Water Research 188 (2021) 116407

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management

Qian Zhang[a,*], Thomas R. Fisher[b], Emily M. Trentacoste[c], Claire Buchanan[d], Anne B. Gustafson[b], Renee Karrh[e], Rebecca R. Murphy[a], Jennifer Keisman[f], Cuiyin Wu[g], Richard Tian[a], Jeremy M. Testa[h], Peter J. Tango[i]

[a] University of Maryland Center for Environmental Science / Chesapeake Bay Program, 410 Severn Avenue, Annapolis, MD 21403, USA
[b] University of Maryland Center for Environmental Science, Horn Point Laboratory, 2020 Horns Point Rd, Cambridge, MD 21613, USA
[c] U.S. Environmental Protection Agency, Chesapeake Bay Program Office, 410 Severn Avenue, Annapolis, MD 21403, USA
[d] Interstate Commission on the Potomac River Basin, 30 West Gude Drive, Suite 450, Rockville, MD 20850, USA
[e] Maryland Department of Natural Resources, 580 Taylor Ave, Annapolis, MD 21401, USA
[f] U.S. Geological Survey, MD-DE-DC Water Science Center, 5522 Research Park Drive, Catonsville, MD 21228, USA
[g] Chesapeake Research Consortium / Chesapeake Bay Program, 410 Severn Avenue, Annapolis, MD 21403, USA
[h] University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, 146 Williams Street, Solomons, MD 20688, USA
[i] U.S. Geological Survey / Chesapeake Bay Program, 410 Severn Avenue, Annapolis, MD 21403, USA

ARTICLE INFO

Article history:
Received 20 May 2020
Revised 4 September 2020
Accepted 6 September 2020
Available online 7 September 2020

Keywords:
Nutrient limitation
Phytoplankton
Bioassay
Water quality
Long-term monitoring
Nutrient management

ABSTRACT

Understanding the temporal and spatial roles of nutrient limitation on phytoplankton growth is necessary for developing successful management strategies. Chesapeake Bay has well-documented seasonal and spatial variations in nutrient limitation, but it remains unknown whether these patterns of nutrient limitation have changed in response to nutrient management efforts. We analyzed historical data from nutrient bioassay experiments (1992–2002) and data from long-term, fixed-site water-quality monitoring program (1990–2017) to develop empirical approaches for predicting nutrient limitation in the surface waters of the mainstem Bay. Results from classification and regression trees (CART) matched the seasonal and spatial patterns of bioassay-based nutrient limitation in the 1992–2002 period much better than two simpler, non-statistical approaches. An ensemble approach of three selected CART models satisfactorily reproduced the bioassay-based results (classification rate = 99%). This empirical approach can be used to characterize nutrient limitation from long-term water-quality monitoring data on much broader geographic and temporal scales than would be feasible using bioassays, providing a new tool for informing water-quality management. Results from our application of the approach to 21 tidal monitoring stations for the period of 2007–2017 showed modest changes in nutrient limitation patterns, with expanded areas of nitrogen-limitation and contracted areas of nutrient saturation (i.e., not limited by nitrogen or phosphorus). These changes imply that long-term reductions in nitrogen load have led to expanded areas with nutrient-limited phytoplankton growth in the Bay, reflecting long-term water-quality improvements in the context of nutrient enrichment. However, nutrient limitation patterns remain unchanged in the majority of the mainstem, suggesting that nutrient loads should be further reduced to achieve a less nutrient-saturated ecosystem.

1. Introduction

Nutrient enrichment has become a major issue in many coastal ecosystems around the world (Boesch et al., 2001; Cloern, 2001; Kemp et al., 2009; Boesch, 2019; Malone and Newton, 2020). Chesapeake Bay, the largest estuary in North America, has a long history of cultural eutrophication accompanied by excessive algal growth and seasonal hypoxia (Boesch et al., 2001; Hagy et al., 2004; Kemp et al., 2005). Such symptoms have been attributed in part to excessive nutrient and sediment loads from the watershed (Ator et al., 2020; Hagy et al., 2004; Lefcheck et al., 2018; Murphy et al., 2011; Noe et al., 2020; Zhang et al., 2018), as seen

* Corresponding author.
E-mail addresses: qzhang@chesapeakebay.net, zhangqian0324@gmail.com (Q. Zhang).

https://doi.org/10.1016/j.watres.2020.116407

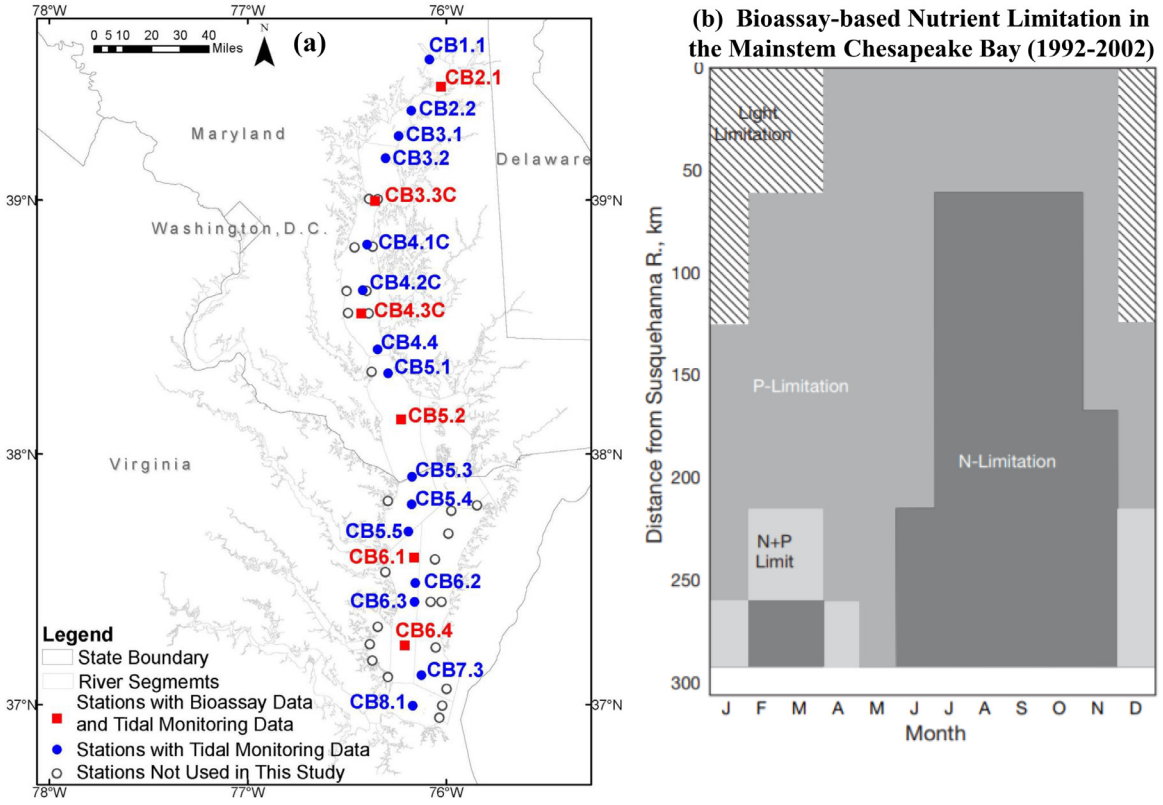

Fig. 1. Maps showing (a) the Chesapeake Bay Program's long-term tidal water-quality monitoring network and (b) nutrient limitation of phytoplankton growth in the mainstem Chesapeake Bay, as based on bioassay data collected between 1992 and 2002 (reproduced from Kemp et al. (2005) with permission). For clarity, only mainstem stations are shown in panel (a). Stations analyzed in this work are shown as filled squares and filled circles.

in many other coastal systems (Boesch, 2019). Consequently, coordinated management efforts among the Bay state jurisdictions began in the 1980s to reduce nutrient and sediment loads from the watershed. Furthermore, the Chesapeake Bay Total Maximum Daily Load (TMDL) was established in 2010 to enforce load reductions in order to improve habitat health that can fully support living resource survival, growth, and reproduction (U.S. Environmental Protection Agency, 2010). The TMDL is designed to ensure that all pollution control measures needed to fully restore the Bay and its tidal tributaries are in place by 2025.

To control algal growth in aquatic ecosystems, it is critical to understand growth limiting factors, which may be nitrogen (N), phosphorus (P), light (L), silicon (Si), temperature, or others (Cloern et al., 2014; Conley, 1999; Elser et al., 2007; Hecky and Kilham, 1988). In general, fresh waters are often considered to be P-limited due to a high N:P ratio in riverine sources (e.g., Fennel and Testa, 2019) and N fixation (Hecky and Kilham, 1988; Schindler, 1974), whereas marine waters are often considered to be primarily N-limited due to sediment release of P, slow N fixation, and high denitrification (Hecky and Kilham, 1988; Paerl, 2018). Estuaries are in the transitional zones between freshwater and marine systems, thereby often exhibiting more complicated patterns in nutrient limitation (Cloern et al., 2014; Elser et al., 2007). For example, tidal waters of Chesapeake Bay have shown occurrences of N-, P-, L-, and Si-limitation that vary in space and time (Conley and Malone, 1992; Fisher et al., 1999, 1992; Malone et al., 1996).

To understand the seasonal and spatial variations in nutrient limitation in Chesapeake Bay, bioassay experiments were conducted from a set of locations (Fig. 1a) between 1992 and 2002 under controlled, near-surface light conditions (Fisher and Gustafson, 2003,2005; Fisher et al., 1999). The results were synthesized in Kemp et al. (2005) to illustrate the seasonal and spatial variations in nutrient limitation in that period (Fig. 1b). Such patterns demonstrate the necessity of a dual nutrient management strategy (U.S. Environmental Protection Agency, 2010) and have informed the development of the Chesapeake Bay Estuary Model (Cerco and Noel, 2017).

In the Chesapeake Bay watershed, large-scale dual nutrient reduction goals have been in place for decades, leading to an overall long-term reduction in total N (TN) load to the Bay (Ator et al., 2019; Chanat et al., 2016; Chanat and Yang, 2018; Hirsch et al., 2010; Zhang et al., 2016a, 2015). Total P (TP) load to the Bay declined in the 1980s to early 1990s, due to improved wastewater treatment and phosphate-detergent bans. However, TP load increased dramatically since the mid-1990s, largely due to reduced sediment trapping as sediment storage has neared full capacity in Conowingo Reservoir on the Susquehanna River, the largest tributary to the Bay (Hirsch, 2012; Langland, 2015; Zhang et al., 2013, 2016b). While the N load reduction is primarily associated with nitrate, the P load increase from the Susquehanna River is primarily associated with particulate P, although dissolved orthophosphate load from the Susquehanna has also increased substantially in recent years (Fanelli et al., 2019). This contrast in the direction of N and P trends and the form-specific differences suggests that there have been changes in the ratios of nutrients for phytoplankton growth and thus changes in nutrient limitation on a Bay-wide scale. Unfortunately, bioassay experiments have not been conducted since 2005 to characterize recent conditions of nutrient limitation.

