

# Long-Term Trends in Chesapeake Bay Seasonal Hypoxia, Stratification, and Nutrient Loading

Rebecca R. Murphy · W. Michael Kemp · William P. Ball

Received: 25 October 2010 / Revised: 3 March 2011 / Accepted: 9 May 2011 / Published online: 24 May 2011  
© Coastal and Estuarine Research Federation 2011

**Abstract** A previously observed shift in the relationship between Chesapeake Bay hypoxia and nitrogen loading has pressing implications on the efficacy of nutrient management. Detailed temporal analyses of long-term hypoxia, nitrogen loads, and stratification were conducted to reveal different within-summer trends and understand more clearly the relative role of physical conditions. Evaluation of a 60-year record of hypoxic volumes demonstrated significant increases in early summer hypoxia, but a slight decrease in late summer hypoxia. The early summer hypoxia trend is related to an increase in Bay stratification strength during June from 1985 to 2009, while the late summer hypoxia trend matches the recently decreasing nitrogen loads. Additional results show how the duration of summertime hypoxia is significantly related to nitrogen loading, and how large-scale climatic forces may be responsible for the early summer increases. Thus, despite intra-summer differences in primary controls on hypoxia, continuing nutrient reduction remains critically important for achieving improvements in Bay water quality.

**Keywords** Hypoxia · Stratification · Nutrients · Chesapeake Bay · Kriging · Long-term trends

## Introduction

Hypoxia, or depleted dissolved oxygen (DO), has been a recurring summertime condition in deeper regions of the mainstem Chesapeake Bay for decades (e.g., Hagy et al. 2004), and is an expanding problem in coastal regions around the world (Diaz and Rosenberg 2008). The negative effects of hypoxia on living resources include decreased habitat for fish and invertebrates (Ludsin et al. 2009; Seitz et al. 2009) in both deep and shallow waters (Breitburg 1992, 2002) and overall shifts in trophic energy transfer and production (Diaz and Rosenberg 2008). The causes of hypoxia are generally agreed to be a combination of both biological and physical factors (Boynton and Kemp 2000; Kemp et al. 1992). While physical factors can set up local environments that minimize oxygen transfer and mixing (Boicourt 1992), biological processes are involved in that excessive nutrient loads can cause increased oxygen consumption via phytoplankton blooms and decay (Malone et al. 1988; Malone 1992; Officer et al. 1984).

Long-term trends in hypoxic volume of the main channel of Chesapeake Bay over a 50-year period were found to be related to Susquehanna River nitrogen loads and freshwater flows (Hagy et al. 2004). Both nitrogen, the primary limiting nutrient (Fisher et al. 1999), and freshwater flow, a major factor in regulating stratification (Boicourt 1992), were predictors of summertime hypoxic volume in the main channel, but these variables did not fully explain the sharp increase of average July hypoxic volume from 1950 to 2001 (Hagy et al. 2004). In particular, for the same amount of nitrogen load or flow, the analysis of Hagy et al. revealed

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s12237-011-9413-7) contains supplementary material, which is available to authorized users.

---

R. R. Murphy (✉) · W. P. Ball  
Department of Geography and Environmental Engineering,  
Johns Hopkins University,  
3400 N. Charles St,  
Baltimore, MD 21218, USA  
e-mail: rebecca.murphy@jhu.edu

W. M. Kemp  
Horn Point Laboratory, Center for Environmental Science,  
University of Maryland,  
2020 Horns Point Road,  
Cambridge, MD 21613, USA

more hypoxic volume in July in recent years than in years at the beginning of the data set. Using the same time series, Kemp et al. (2005) found that two separate significant trend lines described the relationship between hypoxia and nitrate loading for 1950–1979 and 1980–2001 and suggested that the Bay has become less able to assimilate nitrogen inputs. Using a change-point analysis, Conley et al. (2009) found that 1986 was a significant breakpoint in the same Bay hypoxia trend. These studies point to the possibility that a shift may have taken place in the hypoxia–nutrient relationship in Chesapeake Bay. Similar trends of recently increasing hypoxia for a given nutrient loading have been observed in other locations (Conley et al. 2009; Kemp et al. 2009) with reduced response of hypoxia to inter-annual nutrient load reductions compared to previous conditions in systems such as Danish coastal waters (Conley et al. 2007) and the northern Gulf of Mexico (Turner et al. 2008).

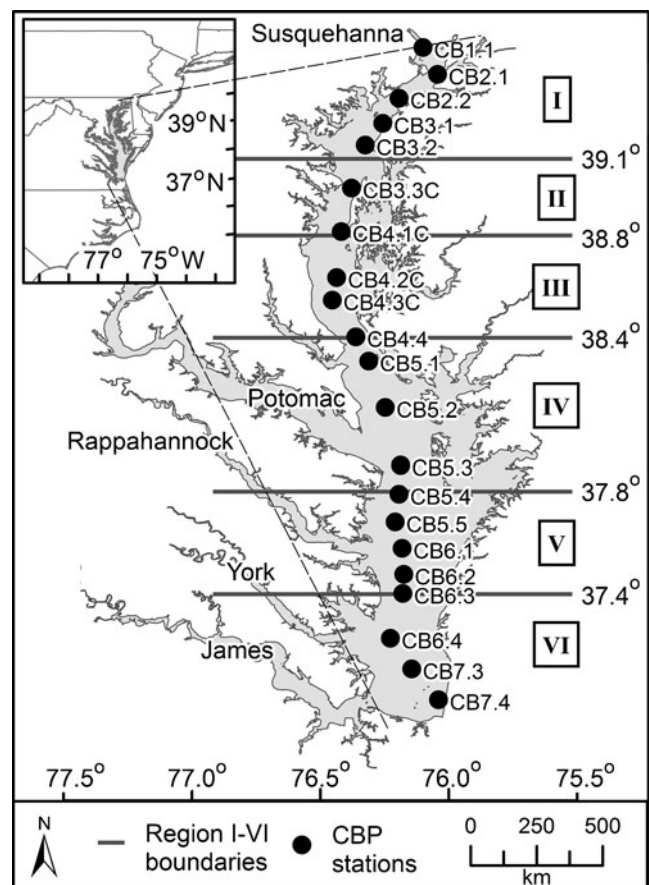
In Chesapeake Bay, a few recent studies have investigated possible reasons for the observed hypoxic volume trends. A change in the predominant summertime wind direction around the mid-1980s was found to statistically explain the increase in average July hypoxic volume unaccounted for by nitrogen loads (Scully 2010a), based on the findings that winds from different directions can impact Bay salinity structure and circulation (Guo and Valle-Levinson 2008; Scully 2010b). Other hypotheses for the unexpected shift in this hypoxia–nutrient relationship include increased benthic recycling of nitrogen or changes in atmospheric forcing and continental shelf circulation (Kemp et al. 2009). The Chesapeake Bay Environmental Observatory (CBEO) team investigated these and other hypotheses to explain why the hypoxia–nutrient shift occurred, making use of Bay-wide data accessible for analysis through the CBEO’s prototypical environmental observatory testbed (CBEO Project Team 2008).

This paper presents results from investigation of one of CBEO team’s hypotheses—that long-term trends of increasing Bay stratification are a driving force behind the observed increases in hypoxic volume per nitrogen loading. Our approach involved re-evaluating data on long-term hypoxic volume with more temporal resolution and calculating Bay stratification strength with similar temporal resolution. This analysis revealed the unanticipated finding that the previously reported change in the hypoxia–loading relationship is occurring only in the early summer, corresponding with an increase in early summer Bay stratification. Freshwater flow explains some of the observed inter-annual variations in stratification; however, there is no evidence to suggest a long-term trend in spring freshwater flow to the Bay. We therefore explored other physical factors that influence stratification including wind (Goodrich et al. 1987), temperature, and salinity (Hilton et al. 2008; Preston 2004).

