AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER

10.1002/2016GL069254

Key Points:
- Climate variability promotes regional biogeochemical extremes and negatively affects water quality
- Coupled land-river model captures dynamics of dissolved N loads in nine sequences of climate extremes
- Dry years promote N accumulation and then enhanced transport to rivers in wetter years

Supporting Information:
- Supporting Information S1

Correspondence to:
M. Lee,
minjinl@princeton.edu

Citation:
Lee, M., E. Shevliakova, S. Malyshev, P. C. D. Milly, and P. R. Jaffé (2016), Climate variability and extremes, interacting with nitrogen storage, amplify eutrophication risk, *Geophys. Res. Lett.*, *43*, 7520–7528, doi:10.1002/2016GL069254.

Received 19 APR 2016
Accepted 21 JUN 2016
Accepted article online 23 JUN 2016
Published online 16 JUL 2016

Climate variability and extremes, interacting with nitrogen storage, amplify eutrophication risk

Minjin Lee[1], Elena Shevliakova[2], Sergey Malyshev[2], P. C. D. Milly[3], and Peter R. Jaffé[4]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA, [2]NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA, [3]U.S. Geological Survey and NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA, [4]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA

Abstract Despite 30 years of basin-wide nutrient-reduction efforts, severe hypoxia continues to be observed in the Chesapeake Bay. Here we demonstrate the critical influence of climate variability, interacting with accumulated nitrogen (N) over multidecades, on Susquehanna River dissolved nitrogen (DN) loads, known precursors of the hypoxia in the Bay. We used the process model LM3-TAN (Terrestrial and Aquatic Nitrogen), which is capable of capturing both seasonal and decadal-to-century changes in vegetation-soil-river N storage, and produced nine scenarios of DN-load distributions under different short-term scenarios of climate variability and extremes. We illustrate that after 1 to 3 yearlong dry spells, the likelihood of exceeding a threshold DN load (56 kt yr^{-1}) increases by 40 to 65% due to flushing of N accumulated throughout the dry spells and altered microbial processes. Our analyses suggest that possible future increases in climate variability/extremes—specifically, high precipitation occurring after multiyear dry spells—could likely lead to high DN-load anomalies and hypoxia.

1. Introduction

Eutrophication (i.e., algae overgrowth often due to excessive nutrients) is a pressing environmental problem facing coastal waters worldwide. The two most acute ecosystem responses to eutrophication are the proliferation of harmful algal species (i.e., harmful algal blooms (HABs)) [*Anderson et al.*, 2002] and hypoxic "dead" zones, whose numbers have increased around the world since the 1960s [*Diaz and Rosenberg*, 2008]. Hypoxia occurs when overgrown planktonic algae, whose growth had been stimulated by excessive nutrients, die, settle to the floor of the estuary, and fuel microbial respiration, which in turn leads to dissolved oxygen (DO) levels below 2 mg l^{-1} in bottom waters [*Hagy et al.*, 2004]. HABs and hypoxia are among the key stressors of coastal ecosystems and are linked to harmful impacts on human health, benthic community, and fisheries [*Anderson et al.*, 2002; *Breitburg*, 2002; *Diaz and Rosenberg*, 1995], causing considerable economic damage [e.g., *Lu and Hodgkiss* [2004]].

Efforts have been made to restore water quality and ecological health of large estuaries and coastal ecosystems, such as the Chesapeake Bay [*National Research Council*, 2011], the Gulf of Mexico [*Mississippi River/Gulf of Mexico Watershed Nutrient Task Force*, 2001], and the Baltic Sea [*Baltic Marine Environment Protection Commission*, 2007], but these efforts have not achieved the intended improvements, and hypoxic zones persist [*Murphy et al.*, 2011; *Turner et al.*, 2008; *Conley et al.*, 2009]. For example, despite basin-wide nutrient-reduction efforts since 1983 [*National Research Council*, 2011], only 12% of Chesapeake tidal waters met desirable DO levels during the summer of 2005–2007 and only 26% met acceptable chlorophyll *a* levels in 2007 [*Ator et al.*, 2011; *U.S. Environmental Protection Agency*, 2008]. Substantial ecosystem (i.e., vegetation, soil, river, and groundwater) nutrient storage in basins can explain such delays in improvements [e.g., *Sanford and Pope*, 2013], as well as observed lags between increases in fertilizer applications and DO decline in coastal waters worldwide [*Diaz and Rosenberg*, 2008; *Galloway et al.*, 2008]. However, recent increasingly severe and extensive hypoxia [e.g., *Hagy et al.*, 2004; *Turner et al.*, 2008; *Conley et al.*, 2009] cannot be completely explained by the nutrient storage and associated lags. Recent studies of estuaries or coastal systems proposed additional mechanisms and suggested, for example, that the ecosystems themselves have become more susceptible to eutrophication due to combinations of potential factors, including (1) increased phosphorus (P) and nitrogen (N) recycling under repeated or sustained hypoxic conditions [*Conley*, 1999; *Smith and Hollibaugh*, 1989; *Conley et al.*, 2009; *Kemp et al.*, 2005]; (2) strengthened stratification, which prevents oxygen replenishment from the surface to bottom waters [*Murphy et al.*, 2011]; and (3) enhanced microbial respiration and reduced oxygen transfer rates from the air to water caused by increased water temperature [*Conley et al.*, 2009].

In contrast to previous studies, here we focus on the role of interactions between the climate and the entire watershed system (i.e., vegetation-soil-river) and investigate what may drive unusually high river DN loading to estuaries or coastal ecosystems as well as DN loading variability. Although other factors (e.g., degradation of coastal wetlands) may play a role and other nutrients like P can limit or colimit phytoplankton productivity [*Fisher et al.*, 1992; *Kemp et al.*, 2005], numerous studies have highlighted the significance of controlling N loading from large rivers, based on strong relationships with phytoplankton production or the onset and duration of hypoxia in estuarine or coastal waters (e.g., Susquehanna River/Chesapeake Bay [*Hagy et al.*, 2004; *Murphy et al.*, 2011] and Mississippi River/Gulf of Mexico [*Rabalais et al.*, 1999, 2002, 2007]).

