed in southeastern New Hampshire, USA (Figure 1). The watershed (defined at the head of tide dam) is characterized by a mixture of developed (17.3%), agricultural (11.1%) and forested (59.1%) land cover (site = OMPD in Table 1 and Figure 1) (Homer et al., 2015). Developed includes all all nonagricultural human land cover as defined in the National Land Cover Dataset. Wetlands and open water are 11.6% and 1% of the watershed area, respectively. Impervious surfaces are 8%. The head of tide dam (which prevents tidal influence) creates the small Mill Pond (pond surface area = 30,000 m^2) located at the basin mouth. Most of the anthropogenic land cover (developed + agricultural) is skewed toward the basin mouth (skewness index = 0.77). The skewness index is an index of how land cover is distributed relative to the average flow path in the watershed, with values less than one indicating anthropogenic land cover is located downstream in the watershed (Mineau et al., 2015). Mean annual rainfall is 1,280 mm yr^{-1}, mean annual runoff at the USGS gage located ~6 km upstream of the head of tide dam is 582 mm yr^{-1}, and average annual air temperature is 8.9°C. The period of study (2013–2015) had below average

Table 1
Land Cover Characteristics of the Intensively Studied Watersheds

Station	Name	Area (km^2)	Impervious (%)	Developed (%)	Agriculture (%)	Forest (%)	Wetlands (%)	Other (%)	Sensor time period
BRDS	Beard's Creek	4.8	14.0	37.4	13.2	43.9	4.9	0.6	2013–2015 w/NO_3 (2013)
CHSB	Chesley Brook	4.0	6.2	12.9	24.5	48.9	13.8	0.0	2013–2015 w/NO_3 (2014)
CLGB	College Brook	2.3	28.4	68.7	9.8	20.8	0.7	0.0	2013–2015 w/NO_3
CLGB.AG	College Brook	0.64	27.9	77.2	20.1	2.8	0.0	0.0	2014–2015
CLGB.UPPER	College Brook	1.7	20.3	57.7	13.5	27.9	0.9	0.0	2014–2015
DBB	Dube Brook	3.3	4.8	7.9	15.4	59.4	17.3	0.0	2013–2015
MOORE	Moore Stream	0.14	2.0	11.0	27.1	32.9	29.0	0.0	2014–2015
OGS	Oyster River @ USGS Gage	30.3	6.2	11.2	7.5	66.3	13.7	1.3	2013–2015
OMPD	Oyster River @ Mill Pond	50.6	8.0	17.3	11.1	59.1	11.6	0.9	2013–2015 w/NO_3
PTEB	Pettee Brook	3.0	23.0	52.6	10.4	29.4	6.6	0.9	2013–2015

Note. Land cover is based on the 2011 National Land Cover Dataset. Developed includes all nonagricultural human dominated land cover classified by NLDC. Anthropogenic land cover is Developed + Agriculture. All sites included stage and conductivity loggers throughout their period of deployment.

rainfall, with mean precipitation = 1,066 and 1,069 mm yr^{-1} and annual runoff = 586 and 448 mm yr^{-1}, corresponding to the June–May period for 2013–2014 and 2014–2015, respectively. Rainfall is typically distributed evenly throughout the year, while runoff is much lower during the summer due to high evapotranspiration. However, storms that occur throughout the year result in elevated runoff, particularly in catchments with high impervious surface cover. The watershed lies in a formerly glaciated region with shallow soils, and is relatively shallow sloped, with maximum elevation ~100 m.

2.2. Overall Measurement Design

We deployed a network of high frequency sensors in headwater streams (< 5 km^2) and at the basin mouth (50.6 km^2) over a 2 year period (June 2013 to May 2015). We had two sets of high frequency sites. One set (n = 6) included stage, water temperature, and conductivity loggers, along with weekly grab sampling (less-intensive high frequency sites). The other set (n = 4) also included high frequency nitrate measurements (nitrate high frequency sites). In this study we focus mainly on the stage, conductivity, and nitrate sensor, and grab sample results, but also estimate total nitrogen (TN) annual fluxes using grab samples. We also conducted spatially extensive synoptic grab measurements (n = 34) throughout the watershed seasonally (synoptic sites) to evaluate representativeness of the high frequency sites.

2.3. High Frequency Measurements

All ten high frequency sites were located in nontidal locations across a range of land covers (forest, agriculture, and residential) and river size (headwaters versus main stem) (Figure 1 and Table 1). Each year, three of the sites had the nitrate sensors (nitrate high frequency) while four to seven sites were less-intensive. The nitrate high frequency sites for 2013 were Beards Cr. (BRDS), College Br. (CLGB), and the Oyster R. at Mill Pond (OMPD). In spring 2014, we moved the nitrate high frequency site at BRDS to Chesley Br. (CHSB) in order to increase the number of headwater sites for which we could characterize nitrate flux versus storm size relationships. The less-intensive sites are Dube Br. (DBB), the Oyster R. at the USGS gaging station (OGS), Pettee Br. (PTEB), and either CHSB (2013–2014) or BRDS (2014–2015). For 2014–2015, we added three additional small stream sites, two of which were nested within CLGB to better understand causes of impairment in this mixed-use watershed. CLGB agricultural site (CLGB.AG) is located in the headwaters of College Br., downstream of a conventional dairy with hay fields. CLGB.UPPER is located downstream of CLGB.AG and a remnant forest (College Woods), about 1 km upstream of the CLGB location. MOORE drains a mixture of cornfields and developed land. Land cover of each site is summarized in Table 1.

The nitrate high frequency sites included a Submersible Ultraviolet Nitrate Analyzer (SUNA, Satlantic Inc.) in addition to the Hobo stage and conductivity sensors (Onset Inc.) which were located at all high frequency sites. Due to instrument failure or ice conditions, we did not have continuous nitrate measurements

throughout the whole year. The percent of days in each of the 2 years with SUNA nitrate data were OMPD = 32% and 99%; CLGB = 49% and 48%, BRDS = 26% (2013–2014), and CHSB = 51% (2014–2015), with higher percent coverage during summer - fall periods (50–98%). Conductivity and stage loggers were deployed all year, resulting in greater data coverage for the year. All instruments were set to log every 15 min.

All sondes were calibrated at the beginning of the deployment and periodically throughout the year. Weekly site visits at the nitrate high frequency sites were used to maintain and clean the sensors to minimize the effect of biofouling. Although the sensors were equipped with wipers that minimize biofouling, regular cleaning is needed to ensure high quality data.

SUNA sensor estimates of nitrate were adjusted based on weekly grab samples analyzed for nitrate analytically at each site using simple linear regression (supporting information Figure S1). Raw SUNA and grab nitrate were highly correlated at all sites ($R^2 > 0.65$, $p < 0.001$) and generally near the 1:1 line, except at OMPD which had lower analytical nitrate than did the raw SUNA estimate (~ 0.1 mg L^{-1} offset, supporting information Figure S1d). See supporting information section S2 for more detailed results of the nitrate sensor adjustments. For chloride, we applied Robust Linear Regression (RLR) in the MASS package in R (R Core Team, 2015) to determine proxies for chloride from continuous specific conductance (supporting information Figure S2). We applied RLR to discount the weight of extreme outliers on the final regression.

2.4. Grab Samples

Weekly grab samples were collected at the high frequency monitoring sites to calibrate nitrate sensor measurements and to develop proxies for chloride from conductivity. We also targeted grab sampling during several storm events to validate concentration changes detected by the sensors during high flows when the optical matrix may differ. All grab samples were filtered in the field through GF/F filters and then frozen until analysis. Dissolved solutes were analyzed for nitrate-N (NO_3^--N), ammonium-N (NH_4^+-N), total dissolved N (TDN; DON = TDN – NO_3^--N – NH_4^+-N), and chloride (Cl^-). An additional 1L sample was collected to filter for total suspended solids (TSS) and particulate organic nitrogen (PON). Lab analysis was conducted by the UNH Water Quality Analysis Laboratory following standard procedures.

2.5. Discharge

Continuous stage height was measured with HOBO stage loggers deployed at fixed depth in deeper sections of the channel. Flows were measured at six sites (BRDS, CHSB, CLGB, DBB, OMPD, PTEB) between March and June 2014 by the USGS (Rick Kiah, USGS New England Water Science Center, Pembroke, NH). Measurements during this period included the highest flows measured over the 2 year period. UNH personnel performed additional flow measurements at the six sites, and three additional sites (CLGB.AG, CLGB.UPPER, MOORE). Flow information from USGS gage (USGS site id 01073000, OGS in Table 1) was also used. Flows were measured using the area-velocity method with a Flowtracker velocity meter (Sontek Inc.) or salt dilution gaging (for low flows). USGS developed rating curves are available on the NWIS website (see supporting information Text S1). At the Oyster R. at Mill Pond, which is located at a dam, low flows were measured and 10, 50, 100, and 150 year flow events were modeled (Kennedy, 1984), and then related to stage measurements. All discharge in volumetric units was converted to runoff in depth units by dividing by watershed area to standardize flows across sites.

2.6. Spatially Extensive Measurements

Samples were collected synoptically throughout the Oyster River watershed at n = 34 sites on a single day at four times across seasons: 8 August 2013; 5 November 2013; 22 April 2014; 17 June 2014. Samples were distributed across different upper headwater catchments (n = 16, <2 km^2), larger tributaries (n = 9, <5.5 km^2), which include the high frequency sites, and along the Oyster River main stem (n = 9) (Figure 1). Each synoptic campaign was conducted at base flow typical of each season to ensure relatively uniform flow conditions throughout the watershed. Discharge as measured at the USGS gaging station (OGS) was 0.045 m^3 s^{-1}, 0.19 m^3 s^{-1}, 0.68 m^3 s^{-1} and 0.14 m^3 s^{-1} for each of the four campaigns, respectively. Samples were filtered in the field and frozen upon return to the lab for later analysis as described above.