## Methods

### Study Area and Data

The Chesapeake Bay is a large, partially mixed estuary extending approximately 320 km from the mouth of the Susquehanna River at Havre de Grace, Maryland, to the Atlantic Ocean near Virginia Beach, Virginia (Fig. 1). For this study, we used data for salinity, temperature, and DO from along the Bay’s main deep channel. The data from 1984 to 2009 were collected by the Chesapeake Bay Program and their collaborators (Chesapeake Bay Program 2010), and the data from 1949 to 1980 were collected predominantly by the Chesapeake Bay Institute (Chesapeake Bay Program 2008). Sufficient main channel data from 1981 to 1983 were not available for proper analysis, and these years have therefore been excluded from this study. All data were compiled and accessed through the CBEO testbed, which is the research and development prototype for data storage and tool development within the CBEO project (CBEO Project Team 2008). This compila-



**Fig. 1** Chesapeake Bay location, 1984–2009 main channel sampling stations, and regional boundaries used in this study. The main channel can be identified as a line connecting the 21 stations shown

tion step was needed in order to run multiple data analysis scripts across different data sets.

Time periods were identified during which sufficient data had been collected to provide a comprehensive spatial “snapshot” of water quality. For the recent (post-1984) data, such snapshots were essentially defined by the nature of the CBP sampling, which is organized into cruises that consist of approximately 4–7-day periods of data collection. For summertime CBP cruises from July 1984 through 2009, the same 21 main channel station locations were normally sampled (Fig. 1). In some cases, stations were skipped, and for cruises with 15 or fewer sampled stations, we examined the sampling distribution on a map before including the cruise in our studies. For data collected prior to July 1984, data sets were selected by identifying groups of consecutive days with samples in each of six regions along the main channel to generate a complete main channel picture. This process resulted in 20 data sets from 4 sources from 1949 to 1980 in July (or in June or August of years with no July sampling). Sources and sample dates are listed in Supplemental Materials (SM), Table S1.

Fall-line nitrogen and phosphorus loads from 1981 to 2009 were downloaded from the USGS Chesapeake Bay River Input Monitoring Program website (U.S. Geological Survey 2010a). Monthly average total nitrogen (TN) and total phosphorus (TP) loads in units of kilogram per day were estimated from streamflow and nutrient concentrations by the USGS for this website. TN loads at the Susquehanna River fall line (i.e., Conowingo dam) before 1981 were estimated by Dr. James Hagy who developed a model to estimate monthly nitrate concentrations based on intermittent observations at the upstream USGS station in Harrisburg, PA from 1945 to 1995 (Hagy et al. 2004; equation 2) and then used the modeled nitrate concentrations at Harrisburg, observed river flows, and a regression model to estimate TN loading at Conowingo for the years it was not measured. The regression model (Hagy et al. 2004; equation 7) was fit using winter–spring average loads calculated from observations at both locations in the same years. The previously unpublished Conowingo TN loading estimates were provided by Hagy and these are used in our analyses.

### Hypoxic Volume Calculation

Hypoxic volumes were calculated by summing the volume of water with DO concentrations below three levels:  $DO < 0.2$  mg/L (near-anoxia),  $DO < 1$  mg/L (severe hypoxia), and  $DO < 2$  mg/L (moderate hypoxia). These three levels were used for consistency with Hagy et al. (2004), who considered these three levels because hypoxia-related ecological effects vary across the continuum of low DO concentrations (Hagy et al. 2004). For computing anoxic volume, the near-anoxia definition ( $DO < 0.2$  mg/L) was

used as opposed to complete anoxia ( $DO = 0.0$  mg/L) because 0.2 mg/L was the calibration accuracy for some of the DO observations (Chesapeake Bay Program 1993). For brevity, we occasionally refer to all three levels generally as “hypoxic volume.” Although some figures and tabular results in the main text are presented only for  $DO < 1$  mg/L, details of most analyses at the other two levels are provided in the Supplemental Materials (SM).

The statistical interpolation method of kriging (e.g., Cressie 1993; Diggle and Ribeiro 2007) was used to interpolate spatially the main channel DO observations to a two-dimensional depth–length grid along the main channel of the Bay. To calculate hypoxic volumes, we assumed that interpolated DO concentrations were uniform laterally across the mainstem portion of the Bay (i.e., to the east and west), and we used tabulated cross-sectional volumes (Cronin and Pritchard 1975) to calculate the volume of water with DO less than 0.2, 1, and 2 mg/L for each cruise. This assumption was consistent with Hagy et al. (2004), but we also performed an additional check by using the Bay Program Interpolator Tool, VOL3D (Bahner 2006) to interpolate DO concentrations throughout the mainstem of the Bay for both early and late July from 1984 to 2009. This tool uses all mainstem data, including that collected to the east and west of the main channel. Although there are slight differences between the results generated from our kriging method and the Interpolator Tool’s inverse distance weighting method (results not shown), the long-term trends and general patterns are the same, giving us confidence that the trends we describe below are not sensitive to the assumption of constant lateral isopleths.

All interpolations were performed using the statistical package R (R Development Core Team 2008) and the geoR contributed package (Ribeiro and Diggle 2008). The RODBC contributed package (Ripley and Lapsley 2008) was used to import data from the CBEO testbed SQL server to the R computing environment. The use of the kriging method for spatial interpolation provided a re-examination of trends reported by Hagy et al. (2004) through the use of a different interpolation method (CBEO Project Team 2008). In addition, kriging allowed us to experiment with data transformations and anisotropy fits to find the most appropriate method to interpolate the data both in years with sparse sampling (pre-1984) and years with more samples (post-1984).

One of the key steps in kriging is to build a model of the spatial correlation of the observed parameter of interest (i.e., DO) as a function of distance. This model, the variogram, is then used in the kriging interpolation procedure. For this application, we followed a variogram fitting process similar to the one used in our prior work (Murphy et al. 2010), but with three major differences: (1) we fixed the nugget to zero for each variogram (thus reflecting the relatively small uncer-

tainty associated with the observations); (2) we analyzed a two-dimensional region defined vertically along the main channel (with dimensions of length and depth), and with allowance for geometric anisotropy in the variogram fit to account for the extreme differences of correlation in the two directions; and (3) we experimented with data transformations and utilized the Box–Cox family of transformations for each DO interpolation (Box and Cox 1964). Different Box–Cox transformations were applied to each data set in order to ensure that each data set properly met the Gaussian assumption of our interpolations (useful because we used restricted maximum likelihood fitting and beneficial in stabilizing the variances), thus facilitating comparisons.

After the optimal variogram model was selected for each data set, ordinary kriging was performed for all DO interpolations. Samples are not generally collected below 32 m, so we set this to be the maximum interpolation depth to prevent possible unrealistic extrapolations. All interpolations were evaluated similar to interpolations in Murphy et al. (2010) with cross validation (Cressie 1993). When cross validation results or variogram parameters fell outside the normal range, we visually inspected the fitted variogram. In addition, we visually inspected every DO interpolation and manually adjusted the variogram parameters for six of the pre-1984 data sets where we discovered poor fits from the automated method.