While evidence suggests that anthropogenic climate change is associated with more frequent and intense extreme weather events [*Blunden and Arndt*, 2014; *Coumou and Rahmstorf*, 2012; *Intergovernmental Panel on Climate Change*, 2013; *Najjar et al.*, 2010], the mechanisms by which the changing climate and its variability influence N transformations and transport in a vegetation-soil-river system remain largely unexplored. These mechanisms are complex and nonlinear, and their long-term responses to climate could not be captured by largely empirical modeling frameworks, such as Spatially Referenced Regression on Watershed Attributes [*Ator et al.*, 2011] and Global Nutrient Export from Watersheds 2 [*Mayorga et al.*, 2010], or by most process-based watershed models designed for short-term simulations, such as Soil and Water Assessment Tool [*Gassman et al.*, 2007] and Regional Hydro-Ecological Simulation System [*Tague and Band*, 2004], which ignore long-term changes in the state of land, simplify, and/or neglect key ecological mechanisms (e.g., forest clearing for agriculture, wood harvesting, and forest regrowth after harvesting) that affect changes in vegetation and soil N storage.

A number of the aforementioned limitations of previous approaches to watershed N modeling were addressed by the coupled terrestrial-river model LM3-TAN (Terrestrial and Aquatic Nitrogen) [*Lee et al.*, 2014], which links vegetation, soil, and river biogeochemical, ecological, and hydrological processes in a consistent and unified framework [*Shevliakova et al.*, 2009; *Gerber et al.*, 2010; *Milly et al.*, 2014; *Lee et al.*, 2014]. LM3-TAN simulates key processes and feedback describing N dynamics across the vegetation-soil-river system [*Lee et al.*, 2014]. Specifically, it is capable of capturing decadal-to-century changes in ecosystem (i.e., vegetation, soil, and river) N storage in response to climate change and variability and land use and land cover changes (LULCC).

This study reveals the critical influence of climate variability and extremes on long-term dynamics of N storage and resultant Susquehanna River DN-load trends and variability. The Susquehanna River drains an area of 71,220 km^2, contributing to about half of the annual N loading to the main stem of the Chesapeake Bay, the largest estuary in the U.S., which provides vital ecological and economic resources [*U.S. Environmental Protection Agency* (*USEPA*), 2008]. The processes and interactions examined here are likely to play an important role in other large river basins experiencing LULCC and climate change. In this manuscript we analyze long-term reported river flows (1932–2013) and crests (1889–2011) and demonstrate historically increased hydroclimate variability and extremes in the Susquehanna Watershed (section 3.1). Unlike flow data, observations of aqueous chemical constituents are seldom available for long-term periods to analyze impacts of changing climate on water quality. Thus, we first analyze available (1987–2009) reported DN loads (section 3.2 and 3.3) and then use the process model LM3-TAN to explore implications of increased climate variability and extremes for DN-load anomalies (section 3.4). We discuss implications of our findings for effective land use management, particularly in the absence of long-term water-quality observations.

2. Methods

2.1. Metrics for Annual Trends and Seasonal Cycle of River Flows and Water Quality

To characterize relationships between river flows and water quality, we used a combination of observations and model LM3-TAN results at the Marietta station, the last downstream Susquehanna River Basin Commission (SRBC) station (40.02′N, 76.32′W), whose upstream subbasin covers 95% (67,314 km^2) of the Susquehanna Watershed. Thus, contributions of almost the entire watershed to the flows and DN loads can be assessed at this station. See *Lee et al.* [2014] for watershed and station descriptions.

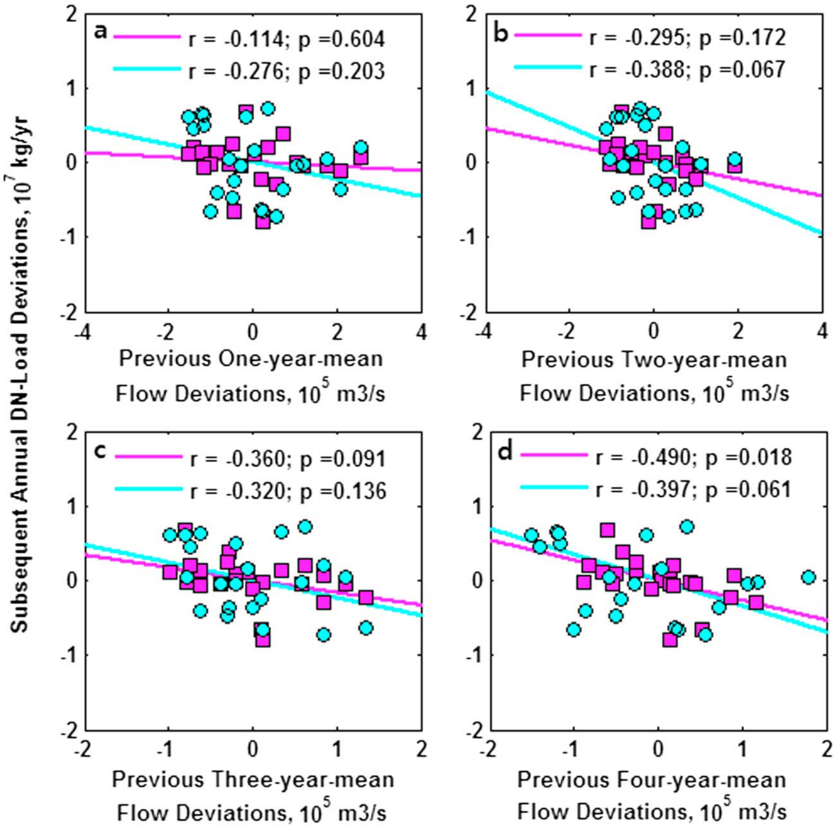

Figure 1. "Annual and multiyear memory effects" on annual DN loads. Regressions (magenta, squares: reported; cyan, circles: simulated by LM3-TAN) between previous annual and multiyear-mean flow deviations and annual DN-load deviations at Marietta for 1987–2009, with Pearson's correlations (*r*) and *P* values (*p*).