Our main objective was to test the above hypothesis through an analysis of historical data from nutrient bioassay experiments (1992–2002) and surface-collected data from the Chesapeake Bay Program (CBP) long-term, fixed-site water-quality monitoring network (1990–2017). The specific goals are as follows:

(1) To develop empirical approaches to predict bioassay-based measures of nutrient limitation ("response variable") using tidal water-quality monitoring data in the concurrent period of 1992–2002 ("Goal 1"), and
(2) To apply the selected approach to tidal water-quality monitoring data in more recent periods without bioassay data to predict nutrient limitation and explore potential changes in response to altered nutrient loading ("Goal 2").

This research provides updated information specific to Chesapeake Bay management and research communities regarding the spatial and seasonal patterns in nutrient limitation. More broadly, we propose that this approach can be applied in any estuary with similar long-term monitoring data to address similar questions.

2. Data

2.1. Bioassays

Bioassay experiments were conducted between 1992 and 2002 to characterize the response of algal growth or biomass accumulation to nutrient additions at six mainstem stations (Fisher and Gustafson, 2003,2005; Fisher et al., 1999). These stations, namely, CB2.1, CB3.3C, CB4.3C, CB5.2, CB6.1, and CB6.4, had 18, 245, 264, 265, 13, and 13 bioassay samples, respectively. Kemp et al. (2005) synthesized these bioassay data to characterize the nutrient limitation class for each station-month combination for the period of 1992–2002 (Fig. 1b). This data set of nutrient limitation has 72 records (i.e., 6 stations x 12 months = 72 records).

Among the limitation classes, N- and P-limitations were defined as increases in phytoplankton productivity (PP) and biomass (PB) with N and P additions, respectively, normalized to controls. NP-limitation was defined as increases in PP and PB only with simultaneous additions of N and P. L-limitation was defined as no differential growth response between the controls and NP treatments, although all grew equally well when held at 60% light due to high ambient N and P concentrations. To be more precise, we changed the term "L-limitation" to "NoR" (i.e., No significant Response to nutrient additions)". This category contains, but is not limited to, L-limitation, although other limitation factors (e.g., trace metals and Si) were unlikely. Hereafter, the response categories reported are NoR, N-limitation, P-limitation, and NP-limitation.

2.2. Water-quality monitoring data

The CBP partnership maintains a long-term tidal water-quality monitoring program (Fig. 1a) at 100+ stations distributed along the mainstem Bay and its tidal tributaries (Tango and Batiuk, 2016; U.S. Environmental Protection Agency, 2010). These stations have been sampled consistently since 1985. The field and laboratory methods and quality-control protocols are documented in Chesapeake Bay Program (2017). We downloaded tidal water-quality monitoring data for the period of 1990–2017 from the CBP Water Quality Database (https://datahub.chesapeakebay.net/), which consists of more than three million values of various physical, chemical, and biological parameters. These downloaded data, as well as subsequently processed data, are archived in Zhang (2020). The availability of initially selected parameters at the mainstem stations is summarized in **Figure S1**.

Downloaded data in the period of 1992–2002 were processed to produce an aggregated data set for each station-month combination (i.e., 6 stations x 12 months = 72 records) – see "Data Manipulation" in Supplementary Materials for details. This data set and the bioassay-based data set, both having 72 records, were used in our empirical approach development (Goal 1).

Similarly, downloaded data in the period of 2007–2017 were processed to produce an aggregated data set for analysis for Goal 2. This period was chosen because it has an average annual flow (77,850 $m^3\ s^{-1}$) comparable to 1992–2002 (77,570 $m^3\ s^{-1}$), thereby allowing for a fair comparison between the two periods regarding nutrient limitation given the strong association of river flow and nutrient availability. Moreover, the above data manipulations for both periods were expanded to a larger suite of stations (n = 21), including the six stations that had bioassay data (Fig. 1a).

3. Methods

For Goal 1, three empirical approaches were developed to relate tidal monitoring data collected during the period from 1992 to 2002 to bioassay-based nutrient limitation classes at the six mainstem stations developed using bioassay data from the same period (Sections 3.1-3.3). We considered two non-statistical approaches (A1 and A2) that used only dissolved inorganic N (DIN) and dissolved inorganic P (DIP) data and one statistical approach (A3) that used DIN, DIP, and other monitored variables – see Fig. 2 for an overview of these approaches and associated analyses. For Goal 2, the selected approach was applied to tidal monitoring data from 2007 to 2017 to explore potential changes in nutrient limitation (Section 3.4).

3.1. Probability-based approach (A1)

The first approach is a probability-based approach (A1). For each of the 72 station-month pairs, all DIN and DIP concentration data from 1992 to 2002 were extracted from the originally downloaded data. These concentrations were compared with the bloom-limiting thresholds for DIN (0.07 mg l^{-1}) and DIP (0.007 mg l^{-1}), respectively, to compute the probabilities of below-threshold DIN and DIP for each station and month:

(1) Probability of below-threshold DIN = count of DIN values less than 0.07 mg l^{-1} / count of all DIN values;
(2) Probability of below-threshold DIP = count of DIP values less than 0.007 mg l^{-1} / count of all DIP values.

It was previously documented that concentrations above these thresholds are associated with a lack of response to nutrient additions (Buchanan et al., 2005; Fisher and Gustafson, 2003). These thresholds were derived from bioassays as general reference values for algal response, but they may not lead to accurate predictions of nutrient limitation at all locations or in all months due to spatial and temporal variability in controls of nutrient limitation.

The computed probabilities were then converted to nutrient limitation classes, as follows:

(1) If the probability of below-threshold DIN (i.e., DIN < 0.07 mg l^{-1}) ≥ 50%, class = N;
(2) If the probability of below-threshold DIP (i.e., DIP < 0.007 mg l^{-1}) ≥ 50%, class = P;
(3) If both probabilities ≥ 40%, class = NP;
(4) Else, class = NoR.

These probability-based classes were compared with the bioassay-based classes (n = 72) to compute the classification rate (i.e., [number of matched cases / 72] x 100), which is served as a quantitative measure of the performance of approach A1.

3.2. Index-based approach (A2)

The second approach is an index-based approach (A2). Like A1, for each of the 72 station-month pairs, all DIN and DIP concentration data from 1992 to 2002 were extracted from the originally downloaded data. These concentrations were compared with

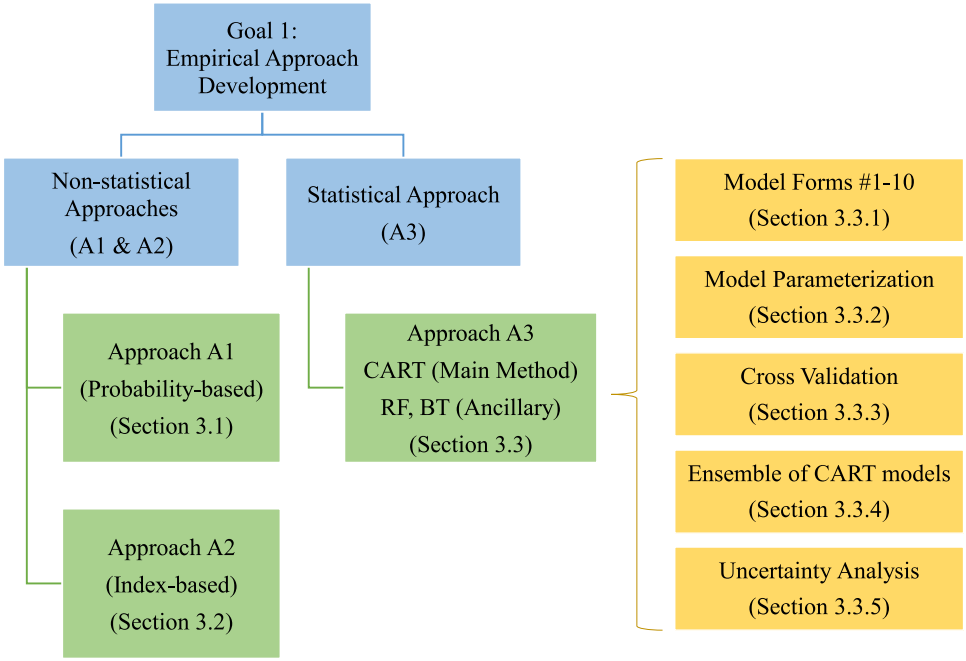

Fig. 2. Overview of the three empirical approaches (A1, A2, A3) considered under Goal 1. For approach A3, classification and regression trees (CART), random forest (RF), and boosted trees (BT) were evaluated for ten different model forms, respectively (see Table 1).

the bloom-limiting thresholds to compute N and P indices, as follows:

(1) If DIN $<$ 0.07 mg l^{-1}, N index = 1 and P index = 0;
(2) If DIP $<$ 0.007 mg l^{-1}, N index = 0 and P index = 1;
(3) If both conditions are met, N index = 0.5 and P index = 0.5;
(4) If neither condition is met, N index = 0 and P index = 0.

For each station-month pair, the relevant indices were averaged over the period of record to compute the aggregated indices, which were converted to nutrient limitation classes, as follows:

(1) If N index $\geq$ 0.5, class = N;
(2) If P index $\geq$ 0.5, class = P;
(3) If both indices $\geq$ 0.4, class = NP;
(4) Else, class = NoR.

These index-based classes were compared with the bioassay-based classes (n = 72) to compute the classification rate of approach A2.

3.3. Tree-based approach (A3)

The third approach is a statistical approach (A3), which is centered on the classification and regression tree (CART). CART has three unique advantages: (1) it is a well-established machine learning method for classification (Breiman, 1984; Loh, 2014) and has been widely used in hydrological and marine studies (e.g., Testa et al., 2019; Zhang et al., 2019), (2) it can consider multiple independent variables simultaneously and thus utilize information beyond DIN and DIP concentrations or indices (e.g., water temperature), and (3) it allows more flexibility in relating the bioassay-based nutrient limitation classes (our response variable) to the independent variables, including multiple cutoffs and interactions (e.g., a variable can appear multiple times on the tree with different splitting thresholds).

We also considered more advanced tree approaches, namely, random forest (RF) and boosted trees (BT). RF builds multiple trees and glues them together to get a potentially more stable prediction. By contrast, BT grows trees sequentially – i.e., each tree is grown using information from previously grown trees. Our results showed that RF and BT did not have higher classification rates than CART (**Figure S3**). Thus, CART was chosen as A3 for its ability to illustrate decision trees.

