Article

Nutrient Improvements in Chesapeake Bay: Direct Effect of Load Reductions and Implications for Coastal Management

Rebecca R. Murphy,* Jennifer Keisman, Jon Harcum, Renee R. Karrh, Mike Lane, Elgin S. Perry, and Qian Zhang

Cite This: *Environ. Sci. Technol.* 2022, 56, 260−270
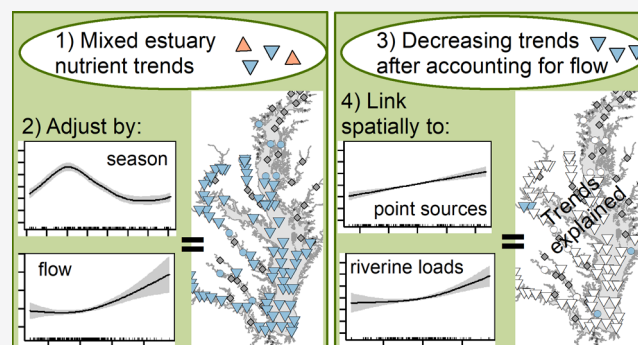
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In Chesapeake Bay in the United States, decades of management efforts have resulted in modest reductions of nutrient loads from the watershed, but the corresponding improvements in estuarine water quality have not consistently followed. Generalized additive models were used to directly link river flows and nutrient loads from the watershed to nutrient trends in the estuary on a station-by-station basis, which allowed for identification of exactly when and where responses are happening. Results show that Chesapeake Bay's total nitrogen and total phosphorus conditions are mostly improving after accounting for variation in freshwater flow. Almost all of these improving nutrient concentrations in the estuary can be explained by reductions in watershed loads entering through 16 rivers and 145 nearby point sources, with the nearby point source reductions being slightly more effective at explaining estuarine nutrient trends. Overall, these two major types of loads from multiple locations across the watershed are together necessary and responsible for the improving estuarine nutrient conditions, a finding that is highly relevant to managing valuable estuarine resources worldwide.

KEYWORDS: *water quality, estuary, eutrophication, generalized additive models, trends*

1) Mixed estuary nutrient trends
2) Adjust by: season / flow
3) Decreasing trends after accounting for flow
4) Link spatially to: point sources / riverine loads = Trends explained

INTRODUCTION

Nutrient over-enrichment of coastal ecosystems has many detrimental consequences including zones of depleted oxygen,[1,2] habitat degradation,[3,4] and toxic algae blooms.[5] Major collaborative efforts have aimed to reduce nutrient pollution into coastal waters and mitigate these negative effects.[6,7] Some successes have been achieved, particularly in point-source-dominated systems,[8,9] while in many regions of the world, development or changing land uses are still resulting in dramatic increases of some nutrient loads to coastal waters.[10,11]

In places where nutrient load reductions have been successfully implemented, scientists have not always observed the expected water quality responses.[12−14] Many possible factors could explain these counter-intuitive findings, including, but not limited to climate change, changes in phytoplankton communities, sediment release of nutrients, changes in light limitation, and filter-feeder abundance.[14−16] Another possible factor is that often large-scale annually or seasonally aggregated metrics such as summer hypoxic volume are the focus of studies looking for response to nutrient reductions[17,18] and actual responses to nutrient reductions are limited in time and place.[19−21] Difficulty in measuring progress and/or explaining observed responses to costly nutrient reduction efforts can undermine continued support for funding. Therefore, identifying progress when and where it has occurred is critical.[22]

Chesapeake Bay, a 11 600 km^2 estuary on the east coast of the United States, is an ideal system for evaluating the response to watershed nutrient load reductions. Nutrient loads and yields to Chesapeake Bay and its tributaries are in the middle compared to other coastal systems;[23,24] however, the overall eutrophic condition is relatively poor when metrics such as depleted oxygen, chlorophyll *a* concentrations, and toxic algae blooms are compared between systems.[23,25] Chesapeake Bay has a strong history of regional collaborations and agreements over nearly four decades to reduce nutrient loads to the estuary. To date, however, the responses to nutrient reduction efforts in the watershed have been complex, non-linear, and season- or location-specific.[26−30] This study evaluates the responses of nutrient concentrations in Chesapeake Bay to changes in monitored watershed nutrient loads at more than 100 specific locations spread throughout the bay, thereby providing missing information for understanding the direct responses to nutrient management efforts. This study also

presents a data-driven method for environmental change analysis that is versatile and flexible enough to be used in other studies where interacting forces result in non-linear change over time.

■ MATERIALS AND METHODS

Data. Collaborative teams in Maryland and Virginia have collected total nitrogen (TN) and total phosphorus (TP) samples bi-monthly or monthly since 1984 at more than 130 locations throughout the Chesapeake Bay (Figure 1, details in
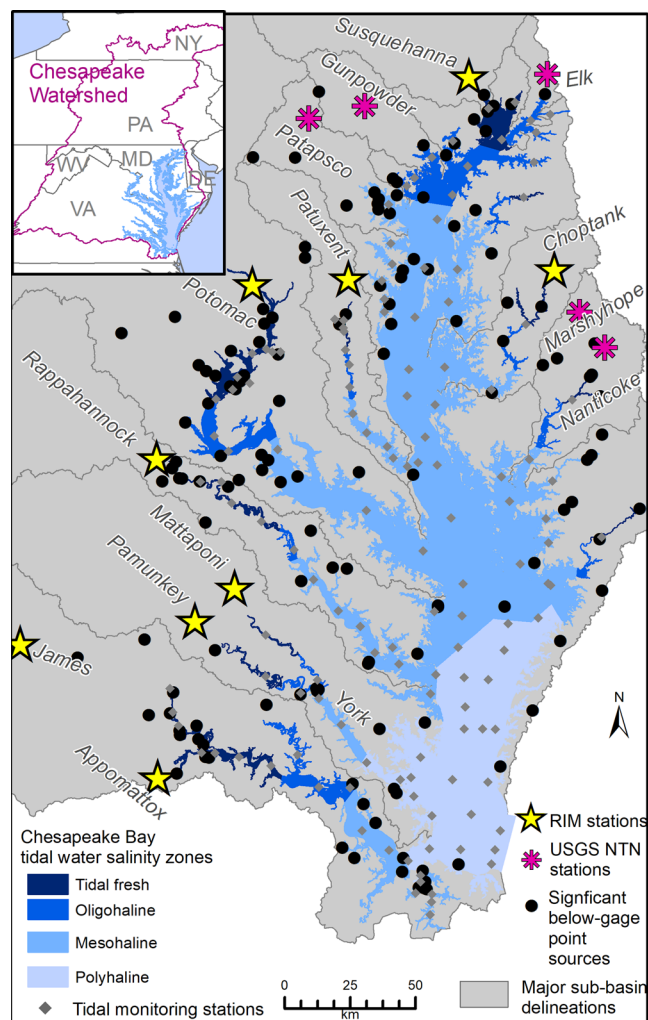
Figure 1. Chesapeake Bay location, salinity zones, watershed, and monitoring stations used in this study. Note the labels on tidal tributaries that are referred to in the study.

Figure S1). Field crews take *in situ* observations at fixed stations every meter or two in the water column for physical parameters and collect samples at multiple depths for laboratory analysis of nutrients and other constituents. Field and laboratory methods are consistent between the teams,[31] and rigorous laboratory comparison studies are routinely conducted to ensure data comparability including split-samples from two different locations analyzed separately by each of the laboratories four times a year to track the inter-laboratory performance. The data are available through the Chesapeake Environmental Data Repository.[32]

Method and laboratory changes have occurred over the more than 30 years of the program. The monitoring teams have documented these in informal documents and somewhat in the data user guide;[33] however, to aid future researchers, we have consolidated a list of method changes for TN and TP (Table S1). In preliminary work, we found it very difficult to distinguish between method changes and load reductions when they occurred at almost the same time. Therefore, in this study, we analyze two different time periods. First, we report long-term trends from 1985 (or 1986) to 2018 and account for method changes where necessary using an intervention approach (see Methods). In the second phase of the analysis, we restrict our time period to 1999–2018 to link nutrient loads to estuarine concentrations during a period for which there are no major method changes of concern.