Total annual river flows at Marietta from 1932 to 2013 (Figure S1a in the supporting information) were calculated by summing daily mean river flows reported by the U.S. Geological Survey (USGS). The total annual flows were detrended by subtracting the linear trend (shown in Figure S1a) and then used to compute 20 year moving variances (Figure S1b). Historical crests (highest river levels above a flood stage) at Marietta since the late nineteenth century (Table S1 in the supporting information) are based on observations from the NOAA National Weather Service (NWS). Anthropogenic N inputs to the Susquehanna Watershed (Figure S2a) are estimates from the U.S. Environmental Protection Agency [*USEPA,* 2010]. Seasonal and annual river flows (i.e., total flows in $m^3\,s^{-1}$) and DN loads (the sum of dissolved nitrate, dissolved ammonium, and dissolved organic N loads in kg $season^{-1}$ and kg yr^{-1}; because LM3-TAN does not simulate particulate N, we used the DN instead of total N loads) at Marietta from 1987 to 2009 (Figures 1, S2b, S3, and S4) are estimates by using the USGS minimum variance unbiased estimator (MVUE) model [*Cohn et al.*, 1989] from the SRBC Sediment and Nutrient Assessment Program (SRBC SNAP). Simulated seasonal and annual river DN loads at Marietta from 1987 to 2009 (Figures 1, S2b, and S3b) are the model LM3-TAN results.

2.2. Capturing Memory Effects of River Flows on DN-Load Anomalies

To characterize the annual and multiyear memory effects of previous years' river flows on DN-load anomalies, we introduced five variables (Figure 1). *Previous annual and mult year-mean flow deviation* (PY_LFD_n) is annual (i.e., $L = 1$) and multiyear (i.e., $L = \{2,3,4\}$) running-mean flow for year *n* minus long-term mean flow:

$$PY_LFD_n = \frac{1}{L}\sum_{q=1}^{L} F_{a,n-q} - \frac{1}{M}\sum_{r=1}^{M} F_{a,r}, \quad n = 5, \ldots, M \tag{1}$$

where $F_{a,r}$ (total flows in $m^3\,s^{-1}$) is annual flow for year *r*, *L* (=1, …, 4) is the length of running-mean years, and *M* (=27) is the total number of years. For example, PY_1FD_5 is computed as 1986 flow—(1987 flow + … + 2009 flow)/23 and PY_4FD_5 is (1983 flow + 1984 flow +1985 flow + 1986 flow)/4—(1987 flow + … + 2009 flow)/23.

Table 1. 9 "Alternative" Climate Time Series Simulations[a]

Driest				Normal				Wettest
D4	D3	D2	D1	N_o	W1	W2	W3	W4
4 year long dry spell	3 year long dry spell	2 year long dry spell	One dry year (lowest precip.)	One normal year (avg. precip.)	One wet year (intense precip., Hurricane Agnes)	One wet + one normal years	One wet + two normal years	One wet + two normal years + one wetter year
1963	1963	1963	1963	1954	1972	1972	1972	1972
1964	1964	1964				1973	1973	1973
1965	1965						1974	1974
1966								1975

[a]Four dry simulations D1–D4, four wet simulations W1–W4, and one normal simulation N_o.

The four different panels of Figure 1 correspond to the different length (L) of the running-mean years, from annual (upper left-hand) to four-year (lower right-hand) average.

Annual DN-load deviation (YDD_n) is the departure of annual DN loads ($DN_{a,r}$, kg yr^{-1}) from an estimate based on regression between the annual DN loads and flows:

$$\mathrm{YDD}_n = \mathrm{DN}_{a,\,n} - \left(cF_{a,n} + d\right), \quad n = 5, \ldots, M \tag{2}$$

where c (=128.23) and d (=− 9196.2) are the regression slope and intercept.

2.3. Model Simulations

NOAA's Geophysical Fluid Dynamics Laboratory Land Model LM3-TAN (Terrestrial and Aquatic Nitrogen) [*Lee et al.*, 2014] integrates hydrological [*Milly et al.*, 2014], biophysical, biogeographical, and biogeochemical [*Shevliakova et al.*, 2009; *Gerber et al.*, 2010] processes to capture key controls of transport and fate of N in the vegetation-soil-river system. A previous study of the Susquehanna Watershed with LM3-TAN [*Lee et al.*, 2014] at 1/8° resolution successfully established the model fidelity in capturing coupling between vegetation-soil-river interactions, including comparisons of simulated river dissolved nitrate, ammonium, and organic N loads with observations at 16 monitoring stations from 1986 to 2005 and the simulated terrestrial-aquatic N budgets (e.g., soil and river denitrification, harvest, and river N loads) with published estimates from 1998 to 2002 [e.g., *Boyer et al.*, 2002; *Seitzinger et al.*, 2002; *Van Breemen et al.*, 2002]. Here we recalibrated the model by modifying six parameters (Table S2) and used results from 1987 to 2009 to produce Figures 1, S2b, and S3b.

The simulations were performed from 1704 to 2009 with a 30 min time step. As in *Lee et al.* [2014], the model was forced by 3-hourly observation-based global near-surface meteorology data [*Sheffield et al.*, 2006]. For the 1704–1947 period, the available 1948–2008 forcing was repeatedly used (see Text S1 in the supporting information for full forcing description). The probability density function (PDF) of basin-wide mean annual precipitation from the 1948–2008 forcing data was fitted with a gamma distribution (Text S2), and its lower and upper 5th, 10th, and 15th percentiles were used as dry and wet climate indices (Figure S5). To examine legacy effects of prior climate conditions, we simulated nine "alternative" climate time series by using particular combinations of dry or wet years from the Sheffield forcing and holding all the other inputs (e.g., anthropogenic N) the same as in 2005: four "dry" simulations forced by climate with dry spells lasting 1 to 4 consecutive years (experiments D1, D2, D3, and D4); four "wet" simulations forced by climate with wet years, including a year (1972) with extremely high precipitation (experiments W1, W2, W3, and W4), and one simulation with average precipitation (year 1954, experiment N_o) (Table 1 and Figure 2). For example, the dry experiment D1 was run using the 1963 forcing with the lowest annual precipitation in the 1948–2008 period. The wet experiment W4 was run using the 1972 climate with the most intense summer precipitation (from hurricane Agnes) and the second highest annual precipitation. We continued each of the nine alternative climate trajectories for another year by using 61 separate years from the 1948–2008 forcing and thus obtain 549 different yearlong simulations, which followed nine different climate conditions. We used these 549 simulations and constructed nine PDFs of annual river DN loads at Marietta; each PDF was produced using the 61 annual DN loads from the 61 separate 1 year simulations (Figure 3).