3.3.1. Model forms

All monitored parameters in the processed data set (Section 2.2) were evaluated for collinearity and relative importance. Correspondingly, ten model forms were proposed, with increasing levels of complexity (i.e., covariates were added sequentially) (Table 1). In model 10 ("full model"), the covariates are above-mentioned N index (indexN) and P index (indexP), water temperature (WTEMP, °C), TN:TP molar ratio (TNTP.ratio), DIP concentration (DIP, mg l^{-1}), salinity (SALINITY, ppt), Secchi disk depth (SECCHI, m), chlorophyll-a (CHLA, μg l^{-1}), dissolved silicate (SIF, mg l^{-1}), January-May average streamflow from the Susquehanna River (Q15, m^3 s^{-1}; where Q = flow, 1 = January, and 5 = May), and Season. We excluded other variables (e.g., DIN, TN, TP, DIN:DIP, and dissolved oxygen) because they had (a) low variable importance, (b) high correlations (i.e., Spearman correlation $>$ 0.85) with one or more of the variables already included, or (c) large presence of below detection-limit values (in the case of DIN; **Figure S2**). Note that WTEMP is a numerical variable that accounts for seasonal changes in water temperature, whereas Season is a categorical (i.e., discrete) variable for calendar month that characterizes features of the annual cycle affecting nutrient limitation (e.g., day length and daylight).

3.3.2. Model parameterization

The CART tree includes the top node (which is at the top of the tree, containing all data records), terminal nodes (which are at the bottom of the tree), and intermediate nodes. At every node except the terminal nodes, CART evaluates all possible split points for each predictor variable and selects the predictor variable and split point that leads to two most dissimilar subgroups. The selected variable and split point correspond to the splitting variable and the splitting rule, respectively. Ideal CART trees should not

Table 1
Model forms used for the tree-based approaches.[a]

Model No.	Model form
#1	*Class ~ indexN + indexP*
#2	*Class ~ indexN + indexP + WTEMP*
#3	*Class ~ indexN + indexP + WTEMP + TNTP.ratio*
#4	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP*
#5	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY*
#6	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY + SECCHI*
#7	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY + SECCHI + CHLA*
#8	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY + SECCHI + CHLA + SIF*
#9	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY + SECCHI + CHLA + SIF + Q15*
#10	*Class ~ indexN + indexP + WTEMP + TNTP.ratio + DIP + SALINITY + SECCHI + CHLA + SIF + Q15 + Season*

[a] In model 10 ("full model"), the covariates are N index (indexN), P index (indexP), water temperature (WTEMP, °C), TN:TP molar ratio (TNTP.ratio), DIP concentration (DIP, mg l^{-1}), salinity (SALINITY, ppt), Secchi disk depth (SECCHI, m), chlorophyll-a (CHLA, μg l^{-1}), dissolved silicate (SIF, mg l^{-1}), January-May average streamflow from Susquehanna River (Q15, m^3 s^{-1}; where Q = flow, 1 = January, and 5 = May), and Season. Parameters not displayed in model 10 were excluded because they had (a) low variable importance, (b) high correlations (i.e., spearman correlation > 0.85) with one or more of the variables already included, or (c) large presence of below detection-limit values (in the case of DIN; Figure S2).

be too large (over-fit) or too small (poorly-fit), and should have informative splitting rules and pure terminal nodes. Because our data set has only 72 records (6 stations x 12 months), we set the CART parameter *minsplit* to six and *maxdepth* to five. *minsplit* defines the minimum number of observations that must exist in a tree node in order for a split to be attempted. *maxdepth* defines the maximum depth of any node of the final tree (root node = depth zero). We set the complexity parameter *cp* to its default value, 0.01. This means that only splits that decreased the overall lack of fit by a factor of 0.01 or more were attempted. CART analyses were conducted using the "rpart" package (Therneau and Atkinson, 2019).

3.3.3. Cross validation

To quantify the performance of the proposed model forms, K-fold cross-validation analysis was conducted, which can give insights on how well the model generalizes to a new data set. K defines the number of subgroups that the entire data set is split into:

(1) $K = 72$: leave 1 of 72 out each time, essentially the leave-one-out cross-validation (LOOCV) approach. In each iteration, one data record was left out and treated as "new data". The remaining 71 data records were used to build a model and that model was used to make the prediction for the left-out data. In total, 72 models were built. From each of the 72 models, prediction was made for the left-out case, resulting in a vector of 72 predicted nutrient limitation classes.
(2) $K = 24$: leave 3 of 72 out each time; 24 models were built.
(3) $K = 12$: leave 6 of 72 out each time; 12 models were built.
(4) $K = 6$: leave 12 of 72 out each time; 6 models were built.

For each K, predictions were compared with the bioassay-based nutrient limitation classes ($n = 72$) to compute the classification rate for each model form to determine the best model forms for CART.

3.3.4. Ensemble of CART models

The selected CART model forms have different decision trees and different best-performing zones (locations and months). Thus, an ensemble approach was developed to combine the strengths of the candidate models. Essentially, this approach compares predictions of nutrient limitation for each station-month pair from each candidate model and selects the predicted nutrient limitation class that has the highest probability. For cases where predictions from the three candidate models are identical, it reports that prediction. For cases where model predictions diverge, it extracts the probabilities of each model's prediction and reports the prediction with the highest probability.

3.3.5. Uncertainty analysis

To quantify the uncertainties associated with the predictions of each CART model form as well as the ensemble model, predictions from the LOOCV approach were utilized. Specifically, each set of predictions from each of the 72 LOOCV models was considered one realization. The likelihood of the predicted nutrient limitation class was computed as "(count of occurrence / 72) x 100". If all 72 LOOCV models predict the same class, the likelihood would be (72/72) x 100 = 100%. In this work, a predicted nutrient limitation class is considered "uncertain" if its likelihood is less than 70%.

3.4. Application to new periods

The ensemble approach of CART was applied to new periods and locations to predict nutrient limitations. Specifically,

- 2007–2017 for the six stations that had bioassay data. This period was chosen because it has an average annual freshwater flow entering the Bay (77,850 m^3 s^{-1}) comparable to 1992–2002 (77,570 m^3 s^{-1}), according to the annual mean freshwater flow entering the Bay (https://www.usgs.gov/centers/cba/science/freshwater-flow-chesapeake-bay).
- 2007–2017 for the full set of mainstem stations ($n = 21$) (Fig. 1a).

In addition, both hydrologic conditions and temporal periods are expected to affect nutrient availability, thereby leading to changes in nutrient limitation in the Bay. To isolate these effects, we conducted two controlled experiments with subsets of the monitoring data representative of different conditions:

- In the first experiment, we fixed the period to 2003–2017 and varied the hydrologic condition. Each year in 2003–2017 was classified as wet, dry, or average, depending on the annual mean freshwater flow entering the Bay (https://www.usgs.gov/centers/cba/science/freshwater-flow-chesapeake-bay). Average years have annual flows falling between the 25th and 75th percentiles of the entire record.
- In the second experiment, we fixed the hydrologic condition by selecting four 2-year periods that had similar annual mean freshwater flows to the Bay, namely, 1990–1991 (79,800 m^3 s^{-1}), 1998–1999 (75,800 m^3 s^{-1}), 2007–2008 (76,900 m^3 s^{-1}), and 2013–2014 (78,650 m^3 s^{-1}).

Finally, the CART-predicted nutrient limitation classes from these analyses were converted to nutrient limitation maps, using a rectangular-grid approach that is illustrated in **Figure S4**.

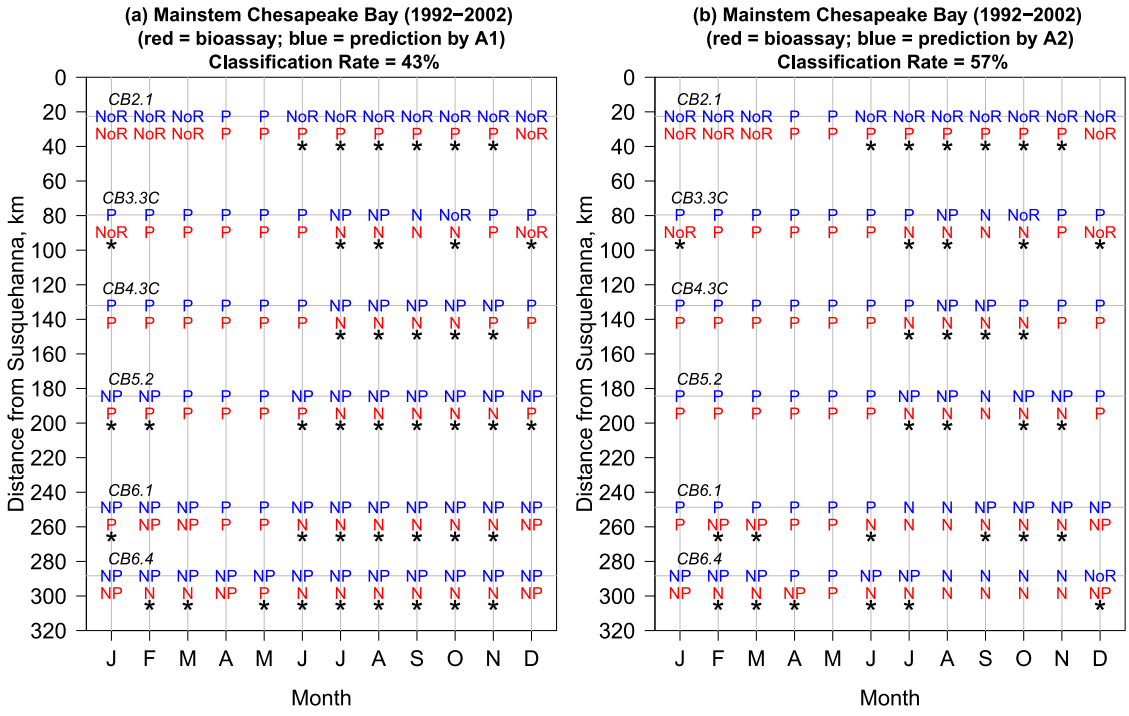

Fig. 3. Nutrient limitation diagrams for the six mainstem stations of Chesapeake Bay, comparing bioassay-based nutrient limitation classes and (a) probability-based classes (approach A1) or (b) index-based classes (approach A2). Mismatches are marked by "*".

4. Results

4.1. Performances of A1 and A2 (Goal 1)

The probability-based approach (A1) showed a classification rate of 43%, where 31 predictions based on tidal monitoring data matched bioassay-based limitation classes out of 72 cases (Fig. 3a). This approach matched most of the bioassay-based classes at the six stations in January through May but matched very poorly in July through November. The index-based approach (A2) showed a classification rate of 57% (i.e., 41 matches) (Fig. 3b). While A2 had a higher classification rate than A1, it also showed widespread mismatches in July through November.

4.2. Selection of tree models (Goal 1)

CART model forms were selected based on the cross-validation analyses. Models 4 and 7 both had the highest classification rate (85%) based on LOOCV (**Figure S3**). Other K-fold cross-validation analyses showed that models 4, 7, and 10 were generally better than or comparable to the other model forms (**Figure S5**). Thus, these three models were selected as our candidate models for CART.