2.7. Analysis

Flow-weighted mean concentrations (FWMC) and areal fluxes for nitrate were estimated at each high frequency site using three approaches of increasing effort, cost, and accuracy. In Method 1, FWMCs were estimated using only infrequent grab samples and instantaneous runoff from the nearest USGS gage (OGS in Table 1) at the time of sampling. Fluxes were estimated using FWMC multiplied by total runoff at the USGS gage for the time period of interest. In Method 2, FWMCs were estimated using infrequent grab samples and runoff measured locally at each site. Annual discharge measurements in small headwater catchments are relatively uncommon and difficult, so this is a significant added expense in monitoring programs. Areal fluxes were then estimated using this FWMC and total local runoff. Because local measurements had gaps in the record over the year (supporting information Figure S3), we assumed the runoff coefficients determined from the period of record could be applied to precipitation occurring during the unmonitored period. Runoff coefficients were estimated as total runoff divided by total precipitation for days where we could estimate daily flow. Sites had relatively short temporal gaps in discharge records during the first year (ranging from 0 to 23% of the period), with the exception of DBB, where beaver activity precluded accurate measurements most of the year (70% of year 1; supporting information Figure S3a). At this site, we continued to use runoff from the nearby USGS gage during periods when local runoff was unavailable. In Method 3, used only for the June to December period, FWMCs and areal fluxes were estimated using the high frequency derived measurements of nitrate and runoff measured locally at each nitrate high frequency site. Chloride fluxes were also measured using these approaches.

Estimates of annual nitrate fluxes were compared using Methods 1 and 2 for both 2013–2014, and 2014–2015. The third method was compared to Methods 1 and 2 during the June to December period in each year (mean 60% of days with high frequency nitrate), when data gaps were shorter and took place during representative periods. We could not make reasonable comparisons using Method 3 over the whole year because the nitrate sensors were pulled during winter periods, during which time seasonal increases occurred at all sites, which would lead to bias of annual estimates compared to grabs. Annual TN fluxes were also estimated using Method 2, using TDN and PON grab data.

Storm event scale FWMCs and fluxes were estimated at the nitrate high frequency sites (three headwaters, one basin mouth) using the high frequency nitrate and chloride data between June and December each year. Storm event scale nitrate and chloride were also estimated at the forested headwater site (DBB) to provide a low impact end member. At the forested end member (DBB), storm event chloride concentration derived from the conductivity sensor indicated little chloride concentration response to changes in flow (slope = 0.16, $r^2 = 0.01$, $p < 0.001$). For nitrate at DBB, where there were no high frequency measurements, we assumed a constant nitrate concentration (mean = 0.091 mg L^{-1}, S.E. = 0.012) applied throughout the storm since there was no evidence of storm responses from grab samples between the June and December period (Figure 4a) (slope = −0.004, $r^2 = 0.002$, $p = 0.77$, supporting information Table S4). The lack of a nitrate concentration versus storm discharge relationship at DBB is consistent with previous findings in a different forested catchment in the nearby Lamprey watershed where a nitrate sensor was deployed (Price, 2014).

The number of storms captured by high frequency flow, nitrate, and chloride over the 2 year period were CLGB = 33, BRDS = 12, CHSB = 5, OMPD = 17, and DBB = 13. All of these storms were from the period when high frequency nitrate sensors were deployed between June and December. Storm events were manually identified based on when discharge increased above stable or declining base flow and corresponded with measurable precipitation at a nearby precipitation gage (http://www.weather.unh.edu/). The gage was centrally located in the study area (near CLGB.Upper in Figure 1). The R package EcoHydRology, BaseFlowSeparation function was then used to identify the specific start and end of individual storm events based on first generation of quick flow and return to prestorm or stable quick flow. Between the start and end of the storm period, total runoff, total flux of NO_3^--N and Cl^-, and FWMCs of NO_3^--N and Cl^- were calculated. Storm FWMCs and fluxes (kg km^{-2} $storm^{-1}$) versus storm runoff (mm $storm^{-1}$) regressions were developed using the R statistical package (R Studio Team, 2015). We used analysis of covariance (ANCOVA) and Tukeys post hoc test using the glht function in R to simultaneously test the differences among slopes across sites.

2.8. River Network Scale Nitrate Retention Across Flows

River network scale nitrate retention across different flow conditions was estimated using an end member mixing approach based on NO_3^--N to Cl^- flux ratios in the headwaters compared to observations at the basin

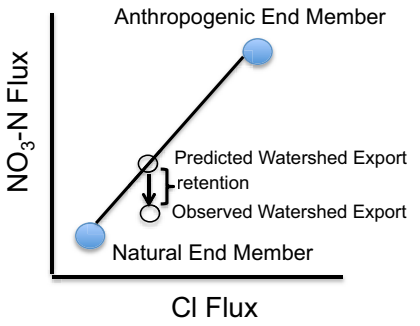

Figure 2. Conceptual model of the end member mixing analysis using nitrate and chloride fluxes, for a single storm. Divergence between predicted (from weighted inputs from anthropogenic and natural areas) and observed in watershed export is due to retention. This analysis is applied across storm levels using the regressions of storm flux versus storm size.

mouth at storm event scales (Figure 2, for one storm event). The following steps were followed to derive network-scale retention for each storm size:

1. Calculate observed storm-scale flux (kg km^{-2} $storm^{-1}$) versus storm event runoff relationships for each site for both nitrate and chloride separately (section 2.7).
2. Estimate anthropogenic (developed + agriculture) end member flux ($Flux_{anthro}$; kg km^{-2} $storm^{-1}$) versus storm event runoff (mm $storm^{-1}$) from each headwater site classified as anthropogenic (CLGB, CHSB, BRDS) for both nitrate and chloride (n = 3). Each site has some proportion of nonanthropogenic land (LU_{undev}), so the end member estimate accounts for the anthropogenic only signal by subtracting out nonanthropogenic end member (DBB). Although DBB contains some anthropogenic lands (Table 1), we see no anthropogenic signal of impairment in the data (low chloride, low nitrate, no response to flow of either chloride or nitrate), similar to other forested watersheds in the nearby Lamprey River watershed (Price, 2014). Further, DBB is very similar to other forested streams within the watershed based on the synoptic survey (supporting information Figures S6 and S7). Using the regressions from step 1, $Flux_{anthro}$ is determined for each flow level i:

$$Flux_{anthro\_i} = (Flux_{Total\_i} - LU_{undev} * Flux_{undev\_i})/(LU_{anthro}) \quad (1)$$

where, $Flux_{Total\_i}$ is the total flux for a given storm size i in the headwater catchment (kg km^{-2} $storm^{-1}$), $Flux_{undev\_i}$ is the total flux for a given storm size i in the undeveloped headwater (kg km^{-2} $storm^{-1}$), and LU_{undev} and LU_{anthro} are the proportion of land cover that is undeveloped and anthropogenically dominated (developed plus agriculture) in the headwater catchment.

3. Calculate predicted nitrate and chloride storm fluxes at the basin mouth ($Flux_{mouth\_pred}$) for each storm size i using the anthropogenic (EQN 1) and undeveloped end members (Table 2, Row 1) weighted by the proportion of anthropogenic versus nonanthropogenic land cover at the mouth of the watershed as a whole (for OMPD, Table 1):

$$Flux_{mouth\_pred\_i} = Flux_{anthro\_i} * LU_{anthro\_mouth} + Flux_{undev\_i} * LU_{undev\_mouth} \quad (2)$$

4. Calculate the predicted storm flux nitrate to chloride ratio at the basin mouth ($N{:}Cl_{mouth\_pred}$) for each storm size. This ratio represents the expected flux ratio at the basin mouth without any river network retention, if a particular end member is representative of the entire watershed.

Table 2
Regressions Between Nitrate and Chloride Areal Fluxes (kg km^{-2} $storm^{-1}$) Versus Total Storm Runoff (mm $storm^{-1}$) Used in the End Member Mixing Analysis

Site	Nitrate-N storm flux versus runoff (kg km^{-2} $storm^{-1}$)	Nitrate flux p-value; R^2	Chloride storm flux versus runoff (kg km^{-2} $storm^{-1}$)	Chloride flux p-value; R^2
DBB	$0.1\ RO^{1.0}$	NA	$11.2\ RO^{0.96}$	$p < 0.001$; $R^2 = 0.99$
CLGB		$0.51\ RO^{0.79}$	$213.8\ RO^{0.78}$	$p < 0.001$; $R^2 = 0.94$
CHSB	$0.45\ RO^{0.57}$	$p = 0.01$; $R^2 = 0.88$	$33.1\ RO^{0.65}$	$p = 0.003$; $R^2 = 0.95$
BRDS	$0.20\ RO^{0.84}$	$p < 0.001$; $R^2 = 0.97$	$50.1\ RO^{0.6}$	$p < 0.001$; $R^2 = 0.75$
OMPD	$0.17\ RO^{1.13}$	$p < 0.001$; $R^2 = 0.96$	$60.3\ RO^{0.81}$	$p < 0.001$; $R^2 = 0.89$

5. Calculate the network-scale retention by comparing the predicted versus observed flux ratio at the basin mouth using the end member derived from each anthropogenic headwater. Network scale proportion of nitrate retained ® for storm size i was estimated as:

$$R_i = \left(N : Cl_{mouth\_pred\_i} - N : Cl_{mouth\_obs\_i}\right)/N : Cl_{mouth\_pred\_i} \quad (3)$$

Where, $N:Cl_{mouth\_obs\_i}$ is the observed nitrate to chloride flux ratio at OMPD for storm size I (step 1) and $N:Cl_{mouth\_pred\_i}$ is the expected flux at OMPD for storm size i (step 4).