GAMs for Temporal Change Analysis. A GAM, or generalized additive model, has the following basic structure[34]

$$g(E[Y]) = \mu + f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m) \tag{1}$$

in which a function $g()$ is applied to the expected value E of dependent variable Y. The term μ is the model intercept and f_1 to f_m are smooth functions of covariates x_1 to x_m. A spline function is used in this application to approximate the functions f_1 to f_m[34] and is an important part of this study in that it allows for the relationships between parameters to be determined by the data, not by any *a priori* expectations such as a linear or quadradic relationships.

Our team developed a GAM implementation[35] and built it into an R package, 'baytrends' (https://CRAN.R-project.org/package=baytrends). For this study, we fit GAMs to surface mixed layer TN and TP at each of 135 stations for the entire data record, mostly starting in 1985 or 1986 (Table S1). One additional station starts in 1999 because sampling did not begin until that year. In locations where vertical stratification does not commonly occur, the surface observation was used. In locations where stratification is common, an above-pycnocline (AP) observation was also taken, and the surface and AP values were averaged to get the surface-mixed value. Results were also generated for the bottom mixed layer, which was defined with a bottom sample in regions without stratification and as the average of the bottom and below-pycnocline samples in regions with stratification.[35]

Within 'baytrends', the 'mgcv' R package (https://CRAN.R-project.org/package=mgcv) is used to fit GAMs, and the format shown here is in the syntax for that package for ease of application in other studies. More details are available from our previous work[35] in which we devised and tested different options for evaluating seasonal and long-term changes in estuarine water quality at the station level with GAMs. We use the final equations from that previous research study here as the first phase in this study:

$$\text{gam}(y \sim \text{cyear} + s(\text{cyear}, k = \text{gamK}_1)$$
$$+ s(\text{doy}, \text{bs} = \text{'cc'}) + \text{ti}(\text{cyear}, \text{doy}, \text{bs} = c(\text{'tp'}, \text{'cc'}))$$
$$, \text{knots} = \text{list}(\text{doy} = c(1,366)), \text{select} = \text{TRUE}) \tag{2}$$

$$\text{gam}(y \sim \text{intervention} + \text{cyear} + s(\text{cyear}, k = \text{gamK}_1)$$
$$+ s(\text{doy}, \text{bs} = \text{'cc'}) + \text{ti}(\text{cyear}, \text{doy}, \text{bs} = c(\text{'tp'}, \text{'cc'}))$$
$$, \text{knots} = \text{list}(\text{doy} = c(1,366)), \text{select} = \text{TRUE}) \tag{3}$$
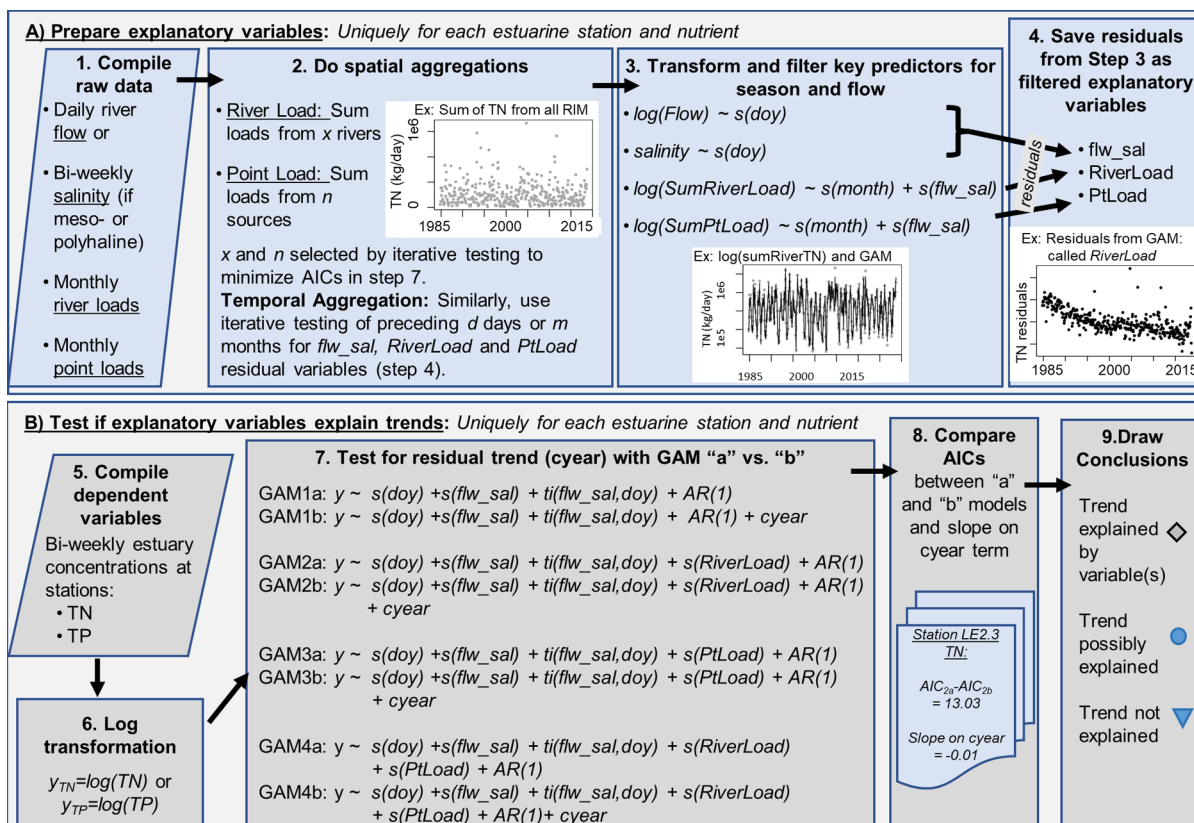
Figure 2. Analysis steps for tidal Chesapeake Bay stations using GAM1a to GAM4b at each station for TN and TP, separately. Variable *doy* refers to day of year of the sample collection, *cyear* is the date in decimal format, $s()$ refers to a smooth spline function, $ti()$ refers to an interacting spline function, *flw_sal* is either flow or salinity depending on the location, and $AR(1)$ indicates our use of a generalized additive mixed model approach ("gamm" function in 'mgcv' R package) which allowed for incorporation of residual autocorrelation with a lag of 1 time step.

where the k-value on the *cyear* term is set to $gamK_1$ which is the larger of 10 or 2/3 times the number of years in the record being analyzed. The variable *cyear* is a date in decimal form, *doy* is the day of the year, $s()$ indicates a spline function of the variable, $bs = 'tp'$ specifies a penalized thin plate regression spline and $bs = 'cc'$ specifies a cyclic penalized cubic regression spline. If the spline is not specified, the default $'tp'$ was used. The $ti()$ format specifies a tensor product of two smoothers to account for the interacting effects of these two variables.[36]

The two equations are only different due to the factor variable *intervention* in eq 3. We used this structure when we found a method or laboratory change to have a significant impact on the data values (Table S1). If so, *intervention* was set to one value before the date of the change and another after. This allowed for there to be a data-fitted step adjustment applied to values before the intervention.[35]

The GAMs (eqs 2 and 3) were fit to the monitoring data for each station and constituent to generate a series of mean predictions over time. The difference between the predictions generated for any two times periods is the mean percent change over time. The standard error in the computation gives a measure of uncertainty (as documented in previous work[35]). We computed the percent change between 2017−2018 and (1) 1985−1986 as the first two years for the "long-term change", (2) 1999−2000 that represents the longest portion of the record without method change concerns and as the period used in the part of this study linking concentrations to loads, and (3) 2009−2010 for a short-term 10-year change estimate.

Watershed Nutrient Load Sources. The U.S. Geological Survey (USGS) computes monthly TN and TP loads at nine River Input Monitoring (RIM) stations and five additional non-tidal network stations used in this study (Figure 1, Table S2).[37] These are referred to as "river loads" and represent the loads from all source types including non-point, point, atmospheric, and natural mostly upstream of the tidal extent of the bay. A recent estimate shows that the majority of the river flow (78.3%), TN (61.8%), and TP (60.4%) from the Chesapeake Bay watershed flows through the nine RIM stations (Q. Zhang, personal communication, Oct. 22, 2021).

We also used monthly TN and TP loads from major point sources located downstream of the portion of the watershed monitored by the USGS gages (Figure 1) (available at https://cast.chesapeakebay.net/). These are referred to as "below-gage point loads" and were located using a spatial discretization of the watershed developed for watershed modeling of Chesapeake Bay.[38] This process allowed us to identify which segment of the estuary (Figure S1) each point source drains to. The point loads were then summed by month into estuary segment(s).