The candidate models for CART showed much improved performance compared to A1 and A2. With LOOCV, models 4, 7, and 10 showed a classification rate of 85%, 85%, and 82%, respectively (**Figure S6**). Mismatches tended to be on the edges of the diagram (i.e., colder months and salinity extremes) and associated with the inability to discern NoR from P-limitation or NP-limitation from single nutrient limitation. Model 10 matched with bioassay-based classes better than model 4 and model 7 at CB6.1 and CB6.4 (the two southernmost stations) but worse at CB2.1 (the most upstream station). With the full data set, models 4, 7, and 10 showed a classification rate of 96%, 94%, and 97%, with only three, four, and two mismatches, respectively (**Figure S7**).

4.3. Tree plots of selected models (Goal 1)

Tree plots were produced to split the entire data set to more pure subsets. For model 4 (Fig. 4), TNTP.ratio is the top splitting variable, with WTEMP, indexN, and DIP as additional splitting variables. At the top node, cases are split to N-limitation (node 2) if TNTP.ratio < 51 and P-limitation (node 3) if TNTP.ratio ≥ 51. Node 2 is split by WTEMP, with warmer months (≥ 12 °C) classified as N-limitation (node 4) and colder months (< 12 °C) as NP-limitation (node 5). At node 5, cases are split to P-limitation (node 11) if indexN < 0.35 and NP-limitation (node 10) if indexN ≥ 0.35. At node 3, cases are split to three classes when WTEMP < 7.5 °C, depending on DIP. Specifically, these classes are NoR (if DIP ≥ 0.0062 mg l^{-1}), NP-limitation (DIP < 0.0018 mg l^{-1}), and P-limitation (DIP ≥ 0.0018 and < 0.0062 mg l^{-1}). At the bottom, there are seven terminal nodes and five of them are pure (with a unique limitation class). For model 7 (**Figure S8**), TNTP.ratio is also the top splitting variable, with WTEMP, indexP, DIP, and CHLA as additional splitting variables. Like model 4, node 2 is split by WTEMP, with warmer months (≥ 12 °C) classified as N-limitation and colder months (< 12 °C) as NP-limitation. This tree has eight terminal nodes and six of them are pure. For model 10 (**Figure S9**), TNTP.ratio is again the top splitting variable, with Season, indexN, DIP, and CHLA as additional splitting variables. In general, this tree is similar to model 7 and also has eight terminal nodes. One key difference is that WTEMP is replaced by Season as a splitting variable.

4.4. The ensemble approach (Goal 1)

Given that the three candidate models have different decision trees and different best-performing zones (locations and months) (**Figures S6-S7**), the ensemble approach can combine the strengths of the candidate models. The ensemble approach achieved a classification rate of 89% with LOOCV (8 mismatches) and 99% with the full data set (1 mismatch), both of which were higher than any

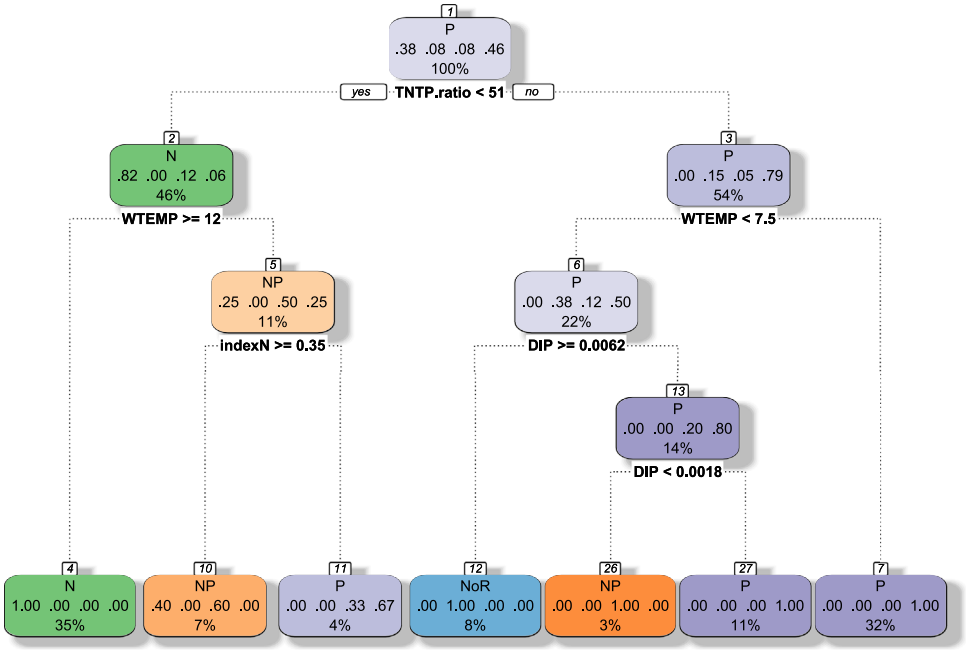

Fig. 4. CART tree plot of model 4 using the full data set. At each node, the four fractional numbers represent the probabilities of N-limitation, NoR, NP-limitation, and P-limitation, respectively, and the percentage number represents the proportion of data in that node (which is 100% at the top node).

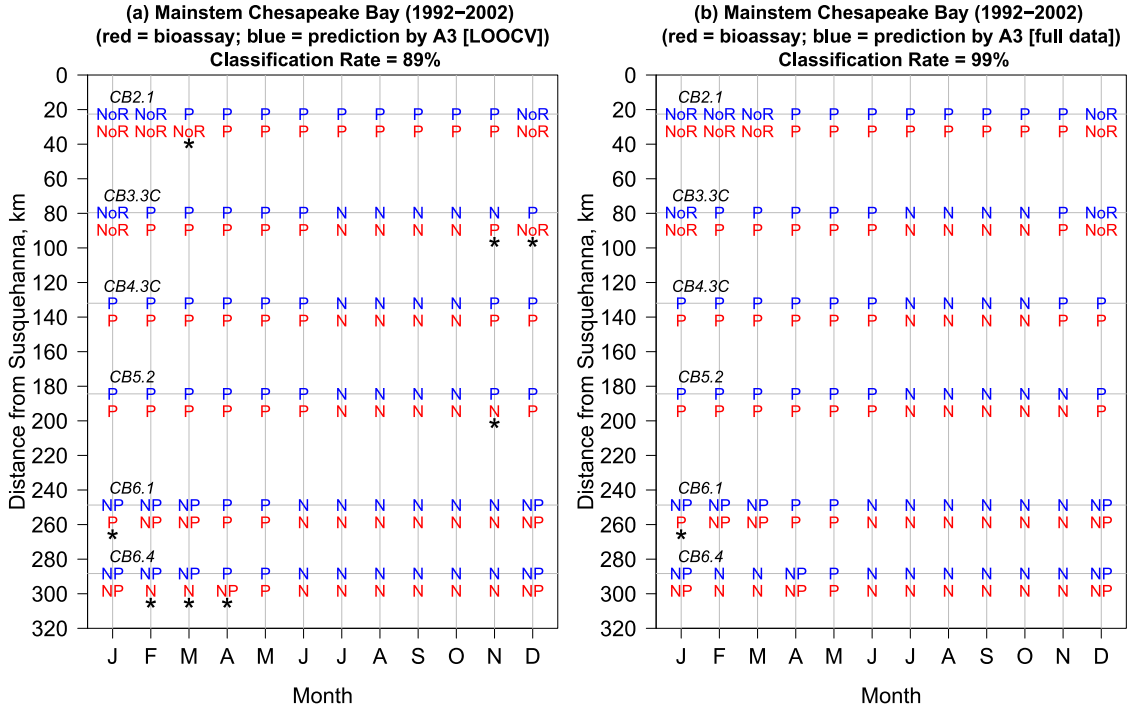

Fig. 5. Nutrient limitation diagrams for the six mainstem stations of Chesapeake Bay, comparing bioassay-based nutrient limitation classes and CART-based classes (approach A3) under (a) LOOCV and (b) the full data set. The CART-predicted classes were obtained from the ensemble approach that incorporated predictions from all three candidate models. Mismatches are marked by "*".

of the candidate models (Fig. 5). Consequently, predictions from the ensemble approach resulted in nutrient limitation maps that closely matched the bioassay-based limitation map for the period of 1992–2002 (Fig. 6).

4.5. Comparison of decadal periods (Goal 2)

For the six stations that had bioassay data, 10 of the 72 station-month pairs showed changes in nutrient limitation (Table 2; **Figure S10**), including switches from NoR to P-limitation ($n = 2$; upper Bay in winter-spring), P-limitation to N-limitation ($n = 4$; upper and middle Bay in summer-fall), NP-limitation to P-limitation ($n = 2$; lower Bay in winter-spring), and N-limitation to NP-limitation ($n = 2$; lower Bay in winter-spring). These cases all have high likelihoods based on the uncertainty analysis (Table 2). Overall, nutrient limitation showed expanded areas of N-limitation and contracted areas of NoR in 2007–2017 compared to 1992–2002 (Fig. 7).