#### Stratification Strength Calculation

Stratification-related parameters were calculated using the CBP data for density and depth for samples collected along the main channel of the Bay in the period 1984–2009. Ideally, we would calculate pycnocline strength (i.e., the strength of stratification in the vertical region where density changes sharply) for the years before 1984 for comparison to the long-term trends in hypoxic volume; however, the vertical sampling density for the data sets before 1984 was found to be of insufficient resolution to capture the maximum pycnocline strength. To determine the strength and depth of the pycnocline for 1984–2009, we first vertically interpolated all density observations to obtain values at 1 m intervals. In most cases, this step was not necessary near the pycnocline because the sampling resolution is typically 1 m in that region. These data were used to calculate the square of the Brunt Väisälä Frequency, or  $N^2$  (e.g., Knauss 1997), as a measure of buoyancy frequency or stratification strength at every depth  $z_i$ :

$$N^2(z_i) = \frac{g}{\rho_i} \frac{\partial \rho}{\partial z}$$

In this equation,  $g$  is the gravitational constant,  $\rho_i$  is the water density (in  $\text{kg m}^{-3}$ ) at the depth  $z_i$ , and  $\partial \rho / \partial z$  is the

density gradient at depth  $z_i$  calculated using a 2-m window around  $z_i$ .

An  $N^2$  value was generated for every sampled depth. For each station, the maximum  $N^2$  ( $\text{max}N^2$ ) and its depth were identified as the pycnocline strength and depth (e.g., SM, Fig. S1). Estimates of pycnocline strength and depth were then interpolated along the main channel of the Bay using ordinary kriging, with a similar variogram fitting procedure as DO.

#### Statistical Analyses

Simple linear regression and multiple regression models were used to examine hypotheses for the observed trends in hypoxic volume and stratification. For the multiple regression models, each variable was mean-centered and scaled to unit standard error so that coefficients ( $\beta$ s) of the variables could be meaningfully compared. The residuals of the finally proposed regression models (i.e., those that did not still show a temporal trend) were tested for autocorrelation (Durbin–Watson test) and heteroscedasticity (Breusch–Pagan test). In addition, for each multiple regression model, we checked whether the variables were significantly correlated with each other before fitting the model. In one case noted in the “Results”, there were significant correlations among the variables, so we transformed the variables using principle component analysis (Jolliffe 2002). Using the principle components as regression variables, the coefficients, or “dependence” values, were then determined from the results.

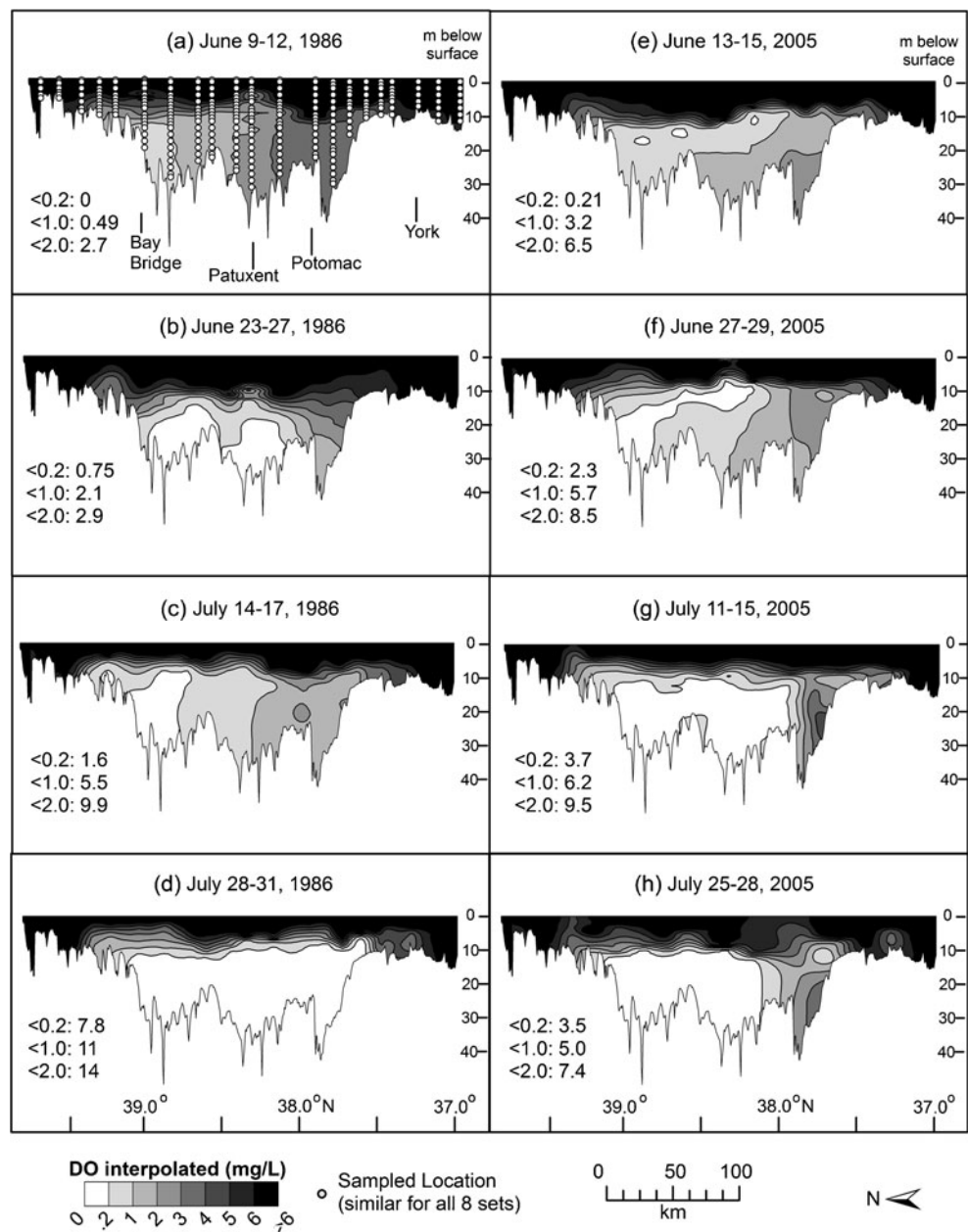
A center of volume (COV) date analysis (e.g., Hodgkins and Dudley 2006) was used to examine whether there have been any relevant shifts in timing of the spring freshwater flow. This analysis involved cumulatively summing the daily flow volumes at the Susquehanna River Harrisburg Station and the Potomac River Little Falls Pump Station (U.S. Geological Survey 2010b) and identifying the date when half of the seasonal flow volume had been recorded in each year for each of these rivers.

## Results

### Hypoxic Volume Trends

Examination of the DO interpolations based on the CBP data (1984 to 2009) revealed that hypoxic conditions were present in Chesapeake Bay from early June through September in almost every year, with large fortnightly variability and clear inter-annual differences in the timing of hypoxia onset. June and July DO interpolations from 1986 and 2005 are presented as examples (Fig. 2). These 2 years had similar, relatively large, January to May

**Fig. 2 a–h** Interpolated DO for each June and July cruise in 1986 and 2005. Sampling locations are noted in **a** and are approximately the same for each of the cruises. Maps are oriented with the Susquehanna River on the left and the Atlantic Ocean on the right. Volumes of water (cubic kilometer) with DO less than 0.2, 1, and 2 mg/L are listed with each map

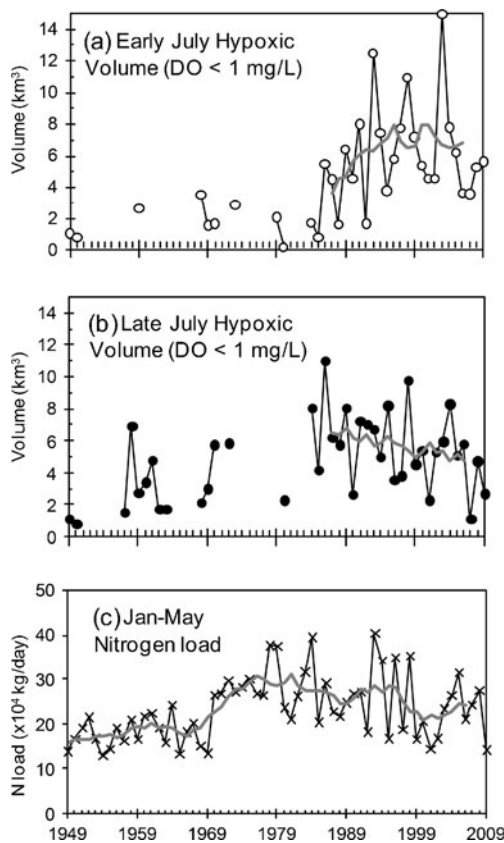


Susquehanna TN loads ( $2.90 \times 10^5$  and  $3.12 \times 10^5$  kg/day, respectively), ranking 6th and 5th highest out of the 25 years. Near-anoxic (DO<0.2) and hypoxic (DO<1 and 2 mg/L) water volumes for these years were also above average. Hypoxia affected more of the Bay earlier in 2005 than in 1986.