GAMs Linking Watershed to Estuary. Different explanatory variables were tested to try to explain the patterns in TN or TP concentrations over time using a two-step model fitting process. The general process involved fitting eqs 4 and 5 to the same station, parameter, and explanatory variables:

GAMa:
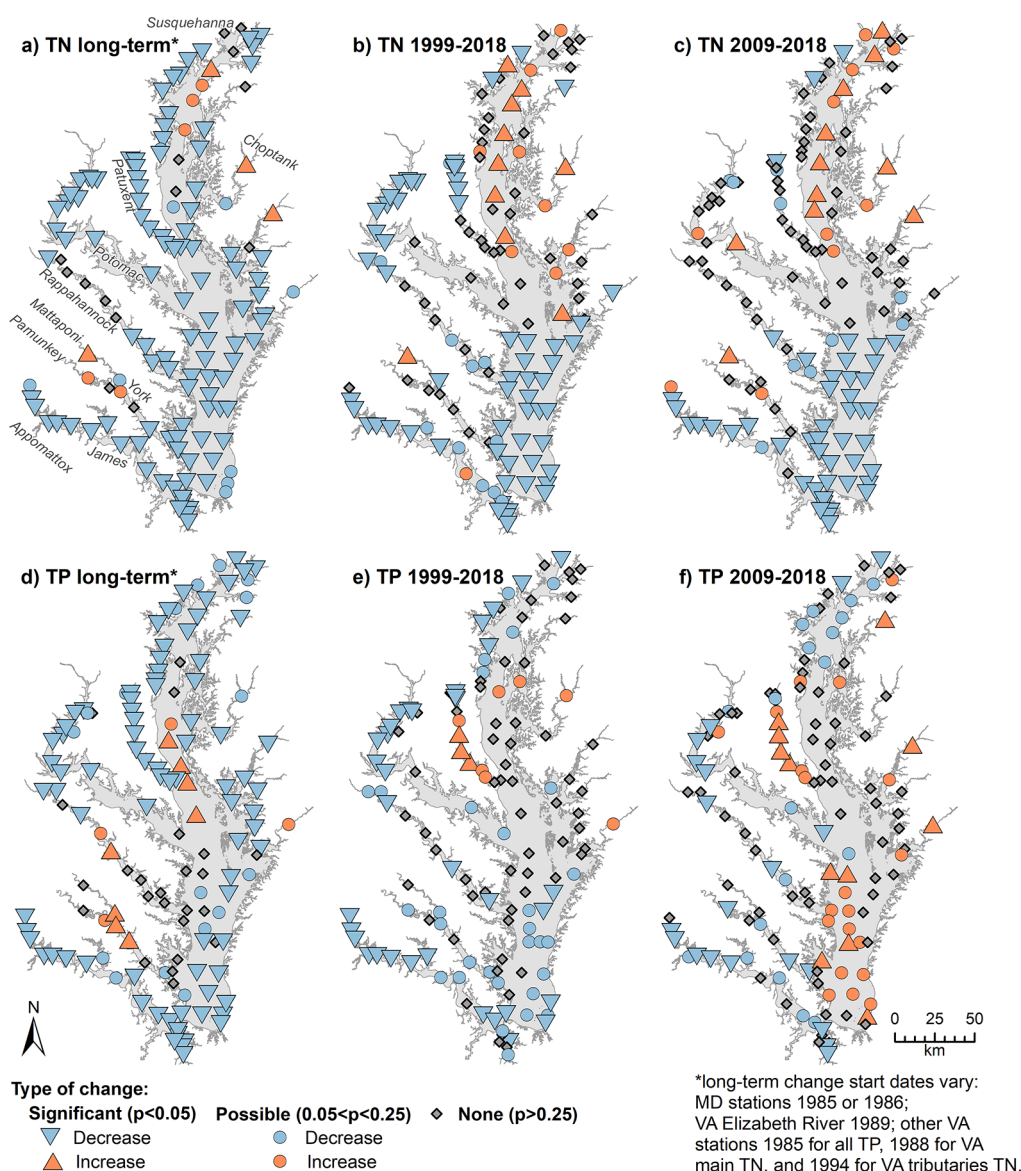
Figure 3. Chesapeake Bay tidal station categorical results for mean change in surface TN (a−c) and TP (d−f) over three time periods computed using temporal GAM fits (eqs 2 and 3) but not filtering for flow or any other explanatory variable.

$$gam(y \sim s(doy) + s(\text{explanatory variable}_1) + \dots s(\text{explanatory variable}_n)) \quad (4)$$

GAMb:

$$gam(y \sim s(doy) + s(\text{explanatory variable}_1) + \dots s(\text{explanatory variable}_n) + cyear) \quad (5)$$

The only difference between eqs 4 and 5 for each test is the *cyear* term in the "b" GAM (eq 5). This decimal date term will capture a linear trend over time. Therefore, if the "b" model was a substantially better fit than the "a" model, we concluded there was an unexplained temporal change that was not accounted for by the explanatory variables. Conversely, if the fit did not improve between the "a" and "b" models, then the explanatory variables were sufficient to explain the temporal change in *y*. The corresponding "a" and "b" GAM fits were compared for each data set in two ways: (1) taking the

difference in the Akaike information criteria (AICs) between the models and (2) examining the slopes on the *cyear* term.

There is no single cutoff for a difference in AIC values that is deemed "significant" (e.g., refs 39 and 40), but instead practitioners often consider AIC difference ranges in evaluating whether models are substantially different. Similarly, we used a range of values based on a tabulation of studies[40] and also tested for a trend in the residuals from each fitted "a" model as a double-check of our decision points.

In this study if AIC improvement between model "a" and "b" (i.e., $AIC_a - AIC_b$) was

- >7: there was **Strong** evidence that the trend was not explained with the "a" model;
- Between 4 and 7: there was **Possible** evidence that the trend was not explained with the "a" model; and
- <4: there was **Little** evidence, or no trend after the "a" model was applied.

Four pairs of GAMs with the structure of eqs 4 and 5 were designed to test whether freshwater flow and nutrient loads
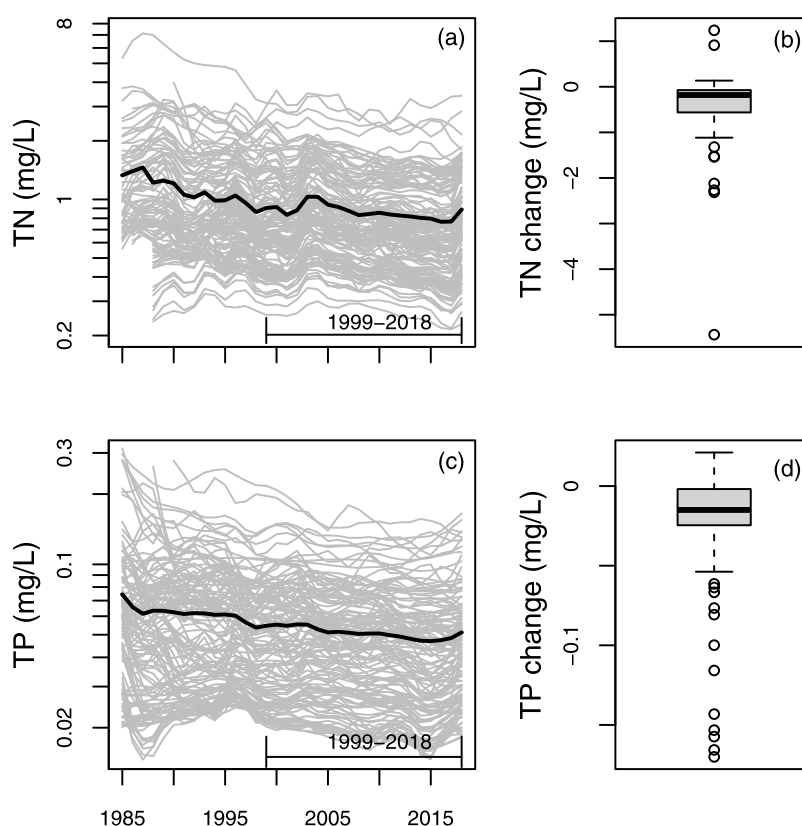
Figure 4. Aggregated surface TN and TP patterns and changes over time for all long-term tidal Chesapeake Bay stations. For surface TN, (a) shows GAM estimates for each station (grey lines, log scale) and average estimate across all stations (black line, log scale) and (b) shows a boxplot of computed change in TN between endpoints of each timeseries. Surface TP summarized similarly (c,d). 1999−2018 is indicated on the graphs as the period of focus in the next section of Results.

could explain the temporal trends at the estuary stations. The process to implement these tests is described in Figure 2. The steps are prepare the explanatory variables (steps 1−4), set up the dependent variables of at each station (steps 5−6), fit the paired GAMs labeled GAM1a to GAM4b (step 7), and draw conclusions (steps 8−9) for each station. An example is presented in Figure S2.