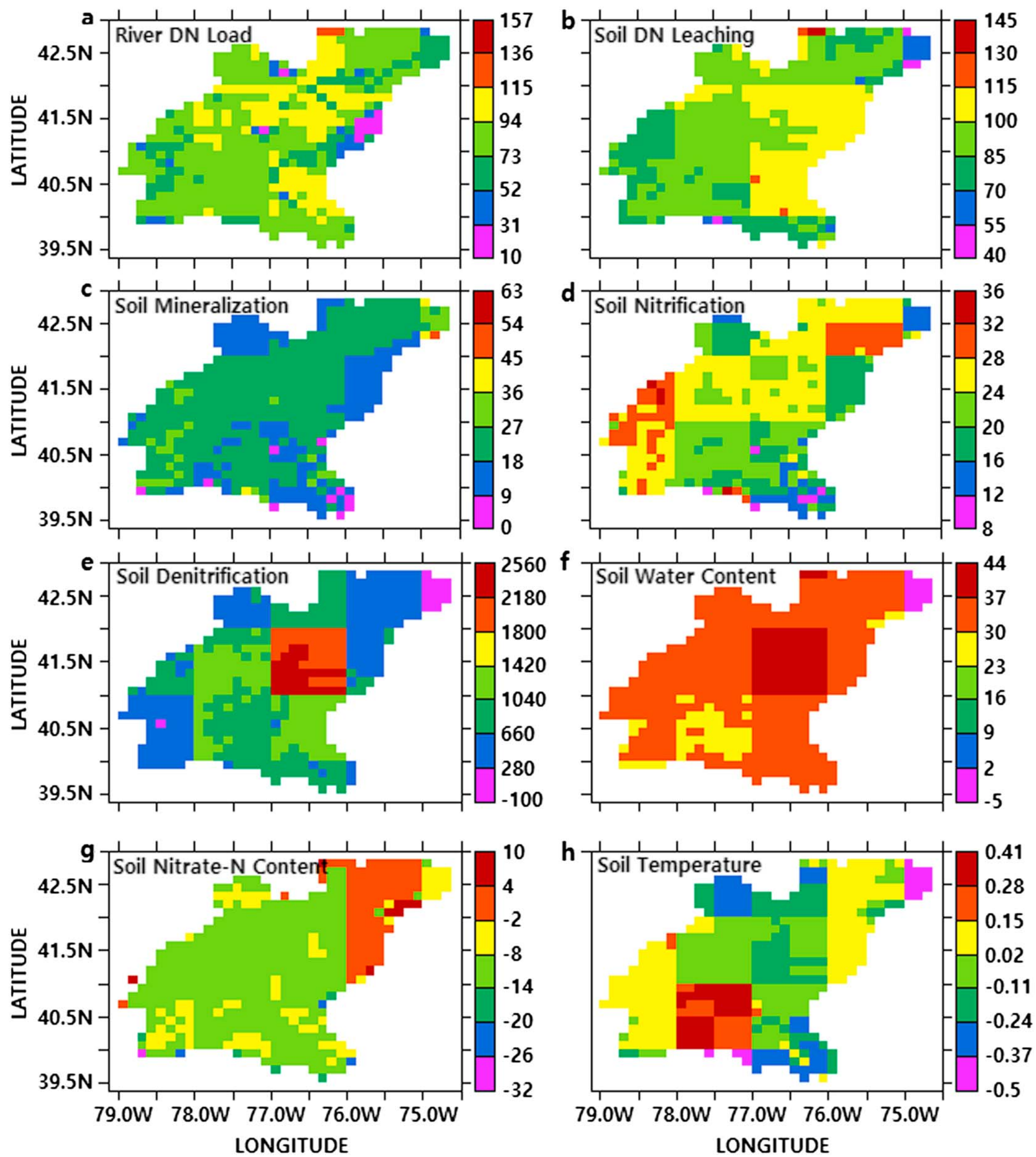

Figure 2. Difference between wet and dry preconditioning. The model results in the dry (D1) and wet (W1) simulations are compared by [(W1 − D1)/D1] * 100 for basin-wide mean annual (a) river DN loads, (b) soil DN leaching, (c) soil mineralization, (d) soil nitrification, (e) soil denitrification, (f) soil water content, and (g) soil nitrate-N content and by W1 – D1 for (h) soil temperature (%, unit for the color legends in Figures 2a–2g; °C, unit for the color legends in Figure 2h). The model is implemented at 1/8° resolution, and the blocks on 1° resolution are produced by the applied forcing data [*Sheffield et al.*, 2006] (see Text S1 for the forcing data description).

3. Results and Discussion

3.1. Increased Hydroclimate Variability and Extremes

During the last eight decades (1932–2013), annual river flow variability at the Marietta station (Figure S1) had been shaped by climate variability and extremes over the entire Susquehanna Watershed: (1) dry spells in the 1960s (1965, 1963, 1969, and 1966; lower 15th percentile of the flow distribution; Text S2), (2) prolonged wet periods in the 1970s (1972, 1975, 1977, and 1979; upper 15th percentile), and (3) increased variability in the 1990s (1995, 1999, 1991/1996, 1994, and 1993; lower/upper 15th percentiles). During this period long-term variability had grown—the 20 year moving variances of the detrended flows had increased 3.4 times from 4.54 to 15.6, $10^9\ m^6\ s^{-2}$. In addition, hydroclimate extremes, such as historical crests