Similar analyses of the available early and late July data for the entire 1949–2009 period were conducted, using data from the first and second July cruises, which usually fell within July 1–15 and July 16–31, respectively (SM, Table S1). July was the focus of this longer analysis because before the CBP data collection program began in 1984, there was typically only a single summer cruise that was in July. Results revealed differences in the long-term

trends in early versus late July hypoxic volume (Fig. 3a, b; SM, Fig. S2). Early July hypoxic volume (Fig. 3a and SM, Fig. S2a, c, e) increased through the 1980s and 1990s, possibly leveling-out in the 2000s, with the exception of 2003, when a maximum value (for DO<1 mg/L) of  $14.9 \text{ km}^3$  was observed. By contrast, the long-term trend in late July hypoxic volume (Fig. 3b and SM, Fig. S2b, d, f) has been fairly constant or decreased slightly since 1984, with a maximum volume (for DO<1 mg/L) of  $11.0 \text{ km}^3$  in 1986. Large year-to-year variability is present in both sets of results.

Average January to May Susquehanna TN load is presented for comparison (Fig. 3c) because previous research has demonstrated the important role played by



**Fig. 3** Hypoxic volume ( $\text{DO} < 1 \text{ mg/L}$ ) calculated for mainstem of the Bay from data collected in early July (a) and late July (b) compared to winter–spring nitrogen load through the Susquehanna River (c). *Solid smoothed lines* are 7-year moving averages

winter/spring nitrogen flux (Hagy et al. 2004; Malone et al. 1988). TN shows a marked increase around 1970 followed by a more stable trend in the 1970s and 1980s and a slight decreasing pattern, with greater variability, in the 1990s. Some of these patterns are due to higher flow through the Susquehanna in the early 1970s than the 1960s and large fluctuations in river flow in the 1990s. Nonetheless, the decreasing trend of TN from 1970 to 2009 is statistically significant (linear regression,  $R^2=0.11$ ,  $p=0.03$ ).

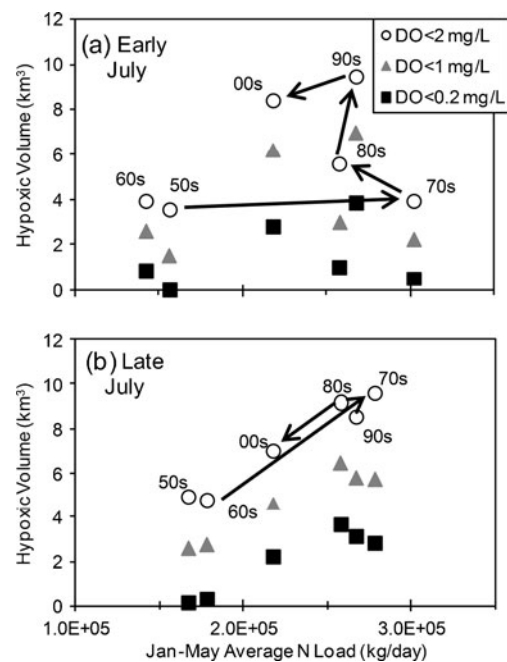
Comparing decade-averaged TN to the hypoxic volumes reveals a more consistent relationship between TN and hypoxic volume in late July (Fig. 4b) than early July (Fig. 4a). Linear regression model results for the entire 1949–2009 period confirm that TN is a more significant variable for explaining late July hypoxic volume than for early July (Table 1). For both time periods, the correlation to TN load is more significant with near-anoxic volume than hypoxic volume (Table 1). The residuals from each regression (Table 1) were regressed with year to determine if, after accounting for TN effects, there has been an increase in hypoxic volume over time. For all definitions of hypoxia in early July, the residuals are significantly increasing over time ( $p < 0.01$ ), while only the residuals

from near-anoxic volume ( $\text{DO} < 0.2 \text{ mg/L}$ ) in late July are increasing over time ( $p=0.04$ ).

There are few data available for other summer months besides July in the pre-1984 data sets; however, examination of the available data from 1984 to 2009 for June, July, and August (SM, Fig. S3) shows that the early July and late July patterns (SM, Fig. 3b, c) are good representations of the early and late summer, respectively. For example, regression results using data from 1984 through 2009 (SM, Table S2) suggest that the June volumes of hypoxic water ( $\text{DO} < 1$  and  $2 \text{ mg/L}$ ) are increasing over time in a manner that TN loadings cannot explain. In contrast, the early August volumes do not exhibit an increasing trend over time and are correlated with TN loading. In these aspects, the June and early August long-term trends are similar to the trends in early and late July, respectively.

#### Hypoxic Volume by Region

The regional long-term trends of early and late July hypoxic volume (Fig. 5 for  $\text{DO} < 1 \text{ mg/L}$ ) followed the same long-term trends observed for the entire mainstem hypoxic volume time series (Fig. 3). As with the data for the entire Bay, the regional hypoxic volumes in early July increased



**Fig. 4** Summary of hypoxic volume (in early (a) and late (b) July) responses to January–May Susquehanna TN, with points representing decadal mean values for three definitions of hypoxic volume:  $\text{DO} < 2 \text{ mg/L}$  (circles),  $\text{DO} < 1 \text{ mg/L}$  (triangles), and  $\text{DO} < 0.2 \text{ mg/L}$  (squares). The decade for which observations are averaged is indicated for  $\text{DO} < 2 \text{ mg/L}$ , and data positions are at the same  $x$ -axis value for each hypoxia measure. The number of values for hypoxic volume and TN load averaged in each symbol varies among decades (SM, Table S1)

**Table 1** Regression models for Chesapeake Bay hypoxic volume from 1949 to 2009 (cubic kilometer)

Time period	DO level (mg/L)	HypoxicVol= $\beta_0 + \beta_1(TN)$			Residuals regressed with year	
		Model fit: $R^2$ and $p$ value	Intercept, $\beta_0$	TN dependence, $\beta_1$	Model fit: $R^2$ , $p$ value	Temporal slope
Early July ( $n=40$ )	<0.2	0.24, $p=0.003^a$	2.2	1.0	0.24, $p=0.003$	0.87
	<1	0.12, $p=0.05^b$	4.8	1.1	0.26, $p=0.002$	1.6
	<2	0.08, $p=0.1^b$	7.1	1.1	0.23, $p=0.004$	1.8
Late July ( $n=42$ )	<0.2	0.33, $p<0.001$	2.2	1.2	0.11, $p=0.04$	0.56
	<1	0.28, $p<0.001$	4.8	1.3	0.07, $p=0.1$	0.56
	<2	0.21, $p=0.003$	7.3	1.4	0.05, $p=0.2$	0.61

The model is: HypoxicVol= $\beta_0 + \beta_1(TN)$ . Residuals are then regressed with year