Pre-processing of each of the explanatory (i.e., x) variables was performed, as indicated in Figure 2 in steps 1−4. For each station, we use one of the two ways to represent the effect of freshwater input to the estuarine system. For stations in tidal fresh regions of the bay (see Figure 1 for salinity zones), river flow from an upstream USGS gage was used. This river flow data is daily, so testing was conducted to find an appropriate number of preceding days before the sample date to average the flow observations. For each station, 1 to 210 days of flow averaging was tested and the period with both a relatively low AIC and spatial consistency with nearby stations was selected (see Table S3 for average periods). For stations in mesohaline or polyhaline stations of the bay, flow was incorporated indirectly using the salinity measured at the same place and time as the TN or TP observation. For oligohaline stations, both salinity and flow were tested and the choice that resulted in the lower AIC was selected.

Freshwater flow variability has a very large impact on riverine loads from the watershed. An example of a GAM fit for total RIM TN load is shown in step 2 of Figure 2. This GAM explaining TN load with flow and season has an R^2 of 0.987. The similar GAM for TP has $R^2 = 0.936$. This suggests that it would be almost impossible to see a signal of nutrient changes

due to management actions in the watershed if the raw observed TN and TP loads were used to explain the estuarine concentrations because of the large impact of flow-derived variability. That is why we pre-processed the *RiverLoad* variables to filter out the impact of flow (steps 3 and 4) but still included the flow impact separately with *flw_sal* in the GAM1a-GAM4b equations. Note that our processing of *RiverLoad* to filter out the effect of flow is conceptually similar to the approach used by USGS to flow-normalize river loads.[37] In fact, we refer to these USGS-derived flow-normalized results later in the discussion of our results. However, we chose to use GAMs in our analysis to simultaneously filter season effects and flow for methodological consistency.

Selecting Spatial and Temporal Matches. Selecting the appropriate river loads and point source loads to link to the nutrient concentrations at each estuary station through the *RiverLoad* and *PtLoad* variables needed careful consideration (Figure 2, step 2). Previous efforts have linked loads from only the largest riverine source to conditions in the mainstem bay[28,41] or aggregated the influence of only the tributaries that are upstream of the location in the estuary.[30] We chose to test multiple combinations of load sources to explain the estuarine concentrations at each station because two-layered circulation creates the potential for mixing of the influence of sources, and modeling studies have demonstrated that loads from multiple basins influence mid-bay water quality.[42] At the same time, some estuary stations are in small tributaries and/or fresh regions far from any gaged river source. Because a major focus of this study is to compare gaged river loads and point loads, we excluded stations that did not have gaged load estimates for

their primary watershed. Therefore, of the 136 stations included in the temporal GAM results (eqs 2 and 3), 112 were evaluated for links to watershed loads. Of those 112 stations, an iterative process was used to link them to river and below-gage point loads. In most cases outside the tidal fresh portions of the estuary, multiple tests were conducted to match the estuary stations to the most explanatory loads. Results showed in many cases that both riverine and point loads from multiple locations, not just the local region, needed to be summed to explain estuarine concentrations. The process and results for this spatial matching are documented in the Supporting Information (text and Figures S3 and S4).

RESULTS AND DISCUSSION

Observed Nutrient Trends Variable but Mostly Decreasing. Long-term TN and TP surface concentrations have decreased at 112 stations for TN and 99 stations for TP out of 135 with records to the 1980s (Figure 3a,d). There are local exceptions, but averaged across all stations, the long-term decrease is visible (Figure 4). The shorter-term changes over 20 and 10 years (Figure 3b,c,e,f) show more stations with increasing concentrations or no change. Bottom TN and TP results (Figure S5) generally show similar changes over time, with the exception that there are fewer TN increases in the mid- to upper bay in the bottom compared to surface. Bottom TP has almost the same patterns as the surface. In several studies where nutrient concentrations were aggregated spatially in the Chesapeake Bay, researchers observed long-term increases in the oligohaline mainstem and decreases in the mesohaline and polyhaline mainstem for dissolved nitrogen[26,27] and phosphorus.[26] These general trends are consistent with our TN long-term trends at most stations (Figure 3a), but not necessarily TP (Figure 3d); possibly due to differences between how dissolved and total P could be trending and/or the time periods of analysis. In the tidal Patuxent,[43] Potomac,[44] and James Rivers,[45] studies have documented long-term decreases in N and P after major wastewater load reductions in the first half of this record, consistent with long-term decreases in those tributaries in Figure 3a,d.

The observed changes over time (Figures 3 and S5) have the potential to be impacted by year-to-year variations in freshwater flow (Figure 5a). Nutrient concentrations in Chesapeake Bay generally increase during high flow conditions and decrease during low flow conditions.[27] A multi-year drought from 1999−2002 and wet period that started in 2018 (Figure 5a) likely had an impact on the shorter-term changes. Despite flow-impacted fluctuations, the long-term picture (Figure 4a,c) demonstrates that decreasing nutrient concentrations have persisted. The average percent change in TN across all stations from the beginning to end of this record is −25% and for TP is −22%, corresponding to an average decrease of 0.39 mg/L TN and 0.022 mg/L TP.

Loads Variable but Mostly Decreasing. Total annual watershed loads of TN and TP aggregated across nine USGS RIM stations and all significant below-gage point sources used in this study demonstrate large year-to-year variability of the river TN and TP loads and a clear decrease in bay-wide point source loads (Figure 5b,c). This graph does not appear to show a long-term trend in riverine TN or TP loads to the bay, and trend tests reveal there are few trends in observed loads. Only the observed loads through the James River (for TN and TP) and Patuxent River (for TN) have long-term significant decreases, and loads through the Choptank River (for TN and
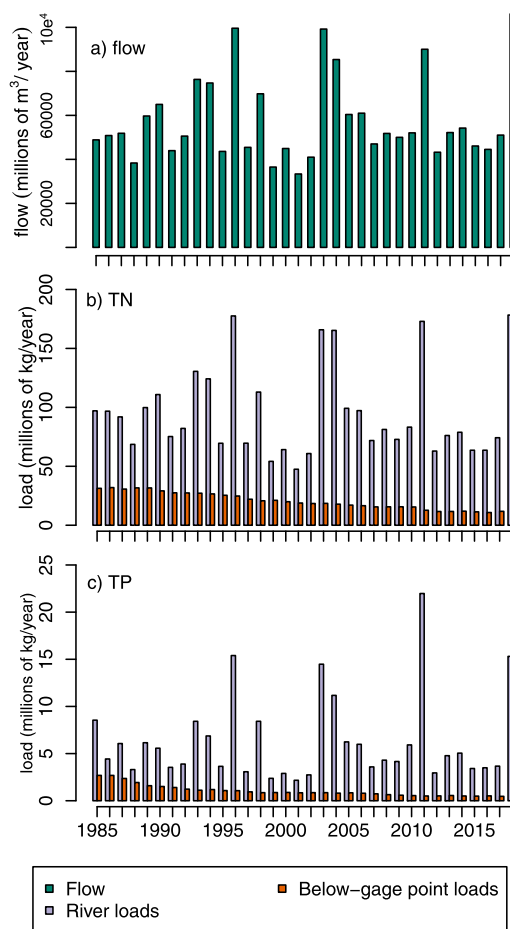

Figure 5. Total annual gaged freshwater flow into Chesapeake Bay (a), TN loads (b), and TP loads (c) from the river loads (purple) and below-gage point sources (orange) summed across watersheds used in this study.

TP) have long-term significant increases (Q. Zhang, personal communication, Oct. 22, 2021). These observed trends in true condition loads represent the actual loads impacting the bay.