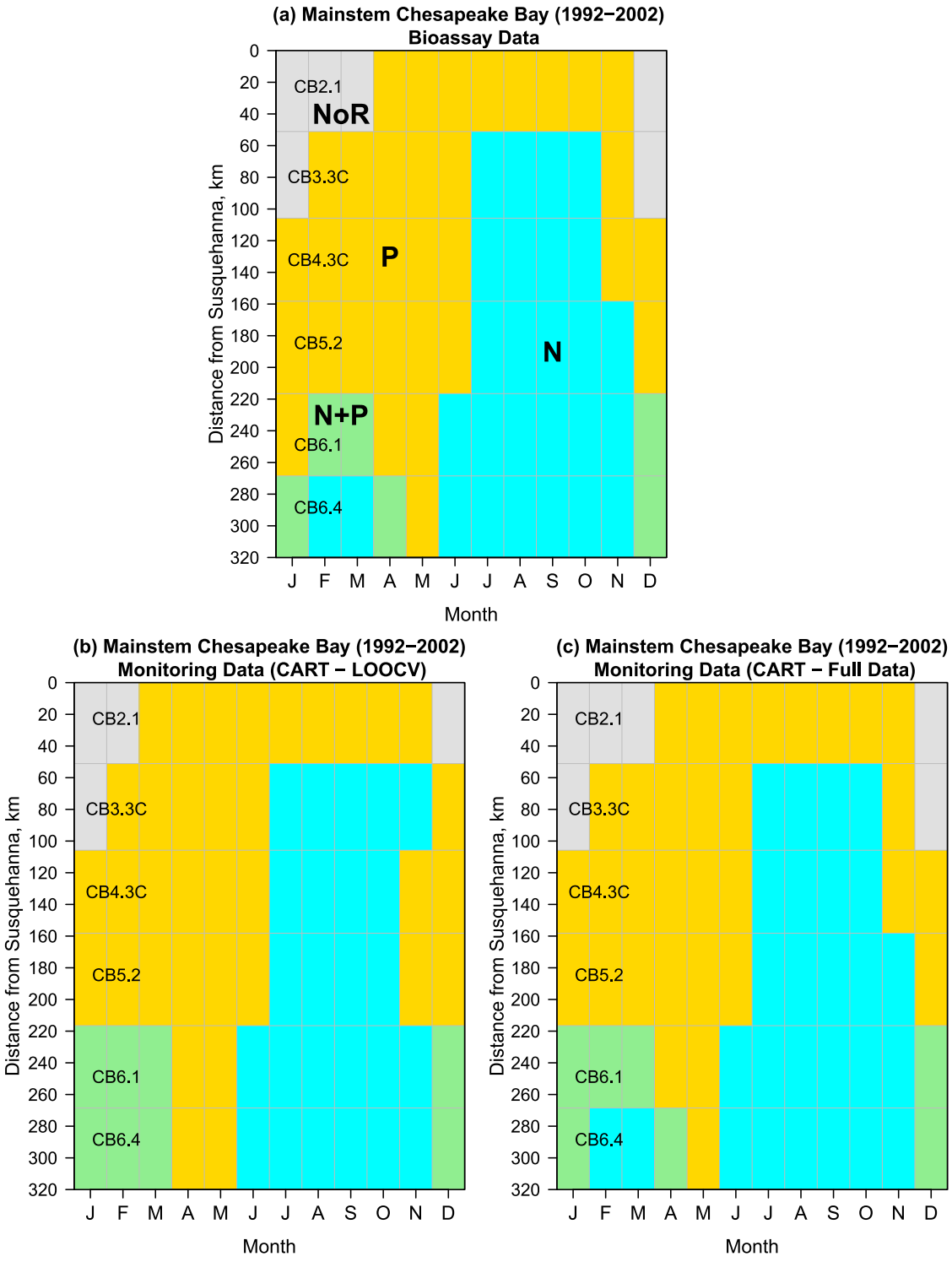

Fig. 6. Nutrient limitation maps for the mainstem Chesapeake Bay for the period of 1992–2002, as based on (a) bioassay data, (b) CART with LOOCV, and (c) CART with the full data set. The CART-based maps were developed based on the ensemble approach that incorporated predictions from all three candidate models.

In our extended analysis to the 21 tidal monitoring stations, including the six stations that had bioassay data, 30 pairs (12% of the 252 station-month pairs) showed changes in nutrient limitation. These changes are generally similar in directions to those reported above (i.e., from NoR to P-limitation, P-limitation to N-limitation, NP-limitation to P-limitation, and N-limitation to NP-limitation). Moreover, these cases all have high likelihoods based on the uncertainty analysis. The resulting maps (**Figure S11**), like Fig. 7, showed expanded areas of N-limitation and contracted areas of NoR in 2007–2017 compared to 1992–2002.

4.6. Comparison of different conditions (Goal 2)

Two controlled experiments were conducted to investigate changes in nutrient limitation as a function of hydrology or temporal periods. In the first experiment, we fixed the analysis period to 2003–2017 and varied the hydrologic condition (Fig. 8). Compared

Table 2
Summary of station-month pairs that showed changes in CART-predicted nutrient limitation class between the 1992–2002 period and the 2007–2017 period. [a]

Station	Season	Estimated nutrient limitation in 1992–2002		Estimated nutrient limitation in 2007–2017		Estimated change (Δ [c]) in DIN between 1992–2002 and 2007–2017		Estimated change (Δ [c]) in DIP between 1992–2002 and 2007–2017	
		Limitation	Likelihood [b]	Limitation	Likelihood [b]	Δ, mg l^{-1}	p-value	Δ, mg l^{-1}	p-value
CB2.1	Aug.	P	100%	N	94%	−0.073	< 0.01	0.000	0.810
CB2.1	Sep.	P	100%	N	96%	−0.080	< 0.01	0.003	0.066
CB3.3C	Jan.	NoR	99%	P	100%	−0.153	< 0.01	−0.003	0.016
CB3.3C	Jun.	P	100%	N	93%	−0.112	< 0.01	−0.002	< 0.01
CB3.3C	Dec.	NoR	99%	P	100%	−0.149	< 0.01	−0.003	0.053
CB4.3C	Nov.	P	100%	N	100%	0.013	0.302	0.001	< 0.01
CB6.1	Dec.	NP	100%	P	86%	−0.013	< 0.01	−0.0001	0.017
CB6.4	Feb.	N	99%	NP	100%	−0.018	< 0.01	−0.0003	< 0.01
CB6.4	Mar.	N	99%	NP	100%	0.0003	< 0.01	−0.0002	0.020
CB6.4	Apr.	NP	99%	P	86%	−0.030	< 0.01	0.0000	0.998

[a] These cases of change are also presented in Figure S10.
[b] Likelihood values of CART-predicted classes were obtained from the uncertainty analysis (Section 3.3.5).
[c] Estimated changes (Δ) between periods in DIN and DIP from Generalized Additive Model fits (Murphy et al., 2019).

with wet years, dry years expanded the areas of NP-limitation (additional 10 out of 252 cases; +4%) and N-limitation (+12%) and reduced the areas of NoR (−6%) and P-limitation (−10%). Specifically, a majority of NoR (in wet years) switched to P-limitation (dry years) in the upper Bay in winter months (December-February). N-limitation extended from the July-September window (wet years) further into November before shifting back to P-limitation in December (dry years) along most of the Bay's axis. Moreover, additional cases of NP-limitation (dry years) occurred primarily in locations and months that were P-limited (wet years).

In the second experiment, we fixed the hydrologic condition by selecting four 2-year periods that had similar annual freshwater flows to the Bay (Fig. 9). Compared with 1990–1991, the latter periods all showed contracted areas of P-limitation (−9%) and expanded areas of NP-limitation (+6% to +10%), with the latter occurring primarily in locations and months that had previously been P-limited.

5. Discussion

Large-scale nutrient reduction goals have been in place for decades as targets to restore water quality and habitat health in Chesapeake Bay and many other estuaries worldwide (Boesch et al., 2001; Cloern, 2001; Kemp et al., 2009; Boesch, 2019; Malone and Newton, 2020). Management actions resulting in reductions of nutrient loads may have led to increases in nutrient limitation of phytoplankton growth in the estuary. We demonstrate that CART can be used to characterize nutrient limitation from long-term water-quality monitoring data on much broader geographic and temporal scales than would be feasible using bioassays, providing a new tool for informing water-quality management. The CART approach can be adapted to other waterbodies where long-term bioassays and water-quality monitoring data sets are available, although the selection of water-quality variables and model forms may vary based on data availability in specific systems. Below, we provide a discussion on the validity of the CART ensemble approach and the applications of the approach specific to Chesapeake Bay.

5.1. CART was selected as the empirical approach to relate tidal water-quality data to bioassay-based measures of nutrient limitation (Goal 1)

DIN and DIP concentrations (A1) or their indices (A2) alone cannot satisfactorily reproduce the bioassay-based nutrient limitation patterns (Fig. 3). Although the improved performance by A2 over A1 implies that nutrient indices are better predictors than nutrient concentrations, both approaches failed to match the bioassay-based results in July through November. This may reflect the effects of temperature and light availability on phytoplankton growth or reveal that the low nutrient concentrations in summer and fall are less reflective of nutrient availability in these seasons with high turnover rates associated with biogeochemical cycling of nutrients (e.g., water-column regeneration, sediment release of P, denitrification) (Conley, 1999; Testa and Kemp, 2012; Testa et al., 2018).

Compared with A1 and A2, tree-based approaches (A3) showed substantially improved performance in matching the bioassay-based patterns because of their ability to incorporate additional variables related to phytoplankton growth (**Figures S6-S7**). The tree plots of the three selected CART models (Figs. 4, **S8, and S9**) illustrated how the original full data were split to more homogeneous subsets and these trees had several consistencies. First, each tree has seven or eight terminal nodes and only two of them are not pure, demonstrating the effectiveness of CART for classifying the response variable (i.e., nutrient limitation class). Second, TNTP.ratio is always the top splitting variable on the trees and cases with

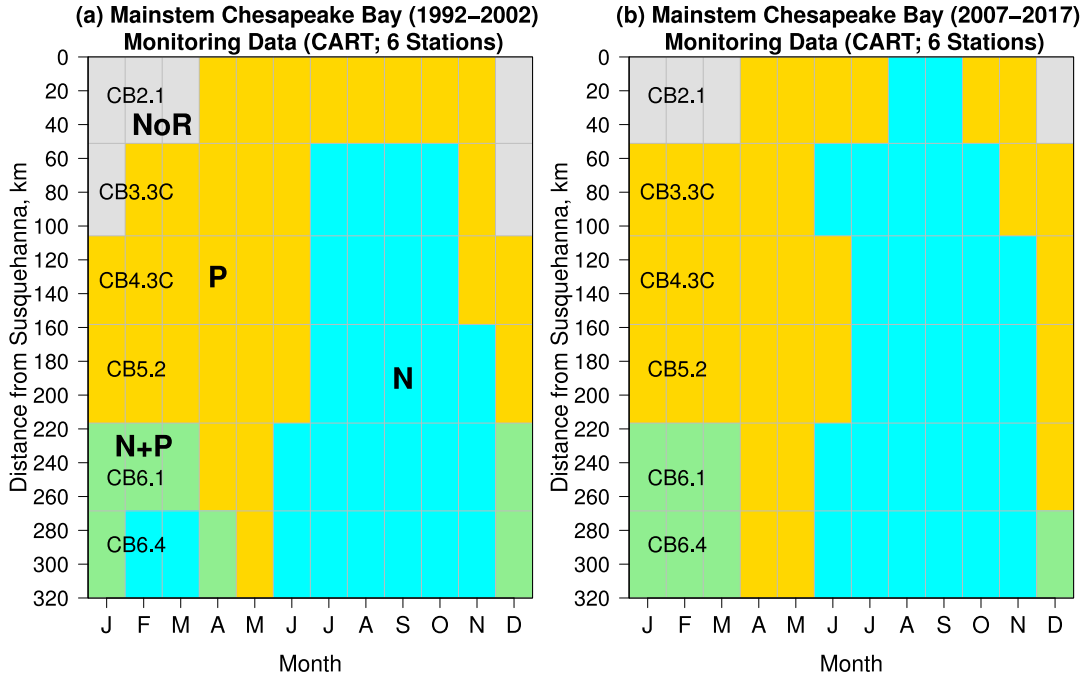

Fig. 7. CART-predicted nutrient limitation maps for the mainstem Chesapeake Bay, comparing two decadal periods: (a) 1992–2002 and (b) 2007–2017. These maps were developed based on the ensemble approach that incorporated predictions from all three candidate models.