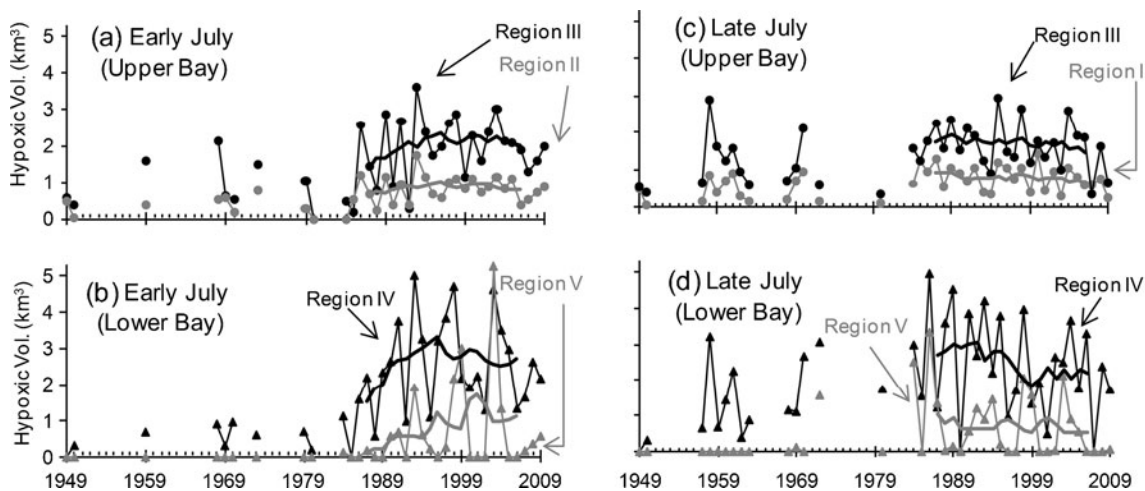
<sup>a</sup> Residuals from this regression are significantly ( $p<0.05$ ) autocorrelated and heteroscedastic, partially because of the temporal trend

<sup>b</sup> Residuals from this regression are significantly ( $p<0.05$ ) autocorrelated because of the temporal trend

from the mid-1980s to a plateau or slight decline since 2000, whereas the volumes in late July were constant or slightly decreasing since the mid-1980s. For early July, this summary reveals large variability in regions II and III from the mid-1980s to mid-1990s (Fig. 5a) and in regions IV and V in the 1990s and 2000s (Fig. 5b). The relatively sparse data from before 1984 suggest that region IV (early July) and region V (both time periods) are locations where the earlier (pre-1984) hypoxic conditions were less severe than those observed for subsequent years.

The sharp peaks in early July hypoxic volume within regions IV and V (Fig. 5b) indicate years when hypoxia extended further south earlier in the summer than it had in the past. These peaks also correspond with peak years in the entire Bay hypoxia trend (Fig. 3a). To evaluate the hypothesis that changes in TN loadings from tributaries other than the Susquehanna are responsible for the increasing trend in the mid- to lower Bay early July

hypoxic volume, we examined the January–May TN loads through each of the major tributaries flowing into the mid- to lower Bay from 1985 to 2009 (U.S. Geological Survey 2010a). During these years, on average, the Susquehanna accounts for 64% of the Bay’s total measured tributary TN load during the January–May period, whereas the Potomac accounts for 26%, and the Rappahannock, James, and York together account for 9%. All three sets of TN loads (SM, Table S3 and Fig. S4) are positively correlated to each other ( $p<0.001$ ) and none of the time series show any significant long-term increase or decrease over the 25-year period. The results do indicate, however, that in years with especially high TN load from the Potomac (e.g., 2003, 2004), mid- to lower Bay hypoxic volumes tend to be larger than expected based on Susquehanna TN loads (SM, Fig. S4). Together, the Potomac and Susquehanna TN loads from January to May are predictive of region IV 1984–2009 early July hypoxic volume (SM, Fig. S5; for DO<1 mg/L:  $R^2=0.29$ ,



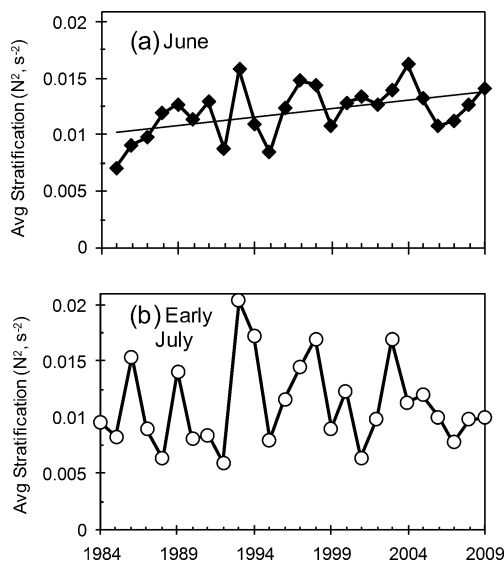
**Fig. 5** a–d Early and late July hypoxic volumes (DO<1 mg/L) by region in the Bay. Regions are identified in Fig. 1 and are the same six salinity-based regions defined in previous studies (Harding 1994; Hilton et al. 2008). Solid smoothed lines are 7-year moving averages

$p=0.005$ ); however, the residuals from this regional regression are significantly increasing over time (for  $\text{DO} < 1 \text{ mg/L}$ :  $p=0.02$ ) in a similar manner as previously observed for the whole Bay. We also examined total phosphorus (TP) loads and found that: (1) TP loads are highly correlated with TN loads through these tributaries, and (2) the peak TP years match peak TN years for each tributary (results not shown).

#### Stratification Strength Temporal Trends

Analysis of long-term trends in calculated pycnocline/stratification strength (e.g., SM, Fig. S6) shows that there was a significant increase in June pycnocline strength from 1985 to 2009 ( $p=0.01$ ), but no trend over time in July (Fig. 6 and SM Fig. S7). Because the Susquehanna and Potomac Rivers together account for 80% of the January–May river flow into the Chesapeake Bay (analysis of USGS data), flows from these two tributaries were used as variables in regression models for June and July pycnocline strength (Table 2). The average residence time of water in the Bay is on the order of 90 to 180 days (Kemp et al. 2005), so we tested whether the flow from the preceding 4 months could explain the variability observed in the stratification trends. After accounting for flow, however, there was still a significant temporal trend in June stratification (Table 2). No temporal trend was observed in early July stratification (Table 2) or late July stratification.

Regional regression results (Table 3) show that the increase in June stratification has occurred throughout the Bay, with significant increases over time ( $p \leq 0.05$ ) in



**Fig. 6** Temporal changes in average Bay pycnocline strength along the main channel in June (a:  $R^2=0.23$ ,  $p=0.01$ ) and early July (b:  $R^2=0.0004$ ,  $p=0.9$ ). Line for June is linear regression fit

regions IV, V, and VI, and a possible increase in region II ( $p=0.1$ ). Region I is the only portion of the Bay where the June stratification temporal slope is not positive; however, this is not surprising given the generally low salinity and weak stratification in this region. In regard to volume of water below the pycnocline (calculated from the average depth of the maximum Brunt Väisälä Frequency), there has been no significant change in June or July (SM, Fig. S8).

Tests were performed to evaluate hypotheses related to flow, temperature, wind, or sea level (as related to salinity) changes on June stratification trends. We observed no significant shift in the timing of the winter/spring freshwater flux (i.e., COV date) through either the Susquehanna or Potomac (SM, Fig. S9). Furthermore, analyses of each month's flow from January to June revealed no long-term change in the fraction of annual flow that occurs in each of these months (results not shown). To examine the potential role of changing water temperatures, we re-calculated pycnocline strength after first removing the effects of temperature on density—i.e., we calculated the pycnocline strength using only salinity. Results (SM, Fig. S10) still show a significant increase in pycnocline strength in June ( $p=0.02$ ), leading us to conclude that changes in water temperature during this period are not the cause of the pycnocline strength increase. Consistent results were obtained when the opposite analysis was performed—we calculated the pycnocline strength using only temperature and failed to see an increasing trend.