6. Calculate the mean and standard error using the end members determined from each of the anthropogenic headwaters (n = 3). Application of each end member assumes that a particular site is representative of the anthropogenic land covers in the watershed as a whole. The standard error provides a measure of uncertainty in the retention estimate assuming the anthropogenic end members from the three headwater sites are a sample of the true loading. We do not make any assumptions about how nitrate and chloride are coupled. The nitrate to chloride ratio is entirely empirical, based on the storm scale flux measurements and can vary for each headwater based on the particular land cover (e.g., the more urban CLGB has lower N:Cl ratio than the more agricultural CHSB). Further, the loading ratios change with storm size because they are based on independent solute flux versus storm size relationship developed for each site.

3. Results

3.1. Concentrations

Across all sites, median annual concentrations of both NO_3^--N and Cl^- using low frequency grab samples increased with increasing anthropogenic land cover (Figure 3). Median NO_3^--N increased from <0.2 to ~1.0 mg L^{-1}, whereas Cl^- increased from <30 to 300 mg L^{-1} from anthropogenic land cover = 20% to nearly 100% (Figure 3). However, because of the variability caused by the three most agricultural sites (CHSB, CLGB.ag, MOORE, all with >20% agricultural land cover), the relationship between median NO_3^--N and land cover was not significant (p= 0.3). If these three agricultural catchments are excluded, NO_3^--N versus anthropogenic land cover becomes significant ($R^2 = 0.93$, $p < 0.001$). In contrast, median Cl^- concentration was highly related to anthropogenic land cover across all sites ($R^2 = 0.83$, $p = < 0.001$). Temporal variability expressed as interquartile range (IQR) also tended to increase with increasing anthropogenic land cover. The increase in NO_3^--N IQR was not statistically significant when all sites are included (n = 10, $R^2 = 0.11$, $p = 0.348$), but was if CHSB was excluded (n = 9, $R^2 = 0.86$, $p < 0.001$) (supporting information Table S1). The increase in Cl^- IQR was statistically significant for all sites (n = 10, $R^2 = 0.59$, $p = 0.01$).

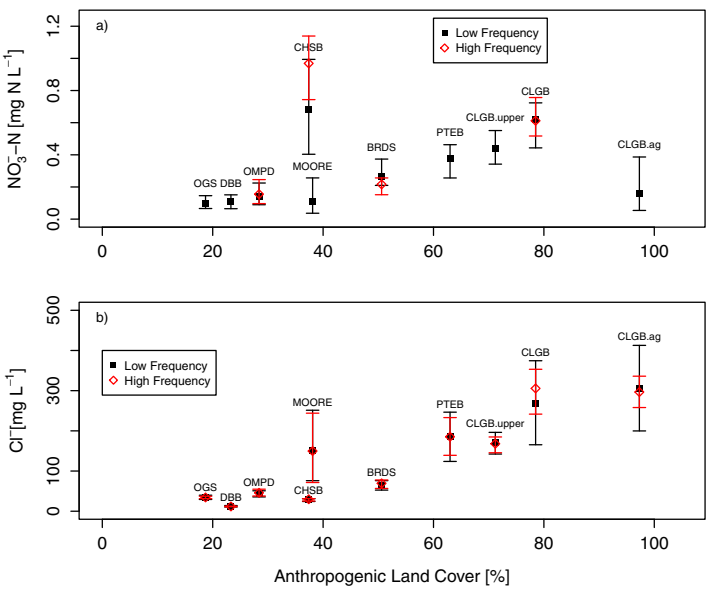

Figure 3. Annual median concentrations of (a) nitrate and (b) chloride versus anthropogenic land cover (developed + agriculture) in each catchment. Abbreviations refer to sites summarized in Table 1. Error bars represent the interquartile range. Solid black points are using infrequent grab samples, while open, diamond points use high frequency nitrate or conductivity sensors. Note that some symbols overlap one another because of minimal differences, especially in panel b). High frequency estimates are from the deployed period only, which for nitrate misses much of the winter, because the SUNA was not deployed.

Median NO_3^--N and Cl^- concentration did not differ significantly using the high frequency sensors compared to the weekly grab samples collected during the period of nitrate sensor deployment (paired t-test: p = 0.34, df = 4). However, in CHSB median NO_3^--N from the sensor was ~30% higher, because of the large number of high base flow concentrations that the sensor measured during the deployment (supporting information Figure S4). Median Cl^- was not significantly different when using infrequent grabs as compared to the high frequency conductivity probes (paired t-test: p = 0.43, df = 10) (Figure 3b).

Nitrate concentrations showed seasonality across all the sites. Highest concentrations occurred in winter, declined during spring, slightly increased in summer (particularly sites with high anthropogenic land cover), declined in fall, and increased again in winter (supporting information Figure S4). The exception was the agricultural site, CHSB, which showed highest NO_3^--N concentrations during low flow summer periods. Infrequent grab samples showed similar seasonal patterns.

The high frequency time series of NO_3^--N revealed considerable shorter-term variability, mainly due to storms, and secondarily diel variation (supporting information Figure S4).

In contrast, Cl^- generally showed less seasonality, but also considerable short-term variability in association with storm events in the more anthropogenic catchments (supporting information Figure S5). Spikes in Cl^- occurred during winter periods associated with snow events (increasing in CLGB from ~300–400 to >800 mg L^{-1}), whereas dilutions occurred during summer storms (in CLGB from ~300–400 to <100 mg L^{-1}). There is a seasonal chloride dilution in spring in two of the more anthropogenic sites (CLGB, CHSB), which we attribute to increased flows associated with snowmelt (supporting information Figures S3 and S5).

Synoptic surveys indicate that generally NO_3^--N and Cl^- concentrations at the high frequency sensor sites are similar to other headwaters in the watershed (supporting information Figures S6 and S7). While CHSB has high NO_3^--N concentrations given the total amount of anthropogenic land cover, several other headwater catchments with similar proportions of these land covers show similar concentrations during the lower flow seasons (supporting information Figures S6a, S6b, and S6d). For additional discussion of the synoptic survey results see supporting information Text S3.

3.2. Concentration Versus Discharge Relationships

Nitrate versus discharge relationships using all high frequency NO_3^--N data indicated considerable variability (Figure 4). Grab samples generally clustered within the cloud of high frequency NO_3^--N points, and similar

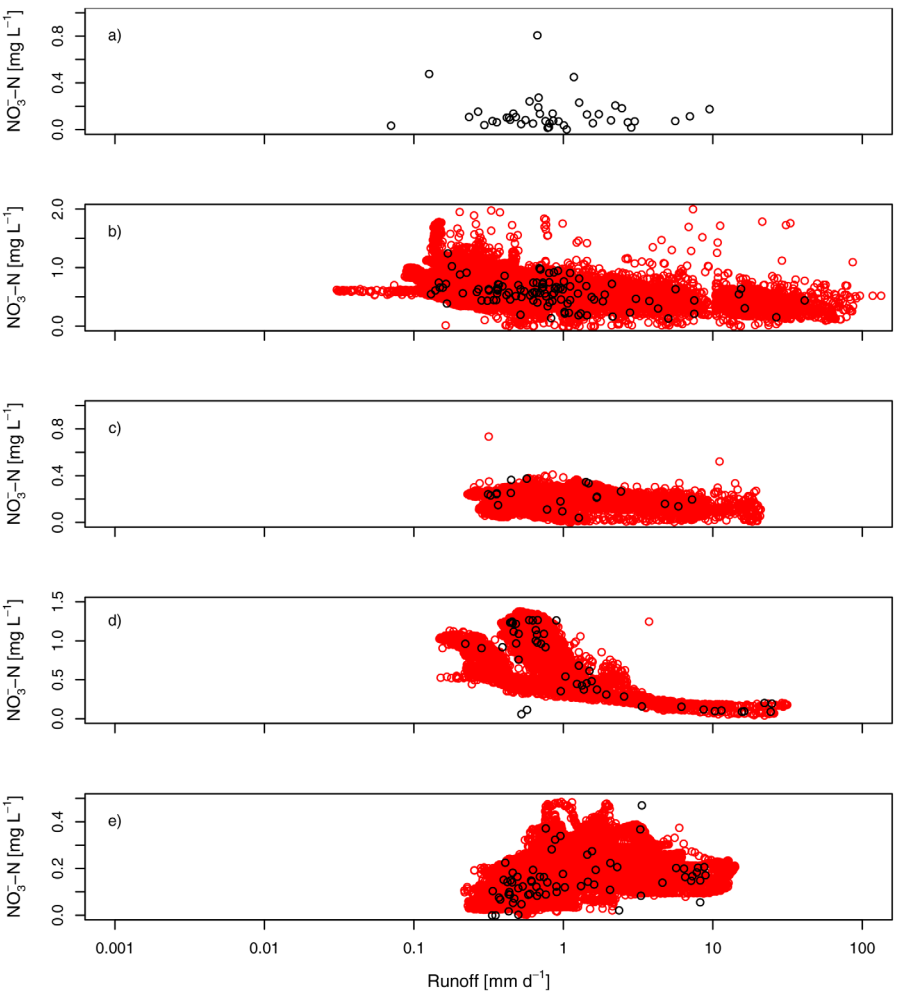

Figure 4. Instantaneous nitrate concentration versus runoff relationships for (a) DBB (forested headwater), (b) CLGB (urban), (c) BRDS (suburban), (d) CHSB (agriculture), and (e) OMPD, the mouth of the watershed. Red dots are from high frequency nitrate sensors, while black dots are from weekly grab samples during the period of sensor deployment.