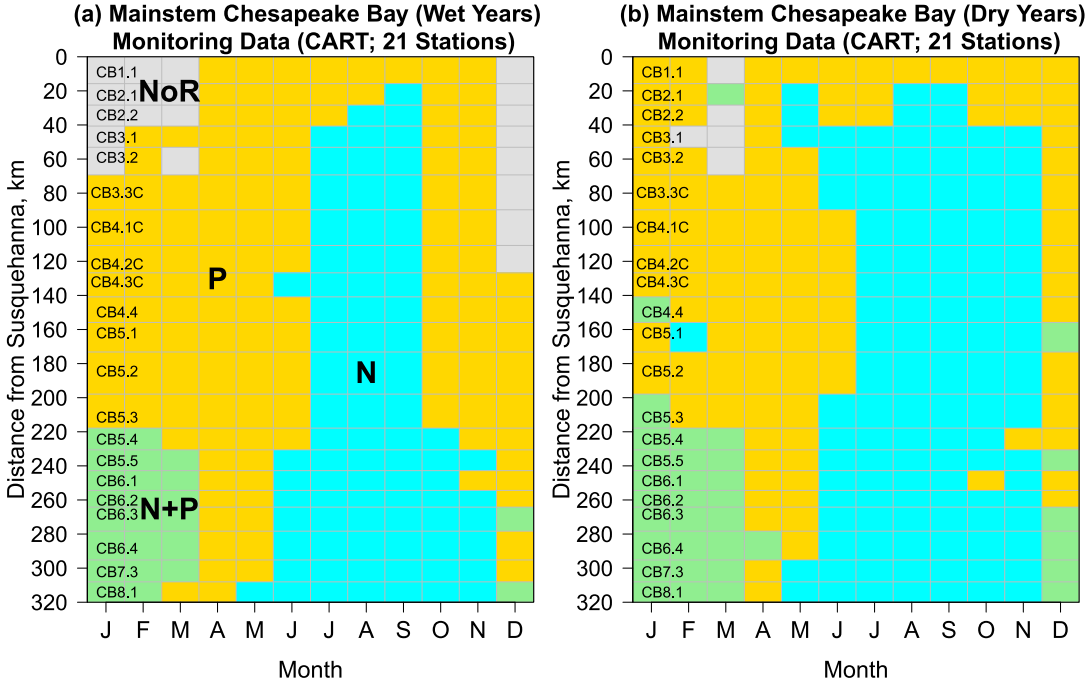

Fig. 8. CART-predicted nutrient limitation maps for the mainstem Chesapeake Bay for the period of 2003–2017, comparing (a) wet years and (b) dry years. These maps were developed based on the ensemble approach that incorporated predictions from all three candidate models.

smaller TNTP.ratio values are classified as N-limitation while cases with larger TNTP.ratio values are classified as P-limitation. This splitting rule makes sense, because a larger TNTP.ratio indicates more N relative to P, leading to a high potential for P-limitation and low potential for N-limitation. Although the splitting threshold (N:P = 51:1) is much larger than the Redfield Ratio (N:P = 16:1) (Redfield, 1958), the discrepancy may reflect the contributions of particulate N and P to TN and TP in regions where external inputs of nutrients are substantial. Although TNTP.ratio is always the first splitting variable on the trees, it alone cannot satisfactorily explain the spatial and temporal variability in nutrient limitation. In this regard, a third consistency is the inclusion of seasonal-

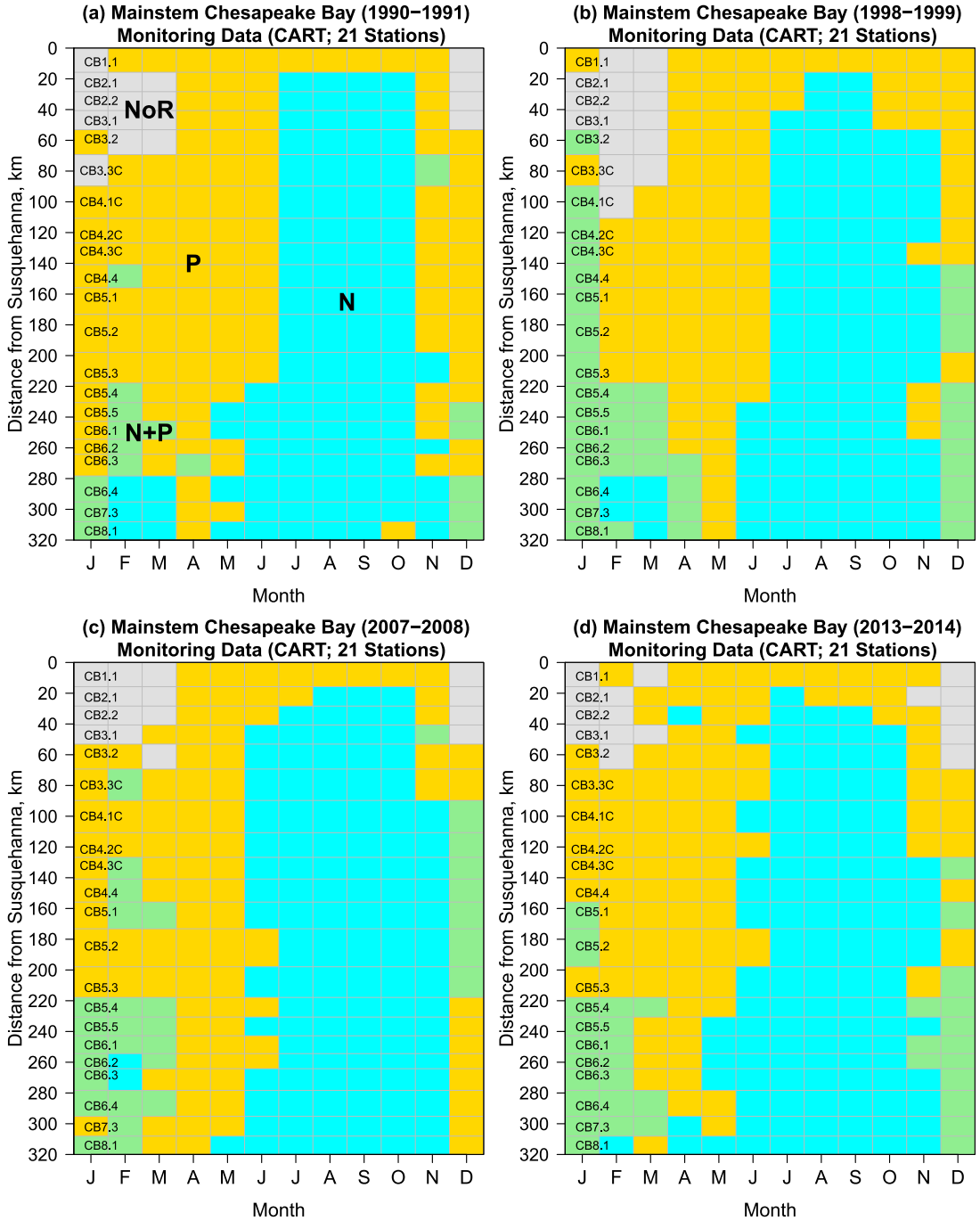

Fig. 9. CART-predicted nutrient limitation maps for the mainstem Chesapeake Bay for four sub-periods that have similar hydrologic conditions: (a) 1990–1991, (b) 1998–1999, (c) 2007–2008, and (d) 2013–2014. These maps were developed based on the ensemble approach that incorporated predictions from all three candidate models.

ity, as represented by either WTEMP (numerical) or Season (categorical). This indicates the importance of seasonal shifts in features that directly or indirectly influence phytoplankton growth. These features may include temperature effect on maximum phytoplankton growth rates, day length and daylight available to support photosynthesis, seasonal changes in stratification and vertical mixing, seasonal shifts in plankton community composition, enhanced grazing at higher temperatures, and seasonal alterations of the light field (Droop, 1983; Fisher et al., 1999, 1992; Harding et al., 2016; Kemp et al., 2005; Malone et al., 1996).

The ensemble approach effectively combined the strengths of the candidate models and had a classification rate higher than any of the candidate models (Figs. 5-6). It is not surprising that mismatches were not eliminated with the ensemble approach (Fig. 5a), which is inevitably constrained by the precision of bioassay experiments, the inherent variability in water samples, and the sample size of each limitation class. The mismatches are located on the edges of the limitation diagram (colder months and salinity extremes), generally in transition zones between more consistently defined limitation factors. Two mismatches are related to P-limitation versus NoR, which are often difficult to distinguish in the bioassays (Fisher and Gustafson, 2003,2005). Four mismatches are related to NP-limitation, which persisted from the candidate models to the ensemble approach, probably because NP-limitation is the least frequent in the bioassay data (6 out of 72 cases) and NP-limitation cases are associated with CB6.1

and CB6.4, which have very limited bioassay samples (13 each, compared to 200+ samples at CB3.3C, CB4.3C, and CB5.2). Moreover, CB6.1 and CB6.4 are most influenced by both riverine and Atlantic Ocean water inputs, and the NP-limitation mismatches occur at the boundaries of seasonal transitions from P to N limitation.

This research demonstrates that CART, when combined with long-term water monitoring data, can be used as an assessment tool for nutrient limitation. While direct bioassays are the gold-standard, they are time intensive and much more costly than water-quality monitoring programs can sustain. Instead, CART can be used to estimate nutrient limitation from parameters routinely part of water-quality monitoring protocols, greatly expanding the geographic and temporal extent of these assessments to guide water-quality management. However, we emphasize that CART should be used to fill in the gaps given limited funding resources for bioassays, rather than replace bioassays. An underlying assumption of CART is the stationarity in the derived relationships between nutrient limitation and explanatory variables. This caveat can inadvertently force predictions for new periods to strictly follow a particular splitting rule to fall into a fixed terminal node at the bottom of the tree, regardless of any changes in variables irrelevant to that splitting rule, thereby reducing the likelihood of detecting changes in nutrient limitation. Furthermore, variables selected by the tree may change over time (e.g., WTEMP), which can confound the accuracy of predictions for new periods. In these regards, new bioassay experiments can be particularly useful for recalibrating and updating the CART models.

5.2. Application of the CART ensemble approach to new periods provided insights on changes in nutrient limitation patterns and effects of nutrient reductions (Goal 2)

Predictions from the ensemble approach showed modest changes (10 out of 72 cases) in nutrient limitation patterns at the six stations that had bioassay data (**Figure S10;** Table 2). Our extended analysis of the 21 tidal stations showed similar changes (**Figure S11**), adding further confidence to these results. We emphasize that these patterns should be viewed as the overall behavior of nutrient limitation for the specified periods and these patterns can change between different hydrologic conditions or temporal periods. These estimated changes in nutrient limitation were further compared with estimated changes in monitored nutrient concentrations between the 1992–2002 period and the 2007–2017 period (Δ), as computed with Generalized Additive Models (Murphy et al., 2019) (Table 2). For example, the switch from P-limitation to N-limitation at CB2.1 in September is consistent with its declining DIN concentration ($\Delta = -0.080$ mg l^{-1}; $p < 0.01$) and increasing DIP concentration ($\Delta = 0.003$ mg l^{-1}; $p = 0.066$). Similarly, the switch from NoR to P-limitation at CB3.3C in both January and December is consistent with its declining DIP concentration ($\Delta = -0.003$ mg l^{-1}, $p \leq 0.053$). These consistencies add further confidence to the estimated changes in nutrient limitation, although nutrient concentrations alone do not always explain the estimated changes in nutrient limitation.