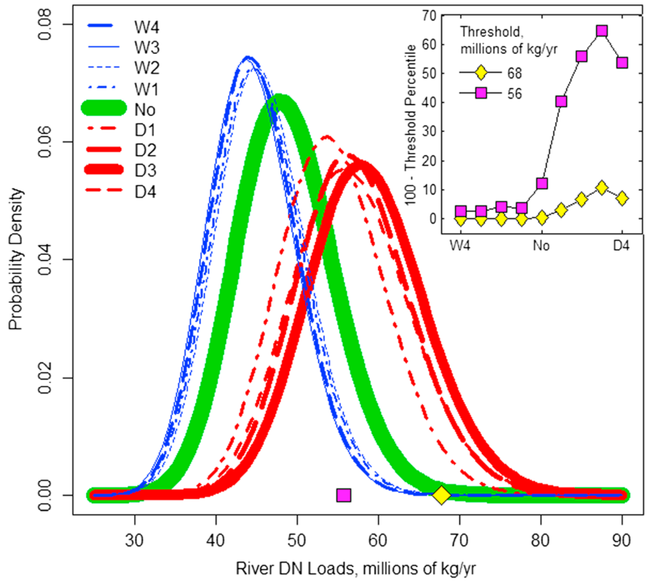

Figure 3. River DN-load responses to interactions between N storage and climate variability/extremes. Nine PDFs of annual river DN loads at Marietta following the climate preconditioning simulations D4 (red dashed), D3 (red solid), D2 (red long dashed), D1 (red dot-dashed), N_o (green solid), W1 (blue dot-dashed), W2 (blue dotted), W3 (blue solid), and W4 (blue long dashed). The insert shows percentages of exceeding the threshold DN loads of 56 kt yr^{-1} and 68 kt yr^{-1} for each distribution. The likelihood of exceeding the thresholds is the lowest for the distribution following W4 and the highest following D3.

at Marietta (1889–2011), had generally become more frequent since the 1990s (Table S1), perhaps as a result of changes not only in climate, but also in land use.

3.2. The Influence of Climate Variability on River DN Loads

As a result of basin-wide nutrient-reduction efforts [*National Research Council*, 2011], the anthropogenic N inputs (including atmospheric N deposition, fertilizer and manure N applications, and point N sources) to the Susquehanna Watershed decreased by 14% (from 3.38 to 2.91, 10^8 kg yr^{-1}) from 1985 to 2005. However, this decreasing input trend is reflected in neither the reported nor the simulated annual river DN loads at Marietta from 1987 to 2009 (Figure S2). This comparison and the strong relationships of the reported (Pearson's correlation, $r = 0.979$) and simulated ($r = 0.834$) annual DN loads with annual flows for 1987–2009 indicate that the interannual climate variability and resultant river flows influence the DN loads more than the reduction in anthropogenic N inputs to the watershed over the past two decades. That is, on annual-to-decadal time scales, climate variability can significantly mask the effect of N mitigation for reducing DN loads via large rivers such as the Susquehanna. Thus, efforts for reducing anthropogenic N inputs to large watersheds would need to continue for multiple decades to achieve noticeable progress in mitigating estuarine or coastal eutrophication.

3.3. The Influence of Climate Variability on River DN-Load Anomalies

We find negative relationships between the previous annual and multiyear-mean flow deviations and annual DN-load deviations (Figure 1; see section 2). These relationships demonstrate that lower annual and multiyear flows are likely to result in higher DN loads in the following years than the expected DN loads based on the flow-load relationship (or regression between the annual flows and DN loads). We hypothesize that "dry-wet" transitions from drier years to wetter years in the Susquehanna Watershed lead to the higher DN loads, while "wet-dry" transitions lead to the lower DN loads; we call this "annual and multiyear memory effects' on annual DN loads. In other words, higher runoff or flows generally result in higher DN loads (i.e., flow-load relationships). When this fact is combined with the sensitivity to dry preconditioning, we can make the generalization that climate variability and extremes (or dry-wet transitions) act to increase DN-load anomalies.

Although the annual memory effect does not appear to be statistically significant at the 0.2 level (Figure 1a), many previous observational studies demonstrated that just one year or even shorter dry periods followed by high rainfalls can lead to high DN-load anomalies and eutrophication, such as marked rises of river nitrate-N concentrations or loads [e.g., *Foster and Walling*, 1978; *Morecroft et al.*, 2000; *Kaushal et al.*, 2008] and phytoplankton production [e.g., *Acker et al.*, 2005]. Furthermore, the statistical significance may be low due to the short-term (1987–2009) N data availability, which excludes historically the most dry (1960s) and wet (1970s) periods. We further investigate shorter period memory effects (i.e., "one- and four-season memory effects") in Text S3 and Figures S3 and S4 in the supporting information).

3.4. The Influence of Interactions Between Climate Variability/Extremes and N Storage on River DN-Load Anomalies

The reported and simulated annual river DN loads at Marietta for 1987–2009 agreed well ($r = 0.9$, $p \ll 0.0001$; Figure S2b) [*Lee et al.*, 2014]. Unlike previous process-based modeling studies [e.g., *Lee et al.*, 2014; *Yang et al.*, 2015], in section 3.3 we demonstrate that the vegetation-soil-river N system exhibits the memory effects. Here we used the simulated DN loads to model the annual and multiyear memory effects. Overall, the negative slopes of the simulated relationships (cyan lines in Figure 1) show that LM3-TAN can generally capture the DN-load anomalies in response to climate variability (or "dry-wet/wet-dry" transitions), although the model has a stronger sensitivity of the DN loads to previous year flows than the reported data indicate. That is, the simulated (cyan) regression slopes are steeper compared to the reported (magenta) slopes with $p \ll 0.0001$ in Figure 1. One the one hand, the stronger simulated sensitivity may be due to unresolved processes in the model (e.g., microbial interactions, groundwater storage, and vertical distribution of N), as noted in *Lee et al.* [2014]. On the other hand, the weaker sensitivity in the reported data may be attributed to the method of annual DN-load estimation from observed river flows, using the statistical model MVUE [*Cohn et al.*, 1989] which is designed to smooth out the memory effects of the vegetation-soil-river N system. Thus, using estimated DN loads with an enhanced empirical model that accounts for the memory effects (e.g., *Green et al.* [2014] for Iowa Basins) may produce a stronger sensitivity similar to one from LM3-TAN. The model captures well that the extent and significance of the annual and multiyear memory effects generally increase with the averaging period length (L) (e.g., $L = 1$ year, $r = -0.276$, $p = 0.203$; $L = 4$ years, $r = -0.397$, $p = 0.061$), similar to the reported data (e.g., $L = 1$ year, $r = -0.114$, $p = 0.604$; $L = 4$ years, $r = -0.490$, $p = 0.018$). That is, the reported and simulated data both indicate that longer dry spells likely lead to higher DN loads in the following years with higher confidence. Thus, our focus on the multiyear memory effects in the following model application allows us to gain novel insights into the consequences of climate variability and extremes on water quality if severe multiyear droughts, as ones observed in the 1960s, will occur again in the future or become more common.