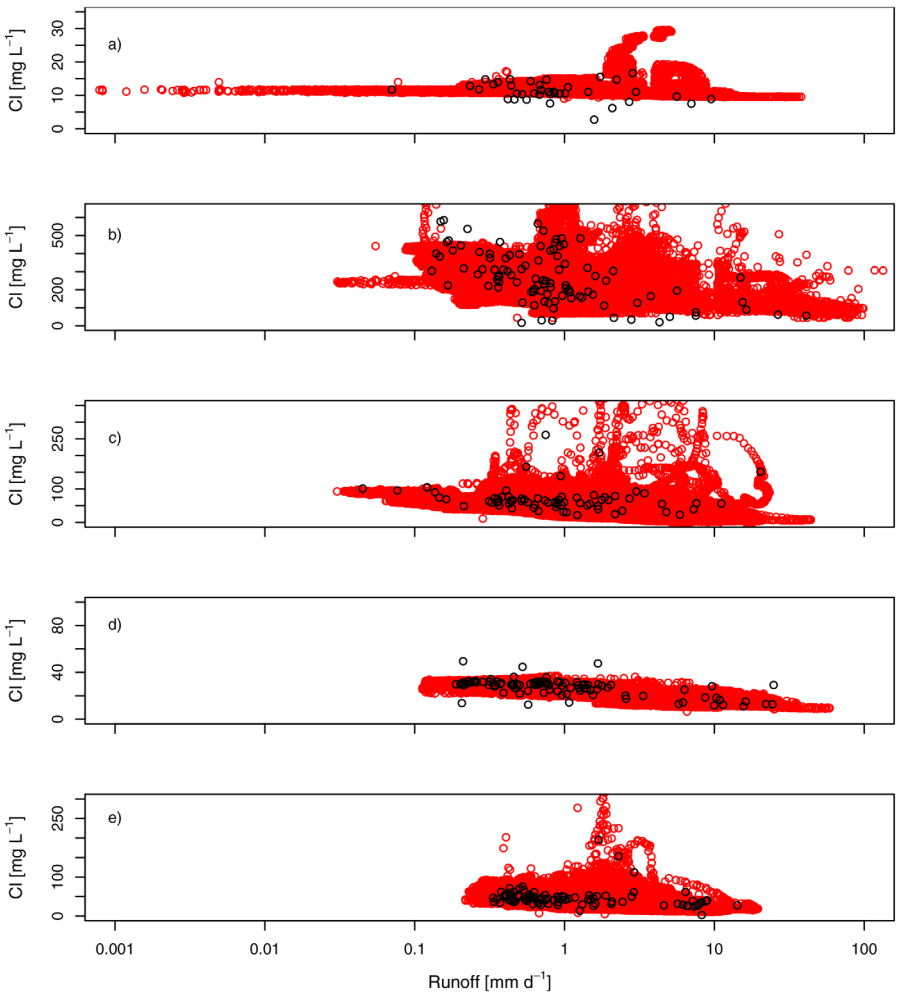

Figure 5. Instantaneous chloride concentration versus runoff relationships for (a) DBB (forested headwater, (b) CLGB (urban), (c) BRDS (suburban), (d) CHSB (agriculture), and (e) OMPD, the mouth of the watershed. Red dots are from high frequency conductivity sensors (using chloride versus specific conductance proxy), while black dots are from weekly grab samples during the period of sensor deployment.

trends with changes in discharge. Dilution of NO_3^--N concentrations with increasing flow was evident at the more anthropogenic headwater watersheds (using high frequency nitrate data: CLGB: slope = −0.0088, $p < 0.001$; CHSB: slope = −0.081, $p < 0.001$, BRDS: slope = −0.0074, $p < 0.001$), whereas chemostatic conditions were evident in forested headwaters (DBB grabs: slope = −0.0035, $p = 0.77$). In contrast, NO_3^--N at the basin mouth increased with flow (OMPD: slope = 0.0138, $p < 0.001$). At OMPD, the highest NO_3^--N concentrations occurred at intermediate runoff (reaching 0.5 mg N L^{-1}), but continued to remain elevated at the highest flows. Thus, despite a dilution effect in streams draining the major NO_3^--N source areas within the watershed, the river mouth showed increasing concentrations during storms, suggesting a potential signal of river network retention or transformations during low flows (see discussion). In contrast, Cl^- diluted with increasing flow at all sites (slopes of −1.66 to −5.66, $p < 0.001$) except DBB where the very low Cl^- increased slightly with flow (slope = 0.15, $p < 0.001$) (Figure 5). Intermediate flows periodically showed elevated Cl^- at all sites but CHSB. The declining slopes of Cl^- versus flow at OMPD (while nitrate increased) further supports the possibility that low nitrate concentrations during low flows occurs because of retention by the river network.

3.3. Annual Fluxes

Runoff coefficients tended to increase with greater anthropogenic land cover, with a stronger relationship in 2013–2014 ($R^2 = 0.54$) than in 2014–2015 ($R^2 = 0.18$), though neither was statistically significant (Figure

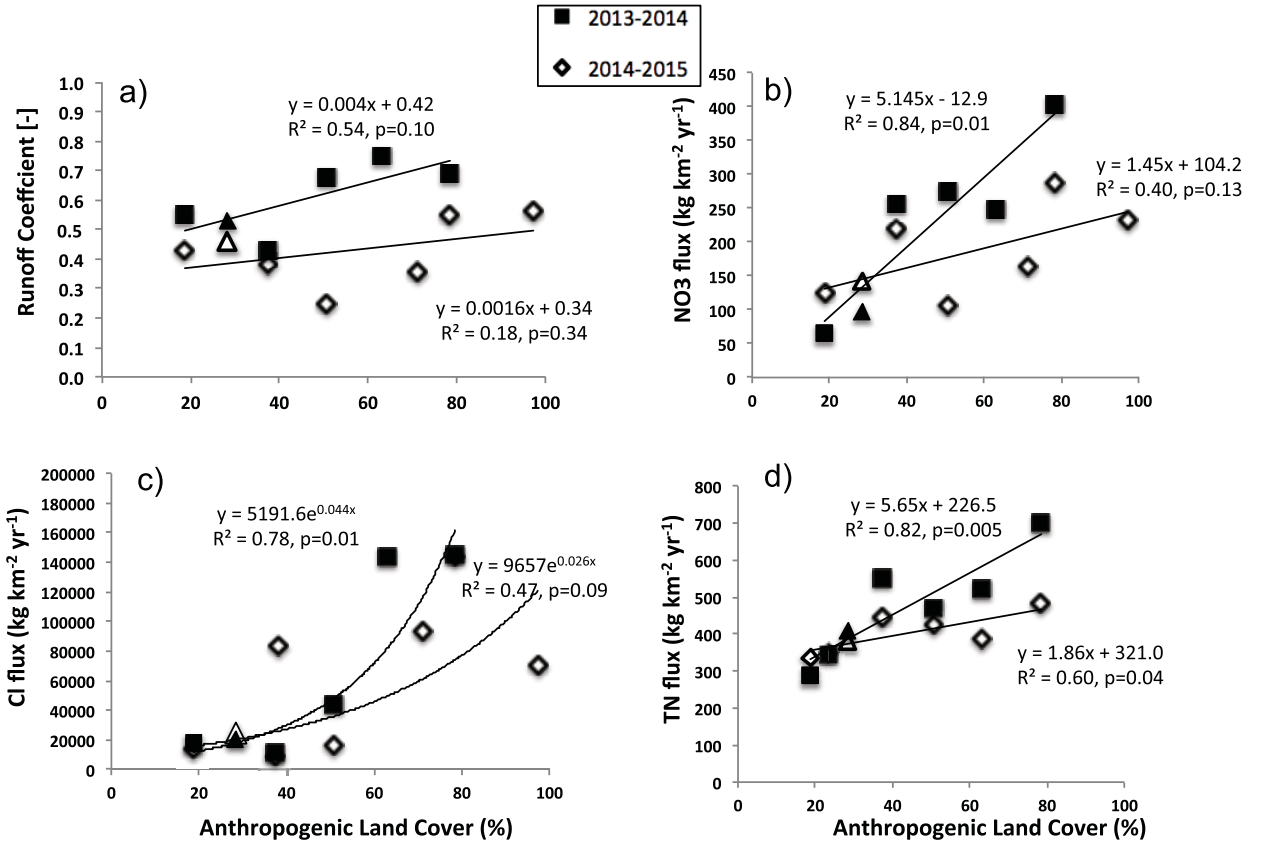

Figure 6. Annual (a) runoff coefficients, (b) nitrate areal fluxes, (c) chloride areal fluxes, and (d) total nitrogen areal fluxes versus anthropogenic land cover for 2013–2014 and for 2014–2015. The open and closed triangle symbol in each plot represents the basin mouth at OMPD. These fluxes were estimated using local discharge and local grab samples in order to compare across all the sites.

6a). Although precipitation was similarly below average in both years, runoff coefficients were lower across all sites in 2014–2015 (25% lower on average, more so in the more developed catchments). Both years had lower than average precipitation (~1,070 mm yr^{-1} versus the long-term average of 1,280 mm yr^{-1}), but while runoff at the USGS gage was close to average during the first period (586 mm yr^{-1} versus long-term average of 582 mm yr^{-1}), it was well below average during the second period (448 mm yr^{-1}), presumably due to compounding effects of two consecutive years of dryer than normal conditions.

Annual NO_3^--N fluxes using FWMC from grabs and runoff from local measurements (Method 2) increased sharply with increasing anthropogenic land cover in the catchment (250% in 2014–2015 versus 700% in 2013–2014) (Figure 6b). Fluxes and slopes were greater during 2013–2014 (slope = 5.15) than 2014–2015 (slope = 1.45), consistent with differences in runoff coefficients. Annual TN fluxes showed similar relationships (increasing 50% in 2014–2015 versus 250% in 2013–2014) (Figure 6d). Fluxes of NO_3^--N or TN at OMPD did not differ appreciably from the overall relationships. Annual Cl^- fluxes increased steeply with percent anthropogenic land cover (Figure 6c). As with NO_3^--N, annual Cl^- fluxes in impacted watersheds tended to be lower during 2014–2015 than 2013–2014.