In addition to knowing the amount of nutrients truly entering the bay each year, to evaluate the impact of upstream management actions, it is useful to estimate what the trends in nutrient loads are after removing the effect of the year-to-year flow variations. The USGS annually provides flow-normalized predictions and trends for that purpose.[37] From 1985 to 2018, after removing the effect of flow on river loads, USGS reported decreasing flow-normalized trends in TN loads at six of the nine RIM stations (Susquehanna, Patuxent, Potomac, Rappahannock, Mattaponi, and James) and increasing flow-normalized TN trends at three RIM stations (Choptank, Pamunkey, and Appomattox). Because the TN decreases include the three largest rivers, it can be concluded that the total riverine load in Chesapeake Bay of TN after accounting for flow is decreasing. For TP, the story is more mixed with three RIM stations showing flow-normalized decreases (Patuxent, Potomac, and James) and the six remaining (including the Susquehanna) all showing increasing trends.[37] Overall for TN and TP, more trends are apparent in the flow-normalized loads than in the true condition loads.

Estuarine Trends Explained by Watershed Inputs. The freshwater flow effect (i.e., *flw_sal*) is highly significant at most stations (Table S3). Across the flow-based GAM results
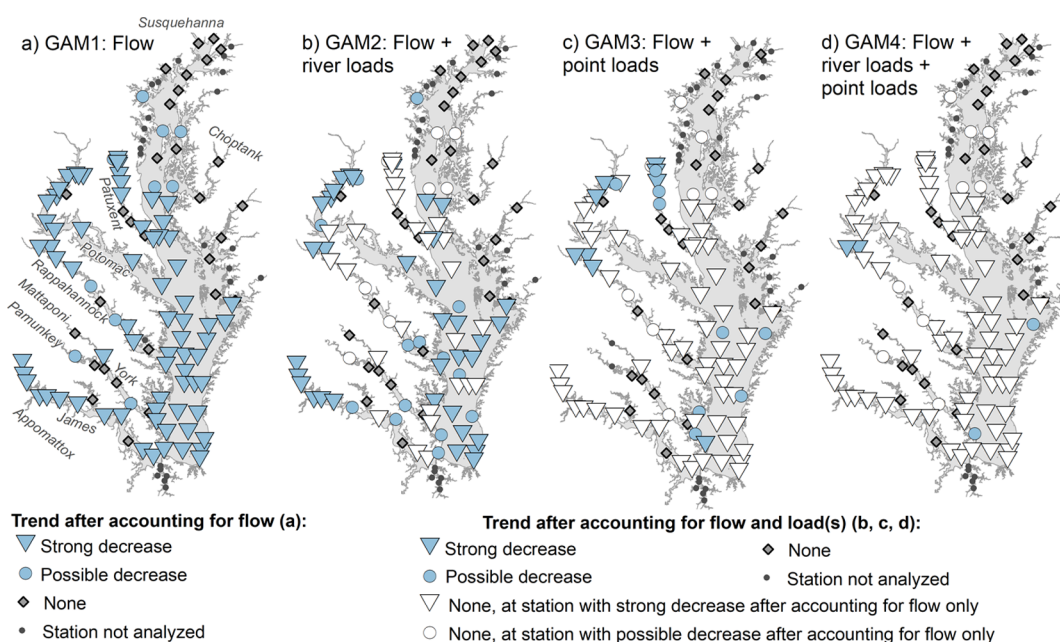
Figure 6. Chesapeake Bay surface TN 1999−2018 analysis summarizing remaining trend after accounting for freshwater flow (a), river load (b), below-gage point load (c) and both loads (d). Filled symbols indicate trend was not explained by the variable(s), open symbols indicate trend was explained by the variable(s) added in that equation.

for TN (GAM1a, Figure 2), the p values on at least one of the *flw_sal* terms is <0.05 at 97% of the stations. For TP, 79% of the stations' models have $p < 0.05$ for one of the *flw_sal* terms. On average, flow explains slightly more variability for TN than for TP with the average R^2 of the GAM1a fits across all stations of 0.53 for TN and 0.43 for TP.

We conducted an analysis to spatially link the most influential load sources to each estuarine station and that is documented in Supporting Information. The findings (Figure S3 and Table S3) demonstrate the large spatial influence of loads from many parts of the watershed to regions in the estuary, and that reductions from only one source type or subbasin will not be sufficient to reduce nutrient concentrations bay-wide. Multiple loads flowing through the major tributaries influence nutrient concentrations in the mainstem, and likewise some of the tributary waters are influenced by loads transported first through the mainstem of the bay (Figure S3).

After these spatial matches were made, when river loads are included in the GAM fits (GAM2a, Figure 2), the average R^2 increases to 0.59 for TN and 0.46 for TP. When only point loads are included (GAM3a, Figure 2), the average R^2 is 0.61 for TN and 0.47 for TP. Finally, for models including both types of loads and flow or salinity (GAM4a, Figure 2), the average R^2 is 0.62 for TN and 0.48 for TP. Even though these average R^2 values indicate that some variability is unexplained, the loads improve the model fits for most of the stations. For TN with the GAM4a model, the s(PtLoad) term or the s(RiverLoad) term had a low p-value (<0.05) at 92% of the stations. For TP, one or the other of the loads is highly explanatory at fewer stations with 56% of the stations meeting this criterion.

For surface TN, river flow or salinity explained all of the increasing trends observed in the data from 1999−2018 (orange symbols in Figure 3b compared to Figure 6a). Specifically, every orange symbol in Figure 3b became either a blue symbol or gray dot indicating that after controlling for

flow, TN is decreasing or not trending at these stations. The reason for this is likely the timing of the analysis. This 20-year period started in a drought and ended with a wet year, resulting in lower nutrient concentrations than normal at the beginning of the record (due to less freshwater input) and higher concentrations than normal at the end of the record (due to higher freshwater input). Overall, after accounting for river flow, 73.2% of the stations show a residual decrease in surface TN from 1999 to 2018 (Table S4). The stations with no trend in TN concentration after accounting for flow are spread throughout the bay, with clusters of more than one station along the eastern shore, upper mainstem, the Patuxent River, and the York River. These findings have some consistency with regions where USGS RIM gages[37] show either degradation or no trend over both the short- and long-term (Choptank on the eastern shore and Pamunkey flowing into the York River).

Comparing the impact of the two types of sources of nutrient loads analyzed here, river loads alone explain fewer of the decreasing trends in TN (Figure 6b) than below-gage point loads alone (Figure 6c, Table S4). Below-gage point loads and river loads together fully explain most TN trends; however, there are some regions where one source or the other appears more influential. River loads explain all the TN trends in the Patuxent River (see Figure 6b). Many of the major wastewater treatment plants in this basin are upstream of the Patuxent RIM gage, and load reductions in upper basin loads have been attributed to upgrades at these wastewater facilities in previous work.[46] In contrast, the point loads explain the TN trends in the James River (Figure 6c). In this case, major wastewater treatment contributions are downstream of the James RIM gage and those point source contributions have declined over both the long- and short-term.[47] Both load sources together, however, explain the largest number of the TN trends (Figure 6d). Of the 82 stations with decreasing TN after accounting for flow, 78 of those decreases can be fully explained by both river and point loads combined. Of all the stations analyzed in this
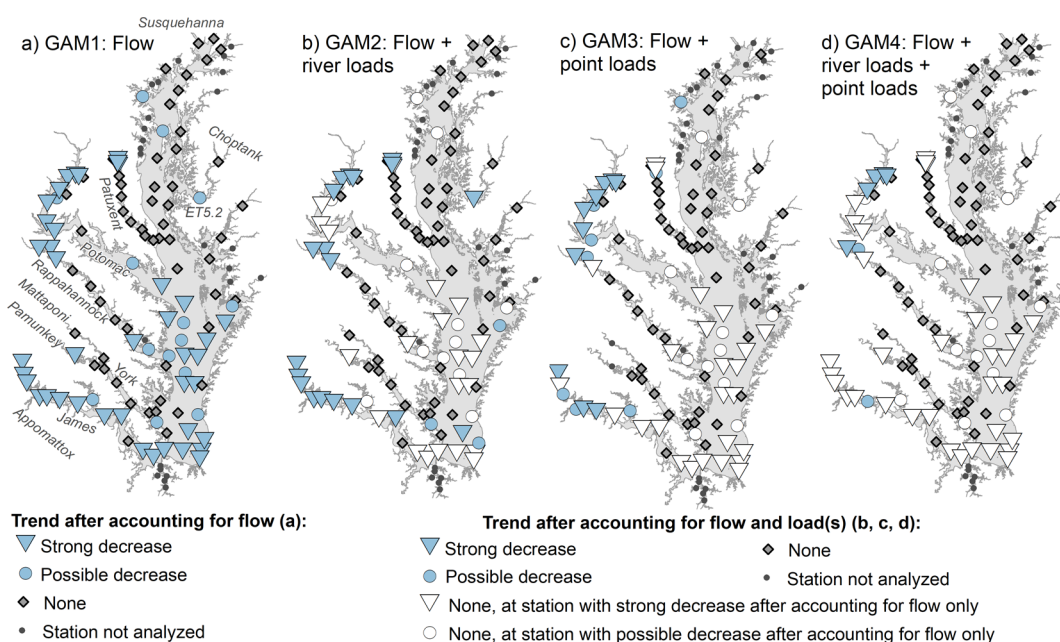
Figure 7. Chesapeake Bay surface TP 1999−2018 analysis summarizing remaining trend after accounting for freshwater flow (a), river load (b), below-gage point load (c), and both loads (d). Filled symbols indicate trend was not explained by the variable(s), open symbols indicate trend was explained by the variable(s) added in that equation. One example station (ET5.2) is labeled in (a).