Regression results show that the frequency of winds from the southeast is a significant variable in predicting June pycnocline strength; however, the regression residuals after inclusion of this variable still show significant temporal increase ( $p=0.002$ ), indicating that the 25-year temporal trend in June stratification cannot be explained by changes in wind direction (Table 4). Mean sea level (MSL), as a factor in Bay salinity, was found to also be a significant predictor of June pycnocline strength ( $p=0.003$ ), and the residuals from a regression including MSL, flow, and wind no longer increase temporally ( $p=0.3$ ; Table 4).

#### Relationships Between Summer Hypoxic Volume, Stratification, Nitrogen Loads, and Temperature

Seasonally, stratification builds up in the spring to a peak value in June, whereas hypoxic volume typically peaks in July (Fig. 7). We hypothesized that stratification impacts on deep water oxygen concentrations could persist for a few weeks to months due to the long residence times and slow mixing in the Bay, so we tested whether pycnocline strengths in the previous few months are correlated with hypoxic volume. The results (SM, Table S4) indicate that early summer (June and early July) hypoxic volumes are



**Table 2** Regression models for Chesapeake Bay pycnocline strength from 1985 to 2009 (inverse square second)

Pycnocline time period	Flow average period	PycStrength= $\beta_0 + \beta_1(\text{flow})$			Residuals regressed with year	
		Model fit: $R^2$ and $p$ value	Intercept, $\beta_0$	River flow dependence, $\beta_1$	Model fit: $R^2$ , $p$ value	Temporal slope
June	February–May	0.21, $p=0.02^a$	0.012	0.0010	0.30, $p=0.004$	0.00015
Early July	March–June	0.64, $p<0.001$	0.011	0.0031	0.006, $p=0.7$	-0.000024

The model is: PycStrength= $\beta_0 + \beta_1(\text{flow})$ . Residuals are then regressed with year

<sup>a</sup>Residuals from this regression are significantly ( $p<0.05$ ) autocorrelated because of the temporal trend

significantly correlated with pycnocline strengths in the current and preceding 2 months. On the other hand, hypoxic volume trends from late July and onward do not show a consistent pattern of correlation with stratification.

Using information from these correlations, we fit linear regression models to the hypoxic volume trends from 1985 to 2009 using the following variables: Susquehanna TN load, pycnocline strength from the current and preceding few weeks, and volume of water below the pycnocline (Table 5). The linear regression results show that the increase in overall June stratification strength can statistically explain the otherwise unexplained portion of the 1985 to 2009 increase in early July hypoxic volume within the mainstem of the Bay (Table 5 and Fig. 8).

We also used the available CBP data from the CBEO testbed to evaluate the effect of nitrogen loads on the temporal persistence of hypoxia in the summer by estimating the number of summer days that DO in the bottom 5 m of water was hypoxic at each station. Using these results, we calculated the correlation coefficient between the number of days that bottom waters are hypoxic and TN loads and found that throughout the mid-Bay there

are significant relationships between hypoxic days and January–May TN loads (Table 6 and Fig. 9).

Finally, we evaluated whether the observed water temperature increase in the Bay between 1949 and 2009 (Preston 2004) may have caused an increase in respiration rates and a decrease in re-oxygenation rates (due to decreased solubility and reaeration of oxygen), thereby enhancing DO depletion and increasing hypoxic volume (e.g., Najjar et al. 2010; Lomas et al. 2002). To assess this hypothesis, we estimated the impact of the observed average temperature increases of 0.14°C/decade for surface waters and 0.34°C/decade for the bottom (Preston 2004) on the March to early July respiration and air–water exchange rates. These months were selected because the observed temperature increases have occurred from winter through summer, and DO depletion does not begin until March. From this analysis (SM, Tables S5, S6, S7), the increase in respiration rates due to temperature increases in March–July would range from 0.10% to 4.4% per decade, and the decrease in air–water exchange would range from 0.11% to 7.4% per decade. The observed DO depletion rate from May to early July (i.e., the period with DO decreases, Fig. S11) increased 28% from 1985–1994 to 1995–2004 (SM, Fig. S12, Table S8).

**Table 3** Slope and significance of temporal trend for pycnocline strength by region

Region	Temporal dependence <sup>a</sup> (s <sup>-2</sup> /year)	$p$ value on temporal slope
I	-0.00001	0.8
II	0.0001	0.1
III	0.0001	0.2
IV	0.0002	0.01
V	0.0003	0.02
VI	0.0002	0.05

Each region’s multiple regression model is: PycStrength= $\beta_0 + \beta_1(\text{flow}) + \beta_2(\text{year})$

<sup>a</sup>Regression variables were not scaled to mean of 0 and unit standard deviation (as they were in other regressions) so that the rate of change with year would be clear

### Discussion

The large year-to-year variabilities in hypoxic volume, nitrogen loads (Fig. 3), and stratification (Fig. 6) appear to be associated with the extremely large variability in freshwater flow through the Bay’s major tributaries (e.g., Hagy et al. 2004). Similar large inter-annual variability due to meteorological events controlling the quantity and timing of freshwater fluxes has been observed for other coastal systems, such as the Neuse River estuary (Paerl et al. 1998). Our results show that, in addition to this expected variability, there has been a previously unreported decadal-scale increase in early summer hypoxic volume, as demonstrated with the data from June (SM, Fig. S3a) and early July (Fig. 3a). This inter-annual

**Table 4** Multiple regression models for June average main channel pycnocline strength ( $\max N^2$ ) from 1985 to 2009 (inverse square second)

June stratification models						Residuals regressed with year: $R^2$ , $p$ value
Equation	Model fit: $R^2$ , $p$ value	Intercept, $\beta_0$	River flow dependence (February–May), $\beta_1$	SE wind <sup>a</sup> dependence (June), $\beta_2$	MSL <sup>b</sup> dependence (March–June), $\beta_3$	
$\max N^2 = \beta_0 + \beta_1(\text{flow}) + \beta_2(\text{SEwind})$	0.37, $p=0.007^c$	0.012	0.0007 ( $p=0.1$ )	-0.001 ( $p=0.03$ )	–	0.34, $p=0.002$
$\max N^2 = \beta_0 + \beta_1(\text{flow}) + \beta_2(\text{SEwind}) + \beta_3(\text{MSL})$	0.59, $p<0.001$	0.012	0.0004 ( $p=0.2$ )	-0.001 ( $p=0.002$ )	0.001 ( $p=0.003$ )	0.06, $p=0.3$

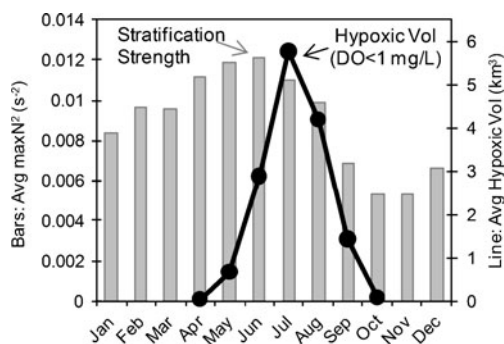
<sup>a</sup> Patuxent Naval Air Station (National Climatic Data Center 2009)

<sup>b</sup> Mean sea level (MSL) 4 months average from Baltimore, Annapolis, Kiptopeke, and Bay Tunnel NOAA stations (National Oceanic and Atmospheric Administration 2010)

<sup>c</sup> Residuals from this regression are significantly ( $p<0.05$ ) autocorrelated because of the temporal trend

increase is only occurring for hypoxia in the early summer; for late July (Fig. 3b) and August (SM, Fig. S3d), declines in hypoxic volume correspond to the slightly decreasing TN loads in recent decades (Fig. 3c). The decreasing TN load we observed is consistent with findings from a USGS analysis that show a decrease in flow-adjusted TN concentrations from 1984 to 2006 at the Conowingo station on the Susquehanna River (Langland et al. 2007). To our knowledge, the recently decreasing TN load trends identified here have not been previously discussed.