3.4. Comparison of NO_3-N Fluxes With and Without Sensors

Annual NO_3^--N fluxes in each year were not consistently affected by the method of obtaining flow estimates from the USGS gage (Method 1) versus local Q measurements (Method 2), even in the flashy headwater sites (Figure 7a). In 2013, fluxes were from 5% lower to 28% higher using local flow compared to USGS-gage derived flow, while in 2014 differences ranged from −40 to +70%. Overall, the difference using the two methods was not statistically significant (paired t-test p= 0.18). We could only compare fluxes between low frequency grabs (Method 2) and high frequency NO_3^--N measurements (Method 3) during the June to December period when the SUNA was deployed for a significant percent of the time. Over this period, flux

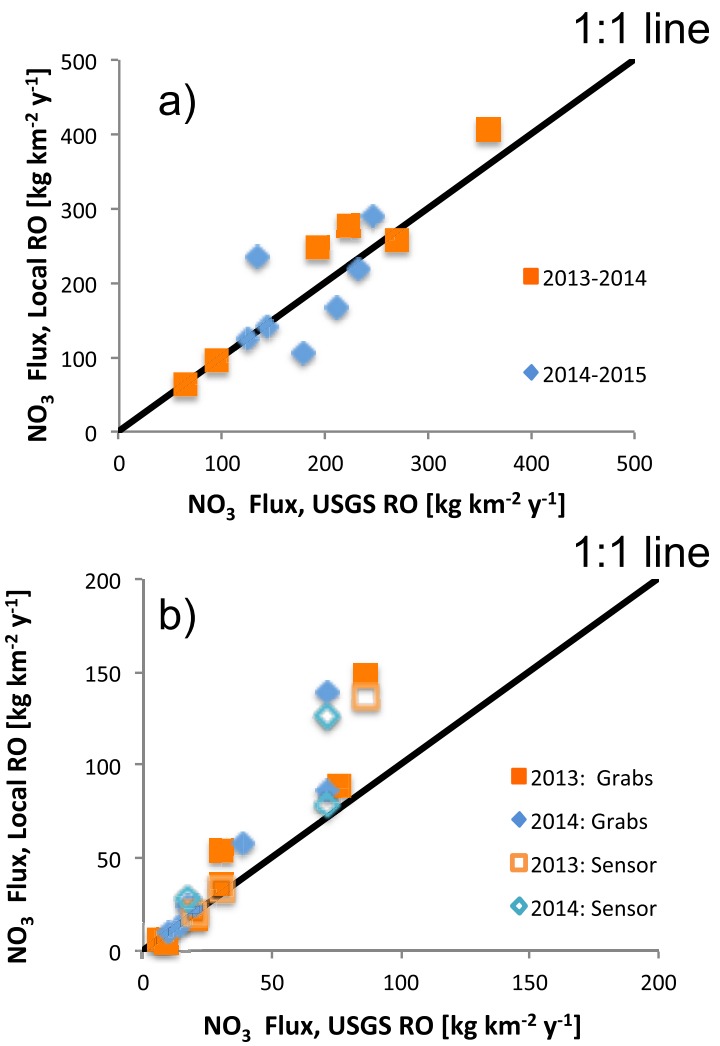

Figure 7. Comparison of nitrate fluxes (a) annually and (b) lower flow season only (June–December) using flow weighted mean concentrations from grab samples and flow estimates derived from the USGS gage (Method 1) or from local flow estimates (Method 2) for each year. For the low flow period in Figure 7b also shown are estimates of flux for the period using high frequency nitrate sensors and local flow (Method 3) (open symbols).

estimates using Method 2 were higher than when using Method 1 (paired t-test $p = 0.015$), with the biggest difference occurring in the highest N-flux sites in both years (Figure 7b). However, Method 3 and Method 2 estimates were small and not statistically significant (paired t-test $p = 0.09$).

3.5. Storm Event Scale Fluxes and Concentrations

In the headwaters, storm event scale NO_3^--N fluxes increased with size of the flow event (Figure 8a and Table 2; CLGB: log-log slope= 0.79, $p < 0.001$; BRDS: log-log slope= 0.84, $p < 0.001$; CHSB: log-log slope= 0.57, $p = 0.011$). At the basin mouth at OMPD, NO_3^--N fluxes increased more steeply than in the headwaters (log-log slope = 1.13, $p < 0.001$). ANCOVA indicated that the slopes of the NO_3^--N versus storm runoff relationships were significantly different among the sites ($p = 0.0004$), with Tukeys post hoc test indicating OMPD slopes were significantly higher than those from CLGB ($p < 0.001$) and CHSB ($p = 0.002$), but not for BRDS ($p = 0.56$). Flow-weighted mean NO_3^--N concentrations in the headwaters declined with storm size (Figure 8b; CLGB: slope= −0.014, $p < 0.001$; BRDS: slope= −0.004, $p = 0.017$; CHSB: slope= −0.017, $p = 0.26$) but increased at OMPD (slope = 0.005; $p = 0.29$).

Storm event scale Cl^-, fluxes showed similar slopes for all sites (OMPD: log-log slope = 0.81; Headwaters: log-log slope = 0.6 - 0.96; Figure 8c and Table 2), while FWM Cl^- diluted strongly at all sites but the forested site (Figure 8d; $p < 0.05$ for all sites but the forested DBB where $p = 0.60$). The ANCOVA indicated that slopes were significantly different among sites ($p = 0.007$), with Tukeys post hoc test indicating that OMPD was not different from CHSB ($p = 0.1$) and BRDS ($p = 0.7$), but was different from both CLGB and DBB ($p < 0.001$).

3.6. Network Scale Retention Across Storm Flow Conditions

Net NO_3^--N retention at network scales estimated using the end member mixing analysis (section 2.8) was highest during small storm events (65% for storm runoff ~ 0.1 mm storm^{-1}), but declined rapidly with increasing storm size (0% net retention for ~10 mm storm^{-1}) (Figure 9). The largest storm flow we measured at OMPD was 9.8 mm storm^{-1}. However, extrapolation of OMPD fluxes beyond 10 mm storm^{-1} using the flux regressions (Figure 8a) would result in OMPD becoming a net source of nitrate for storms > 10 mm storm^{-1} (Figure 9). Given the variability in loading estimates derived from the three anthropogenic-influenced headwaters, NO_3^--N retention during small storms could range from 51 to 78%, and with increasing storm size may drop to 0% for events as small as ~3 mm storm^{-1} using the low error bound (Figure 9). The decline in network NO_3^--N retention occurs because storm event scale NO_3^--N fluxes increase at a faster rate with increasing storm size at OMPD than in the headwaters, while Cl^- increases at a similar rate (Figure 8). The effect is summarized in the conceptual model presented in Figure 10.

4. Discussion

High frequency nutrient sensors offer a new window into watershed biogeochemical dynamics (Rode et al., 2016). The analysis reported here emphasizes a measurement approach using such sensors geared toward improved understanding of nitrate retention within river networks. The approach relies on loading estimates across storm sizes (measured in headwaters) and simultaneous estimates of exports (measured at the basin mouth) at storm event scales. While sensors in this study did not provide significant improvement of estimates of long-term fluxes, even in more flashy, high N headwater catchments, the ability to develop robust storm event relationships is necessary for application of this approach, possible only with high

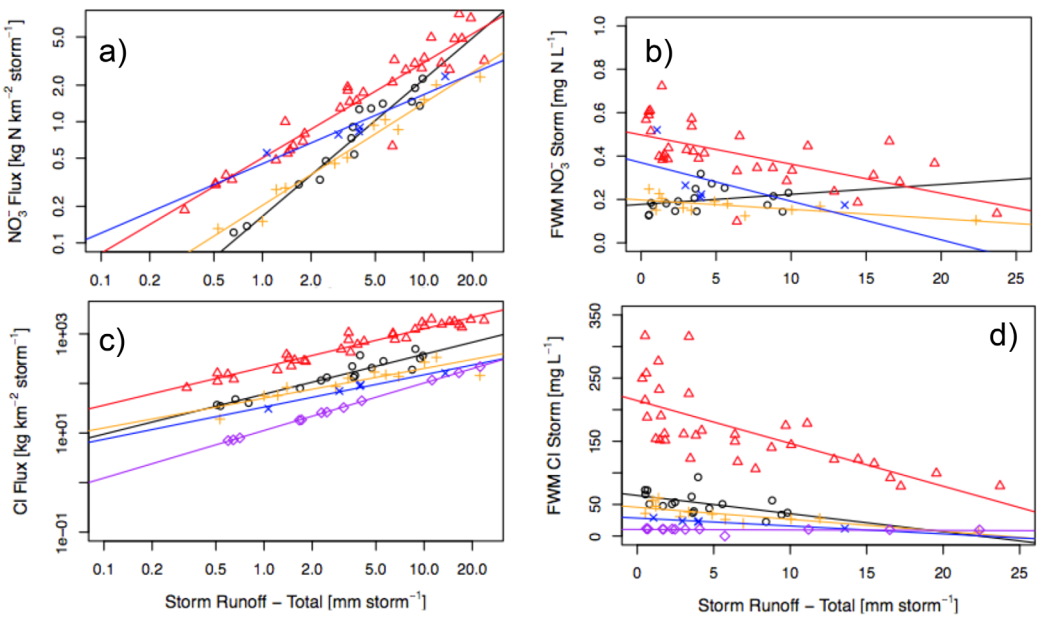

Figure 8. Storm event scale estimates of (a) nitrate fluxes, (b) flow-weighted mean nitrate concentration, (c) chloride fluxes, and (d) flow-weighted mean chloride concentrations for CLGB (red triangle), BRDS (orange +), CHSB (blue X), DBB (purple diamond, chloride only), and OMPD (black circle). Storm fluxes for nitrate in DBB are assumed to have a constant concentration and therefore increase with runoff with a slope =1. Note that the axes for flux in Figures 8a and 8c are log scale, while for concentrations in Figures 8b and 8d are linear scale.

frequency measurements. We note that the Oyster River watershed is relatively little impacted compared to intensive agricultural watersheds where concentrations are much higher and concentration/discharge relationships may differ from those reported here. However, the approach can be readily applied to these more anthropogenic-dominated watersheds as well.