We further investigate the annual and multiyear memory effects and relevant mechanisms by first analyzing the legacy of dry (D1) and wet (W1) years (Table 1; see section 2; Figure 2), and then by analyzing the nine groups of simulations following different dry or wet histories (Figure 3). Simulated basin-wide mean annual river DN load from W1 is 75 % higher than that from D1 (Figure 2a) due to enhanced soil DN leaching (+90%; Figure 2b) and stimulated soil microbial processes, such as mineralization (+20%; Figure 2c) and nitrification (+23%; Figure 2d). Furthermore, elevated soil denitrification in W4 (+836%; Figure 2e) indicates that the extreme wet condition results in substantial N losses from the soil nitrate-N pool. In contrast, simulated basin-wide mean annual soil nitrate-N content in W4, which is 9% less than in D1, indicates that vegetation and soil loose less N during the extreme dry condition of D1 (Figure 2g). The high soil denitrification in W4 is caused by a nonlinear reduction function of soil water content (+34%; Figure 2f [*Lee et al.*, 2014]), rather than soil temperature (not much different; Figure 2h) or soil nitrate-N content (−9%; Figure 2g). This result suggests that lower soil nitrate-N storage in an extreme wet year is not sufficient to attenuate soil denitrification, while increased soil water content enhances denitrification significantly.

The nine constructed PDFs illustrate river DN-load responses following alternative climates in preceding years (Figure 3; see section 2). For example, in Figure 3 the red lines (D1 to D4) are distributions of DN loads in a year following the number of dry years. The longer are the antecedent dry spells (D1, D2, and D3), the higher are the means, standard deviations, and extremes in the DN-load distributions. After the 3 year dry spell, which is based on the period 1963–1965 (D3), the distribution has a 21%, 19%, and 21/20% higher mean, standard deviation, and extremes (i.e., lower/upper 5th percentiles; Table S3) as compared to the distribution following a year with an "average" precipitation (N_o). The confidence also increases with the length of the dry spells. In D4, a relatively wetter year (i.e., 1966 in light red in Figure S5) after the 3 year dry spell (i.e., 1963–1965 in red or dark red) reduces the memory effect, resulting in small decreases in mean and extremes, as compared to those for D3. In contrast, after the extremely wet condition (W1), the distribution has a lower mean (−7%), standard deviation (−8%), and extremes (i.e., lower/upper 5th percentiles, −7%/−7%) compared to those for N_o. Even after precipitation in subsequent 1–3 years returns to a normal or wetter state, the effect of the high N flushing is still evident in the subsequent years; that is, the distributions for W2 (i.e., after one extremely wet and one normal year), W3 (i.e., after one extremely wet and two normal years), and W4 (i.e., after one extremely wet, two normal, and one wetter years) have very similar shapes to that for W1 (i.e., after just one extremely wet year).

Our analysis suggests that the dry conditions in preceding years have a stronger impact on river DN-load anomalies in following years than the wet conditions. In addition, because both means and standard deviations of the DN-load distributions following the dry spells increase with the length of antecedent dry spells, the risks of high DN loads increase, and these likely contribute to severe eutrophication. For example, after the 3 year dry spell, the likelihood of exceeding threshold DN loads of 56 kt yr^{-1} and 68 kt yr^{-1} (reported DN loads at Marietta for the years 1998 and 1993 when Chesapeake Bay summer "dead" zone volumes were much higher than the 1985–2013 average) increases to 65% and 11% (Insert in Figure 3). This result can be explained by the accumulation of soil nitrate-N (and groundwater nitrate-N, accounted by a calibration factor [see *Lee et al.*, 2014]) during the dry spells (e.g., Figure 2g), followed by N flushing (e.g., Figure 2b) and by stimulated soil microbial processes (e.g., Figures 2c and 4d). This result agrees with an empirical analysis [*Kaushal et al.,* 2008] of high annual river nitrate-N loads within Baltimore watersheds in a wet 2003 year following a drought in 2002.

4. Conclusion

Our analysis of legacy effects of prior climate conditions indicates that climate variability (i.e., dry-wet transitions) and extremes (e.g., droughts and hurricanes) strongly interact with N storage accumulated over multiple decades and as a result lead to elevated DN loads compared to DN loads following average precipitation years or wet-dry transitions. The high DN-load anomalies resulting from prolonged or more intense dry spells might explain the increasingly extensive hypoxia in the Chesapeake Bay [*Hagy et al.*, 2004; *Murphy et al.*, 2011], and possibly also in other coastal waters [*Turner et al.*, 2008; *Conley et al.*, 2009], despite the basin-wide nutrient-reduction efforts over decades. If long dry spells as well as hurricane-strength precipitation were to become more prevalent, the risks of extreme N loading and eutrophication would increase. Conversely, a return to less variable conditions would reduce risks. Thus, effective mitigation strategies might benefit from accounting not just for decreasing anthropogenic N inputs and mean climate trends on decadal time scales (section 3.2) but also for changes in interannual climate variability and extremes (sections 3.3,3.4).

Acknowledgments
The daily mean river flows from USGS are available at http://waterdata.usgs.gov/nwis/dv?cb_00060=on&format=rdb&site_no=01576000&referred_module=sw&period=&begin_date=1932-01-01&end_date=2013-12-31; last access: 25 September 2015. The anthropogenic N inputs from U.S. Environmental Protection Agency are available at http://ches.communitymodeling.org/models/CBPhase5/datalibrary/model-input.php; last access: 7 June 2016. The seasonal and annual river flows and DN loads were obtained from SRBC SNAP by personal communication with K. McGonigal. The annual CO_2 concentrations from the NOAA's Earth System Research Laboratory are available at http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html; last access: 25 September 2015. The historical crests from NOAA NWS are available at http://water.weather.gov/ahps2/crests.php?wfo=ctp&gage=mrtp1&crest_type=historic; last access: 25 September 2015. The average summer volumes of "Dead Zone" in Chesapeake Bay from 1985 to 2013 were available at http://mddnr.chesapeakebay.net/eyesonthebay/documents/DeadZoneStatus_Summer2013.pdf, and can be obtained by personal communication, as the website is currently under construction. Support for M. Lee was provided by a Fulbright Scholarship, by the Princeton Environmental Institute at Princeton University through the Mary and Randall Hack '69 Research Fund, by the Korean National Institute of Environmental Research, and by the NOAA (U.S. Department of Commerce) grant NA08OAR4320752.