Although we have grouped the summer hypoxic volume data into bi-weekly and monthly periods, the seasonal cycle of hypoxia development and dissipation is a continuous process from late spring to early autumn. These results indicate that the timing of maximum hypoxic volume is occurring earlier in the summer and that late summer hypoxic volume is not as severe as it once was. We have explored possible causes for these changes by focusing separately on the periods of early and late July because of consistent data availability and because July is the month for which long-term shifts had been previously observed.



**Fig. 7** Average monthly stratification strength and hypoxic volume for each month over the 25-year record (1985–2009) for the main channel data

### Differential Controls on Hypoxia in Early and Late Summer

We explored multiple hypotheses to explain the different early and late July trends in hypoxia, focusing on investigating whether changes have occurred in nutrient loads, phytoplankton growth, respiration and air–water exchange, or stratification. For our nutrient analyses, the focus was on nitrogen loads from the Susquehanna River for similar reasons of previous studies (e.g., Hagy et al. 2004) because: (1) the Susquehanna is the largest source of TN input to the Bay, (2) the Susquehanna discharges directly into the mainstem, and (3) nitrogen is the limiting nutrient for phytoplankton growth in summer (Fisher et al. 1999). Nevertheless, the Potomac River also provides a substantial TN load to the Bay, and we hypothesized that large changes in TN loads from the Potomac could explain increases in mid- to lower Bay hypoxic volume (Fig. 5b). Although variation in Potomac River TN load explains some of the variations in region IV hypoxic volume, a significant underlying trend of increasing hypoxic volume remains (SM, Fig. S5), indicating the importance of other variables. Not surprisingly, Potomac TN loading was not significantly related to hypoxic volume in the mid- to upper Bay and thus did not improve the regression fits for total Bay hypoxic volume presented in Table 5.

We investigated the hypothesis that the observed trends in early summer hypoxic volumes might reflect a seasonal shift in river flow and associated nutrient inputs to the Bay. Our analysis of daily stream flow, however, indicates that there has been no shift in timing of the winter/spring freshwater flux (SM, Fig. S9), and thus probably no shift in TN load. In other water bodies, observed shifts in the timing of seasonal phytoplankton blooms have been attributed to water temperature increases (e.g., Edwards and Richardson 2004; Kromkamp and Engeland 2010). From the CBP data, however, we found no temporal trend

**Table 5** Multiple regression models for Chesapeake Bay hypoxic volume from 1985 to 2009 (cubic kilometer) fit using principle component regression

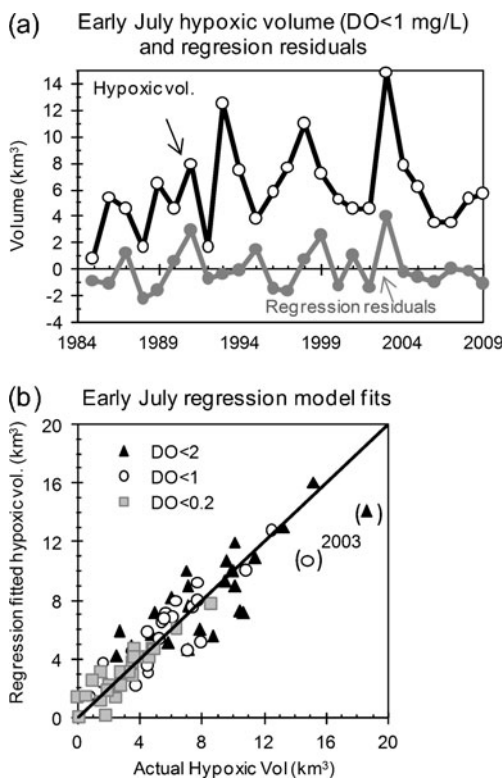
Period	DO level (mg/L)	Model fit: $R^2$ and $p$ value	Sus. TN load dependence (January–May), $\beta_1$	Strength of pycnocline dependence (previous period) <sup>a</sup> , $\beta_2$	Strength of pycnocline dependence (current period) <sup>a</sup> , $\beta_3$	Vol. of sub-pycnocline water dependence (current period) <sup>a</sup> , $\beta_4$
Early July	<0.2	0.85, $p < 0.001$ <sup>b</sup>	0.65	0.88 (June)	0.82 (early July)	<sup>c</sup>
	<1	0.77, $p < 0.001$	<sup>c</sup>	1.2 (June)	1.8 (early July)	0.51
	<2	0.68, $p < 0.001$	<sup>c</sup>	0.90 (June)	2.2 (early July)	0.67
Late July	<0.2	0.37, $p = 0.006$	0.89	<sup>c</sup>	<sup>c</sup>	0.89
	<1	0.46, $p = 0.001$	0.52	0.53 (early July)	<sup>c</sup>	1.4
	<2	0.44, $p = 0.002$	0.48	0.50 (early July)	<sup>c</sup>	1.7

The model is  $\text{HypoxicVol} = \beta_0 + \beta_1(\text{TN}) + \beta_2(\max N_{\text{previous}}^2) + \beta_3(\max N_{\text{current}}^2) + \beta_4(\text{BelowPycVol}_{\text{current}})$ . PCA regression was used because TN and stratification variables are correlated to each other. Additionally, the variables were normalized by mean centering and scaling to unit standard deviation so the  $\beta$ s can be compared

<sup>a</sup>“Previous period” refers to the preceding data collection period for which there are data every year—i.e., for early July cruises it is average June and for late July cruises it is early July. “Current period” refers to the same time period as that in which the hypoxic volume is being regressed

<sup>b</sup>Residuals from this regression are significantly ( $p < 0.05$ ) autocorrelated

<sup>c</sup>Indicates the variable did not improve the regression fit and was excluded from the final model



**Fig. 8** Summary of early July hypoxic volume multiple regression models from Table 5 fit using TN, pycnocline strength, and volume below the pycnocline as variables. The model is:  $\text{HypoxicVol} = \beta_0 + \beta_1(\text{TN}) + \beta_2(\max N_{\text{previous}}^2) + \beta_3(\max N_{\text{current}}^2) + \beta_4(\text{BelowPycVol}_{\text{current}})$  with coefficients fit for early July hypoxic volumes. Hypoxic volume (DO < 1 mg/L) and residuals from the multiple regression model (a) and actual hypoxic volume versus model fit for the three levels of hypoxia (b) are presented; 2003 was identified previously as a year with high nutrient loads from southern Bay tributaries

in timing of the maximum observed surface concentration of chlorophyll-a, which was measured at each main channel station from January to June (data not shown).