Nutrient limitation maps showed expanded areas of N-limitation and contracted areas of NoR in 2007–2017 compared to 1992–2002 (Fig. 7). These changes imply that some parts of the Bay have become less nutrient-saturated, consistent with other observed water-quality improvements that have been linked to nutrient reductions (Harding et al., 2016; Lefcheck et al., 2018; Murphy et al., 2011; Testa et al., 2014; Zhang et al., 2018). This implication is also consistent with the expected impacts of management efforts intended to improve water quality by reducing the human impact on the landscape and addressing legacy pollution as well as pressures from steady population growth in the watershed (Chesapeake Bay Program, 2020). These efforts have been pursued through the implementation of wastewater, agricultural, atmospheric, and urban stormwater source controls to address the pollutant reduction goals outlined in the Chesapeake Bay Partnership Agreements (Chesapeake Executive Council, 1983, 1987, 2000, 2014) and the Chesapeake Bay TMDL (U.S. Environmental Protection Agency, 2010). In response to management actions such as wastewater treatment plant upgrades (Boynton et al., 2008), reductions in fertilizer applications (Keisman et al., 2018), and reductions in N emissions under the Clean Air Act (Eshleman et al., 2013), N load has decreased across the Bay watershed over the last three decades (Ator et al., 2019; Chanat et al., 2016; Chanat and Yang, 2018; Hirsch et al., 2010; Zhang et al., 2016a, 2015), which may have led to expanded areas of N-limitation in the Bay, especially in seasons of low freshwater input.

A strength of the CART approach is its ability to predict nutrient limitation for different hydrologic conditions or temporal periods. Our results showed that, for the same temporal period (2003–2017), dry years had expanded areas of NP-limitation and N-limitation and contracted areas of NoR and P-limitation, compared to wet years (Fig. 8). A majority of NoR (in wet years) switched to P-limitation (dry years) in the upper Bay in winter months (December-February). Moreover, additional cases of NP-limitation (dry years) occurred primarily in locations and months that were P-limited (wet years). These changes reflect the fact that dry conditions are associated with lower nutrient delivery to the Bay (Boynton and Kemp, 2000), thereby creating more opportunities for nutrient limitation to occur. Such an effect of hydrologic condition has been previously observed in bioassay experiments by Fisher et al. (1999) and Fisher and Gustafson (2005). Our results also showed that, under similar hydrologic conditions, more recent periods had expanded areas of NP-limitation and contracted areas of P-limitation, compared to the earliest period (1990–1991) (Fig. 9). The expansion of NP-limitation occurred primarily in locations and months that had previously been P-limited. This indicates that N concentrations have been reduced to its limiting level in those cases, likely due to reductions in N load. The reduction in P-limited areas is consistent with the observations that (1) TP load to the Bay has increased since the mid-1990s, largely due to reduced trapping of sediment and associated P by Conowingo Reservoir as it neared its sediment storage capacity (Hirsch, 2012; Langland, 2015; Zhang et al., 2013, 2016b) and that (2) dissolved orthophosphate to the Bay has increased substantially recently (Fanelli et al., 2019).

Despite these changes, the overall seasonal and spatial patterns of nutrient limitation in Chesapeake Bay remain similar to the 1992–2002 period (Kemp et al., 2005). NoR is common in the upper Bay during winter, when daylight is short and where salinity is low, water is colder and more turbid, and ambient nutrient concentrations are high. (For example, our data at CB2.1 show that winter DIN and DIP concentrations are roughly 14-times and 2-times higher than the bloom-limitation thresholds, respectively.) Algal growth in winter does not appear to be fast enough to reduce the high nutrient concentrations to bloom-limiting levels in either the experimental controls or the NP treatments in the bioassays, leading to no differential growth response between the controls and treatments. P-limitation is common in spring when temperature increases and DIN:DIP ratios are high in river flows. Under these conditions, algae are growing faster and taking up more nutrients, bringing P concentrations to limiting levels. N-limitation is common throughout the Bay in summer and early fall due to depletion of DIN from surface waters and high release rates of DIP from the bottom (anoxic) sediments. NP-limitation only occurs in the high-salinity, oceanic-influenced lower Bay, when DIN and DIP concentrations are both below their limit-

ing levels and hence algal growth is responsive to both N and P additions.

Overall, these patterns in nutrient limitation reinforce a dual nutrient management strategy for controlling nutrient enrichment in Chesapeake Bay. The implications of uneven nutrient management have been seen in the North Sea, where it is reported that a more effective reduction of P than N load has led to a large imbalance in the N:P stoichiometry of coastal waters, resulting in an offshore gradient from P- to N-limitation (Burson et al., 2016). Although estuaries often exhibit strong seasonal and spatial variations in nutrient limitation, detection of shifts in nutrient limitation can provide important indicators of estuarine response to nutrient reductions. For example, it is predicted that P-limited areas at the continental shelf of the northern Gulf of Mexico would extend by over 50% in May and July if DIP concentration in the Mississippi and Atchafalaya Rivers is reduced by half (Laurent et al., 2012). While some of the estimated changes in nutrient limitation in Chesapeake Bay may be an early indication on the impact of nutrient reductions, further reductions are likely needed to reduce nutrient concentrations to limiting levels to achieve a less nutrient-saturated ecosystem.

6. Conclusions

We analyzed historical data from nutrient bioassays and data from the CBP long-term water-quality monitoring network to develop empirical approaches for predicting nutrient limitation. CART reproduced the bioassay-based nutrient limitation patterns in 1992–2002 much better than two non-statistical approaches (A1 and A2), because it can utilize relevant variables beyond DIN and DIP. The ensemble approach of the three selected CART models satisfactorily reproduced the bioassay-based results (classification rate = 99%). Predictions from the ensemble approach showed modest changes in nutrient limitation, with expanded areas of N-limitation and contracted areas of NoR in 2007–2017 compared to 1992–2002. These changes imply that long-term reductions in nitrogen load have led to expanded areas with nutrient-limited phytoplankton growth, reflecting long-term water-quality improvements in the context of nutrient enrichment. However, nutrient limitation patterns remain unchanged in the majority of the mainstem, suggesting that nutrient loads should be further reduced to achieve a less nutrient-saturated ecosystem. These insights can help inform management strategies called for in the Chesapeake Bay TMDL, explain changes in tidal water quality, and facilitate future refinements of Chesapeake Bay estuarine models. These results also indicate the need for conducting additional bioassay experiments for model validation and enhancement. More broadly, this research highlights the value of maintaining a long-term water-quality monitoring network and provides an example on how tidal monitoring data can be assessed in other estuaries.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the USEPA (CBP Technical Support Grant No. 07-5-230480) and the National Science Foundation (Grant No. BET1360415). We thank Zhaoying Wei for making the tidal station map and computing the distances of tidal stations from the Susquehanna River. This work benefited from discussions with the CBP modeling team. This is UMCES contribution CN5905. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. All raw and processed water-quality monitoring data are archived in Zhang (2020).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2020.116407.

References

Ator, S.W., Blomquist, J.D., Webber, J.S., Chanat, J.G., 2020. Factors driving nutrient trends in streams of the Chesapeake Bay watershed. J. Environ. Qual. 49 (4), 812–834.
Ator, S.W., García, A.M., Schwarz, G.E., Blomquist, J.D., Sekellick, A.J., 2019. Toward explaining nitrogen and phosphorus trends in Chesapeake Bay Tributaries, 1992–2012. JAWRA J. Am. Water Resour. Assoc. 55 (5), 1149–1168.
Boesch, D.F., 2019. Barriers and bridges in abating coastal eutrophication. Front. Mar. Sci. 6, 123.
Boesch, D.F., Brinsfield, R.B., Magnien, R.E., 2001. Chesapeake Bay eutrophication: scientific understanding, ecosystem restoration, and challenges for agriculture. J. Environ. Qual. 30 (2), 303–320.
Boynton, W.R., Hagy, J.D., Cornwell, J.C., Kemp, W.M., Greene, S.M., Owens, M.S., Baker, J.E., Larsen, R.K., 2008. Nutrient budgets and management actions in the Patuxent River Estuary, Maryland. Estuaries Coasts 31 (4), 623–651.
Boynton, W.R., Kemp, W.M., 2000. Estuarine Science: a synthesis approach to research and practice. Hobbie, J.E. (ed), Island Press, Washington, D.C., 269–298.
Breiman, L., 1984. Classification and regression trees, New York.
Buchanan, C., Lacouture, R.V., Marshall, H.G., Olson, M., Johnson, J.M., 2005. Phytoplankton reference communities for Chesapeake Bay and its tidal tributaries. Estuaries 28 (1), 138–159.
Burson, A., Stomp, M., Akil, L., Brussaard, C.P.D., Huisman, J., 2016. Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the North Sea. Limnol. Oceanogr. 61 (3), 869–888.
Cerco, C.F., Noel, M.R., 2017. The 2017 Chesapeake Bay Water Quality and Sediment Transport Model: a report to the US Environmental Protection Agency Chesapeake Bay Program, US Army Engineer Research and Development Center, Vicksburg, MS.
Chanat, J.G., Moyer, D.L., Blomquist, J.D., Hyer, K.E., Langland, M.J., 2016. Application of a weighted regression model for reporting nutrient and sediment concentrations, fluxes, and trends in concentration and flux for the Chesapeake Bay Nontidal Water-Quality Monitoring Network, results through water year 2012, U.S. Geological Survey, Reston, VA, p. 76.
Chanat, J.G., Yang, G., 2018. Exploring drivers of regional water-quality change using differential spatially referenced regression—A pilot study in the Chesapeake Bay watershed. Water Resour. Res. 54 (10), 8120–8145.
Chesapeake Bay Program, 2017. Methods and quality assurance for Chesapeake Bay water quality monitoring programs.
Chesapeake Bay Program, 2020. Population.
Chesapeake Executive Council, 1983. The Chesapeake Bay Agreement of 1983.
Chesapeake Executive Council, 1987. 1987 Chesapeake Bay Agreement.
Chesapeake Executive Council, 2000. Chesapeake 2000.
Chesapeake Executive Council, 2014. Chesapeake Bay Watershed Agreement.
Cloern, J.E., 2001. Our evolving conceptual model of the coastal eutrophication problem. Mar. Ecol. Prog. Ser. 210, 223–253.
Cloern, J.E., Foster, S.Q., Kleckner, A.E., 2014. Phytoplankton primary production in the world's estuarine-coastal ecosystems. Biogeosciences 11 (9), 2477–2501.
Conley, D.J., 1999. Biogeochemical nutrient cycles and nutrient management strategies. Hydrobiologia 410 (0), 87–96.
Conley, D.J., Malone, T.C., 1992. Annual cycle of dissolved silicate in Chesapeake Bay: implications for the production and fate of phytoplankton biomass. Mar. Ecol. Prog. Ser. 81, 121–128.
Droop, M.R., 1983. 25 years of algal growth kinetics: a personal view. Botanica Marina 26 (3).
Elser, J.J., Bracken, M.E.S., Cleland, E.E., Gruner, D.S., Harpole, W.S., Hillebrand, H., Ngai, J.T., Seabloom, E.W., Shurin, J.B., Smith, J.E., 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. Ecol. Lett. 10 (12), 1135–1142.
Eshleman, K.N., Sabo, R.D., Kline, K.M., 2013. Surface water quality is improving due to declining atmospheric N deposition. Environ. Sci. Technol. 47 (21), 12193–12200.
Fanelli, R.M., Blomquist, J.D., Hirsch, R.M., 2019. Point sources and agricultural practices control spatial-temporal patterns of orthophosphate in tributaries to Chesapeake Bay. Sci. Total Environ. 652, 422–433.
Fennel, K., Testa, J.M., 2019. Biogeochemical controls on coastal hypoxia. Ann. Rev. Mar. Sci. 11, 105–130.
Fisher, T.R., Gustafson, A.B., 2003. Nutrient-addition bioassays in Chesapeake Bay to assess resources limiting algal growth. Progress report: August 1990 - December 2002. University of Maryland Center for Environmental Science, Cambridge, MD.
Fisher, T.R., Gustafson, A.B., 2005. Nutrient-addition bioassays in Chesapeake Bay to assess resources limiting algal growth. Final Interpretive report: August 1990 -