4.1. Utility of High Frequency Sensors for Longer Term Flux Estimates

The high frequency nitrate sensor data did not improve estimates of nitrate fluxes over seasonal time scales compared to less frequent grab samples only. Minimal differences in longer term nitrate fluxes using the two approaches were also found in larger rivers (Carey et al., 2014; Pellerin et al., 2014). The small difference using high frequency nitrate sensors is surprising in the small, developed headwater streams studied here, given their flashy hydrology coupled with variable storm event concentrations. Typically, storm chemistry is undersampled with traditional weekly grab sampling, because storm flows are relatively short and therefore missed by infrequent sampling. One explanation is that discharge goes up orders of magnitude during storms, while concentrations generally change by <1 order of magnitude (Figure 4). Further, flow-weighted mean nitrate concentrations at storm event scales, while declining more for larger storms, typically varied by less than a factor of three, even in the most urban watershed (CLGB) (Figure 8b). Thus, if infrequent grab sampling encompasses a reasonable distribution in the C versus Q space (Figure 4), flux estimates will be reasonable. In addition, use of flow data from larger, gaged rivers to estimate fluxes in flashy headwater would tend to underestimate high discharges (compare runoff ranges across sites in Figure 4), while infrequent grabs would tend to overestimate concentrations during diluting storms. The net effect is that these errors offset each other. Nevertheless, to better understand shorter time

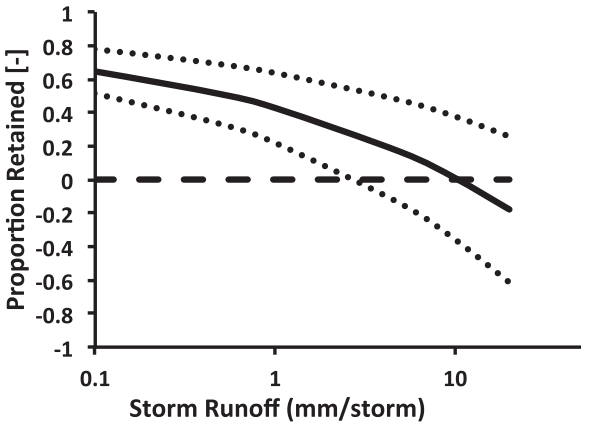

Figure 9. Proportion of nitrate entering the river network that is retained versus storm size (as cumulative storm runoff) using the two end-member mixing analysis. Dotted lines represent plus or minus the standard error of the mean from the three developed catchments that were used to infer the anthropogenic end member. Values for storms greater than 10 mm storm^{-1} are extrapolated.

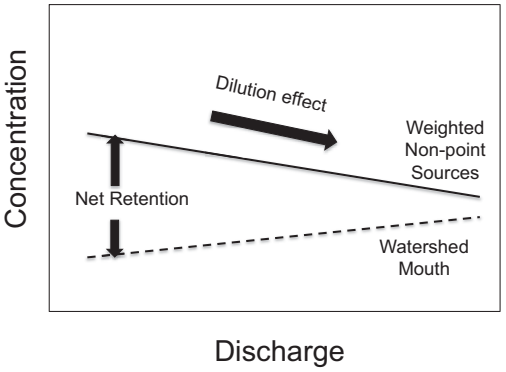

Figure 10. Conceptual model of the interaction between nitrate concentration versus storm event in source areas (weighted based on the proportion of anthropogenic and natural land cover) and at the watershed mouth, accounting for dilution and retention across flow conditions.

scale flux variability among different watersheds, including storm event scales, high frequency sensors provide improved insight (Carey et al., 2014, Figure 8).

4.2. Comparison of Oyster River Watershed Nutrient Fluxes With Other Watersheds

Nutrient fluxes are driven by both water runoff and concentrations. Both the proportion of precipitation running off terrestrial surfaces (runoff coefficients) and nutrient concentrations increased with anthropogenic land covers as seen in many other headwater catchments with similar land cover proportions (Burges et al., 1998; Goodridge & Melack, 2003; Wollheim et al., 2005). As a result, fluxes increased considerably in more anthropogenic catchments. TN fluxes peaked at 500–700 kg N km^{-2} yr^{-1} in the most urban headwater catchments, which is similar to the 380–580 kg N km^{-2} yr^{-1} reported for a suburban headwater near Boston (Wollheim et al., 2005) and to the 450–720 kg N km^{-2} yr^{-1} for suburban sites near Baltimore (Groffman et al., 2004). However, our estimates of TN flux from forested catchments are higher (intercepts to zero anthropogenic land cover = 200–300 kg km^{-2} yr^{-1}) compared to those reported in previous studies of urban gradients (< 100 kg km^{-2} yr^{-1}; Groffman et al., 2004; Wollheim et al., 2005). Runoff coefficients were higher in headwaters of the Oyster River watershed, including in the less developed catchments (40–50%), than in the suburban Boston catchments (20–30%), possibly related to differences in geology or slope. Mean annual nitrate concentrations were somewhat lower in this study (up to 0.7 mg L^{-1}) compared to the more developed suburban Boston catchments (up to 1.1 mg L^{-1}; Wollheim et al., 2005) or urban Baltimore catchments (up to 4 mg L^{-1}; Groffman et al., 2004). Thus, although fluxes from the most anthropogenic headwater catchments in the Oyster are similar to urban catchments further south, they are more dilute.

Annual N export from the mouth of the Oyster River watershed was around 400 kg km^{-2} yr^{-1}, similar to the Ipswich River watershed where TN of 438 kg km^{-2} yr^{-1} was exported on average between 2000 and 2007 (Wollheim et al., 2013). The Ipswich is more impacted (37% anthropogenic in the Ipswich compared to 28% in the Oyster), larger (400 km^2 versus 50 km^2), and impact is skewed more toward the headwaters (Skewness index = 1.09 compared to 0.77 for the Oyster). Since skewness index is correlated with network scale nutrient removal (Mineau et al., 2015), this suggests the relatively high flux from the less developed Oyster R. watershed compared to the Ipswich R. watershed may in part be due to less aquatic retention in the former. Annual fluxes from the Oyster R. watershed are very close to that expected based on the regression line of TN flux versus anthropogenic land cover derived from the headwaters, suggesting relatively little N retention over annual time scales.

4.3. Storm Event Dynamics

Storm size is a major determinant of variability in fluxes of both nitrate and chloride. In headwaters, the rate of nitrate flux increase versus runoff is < 1 because nitrate concentrations dilute during storms, while at the mouth, the rate of flux increase versus runoff is > 1 because concentrations increase during storms. While nitrate in all the developed headwaters in this study diluted during storms, this is not always the pattern seen in other headwaters. Goodridge and Melack (2003) found in watersheds ranging in drainage area from 8 km^2 to 40 km^2 that nitrate concentrations could increase, decrease, or remain constant with increasing flow, while conductivity only diluted (similar to this study). Similarly, Wollheim et al. (2005) found that nitrate in an urban headwater catchment tended to increase during higher flows. However, none of these studies looked at storm event scale FWMCs that is now possible with high frequency nitrate sensors.

Increasing nitrate concentrations during storms, such as those found at the mouth of the Oyster, have also been found in larger rivers. Pellerin et al. (2014) found in the Mississippi R. that mean daily nitrate concentrations increase with flow. The pattern in the Mississippi (Pellerin et al., 2014, Figure 3) indicates lowest concentrations at low flows, highest concentrations at intermediate flows, and high concentrations at the highest flows. This pattern is remarkably similar to that seen in this study in the much smaller Oyster River watershed. The extremely low nitrate concentrations at OMPD during the lowest flows are also consistent with a nitrate retention signal. The pattern of increasing nitrate concentration at OMPD with increasing

storm size, despite greater dilution in the headwaters with larger storms, can be explained by nitrate retention dynamics

4.4. Evidence of Nitrate Retention by the River Network

Sensors deployed hierarchically within a single river network, along with an end-member mixing analysis at storm event scales, provides a new approach for estimating nitrate retention at network scales. The decline in nitrate retention with increasing storm size we quantified is consistent with river network scale models of N removal across flow conditions (Wollheim et al., 2008a) as well as conceptual models of effective discharge of nutrient removal (Doyle, 2005). Miller et al. (2016) used a single station technique that inferred river network nitrate retention in a large watershed (the Potomac) from a single station time series alone, based on the assumption that concentrations during winter are indicative of inputs throughout the year. Divergence from winter concentrations during summers was then assumed to reflect network retention. The end member approach used here avoids the assumption that concentrations loaded to the network are constant over time by looking at the storm event response of nitrate to chloride flux ratios for different flow conditions in source areas coupled with basin mouth responses. The effect of seasonality, particularly during cold winters, could not be evaluated as part of this study because our nested sensors were not deployed in the headwater during the winter, when loading concentrations are generally highest. However, the assumption of conservative mixing during winters as assumed by Miller et al. (2016) could be tested if sensors were deployed in winter.