References

Acker, J. G., L. W. Harding, G. Leptoukh, T. Zhu, and S. Shen (2005), Remotely-sensed chl a at the Chesapeake Bay mouth is correlated with annual freshwater flow to Chesapeake Bay, *Geophys. Res. Lett*, *32*, L05601, doi:10.1029/2004GL021852.

Anderson, D. M., P. M. Glibert, and J. M. Burkholder (2002), Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences, *Estuaries*, *25*(4), 704–726, doi:10.1007/BF02804901.

Ator, S. W., J. W. Brakebill, and J. D. Blomquist (2011), Sources, fate, and transport of nitrogen and phosphorus in the Chesapeake Bay watershed: An empirical model, *U.S. Geol. Surv. Sci. Invest. Rep., 2011–5167*, 27 p.

Baltic Marine Environment Protection Commission - Helsinki Commission (2007), HELCOM Baltic Sea Action Plan, HELCOM Extraordinary Ministerial Meeting, Krakow, Poland.

Blunden, J., and D. S. Arndt (2014), State of the Climate in 2013, *Bull. Am. Meteorol. Soc.*, *95*, S1–S279.

Boyer, E. W., C. L. Goodale, N. A. Jaworski, and R. W. Howarth (2002), Anthropogenic nitrogen sources and relationship to riverine nitrogen export in the northeastern USA, *Biogeochemistry*, *57/58*, 137–169, doi:10.1023/A:1015709302073.

Breitburg, D. (2002), Effects of hypoxia, and the balance between hypoxia and enrichment, on coastal fishes and fisheries, *Estuaries*, *25*(4), 767–781, doi:10.1007/BF02804904.

Cohn, T. A., L. L. DeLong, E. J. Gilroy, R. M. Hirsch, and D. E. Wells (1989), Estimating constituent loads, *Water Resour. Res.*, *25*(5), 937–942, doi:10.1029/WR025i005p00937.

Conley, D. J. (1999), Biogeochemical nutrient cycles and nutrient management strategies, *Hydrobiologia*, *410*, 87–96, doi:10.1023/A:1003784504005.

Conley, D. J., J. Carstensen, R. Vaquer-Sunyer, and C. M. Duarte (2009), Ecosystem thresholds with hypoxia, *Hydrobiologia*, *207*, 21–29, doi:10.1007/s10750-009-9764-2.

Coumou, D., and S. Rahmstorf (2012), A decade of weather extremes, *Nat. Clim. Change*, *2*, 491–496, doi:10.1038/nclimate1452.

Diaz, R. J., and R. Rosenberg (1995), Marine benthic hypoxia: A review of its ecological effects and the behavioural response of benthic macrofauna, *Oceanogr. Mar. Biol. Annu. Rev.*, *33*, 245–303.

Diaz, R. J., and R. Rosenberg (2008), Spreading dead zones and consequences for marine ecosystems, *Science*, *321*(5891), 926–929, doi:10.1126/science.1156401.

Fisher, T. R., E. R. Peele, J. A. Ammerman, and L. W. Harding (1992), Nutrient limitation of phytoplankton in Chesapeake Bay, *Mar. Ecol. Prog. Ser.*, *82*, 51–63.

Foster, I. D. L., and D. E. Walling (1978), The effects of the 1976 drought and autumn rainfall on stream solute levels, *Earth Surf. Process.*, *3*, 393–406, doi:10.1002/esp.3290030407.

Galloway, J. N., A. R. Townsend, J. W. Erisman, M. Bekunda, Z. Cai, J. R. Freney, L. A. Martinelli, S. P. Seitzinger, and M. A. Sutton (2008), Transformation of the nitrogen cycle: Recent trends, questions, and potential Solutions, *Science*, *320*(5878), 889–892, doi:10.1126/science.1136674.

Gassman, P. W., M. R. Reyes, C. H. Green, and J. G. Arnold (2007), The soil and water assessment tool: Historical development, applications, and future research directions, *Am. Soc. Agric. Biol. Eng.*, *50*(4), 1211–1250 ISSN 0001–2351.

Gerber, S., L. O. Hedin, M. Oppenheimer, S. W. Pacala, and E. Shevliakova (2010), Nitrogen cycling and feedbacks in a global dynamic land model, *Global Biogeochem. Cycles.*, *24*, GB1001, doi:10.1029/2008GB003336.

Green, C. T., B. A. Bekins, S. J. Kalkhoff, R. M. Hirsch, L. Liao, and K. K. Barnes (2014), Decadal surface water quality trends under variable climate, land use, and hydrogeochemical setting in Iowa, USA, *Water Resour. Res.*, *50*, 2425–2443, doi:10.1002/2013WR014829.

Hagy, J. D., W. R. Boynton, C. W. Keefe, and K. V. Wood (2004), Hypoxia in Chesapeake Bay, 1950–2001: Long-term change in relation to nutrient loading and river flow, *Estuaries*, *27*(4), 634–658, doi:10.1007/BF02907650.

Intergovernmental Panel on Climate Change (2013), *Climate Change 2013: Detection and Attribution of Climate Change: From Global to Regional: The Physical Science Basis*, edited by F. Stocker et al., Cambridge Univ. Press, Cambridge, U. K.

Kaushal, S., P. M. Groffman, L. E. Band, C. A. Shields, R. P. Morgan, M. A. Palmer, K. T. Belt, C. M. Swan, S. E. G. Findlay, and G. T. Fisher (2008), Interaction between urbanization and climate variability amplifies watershed nitrate export in Maryland, *Environ. Sci. Technol.*, *42*, 5872–5878, doi:10.1021/es800264f.

Kemp, W. M., et al. (2005), Eutrophication of Chesapeake Bay: Historical trends and ecological interactions, *Mar. Ecol. Prog. Ser.*, *303*, 1–29, doi:10.3354/meps303001.