study, only four (3.6%) have residual TN trends that are not explained by flow, point loads, and river loads combined. The two that stand out are in the tidal fresh Rappahannock River where the load-to-concentration relationships in the GAM fits are strongly positive (Table S3); however, there is residual decrease in TN that is not explained by the loads.

For surface TP, river flow also accounts for all the increasing trends from 1999−2018 (compare Figures 7a to 3e). The Patuxent river stations with increases from 1999 to 2018 (Figure 3e) have no trend after filtering out the effect of flow (Figure 7a). Bay-wide, about half of the stations have decreasing flow-filtered trends in TP from 1999−2018, which is a smaller fraction than for TN (Table S4). This is consistent with the fewer reported decreases in flow-normalized loads of TP compared to TN through the major tributaries by USGS, as discussed above.[37] River loads and point loads alone each explain a similar number of trends, which is different from TN for which point loads alone explain more of the trends (Figure 7b,c). This difference may be because in most cases, the largest wastewater reductions of TP occurred before this time period with phosphorus detergent bans.[48] TP trends at seven stations (6.2%) are not fully explained by the combined loads and flow (Figure 7d). The clusters of unexplained trends are in the tidal fresh Potomac and tidal fresh Rappahannock.

Along with the big-picture story, it is possible to zoom in and evaluate station-specific results with this approach. One station where the TP response to the river load is counter-intuitive at first glance is station ET5.2 located in the mesohaline Choptank River (indicated in Figure 7a). TP shows a possible negative trend after accounting for flow (Figure 7a), but a strong negative trend after accounting for river load (Figure 7b). The trend is fully explained by the point source load (Figure 7c). This is a unique location where the TP river load is increasing over time, but major point source reductions have occurred in the watershed.[49]

One gap in this analysis is the lack of monitoring data to account for the nonpoint source loads from the watershed draining downstream of the river load gages. Just like the below-gage point loads were shown to be very influential despite the smaller magnitude of their loads, it is likely that below-gage nonpoint sources and reductions (or increases) in them would be influential to estuarine concentrations. Other major nutrient sources to the Chesapeake Bay that are not accounted for here include atmospheric deposition to tidal waters and nutrients in incoming water from the Atlantic Ocean.[50,51] It is possible that some of these non-monitored loads might play a role in explaining the unexplained trends (Figures 6d and 7d), although is not a very likely explanation in the tidal fresh regions because they are close to the RIM stations. Model R^2s may certainly improve if these non-monitored loads could be included. Other possible explanations for the unexplained TN and TP decreases include an increase in submerged aquatic vegetation (SAV) over recent decades in the Potomac[44] and Rappahannock that could be retaining nutrients in biomass temporarily and then enhancing burial in the sediments.[52] A more methodological explanation is the one- or two-month temporal aggregation of loads applied to these stations may not be a fine enough resolution to explain nutrient concentrations in tidal fresh regions where variability can be large. Future work could include using newly available daily load estimates for the RIM stations (the new WRTDS-K approach[53]) to better explain the short-term variations in RIM load and tidal water quality.

The observed decreases in point source nutrient loads (Figure 5) and flow-normalized trends from the watershed[37] clearly indicate that substantial efforts have already been made by Chesapeake Bay jurisdictions to manage nutrients. The results documented here show that those efforts and investments to date have made a difference in the estuary. Beyond the nutrient reductions we have examined in this study, several promising ecosystem responses have been documented for Chesapeake Bay, including resurgence of

SAV[54] and increases in nutrient limitation in certain areas.[45,55] However, there is clearly a need for continued action because at the same time, low oxygen conditions persist in the summer and many of the water quality criteria for the bay waters are not being met.[18,56] Documented chlorophyll *a* and water clarity trends are unclear and often fail to show an ecosystem response to nutrient reductions.[27,30] Some of the reasons for these less positive observations may be both: (a) the mixed observed trends (Figure 3) and (b) that even in those places where observed nitrogen and phosphorus concentrations are improving, these nutrient concentrations may not be sufficiently low to limit phytoplankton growth.[57] Furthermore, short-term load trends in TN and TP are improving at fewer RIM stations than long-term trends,[37] suggesting that trends are approaching asymptotes in some regions. The impact of freshwater flow variations on observed concentrations in the estuary also needs to be considered as climate-change predictions suggest increased short-term precipitation intensity in the Chesapeake Bay watershed.[58] Our results show that high flow events can easily render undetectable smaller signals from decreasing loads attributable to management actions. Thus, efforts to mitigate the high flow events should be considered in addition to nutrient reduction efforts.

Similar large-scale efforts to those in Chesapeake Bay have been underway to reduce nutrient loads into other estuaries and coastal regions.[14] In estuaries including Narragansett Bay[9] and Tampa Bay[8] in the United States, where large load reductions from the primary nutrient sources were implemented, significant water quality improvements were widespread. The response in Chesapeake Bay is not as clear as in those systems but is more like the more nuanced responses to load reductions in other systems with more diverse nutrient sources. In Roskilde Fjord, Denmark, local immediate improvements in response to nutrient reductions in one region were followed by stabilization and a lack of major change in down-estuary regions,[59] somewhat similar to our long-term decreases plateauing in recent years (Figure 4). In Neuse River estuary in the United States,[60] climate variability, and challenges with managing loads from multiple source types has resulted in slower system response than hoped to a Total Maximum Daily Load for TN, suggesting many similarities to Chesapeake Bay dynamics. These results demonstrate the importance of sustaining those watershed-wide multi-source nutrient reduction efforts and continuing monitoring and analysis to identify change when and where it is happening. Our GAM-based approach allowed us to distinguish between the impact of different nutrient sources even with the large interannual flow variability that usually overwhelms the underlying long-term trends. Furthermore, because responses to nutrient reductions are not always spatially and temporally consistent, spatially targeted analyses such as this one are useful to reveal exactly when and where improvements are occurring or more work is needed.

ASSOCIATED CONTENT

ⓈⒾ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.1c05388.

> Labeled station maps, method change list, river load gage locations, example graphical details for GAM1a-4b model fits, text describing spatial and temporal location matching along with graphics, bottom trend results, and tabulation of percent of station trends explained by each GAM fit (PDF)
>
> Detailed statistical output for each GAM fit to all stations for TN and TP (XLSX)

AUTHOR INFORMATION

Corresponding Author

Rebecca R. Murphy — *University of Maryland Center for Environmental Science, Annapolis, Maryland 21401, United States;* ⓞ orcid.org/0000-0003-3391-1823; Email: rmurphy@chesapeakebay.net

Authors

Jennifer Keisman — *U.S. Geological Survey, Catonsville, Maryland 21228, United States*

Jon Harcum — *Tetra Tech, Inc., Fairfax, Virginia 22030, United States*

Renee R. Karrh — *Maryland Department of Natural Resources, Annapolis, Maryland 21401, United States*

Mike Lane — *Department of Biological Sciences, Old Dominion University, Norfolk, Virginia 23529-0050, United States*

Elgin S. Perry — *Statistics Consultant, Colonial Beach, Virginia 22443, United States*

Qian Zhang — *University of Maryland Center for Environmental Science, Annapolis, Maryland 21401, United States;* ⓞ orcid.org/0000-0003-0500-5655

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.est.1c05388

Notes

The authors declare no competing financial interest.
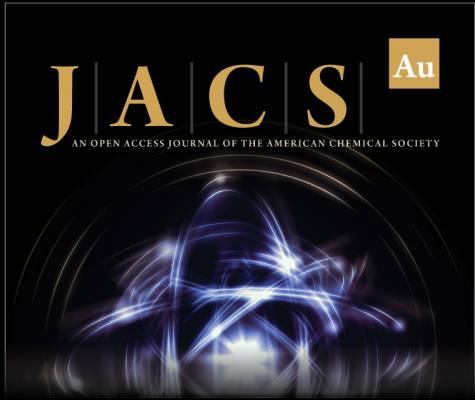
REFERENCES

(1) Diaz, R. J.; Rosenberg, R. Spreading Dead Zones and Consequences for Marine Ecosystems. *Science* **2008**, *321*, 926.