River network retention can explain contrasting storm event response patterns in nitrate concentrations at nested locations within the same watershed. The dilution of nitrate during storms in the headwaters can lead to increasing nitrate at the mouth if retention occurs under lower flow conditions (Figure 10). The forested catchment showed no evidence of dilution during storms, although we did not have storm event scale estimates in this study. Previous studies in forested headwaters in the region and elsewhere also indicate little change in nitrate concentrations during storms (Price, 2014; Wollheim et al., 2005). Flushing of soils and shallow groundwater during storms, leading to higher storm concentrations in headwaters, was not evident in this watershed. For concentrations to increase at the basin mouth, the diluted concentrations of nonpoint sources integrated over the storm must still be higher than what occurs at base flow at the downstream location, as evident over the range of observed storms (compare OMPD and CLGB in Figure 8b). It is possible that nonpoint loading direct to higher order river reaches may have a different loading signal and storm response. However, most terrestrial runoff initially enters the river network into 1st or 2nd order streams (Alexander et al., 2008; Wollheim et al., 2006), so measurements in headwater streams should be representative of most nonpoint inputs.

Declining storm event concentrations in headwaters coupled with a rise in main stem concentrations during storms could also occur if the intensive headwaters were not representative. However, the synoptic surveys across seasons suggest that the intensive headwaters are representative of the watershed as a whole (supporting information Figures S6 and S7). Further, for chloride, both headwaters and basin mouth dilute during storms (slopes similar and < 1 for all sites, Figure 8c, d). Finally, the Oyster R. watershed does not have any point sources that could lead to increasing concentrations during storm events. Thus, we suggest that the pattern of increase during storms at the basin mouth is due to significant nitrate retention by the river network at low flow with retention declining as flows increase (Figure 10).

Our estimates of network-scale nitrate retention may be lower than actual because we are inferring inputs to the network using measurements in relatively large headwater catchments ($\sim$5km^2). Roughly 1 km of stream is located upstream of each of the headwater sample sites, which has additional potential for retention, especially under low flow conditions (Bernhardt et al., 2005a; Peterson et al., 2001). However, removal in headwaters declines rapidly with increasing flow, particularly for moderately reactive nutrients like nitrate (Raymond et al., 2016; Wollheim et al., 2006). Thus, we may be underestimating removal more so during low flows than during high flows when inferring inputs from relatively large headwater streams.

The synoptic surveys provide additional corroborating evidence that the network as a whole is retaining nitrate. Nitrate concentration at the basin mouth (OMPD) is generally in the cloud of points in nitrate versus land cover relationships as determined by its nested headwaters, which suggests little evidence of retention (supporting information Figure S6). However, anthropogenic land cover in the watershed as a whole is at the threshold ($\sim$20%) at which increases in concentration are more likely (supporting information Figure

S6). The highest concentrations tend to occur at these intermediate land cover levels (between 20 and 40% anthropogenic land cover). Yet, there is no indication of high nitrate at OMPD during the synoptic surveys suggesting source hotspots are attenuated by the time flows reach the mouth.

While nitrate retention by the river network is important under low flow conditions, the strength of this process declines at high flows and over annual time periods. Miller et al. (2016) estimated that the river network of Potomac R. watershed (29,950 km^2) retained 23% of annual nitrate loads, while two smaller watersheds within the Potomac (150–250 km^2) retained 11%. While we did not estimate annual retention, estimates of annual fluxes at the mouth of the Oyster did not show evidence of retention based on its land cover (Figure 6b). A reduction of 11% is unlikely to be easily detected at the annual scale given the uncertainties in annual flux estimates and spatial heterogeneity of loading. The river network of a coastal New England watershed draining suburban Boston was predicted to retain 16–33% of annual DIN loads, with runoffs < 0.4 mm d^{-1} retaining 43–71% and runoffs > 2 mm d^{-1} retaining 7–18% (Wollheim et al., 2008a). These model predictions are comparable to the storm event scale estimates we observed in the 50 km^2 Oyster River watershed (Figure 9).

Nitrate retained by the river network may be temporary and eventually exported from the watershed. For example, assimilated nitrogen may be remineralized as ammonium that is nitrified to nitrate, leached as dissolved organic nitrogen or resuspended or sloughed as particulate organic nitrogen (Peterson et al., 2001; Wollheim et al., 2001). Permanent nitrate removal may occur via denitrification or in long-term storage in lake or reservoir sediments (Mulholland et al., 2009; Seitzinger et al., 2006). This study suggests nitrate is retained during smaller storm events without addressing its subsequent fate. As a result, we cannot determine the degree of transformation to other N forms or the associated mechanism, e.g., assimilation followed by remineralization at a later date, or long-term sinks.

4.5. The Potential of Nested Sensor Networks for Management

Increased availability and affordability of high frequency sensors will allow more widespread deployment of sensor networks (Rode et al., 2016). Central to understanding the role of river networks in attenuating pollution fluxes (beyond just dilution) is having robust estimates of nonpoint inputs from land to river systems and simultaneous estimates of downstream fluxes. Knowledge of how storm events influence pollutant inputs is critical, particularly as climate patterns alter the intensity and frequency of storms. More high frequency monitoring in headwaters of different land cover is essential. High frequency measurements in smaller systems remain relatively uncommon because agencies such as USGS generally focus on larger systems.

Due to the nature of river network geomorphology, there are many more small headwater streams than larger rivers. The question of representativeness is therefore central. More sensors deployed in a variety of headwaters could aid in the development of rules that generalize storm event responses of nonpoint inputs as a function of land cover and management, geology, spatial heterogeneity, season, etc. Such rules will be very useful for modeling river network scale processes and will also help identify and prioritize mitigation opportunities. Improved estimates of nonpoint inputs at storm event scales will provide solid baseline estimates that can be used to quantify whether mitigation actions actually work (Bernhardt et al., 2005b). For example, the nitrate flux versus storm size relationship in impacted CLGB provides a robust storm response function under current conditions that can be used to determine whether best management practices impact fluxes, through changes in either the slope or intercept of the response function.

Nested sensor networks in multiple watersheds can be used to generalize the importance of retention by river networks, which could also inform management priorities. The river network retention curve we estimated (Figure 9) may differ considerably across watersheds depending on context. Major factors influencing network scale removal include the size of the watershed (total length of larger rivers) and channel morphology (Raymond et al., 2016), the degree of enrichment (Mulholland et al., 2008), the distribution of inputs relative to the network (skewness index; Mineau et al., 2015), the distribution of flow conditions and storm responses (Doyle, 2005; Wollheim et al., 2008a), and the abundance of fluvial wetlands, ponds, lakes, and reservoirs (Wollheim et al., 2008b). Knowledge of these curves from a variety of watersheds could help identify and prioritize which sources should be mitigated first. For example, land covers associated with nonpoint N sources in the Oyster watershed are distributed relatively near the mouth (skewness index $= 0.77$) so have relatively little opportunity to be removed by network processes, contributing to the magnitude and shape of the curve in Figure 9. In contrast, in the Ipswich River watershed, sources are distributed

further upstream (skewness index = 1.09; Mineau et al., 2015) which allows greater retention and possible different retention curves. In this case, to optimize management with respect nitrate fluxes to the coastal zone regionally, N mitigation in the Oyster River should be targeted first.

Urban storm water is often managed through creation of detention ponds (Walsh et al., 2012). These reduce peak flows and encourage infiltration to ground water, which slowly enters the river system. Even though storms dilute nitrate concentrations in the headwaters, nitrate fluxes are still elevated during storms, during which a large proportion of annual inputs to river networks can occur (Wollheim et al., 2008a). Detention ponds release nitrate more slowly to river systems, and shift inputs from high flow periods when network retention is reduced, to lower flow periods when network retention is much higher. Nonpoint mitigation strategies should consider river network capacity to remove nitrate across flow conditions.

This study was funded in part because of the prospect of implementing an integrated wastewater/storm water permit for the Town of Durham, NH, which comprises much of the developed sections of the Oyster River watershed. The Oyster River drains to the Great Bay estuary, NH, which has been classified by the U.S. EPA as a N-impaired estuary. Wastewater treatment plants (WWTP) run by various towns in the Great Bay watershed are already being upgraded so they release less N. Durham and UNH were exploring whether nonpoint source reduction could be used to avoid further WWTP upgrades. They funded this nested sensor network to provide improved estimates of nonpoint loading and exports. The impact of nonpoint mitigation activities could then be quantified using high frequency sensor measurements. It is likely that to clearly detect improvements, storm event scale measurements will be more sensitive than annual flux estimates. Unfortunately, progress toward this innovative approach of combining waste and storm water permits to optimize watershed scale management has been abandoned for this particular watershed. Nevertheless, studies should be conducted to evaluate whether integrated approaches can lead to improved water quality. This study demonstrates the use of networks of nested high frequency sensors that will help with such evaluations.

5. Conclusion

A revolution in environmental sensing of water resources is underway, which will lead to improved understanding of water quality phenomena and their management. *In situ* sensor-based observations open new worlds of empirical insight (Cohen et al., 2013; Heffernan & Cohen, 2010; Rode et al., 2016). Because river networks are hierarchically organized and highly connected by flows, understanding patterns at any one location requires knowledge of all terrestrial inputs and aquatic transformations upstream. The nested monitoring approach presented here is a step toward improved understanding of whole network dynamics. The key characteristics of a nested monitoring approach are: (1) High frequency measurements of reactive and conservative solutes in representative headwaters (i.e., major land cover types) to isolate the terrestrial loading signal. Robust flow measurements are as important as high frequency biogeochemical sensing. (2) Simultaneous high frequency measurements at the basin mouth which provides the combined terrestrial loading and aquatic transformation signal. Comparison with headwater inputs can be used to infer aquatic transformation alone. If present, point source inputs to larger rivers would also need to be quantified. (3) Spatially extensive but infrequent measurements to place the high frequency locations in context, i.e., to determine the representativeness of the high frequency sites. While we deployed such a network for nitrate, such measurements can be expanded to other nutrients (e.g., phosphorus), organic carbon, sediments and pathogens for which direct sensor measurements or their proxies can be determined, to in each case evaluate network-scale retention. A nested approach in different types of watershed will provide insight into sources and sinks at network scales that can be used to identify the emergent properties of whole river networks as well as management priorities.