May 2005. University of Maryland Center for Environmental Science, Cambridge, MD.
Fisher, T.R., Gustafson, A.B., Sellner, K., Lacouture, R., Haas, L.W., Wetzel, R.L., Magnien, R., Everitt, D., Michaels, B., Karrh, R., 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. Mar. Biol. 133 (4), 763–778.
Fisher, T.R., Peele, E.R., Ammerman, J.W., Harding, L.W., 1992. Nutrient limitation of phytoplankton in Chesapeake Bay. Mar. Ecol. Prog. Ser. 82, 51–63.
Hagy, J.D., Boynton, W.R., Keefe, C.W., Wood, K.V., 2004. Hypoxia in Chesapeake Bay, 1950–2001: long-term change in relation to nutrient loading and river flow. Estuaries 27 (4), 634–658.
Harding, L.W., Gallegos, C.L., Perry, E.S., Miller, W.D., Adolf, J.E., Mallonee, M.E., Paerl, H.W., 2016. Long-term trends of nutrients and phytoplankton in Chesapeake Bay. Estuaries Coasts 39 (3), 664–681.
Hecky, R.E., Kilham, P., 1988. Nutrient limitation of phytoplankton in freshwater and marine environments: a review of recent evidence on the effects of enrichment. Limnol. Oceanogr. 33, 796–822.
Hirsch, R.M., 2012. Flux of nitrogen, phosphorus, and suspended sediment from the Susquehanna river basin to the Chesapeake Bay during Tropical Storm Lee. In: September 2011, As an Indicator of the Effects of Reservoir Sedimentation On Water Quality. U.S. Geological Survey, Reston, VA, p. 17.
Hirsch, R.M., Moyer, D.L., Archfield, S.A., 2010. Weighted Regressions on Time, Discharge, and Season (WRTDS), with an application to Chesapeake Bay river inputs. JAWRA J. Am. Water Resour. Asso. 46 (5), 857–880.
Keisman, J.L.D., Devereux, O.H., LaMotte, A.E., Sekellick, A.J., Blomquist, J.D., 2018. Manure and fertilizer inputs to land in the Chesapeake Bay watershed, 1950–2012, U.S. Geological Survey, Reston, VA, p. 37.
Kemp, W.M., Boynton, W.R., Adolf, J.E., Boesch, D.F., Boicourt, W.C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P.M., Hagy, J.D., Harding, L.W., Houde, E.D., Kimmel, D.G., Miller, W.D., Newell, R.I.E., Roman, M.R., Smith, E.M., Stevenson, J.C., 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. Mar. Ecol. Prog. Ser. 303, 1–29.
Kemp, W.M., Testa, J.M., Conley, D.J., Gilbert, D., Hagy, J.D., 2009. Temporal responses of coastal hypoxia to nutrient loading and physical controls. Biogeosciences 6 (12), 2985–3008.
Langland, M.J., 2015. Sediment transport and capacity change in three reservoirs. In: Lower Susquehanna River Basin, Pennsylvania and Maryland, 1900-2012. U.S. Geological Survey, Reston, VA, p. 18.
Laurent, A., Fennel, K., Hu, J., Hetland, R., 2012. Simulating the effects of phosphorus limitation in the Mississippi and Atchafalaya River plumes. Biogeosciences 9 (11), 4707–4723.
Lefcheck, J.S., Orth, R.J., Dennison, W.C., Wilcox, D.J., Murphy, R.R., Keisman, J., Gurbisz, C., Hannam, M., Landry, J.B., Moore, K.A., Patrick, C.J., Testa, J., Weller, D.E., Batiuk, R.A., 2018. Long-term nutrient reductions lead to the unprecedented recovery of a temperate coastal region. Proc. Natl. Acad. Sci. 115 (14), 3658–3662.
Loh, W.-Y., 2014. Fifty years of classification and regression trees. Int. Stati. Rev. 82 (3), 329–348.
Malone, T.C., Conley, D.J., Fisher, T.R., Glibert, P.M., Harding, L.W., Sellner, K.G., 1996. Scales of nutrient-limited phytoplankton productivity in Chesapeake Bay. Estuaries 19 (2), 371.
Malone, T.C., Newton, A., 2020. The globalization of cultural eutrophication in the coastal ocean: causes and consequences. Front. Mar. Sci. 7, 670.
Murphy, R.R., Kemp, W.M., Ball, W.P., 2011. Long-term trends in Chesapeake Bay seasonal hypoxia, stratification, and nutrient loading. Estuaries Coasts 34 (6), 1293–1309.
Murphy, R.R., Perry, E., Harcum, J., Keisman, J., 2019. A generalized additive model approach to evaluating water quality: chesapeake Bay case study. Environ. Model. Softw. 118, 1–13.
Noe, G.B., Cashman, M.J., Skalak, K., Gellis, A., Hopkins, K.G., Moyer, D., Webber, J., Benthem, A., Maloney, K., Brakebill, J., Sekellick, A., Langland, M., Zhang, Q., Shenk, G., Keisman, J., Hupp, C., 2020. Sediment dynamics and implications for management: state of the science from long-term research in the Chesapeake Bay watershed, USA. WIREs Water 7 (4).
Paerl, H.W., 2018. Why does N-limitation persist in the world's marine waters? Mar. Chem. 206, 1–6.
Redfield, A.C., 1958. The biological control of chemical factors in the environment. Am. Sci. 46 (3), 205–221.
Schindler, D.W., 1974. Eutrophication and recovery in experimental lakes: implications for lake management. Science 184 (4139), 897–899.
Tango, P.J., Batiuk, R.A., 2016. Chesapeake Bay recovery and factors affecting trends: long-term monitoring, indicators, and insights. Regio. Stud. Marine Sci. 4, 12–20.
Testa, J.M., Kemp, W.M., 2012. Hypoxia-induced shifts in nitrogen and phosphorus cycling in Chesapeake Bay. Limnol. Oceanogr. 57 (3), 835–850.
Testa, J.M., Kemp, W.M., Boynton, W.R., 2018. Season-specific trends and linkages of nitrogen and oxygen cycles in Chesapeake Bay. Limnol. Oceanogr. 63 (5), 2045–2064.
Testa, J.M., Li, Y., Lee, Y.J., Li, M., Brady, D.C., Di Toro, D.M., Kemp, W.M., Fitzpatrick, J.J., 2014. Quantifying the effects of nutrient loading on dissolved O2 cycling and hypoxia in Chesapeake Bay using a coupled hydrodynamic–biogeochemical model. J. Mar. Syst. 139, 139–158.
Testa, J.M., Lyubchich, V., Zhang, Q., 2019. Patterns and trends in Secchi disk depth over three decades in the Chesapeake Bay estuarine complex. Estuaries Coasts 42 (4), 927–943.
Therneau, T., Atkinson, B., 2019. rpart: recursive partitioning and regression trees. R package version 4, 1–15. https://CRAN.R-project.org/package=rpart.
U.S. Environmental Protection Agency, 2010. Chesapeake Bay Total Maximum Daily Load for nitrogen, Phosphorus and Sediment. U.S. Environmental Protection Agency, Annapolis, MD.
Zhang, Q., 2020. Data for: Nutrient limitation of phytoplankton in Chesapeake Bay: Development of an empirical approach for water-quality management. Mendeley Data V3. doi:10.17632/8wt3wh2mwr.3.
Zhang, Q., Ball, W.P., Moyer, D.L., 2016a. Decadal-scale export of nitrogen, phosphorus, and sediment from the Susquehanna River basin, USA: analysis and synthesis of temporal and spatial patterns. Sci. Total Environ. 563-564, 1016–1029.
Zhang, Q., Blomquist, J.D., Moyer, D.L., Chanat, J.G., 2019. Estimation bias in water-quality constituent concentrations and fluxes: a Synthesis for Chesapeake Bay rivers and streams. Front. Ecol. Evol. 7, 109.
Zhang, Q., Brady, D.C., Ball, W.P., 2013. Long-term seasonal trends of nitrogen, phosphorus, and suspended sediment load from the non-tidal Susquehanna River Basin to Chesapeake Bay. Sci. Total Environ. 452-453, 208–221.
Zhang, Q., Brady, D.C., Boynton, W.R., Ball, W.P., 2015. Long-term trends of nutrients and sediment from the nontidal Chesapeake watershed: an assessment of progress by river and season. JAWRA J. Am. Water Resour. Asso. 51 (6), 1534–1555.
Zhang, Q., Hirsch, R.M., Ball, W.P., 2016b. Long-term changes in sediment and nutrient delivery from Conowingo Dam to Chesapeake Bay: effects of reservoir sedimentation. Environ. Sci. Technol. 50 (4), 1877–1886.
Zhang, Q., Murphy, R.R., Tian, R., Forsyth, M.K., Trentacoste, E.M., Keisman, J., Tango, P.J., 2018. Chesapeake Bay's water quality condition has been recovering: insights from a multimetric indicator assessment of thirty years of tidal monitoring data. Sci. Total Environ. 637-638, 1617–1625.