(2) Rabalais, N.; Cai, W.-J.; Carstensen, J.; Conley, D.; Fry, B.; Hu, X.; Quiñones-Rivera, Z.; Rosenberg, R.; Slomp, C.; Turner, E.; Voss, M.; Wissel, B.; Zhang, J. Eutrophication-Driven Deoxygenation in the Coastal Ocean. *Oceanography* **2014**, *27*, 172−183.

(3) Breitburg, D. L.; Hondorp, D. W.; Davias, L. A.; Diaz, R. J. Hypoxia, Nitrogen, and Fisheries: Integrating Effects Across Local and Global Landscapes. *Annu. Rev. Mar. Sci.* **2008**, *1*, 329−349.

(4) Rabalais, N. N.; Harper, D. E.; Turner, R. E., Responses of nekton and demersal and benthic fauna to decreasing oxygen concentrations. In *Coastal Hypoxia: Consequences for Living Resources and Ecosystems*, 58 ed.; Rabalais, N. N.; Turner, R. E., Eds. American Geophysical Union, Washington, D.C.: Washington, D.C., 2001; pp 115-128. DOI: 10.1029/ce058p0115

(5) Glibert, P. M.; Mayorga, E.; Seitzinger, S. Prorocentrum minimum tracks anthropogenic nitrogen and phosphorus inputs on a global basis: Application of spatially explicit nutrient export models. *Harmful Algae* **2008**, *8*, 33−38.

(6) Boesch, D. F. Barriers and Bridges in Abating Coastal Eutrophication. *Front. Mar. Sci.* **2019**, *6*, 123.

(7) Gross, C.; Hagy, J. D. Attributes of successful actions to restore lakes and estuaries degraded by nutrient pollution. *J. Environ. Manage.* **2017**, *187*, 122−136.

(8) Greening, H.; Janicki, A.; Sherwood, E. T.; Pribble, R.; Johansson, J. O. R. Ecosystem responses to long-term nutrient management in an urban estuary: Tampa Bay, Florida, USA. *Estuar. Coast Shelf Sci.* **2014**, *151*, A1−A16.

(9) Oviatt, C.; Smith, L.; Krumholz, J.; Coupland, C.; Stoffel, H.; Keller, A.; McManus, M. C.; Reed, L. Managed nutrient reduction impacts on nutrient concentrations, water clarity, primary production, and hypoxia in a north temperate estuary. *Estuar. Coast Shelf Sci.* **2017**, *199*, 25−34.

(10) Sattar, A.; Kroeze, C.; Strokal, M. The increasing impact of food production on nutrient export by rivers to the Bay of Bengal 1970-2050. *Mar. Pollut. Bull.* **2014**, *80*, 168−78.

(11) Tong, Y.; Zhao, Y.; Zhen, G.; Chi, J.; Liu, X.; Lu, Y.; Wang, X.; Yao, R.; Chen, J.; Zhang, W. Nutrient Loads Flowing into Coastal Waters from the Main Rivers of China (2006-2012). *Sci. Rep.* **2015**, *5*, 16678.

(12) Duarte, C. M.; Conley, D. J.; Carstensen, J.; Sánchez-Camacho, M. Return to Neverland: Shifting Baselines Affect Eutrophication Restoration Targets. *Estuar. Coast* **2009**, *32*, 29−36.

(13) Kemp, W. M.; Testa, J. M.; Conley, D. J.; Gilbert, D.; Hagy, J. D. Temporal responses of coastal hypoxia to nutrient loading and physical controls. *Biogeosciences* **2009**, *6*, 2985−3008.

(14) McCrackin, M. L.; Jones, H. P.; Jones, P. C.; Moreno-Mateos, D. Recovery of lakes and coastal marine ecosystems from eutrophication: A global meta-analysis. *Limnol. Oceanogr.* **2017**, *62*, 507−518.

(15) Carstensen, J.; Sánchez-Camacho, M.; Duarte, C. M.; Krause-Jensen, D.; Marbà, N. Connecting the Dots: Responses of Coastal Ecosystems to Changing Nutrient Concentrations. *Environ. Sci. Technol.* **2011**, *45*, 9122−9132.

(16) Cloern, J. E.; Abreu, P. C.; Carstensen, J.; Chauvaud, L.; Elmgren, R.; Grall, J.; Greening, H.; Johansson, J. O. R.; Kahru, M.; Sherwood, E. T.; Xu, J.; Yin, K. Human activities and climate variability drive fast-paced change across the world's estuarine-coastal ecosystems. *Global Change Biol.* **2016**, *22*, 513−529.

(17) Testa, J. M.; Clark, J. B.; Dennison, W. C.; Donovan, E. C.; Fisher, A. W.; Ni, W.; Parker, M.; Scavia, D.; Spitzer, S. E.; Waldrop, A. M.; Vargas, V. M. D.; Ziegler, G. Ecological Forecasting and the Science of Hypoxia in Chesapeake Bay. *BioScience* **2017**, *67*, 614−626.

(18) Zhang, Q.; Murphy, R. R.; Tian, R.; Forsyth, M. K.; Trentacoste, E. M.; Keisman, J.; Tango, P. J. Chesapeake Bay's water quality condition has been recovering: Insights from a multimetric indicator assessment of thirty years of tidal monitoring data. *Sci. Total Environ.* **2018**, *637-638*, 1617−1625.

(19) Langendorf, R. E.; Lyubchich, V.; Testa, J. M.; Zhang, Q. Inferring Controls on Dissolved Oxygen Criterion Attainment in the Chesapeake Bay. *ACS ES&T Water* **2021**, *1*, 1665−1675.

(20) Rowland, F. E.; Stow, C. A.; Johengen, T. H.; Burtner, A. M.; Palladino, D.; Gossiaux, D. C.; Davis, T. W.; Johnson, L. T.; Ruberg, S. Recent Patterns in Lake Erie Phosphorus and Chlorophyll a Concentrations in Response to Changing Loads. *Environ. Sci. Technol.* **2020**, *54*, 835−841.

(21) Testa, J. M.; Murphy, R. R.; Brady, D. C.; Kemp, W. M. Nutrient- and Climate-Induced Shifts in the Phenology of Linked Biogeochemical Cycles in a Temperate Estuary. *Front. Mar. Sci.* **2018**, *5*, 114.

(22) Tango, P. J.; Batiuk, R. A. Chesapeake Bay recovery and factors affecting trends: Long-term monitoring, indicators, and insights. *Reg. Stud. Mar. Sci.* **2016**, *4*, 12−20.

(23) Malone, T. C.; Newton, A. The Globalization of Cultural Eutrophication in the Coastal Ocean: Causes and Consequences. *Front. Mar. Sci.* **2020**, *7*, 670.

(24) Oelsner, G. P.; Stets, E. G. Recent trends in nutrient and sediment loading to coastal areas of the conterminous U.S.: Insights and global context. *Sci. Total Environ.* **2019**, *654*, 1225−1240.

(25) Bricker, S.; Longstaff, B.; Dennison, W.; Jones, A.; Boicourt, K.; Wicks, C.; Woerner, J. *Effects of Nutrient Enrichment in the Nation's Estuaries: A Decade of Change*; National Centers for Coastal Ocean Science, Silver Spring, MD: Silver Spring, MD, 2007; p 328.

(26) Harding, L. W.; Gallegos, C. L.; Perry, E. S.; Miller, W. D.; Adolf, J. E.; Mallonee, M. E.; Paerl, H. W. Long-Term Trends of Nutrients and Phytoplankton in Chesapeake Bay. *Estuar. Coast* **2016**, *39*, 664−681.