Lee, M., S. Malyshev, E. Shevliakova, P. C. D. Milly, and P. R. Jaffé (2014), Capturing interactions between nitrogen and hydrological cycles under historical climate and land use: Susquehanna Watershed analysis with the GFDL Land Model LM3-TAN, *Biogeosciences*, *11*, 5809–5826, doi:10.5194/bg-11-5809-2014.

Lu, S., and I. J. Hodgkiss (2004), Harmful algal bloom causative collected from Hong Kong waters, *Hydrobiologia*, *512*(1–3), 231–238, doi:10.1023/B:HYDR.0000020331.75003.18.

Mayorga, E., S. P. Seitzinger, J. A. Harrison, E. Dumont, A. H. W. Beusen, A. F. Bouwman, B. M. Fekete, C. Kroeze, and G. Van Drecht (2010), Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation, *Environ. Modell. Software*, *25*(7), 837–853, doi:10.1016/j.envsoft.2010.01.007.

Milly, P. C. D., S. Malyshev, E. Shevliakova, K. A. Dunne, K. L. Findell, T. Gleeson, Z. Liang, P. Phillips, R. J. Stouffer, and S. Swenson (2014), An enhanced model of land water and energy for global hydrologic and earth-system studies, *J. Hydrometeorol.*, *15*, 1739–1761.

Mississippi River/Gulf of Mexico Watershed Nutrient Task Force (2001), Action plan for reducing, mitigating, and controlling hypoxia in the Northern Gulf of Mexico, Washington, D. C.

Morecroft, M. D., T. P. Burt, and A. P. Rowland (2000), Effects of the 1995–1997 drought on nitrate leaching in lowland England, *Soil Use Manag.*, *16*, 117–123, doi:10.1111/j.1475-2743.2000.tb00186.x.

Murphy, R. R., W. M. Kemp, and W. P. Ball (2011), Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading, *Estuaries Coasts*, *34*(6), 1293–1309, doi:10.1007/s12237-011-9413-7.

Najjar, R. G., et al. (2010), Potential climate-change impacts on the Chesapeake Bay, *Estuarine Coastal Shelf Sci.*, *86*, 1–20.

National Research Council (2011), *Achieving Nutrient and Sediment Reduction Goals in the Chesapeake Bay: An Evaluation of Program Strategies and Implementation*, The National Academies Press, Washington, D. C.

Rabalais, N. N., R. E. Turner, D. Justić, Q. Dortch, and W. J. Wiseman Jr. (1999), Characterization of Hypoxia: Topic 1 report for the integrated assessment of hypoxia in the Gulf of Mexico. NOAA Coastal Ocean Program Decision Analysis Series No. 15. NOAA Coastal Ocean Program, Silver Spring, Md.

Rabalais, N. N., R. E. Turner, and W. J. Wiseman Jr. (2002), Gulf of Mexico hypoxia, A.K.A. "the dead zone", *Annu. Rev. Ecol. Syst.*, *33*, 235–263, doi:10.1146/annurev.ecolsys.33.010802.150513.

Rabalais, N. N., R. E. Turner, B. K. Sen Gupta, D. F. Boesch, P. Chapman, and M. C. Murrell (2007), Hypoxia in the northern Gulf of Mexico: Does the science support the plan to reduce, mitigate, and control hypoxia?, *Estuaries Coasts*, *30*(5), 753–772, doi:10.1007/BF02841332.

Sanford, W. E., and J. P. Pope (2013), Quantifying groundwater's role in delaying improvements to Chesapeake Bay water quality, *Environ. Sci. Technol.*, *47*(23), 13,330–13,338, doi:10.1021/es401334k.

Seitzinger, S. P., R. V. Styles, E. W. Boyer, R. B. Alexander, G. Billen, R. Howarth, B. Mayer, and N. Van Breemen (2002), Nitrogen retention in rivers: model development and application to watersheds in the eastern US, *Biogeochemstry*, *57/58*, 199–237, doi:10.1023/A:1015745629794.

Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *J. Clim.*, *19*, 3088–3111.

Shevliakova, E., S. W. Pacala, S. Malyshev, G. C. Hurtt, P. C. D. Milly, J. P. Caspersen, L. T. Sentman, J. P. Fisk, C. Wirth, and C. Crevoisier (2009), Carbon cycling under 300 years of land use changes: Importance of the secondary vegetation sink, *Global Biogeochem. Cycles*, *23*, GB2022, doi:10.1029/2007GB003176.

Smith, S. V., and J. T. Hollibaugh (1989), Carbon-controlled nitrogen cycling in a marine 'macrocosm': An ecosystemscale model for managing cultural eutrophication, *Mar. Ecol. Prog. Ser.*, *52*, 103–109.

Tague, C. L., and L. E. Band (2004), RHESSys: Regional Hydro-Ecological Simulation System —An object-oriented approach to spatially distributed modeling of carbon, water, and nutrient cycling, *Earth Interact.*, *8*, 1–42, doi:10.1175/1087-3562(2004)8<1:RRHSSO>2.0.CO;2.

Turner, R. E., N. N. Rabalais, and D. Justic (2008), Gulf of Mexico hypoxia: Alternate states and a legacy, *Environ. Sci. Technol.*, *42*(7), 2323–2327, doi:10.1021/es071617k.

U.S. Environmental Protection Agency (USEPA) (2008), Chesapeake Bay health and restoration assessment—A report to the citizens of the Bay region, U.S. Environmental Protection Agency Publication, CBP/TRS-291-08, EPA-903-R-08-002, 33 p.

U.S. Environmental Protection Agency (USEPA) (2010), Chesapeake Bay Phase 5.3 Community Watershed Model, EPA 903S10002 - CBP/TRS-303-10. U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis.

Van Breemen, N., et al. (2002), Where did all the nitrogen go?, Fate of nitrogen inputs to large watersheds in the northeastern USA, *Biogeochemstry*, *57/58*, 267–293, doi:10.1023/A:1015775225913.

Yang, Q., H. Tian, M. A. M. Friedrichs, C. S. Hopkinson, C. Lu, and R. G. Najjar (2015), Increased nitrogen export from eastern North America to the Atlantic Ocean due to climatic and anthropogenic changes during 1901–2008, *J. Geophys. Res. Biogeosci.*, *120*, 1046–1068, doi:10.1002/2014JG002763.