Acknowledgments
This research was funded by NSF EPSCoR Ecosystems and Society (#EPS 1101245) and the Town of Durham/UNH facilities. Partial funding was provided by the National Sea Grant College Program of the U.S. Department of Commerce's National Oceanic and Atmospheric Administration grant NA10OAR4170082 to the N.H. Sea Grant College Program, and by the New Hampshire Agricultural Experiment Station (NHAES) through USDA National Institute of Food and Agriculture Hatch Project 0225006. This is NHAES Scientific Contribution Number 2713. We also thank Rick Kiah at USGS (NH/VT) for constructing the rating curves, the Oyster River Watershed Association for helping to collect synoptic samples, the Water Quality Analysis Lab at UNH for running the grab sample chemistry, and Dave Cedarholm for sparking the interactions with the Town of Durham and UNH Facilities. The data are available at the University of New Hampshire Data Discovery Center: "https://ddc.unh.edu/ddc_data/variables/list/"with project name "High Intensity Aquatic Network: Oyster River Watershed.

References

Alexander, R. B., Smith, R. A., & Schwarz, G. E. (2000). Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. *Nature*, *403*, 758–761.

Alexander, R. B., Smith, R. A., Schwarz, G. E., Boyer, E. W., Nolan, J. V., & Brakebill, J. W. (2008). Differences in phosphorus and nitrogen delivery to the gulf of Mexico from the Mississippi river basin. *Environmental Science & Technology*, *42*(3), 822–830.

Bende-Michl, U., Verburg, K., & Cresswell, H. P. (2013). High-frequency nutrient monitoring to infer seasonal patterns in catchment source availability, mobilisation and delivery. *Environmental Monitoring and Assessment*, *185*, 9191–9219. https://doi.org/10.1007/s10661-10013-13246-10668

Bernhardt, E. S. Likens, G. E., Hall, R. O., Buso, D. C., Fisher, S. G., Burton, T. M., . . . Lowe, W. H. (2005a). Can't see the forest for the stream? The capacity of instream processing to modify terrestrial nitrogen exports. *BioScience*, *52*, 219–230.

Bernhardt, E. S., Palmer1, M. A., Allan, J. D., Alexander, G., Barnas, K., Brooks, S., . . . Sudduth, E. (2005b). Synthesizing U.S. river restoration efforts. *Science*, *308*, 636–637.

Burges, S. J., Wigmosta, M. S., & Meena, J. M. (1998). Hydrological effects of land-use change in a zero-order catchment. *Journal of Hydrological Engineering*, *3*(2), 86–97.

Carey, R., Wollheim, W. M., & Mulukutla, G. K. (2014). Characterizing storm-event nitrate fluxes in a fifth order suburbanizing watershed using *in situ* sensors. *Environmental Science Technology*, *48*, 7756–7765.

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.

Cohen, M. J., Kurz, M. J., Heffernan, J. B., Martin, J. B., Douglass, R. L., Foster, C. R., & Thomas, R. G. (2013). Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. *Ecological Monographs*, *83*, 155–176.

Doyle, M. W. (2005). Incorporating hydrologic variability into nutrient spiraling. *Journal of Geophysical Research*, *110*, GO1003. https://doi.org/10.1029/2005JG000015

Goodridge, B., & Melack, J. M. (2012). Land use control of stream nitrate concentrations in mountainous coastal California watersheds. *Journal of Geophysical Research: Biogeosciences*, *117*, G02005. https://doi.org/02010.01029/02011JG001833

Groffman, P., Law, N., Belt, K., Band, L., & Fisher, G. (2004). Nitrogen fluxes and retention in urban watershed ecosystems. *Ecosystems*, *7*, 393–403.

Heffernan, J. B., & Cohen, M. (2010). Direct and indirect coupling of primary production and diel nitrate dynamics in asubtropical spring-fed river. *Limnology and Oceanography Methods*, *55*, 677–688.

Helton, A. M. *et al*. (2011). Thinking outside the channel: Modeling nitrogen cycling in networked river ecosystems. *Frontiers in Ecology and the Environment*, *9*, 229–238. https://doi.org/10.1890/080211

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., . . . Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, *81*, 345–354.

Howarth, R. W., & Marino, R. (2006). Nitrogen as the limiting nutrient for eutrophication in coastal marine ecosystems: Evolving views over three decades. *Limnology and Oceanography Methods*, *51*, 364–376.

Kennedy, E. (1984). Discharge ratings at gaging stations. *United States Geological Survey Techniques for Water-Resources Investigations, Book 3*, *Chap*. 10.

Miller, M. P., Tesoriero, A. J., Capel, P. D., Pellerin, B. A., Hyer, K. E., & Burns, D. A. (2016). Quantifying watershed-scale groundwater loading and in-stream fate of nitrate using high-frequency water quality data. *Water Resources Research*, *52*, 330–347. https://doi.org/10.1002/2015WR017753

Mineau, M., Wollheim, W. M., & Stewart, R. (2015). An index to characterize the spatial distribution of land use within watersheds and implications for river network nutrient removal and export. *Geophysical Research Letters*, *42*, 6688–6695. https://doi.org/10.1002/2015GL064965

Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., . . . Thomas, S. M. (2008). Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature*, *452*(7184), 202–205.

Mulholland, P. J., Hall, R. O., Jr., Sobota, D. J., Dodds, W. K., Findlay, S. E. G., Grimm, N. B., . . . Thomasn, S. M. (2009). Nitrate removal in-stream ecosystems measured by (15)N addition experiments: Denitrification. *Limnology and Oceanography Methods*, *54*(3), 666–680.

Pellerin, B. A., Bergamaschi, B. A., Gilliom, R. J., Crawford, C. G., Saraceno, J., Frederick, C. P., . . . Murphy, J. C. (2014). Mississippi river nitrate loads from high frequency sensor measurements and regression-based load estimation. *Environmental Science Technology*, *48*, 12612–−12619. https://doi.org/10.1021/es504029c

Peterson, B. J., Wollheim, W. M., Mulholland, P. J., Webster, J. R., Meyer, J. L., Tank5, J. L., . . . Morrall, D. D. (2001). Control of nitrogen export from watersheds by headwater streams. *Science*, *292*(5514), 86–90.

Price, A. (2014). Nitrate dynamics across temporal scales and land use types on three headwater catchments observed using high-frequency measurements (Masters thesis., 78 pp.). University of New Hampshire.

Raymond, P. A., Saiers, J. E., & Sobczak, W. V. (2016). Hydrological and biogeochemical controls on watershed dissolved organic matter transport: Pulse-shunt concept. *Ecology*, *97*, 5–16.

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc. Boston, MA: R. RStudio, Inc. Retrieved from http://www.rstudio.com/

Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., . . . Jomaa, S. (2016). Sensors in the stream: The high-frequency wave of the present. *Environmental Science Technology*, *50*, 10297–−10307. https://doi.org/10.1021/acs.est.6b02155

Seitzinger, S., Harrison, J. A., Böhlke, J. K., Bouwman, A. F., Lowrance, R., Peterson, B., . . . Bohlke, J. K. (2006). Denitrification across landscapes and waterscapes: A synthesis. *Ecological Applications*, *16*, 2064–2090.

Smith, V. H., Tilman, G. D., & Nekola, J. C. (1999). Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution*, *100*(1–3), 179–196.

Walsh, C. J., Fletcher, T., & Burns, M. (2012). Urban stormwater runoff: A new class of environmental flow problem. *PloS One*, *7*, e45814. https://doi.org/45810.41371/journal.pone.0045814

Wollheim, W. M. (2016). From headwaters to rivers to river networks: Scaling in stream ecology. In J. B. Jones & E. H. Stanley (Eds.), *Stream ecosystems in a changing environment* (pp. 349–388). Amsterdam, The Netherlands: Elsevier.

Wollheim, W. M., Green, M. B., Pellerin, B. A., Morse, N. B., & Hopkinson, C. S. (2013). Causes and consequences of ecosystem service regionalization in a coastal suburban watershed. *Estuaries and Coasts*, 16. https://doi.org/10.1007/s12237-013-9646-8

Wollheim, W. M., Peterson, B. J., Deegan, L. A., Hobbie, J. E., Hooker, B., Bowden, W. B., . . . Hershey, A. E. (2001). Influence of stream size on ammonium and suspended particulate nitrogen processing. *Limnology and Oceanography*, *46*(1), 1–13.

Wollheim, W. M., Pellerin, B. A., Vorosmarty, C. J., & Hopkinson, C. S. (2005). N retention in urbanizing headwater catchments. *Ecosystems*, *8*, 871–884.

Wollheim, W. M., Peterson, B. J., Vorosmarty, C. J., Hopkinson, C. H., & Thomas, S. A. (2008a). Dynamics of N removal over annual time scales in a suburban river network. *Journal of Geophysical Research*, *113*, G03038. https://doi.org/10.1029/2007JG000660

Wollheim, W. M., Vorosmarty, C. J., Bouwman, A. F., Green, P. A., Harrison, J., Linder, E., . . . Syvitski, J. P. M. (2008b). Global N removal by freshwater aquatic systems: A spatially distributed, within-basin approach. *Global Biogeochemical Cycles*, *22*, GB2026. https://doi.org/2010.1029/2007GB002963

Wollheim, W. M., Vorosmarty, C. J., Peterson, B. J., Seitzinger, S. P., Hopkinson, C. S., & Vörösmarty, C. J. (2006). Relationship between river size and nutrient removal. *Geophysical Research Leterst*, *33*, 2–5. https://doi.org/10.1029/2006GL025845