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

(28) Murphy, R. R.; Kemp, W. M.; Ball, W. P. Long-Term Trends in Chesapeake Bay Seasonal Hypoxia, Stratification, and Nutrient Loading. *Estuar. Coast* **2011**, *34*, 1293−1309.

(29) Testa, J. M.; Kemp, W. M.; Boynton, W. R. Season-specific trends and linkages of nitrogen and oxygen cycles in Chesapeake Bay. *Limnol. Oceanogr.* **2018**, *63*, 2045−2064.

(30) Testa, J. M.; Lyubchich, V.; Zhang, Q. Patterns and Trends in Secchi Disk Depth over Three Decades in the Chesapeake Bay Estuarine Complex. *Estuar. Coast* **2019**, *42*, 927−943.

(31) Chesapeake Bay Program. *Methods and Quality Assurance for Chesapeake Bay Water Quality Monitoring Programs*; CBP/TRS-319-17; May 2017, 2017.

(32) Chesapeake Bay Program. CBP Water Quality Database (1984-present) [Online], 2020. https://www.chesapeakebay.net/what/downloads/cbp_water_quality_database_1984_present (accessed July 9, 2019).

(33) Olson, M. *Guide to Using Chesapeake Bay Program Water Quality Monitoring Data*; EPA 903-R-12-001; Chesapeake Bay Program: Annapolis, MD, 2012.

(34) Hastie, T.; Tibshirani, R., Generalized Additive Models. In *Encyclopedia of Statistical Sciences*, Kotz, S.; Read, C. B.; Balakrishnan, N.; Vidakovic, B.; Johnson, N. L., Eds. John Wiley & Sons, Inc.: 2006.

(35) Murphy, R. R.; Perry, E.; Harcum, J.; Keisman, J. A Generalized Additive Model approach to evaluating water quality: Chesapeake Bay case study. *Environ. Model. Software* **2019**, *118*, 1−13.

(36) Wood, S. *Generalized Additive Models: An Introduction with R*, 1 ed.; Chapman and Hall/CRC, Boca Raton, FL: Boca Raton, FL, 2006; p 410.

(37) Moyer, D. L.; Langland, M. J., *Nitrogen, phosphorus, and suspended-sediment loads and trends measured at the Chesapeake Bay Nontidal Network stations: Water years 1985-2018 (ver. 2.0, May 2020): U.S. Geological Survey data release [Online], 2020.* https://doi.org/10.5066/P931M7FT (accessed Jan 15, 2021).

(38) Chesapeake Bay Program, Chesapeake Bay Total Maximum Daily Load (TMDL) Segmentsheds [Online], 2020. https://data-chesbay.opendata.arcgis.com/datasets/ChesBay::chesapeake-bay-total-maximum-daily-load-tmdl-segmentsheds (accessed Oct 20, 2020).

(39) Burnham, K. P.; Anderson, D. R.; Huyvaert, K. P. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav. Ecol. Sociobiol.* **2011**, *65*, 23−35.

(40) Murtaugh, P. A. In defense of P values. *Ecology* **2014**, *95*, 611−617.

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

(42) Wang, P.; Linker, L. C.; Shenk, G. W. Using Geographically Isolated Loading Scenarios to Analyze Nitrogen and Phosphorus Exchanges and Explore Tailored Nutrient Control Strategies for Efficient Management. *Environ. Model. Assess.* **2016**, *21*, 437−454.

(43) Testa, J. M.; Kemp, W. M.; Boynton, W. R.; Hagy, J. D. Long-Term Changes in Water Quality and Productivity in the Patuxent River Estuary: 1985 to 2003. *Estuar. Coast* **2008**, *31*, 1021−1037.

(44) Ruhl, H. A.; Rybicki, N. B. Long-term reductions in anthropogenic nutrients link to improvements in Chesapeake Bay habitat. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 16566.

(45) Wood, J. D.; Bukaveckas, P. A. Increasing Severity of Phytoplankton Nutrient Limitation Following Reductions in Point Source Inputs to the Tidal Freshwater Segment of the James River Estuary. *Estuar. Coast* **2014**, *37*, 1188−1201.

(46) Boynton, W. R.; Hagy, J. D.; Cornwell, J. C.; Kemp, W. M.; Greene, S. M.; Owens, M. S.; Baker, J. E.; Larsen, R. K. Nutrient Budgets and Management Actions in the Patuxent River Estuary, Maryland. *Estuar. Coast* **2008**, *31*, 623−651.

(47) Bukaveckas, P. A.; Beck, M.; Devore, D.; Lee, W. M. Climatic variability and its role in regulating C, N and P retention in the James River Estuary. *Estuar. Coast Shelf Sci.* **2018**, *205*, 161−173.

(48) Litke, D. W. *Review of Phosphorus Control Measures in the United States and Their Effect on Water Quality. Water-Resources Investigations Report 99−4007*; National Water-Quality Assessment Program, U.S. Geological Survey: Denver, Colorado, 1999.

(49) Fisher, T. R.; Fox, R. J.; Gustafson, A. B.; Koontz, E.; Lepori-Bui, M.; Lewis, J. Localized Water Quality Improvement in the Choptank Estuary, a Tributary of Chesapeake Bay. *Estuar. Coast* **2021**, *44*, 1274−1293.

(50) Da, F.; Friedrichs, M. A. M.; St-Laurent, P. Impacts of Atmospheric Nitrogen Deposition and Coastal Nitrogen Fluxes on Oxygen Concentrations in Chesapeake Bay. *J. Geophys. Res.: Oceans* **2018**, *123*, 5004−5025.

(51) Linker, L. C.; Dennis, R.; Shenk, G. W.; Batiuk, R. A.; Grimm, J.; Wang, P. Computing Atmospheric Nutrient Loads to the Chesapeake Bay Watershed and Tidal Waters. *J. Am. Water Resour. Assoc.* **2013**, *49*, 1025−1041.

(52) McGlathery, K.; Sundbäck, K.; Anderson, I. Eutrophication in shallow coastal bays and lagoons: the role of plants in the coastal filter. *Mar. Ecol.: Prog. Ser.* **2007**, *348*, 1−18.

(53) Zhang, Q.; Hirsch, R. M. River Water-Quality Concentration and Flux Estimation Can be Improved by Accounting for Serial Correlation Through an Autoregressive Model. *Water Resour. Res.* **2019**, *55*, 9705−9723.

(54) Lefcheck, J. S.; Orth, R. J.; Dennison, W. C.; Wilcox, D. J.; Murphy, R. R.; Keisman, J.; Gurbisz, C.; Hannam, M.; Landry, J. B.; Moore, K. A.; Patrick, C. J.; Testa, J.; Weller, D. E.; Batiuk, R. A. Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 3658−3662.

(55) Zhang, Q.; Fisher, T. R.; Trentacoste, E. M.; Buchanan, C.; Gustafson, A. B.; Karrh, R.; Murphy, R. R.; Keisman, J.; Wu, C.; Tian, R.; Testa, J. M.; Tango, P. J. Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management. *Water Res.* **2021**, *188*, 116407.

(56) Zhang, Q.; Tango, P. J.; Murphy, R. R.; Forsyth, M. K.; Tian, R.; Keisman, J.; Trentacoste, E. M. Chesapeake Bay Dissolved Oxygen Criterion Attainment Deficit: Three Decades of Temporal and Spatial Patterns. *Front. Mar. Sci.* **2018**, *5*, 422.

(57) Buchanan, C. A Water Quality Binning Method to Infer Phytoplankton Community Structure and Function. *Estuar. Coast* **2020**, *43*, 661−679.

(58) St.Laurent, K. A.; Coles, V. J.; Hood, R. R. Climate Extremes and Variability Surrounding Chesapeake Bay: Past, Present, and Future. *J. Am. Water Resour. Assoc.* **2021**, 1−29 (early view online).

(59) Staehr, P. A.; Testa, J.; Carstensen, J. Decadal Changes in Water Quality and Net Productivity of a Shallow Danish Estuary Following Significant Nutrient Reductions. *Estuar. Coast* **2017**, *40*, 63−79.

(60) Lebo, M. E.; Paerl, H. W.; Peierls, B. L. Evaluation of Progress in Achieving TMDL Mandated Nitrogen Reductions in the Neuse River Basin, North Carolina. *Environ. Manag.* **2012**, *49*, 253−266.