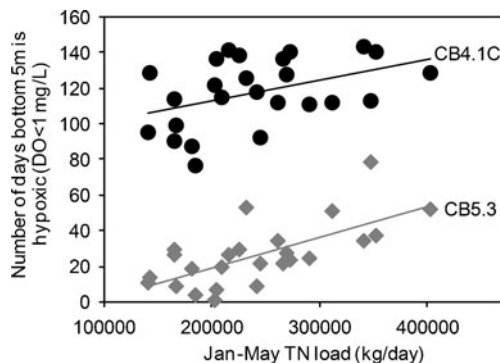
Our analysis to determine whether water temperature increases could have significantly affected respiration and air–water exchange rates (SM, Tables S5, S6, S7, and S8; Figs. S11 and S12) revealed possible increased rates of DO depletion in the spring and early summer months, similar to previously reported calculations (Kemp et al. 2009). Although these rate increases are non-trivial (0.11–7.4% per decade), they are relatively small compared to the actual observed 28% per decade increase in DO depletion from May to early July (SM, Table S8). In addition, if the water temperature increase that has been observed from winter through summer (Preston 2004) was playing a major role in the oxygen depletion, we would expect to see a decrease in DO concentrations throughout the spring and summer months; however, we only observe a decrease in bottom DO concentrations in June and early July (SM, Fig. S11). To be conservative in our analysis, we assumed that the relationship between temperature and planktonic respiration observed for surface waters held for bottom waters, despite actual observations which suggested no relationship between temperature and bottom water respiration (Sampou and Kemp 1994; Kemp et al. 1992). We also gave no consideration to the fact that oxygen production rates (during daylight hours at the surface) would increase with temperature as well. Furthermore, the most significant increases in Bay water temperatures are in the southern Bay where oxygen depletion is less severe than that in the mid- to upper Bay regions (Preston 2004). Thus, based on these observations and computations, we

**Table 6** Pearson correlation coefficients ( $R$ ) between number of days bottom 5 m of water are hypoxic and January–May TN for regions II–IV

Station	DO<0.2 mg/L $R$ ( $p$ value)	DO<1 mg/L $R$ ( $p$ value)	DO<2 mg/L $R$ ( $p$ value)
CB3.3C	0.36 ( $p=0.08$ )	0.12 ( $p=0.6$ )	0.31 ( $p=0.1$ )
CB4.1C	0.42 ( $p=0.04$ )	0.43 ( $p=0.03$ )	0.54 ( $p=0.005$ )
CB4.2C	0.50 ( $p=0.01$ )	0.43 ( $p=0.03$ )	0.54 ( $p=0.005$ )
CB4.3C	0.50 ( $p=0.01$ )	0.46 ( $p=0.02$ )	0.58 ( $p=0.002$ )
CB4.4	0.54 ( $p=0.005$ )	0.55 ( $p=0.005$ )	0.48 ( $p=0.02$ )
CB5.1	0.56 ( $p=0.004$ )	0.44 ( $p=0.03$ )	0.38 ( $p=0.06$ )
CB5.2	0.67 ( $p<0.001$ )	0.47 ( $p=0.02$ )	0.31 ( $p=0.1$ )
CB5.3	0.36 ( $p=0.07$ )	0.68 ( $p<0.001$ )	0.40 ( $p=0.05$ )

conclude that temperature-induced increases in respiration and air–water exchange could explain only a relatively small fraction of the observed increase of early summer hypoxic volume.

Another hypoxia-related hypothesis we investigated is whether stratification has changed during the early summer, thereby restricting mixing and resulting in more hypoxia. Indeed, our results suggest that increasing stratification in June has led to lower bottom-water DO concentrations in June (SM, Fig. S3a), thus enhancing the extent of DO depletion from sub-pycnocline water in early July. Despite the fact that the early July stratification trend is not increasing, the lower initial concentrations of DO at the beginning of the month (from June) provide a lower starting point for DO concentrations in early July, leading to higher hypoxic volumes (Fig. 3a). Inter-annual variation in stratification strength for late July (SM, Fig. S7) exhibits no trend, and long-term trends in later summer hypoxia are more likely controlled by other factors (Table 5; SM, Table S4). In addition to pycnocline strength, the pycnocline depth plays a role in constraining hypoxic volume because it determines how much water is susceptible to hypoxia and was shown to be a factor in the amount of hypoxic volume



**Fig. 9** Hypoxia duration in summer versus Susquehanna January–May TN load for two example stations. See Table 6 for correlation coefficients and other stations

(Table 5), but was not trending throughout the summer (SM, Fig. S8). Similar to these findings, observations on the northwest shelf of the Black Sea indicate that years with earlier onset of stratification (due to warming temperatures and increased freshwater flows) are years when hypoxia area is larger (Ukrainskii and Popov 2009).

Although variations in TN loads may not be responsible for increases in early summer hypoxia, they are strongly correlated with DO depletion throughout the entire summer. Across the three levels of DO depletion studied here, TN loads most significantly explain near-anoxic volume (DO<0.2 mg/L) changes in *both* early and late July (Tables 1 and 5), implying that decreases in TN loads are likely to have the most immediate impact on this severe condition. As the DO level increases, the volume of sub-pycnocline water and stratification strength in the current period become more important variables (Table 5), likely because changes to the depth or strength of the pycnocline will have the most immediate effect on the relatively higher DO conditions closest to the pycnocline. Near-anoxic waters tend to occur at the deeper depths where they are more physically separated from mixing and thus relatively more influenced by nutrient-enhanced degradation of organic matter. Further evaluations of nitrogen loads revealed that they significantly explained variations in the duration of summer hypoxia, suggesting that nitrogen loads have a significant impact on the seasonal persistence of hypoxia (Table 6 and Fig. 9). Previous research has demonstrated the cycle by which the decay products of the spring phytoplankton bloom in the Bay (fed by late winter/spring nutrient fluxes) serve as an internal source of recycled nutrients supporting the growth of summer phytoplankton (Kemp and Boynton 1984; Malone et al. 1988). Thus, the January–May nutrient loads feed a cycle of growth, decay, and recycling that continues throughout the summer, explaining how the duration of summer hypoxia can be correlated with January–May TN loads (Table 6). Similarly, in the Narragansett Bay, which has a much shorter water residence time of 10–40 days, “season-cumulative hypoxia severity” was found to be highly correlated with both the June river flow (which carries nutrients) and June stratification (Codiga et al. 2009).

Overall, the results presented here lead us to conclude that long-term increases in the extent of hypoxia in early summer have been controlled in large part by stratification, while both the extent and persistence of hypoxia during the later summer have been controlled more by nutrient loadings to the Bay (Tables 5 and 6). The idea of a difference in the relative contribution of physical or biological processes in regulating early versus late summer Bay hypoxia was suggested previously (Hagy et al. 2004) and is supported by these results. Research is underway to examine whether similar trends are observed in the Chesapeake Bay tributaries.

## Mechanisms Linking Early Summer Hypoxic Volume and Stratification

We hypothesize that the increasing June pycnocline strength has resulted in a trend of decreasing early summer water-column mixing. This effect is suggested by contrasting the June–July sequence of vertical profiles of density and DO between two different years (1988 and 2000) at one station (Fig. 10). These 2 years had similar spring river flow and nutrient loads (i.e., Susquehanna January–May TN loads in 1988 of  $2.2 \times 10^5$  kg/day and in 2000 of  $2.0 \times 10^5$  kg/day), but very different early July hypoxic volumes (i.e., volume with  $\text{DO} < 1$  mg/L in 1988 was  $1.6 \text{ km}^3$  and in 2000 was  $5.4 \text{ km}^3$ ). The late July hypoxic volumes in these 2 years were more consistent with the respective nutrient loads ( $5.7$  and  $5.4 \text{ km}^3$ ). The pycnocline was at a similar depth (Fig. 10a) in June of these 2 years, but was stronger in June 2000 than in June 1988. The June density profile shows a more gradual change with depth in 1988 than in 2000, and we suggest that this caused the higher deep DO in June and early July of 1988 relative to that seen in 2000. By late July, there was actually slightly more hypoxia in 1988 than 2000, which is consistent with the nitrogen loads.

It is important to note that small changes in pycnocline strength will not cause a major decrease of direct vertical mixing in regimes where the Richardson number, a measure of mixing potential in a stratified environment, remains consistently higher than the value below which vertical mixing is possible (e.g.,  $> 0.25$ ; Miles 1961). The Richardson numbers through the pycnocline are frequently much higher than this cutoff in mid-Bay regions of the Chesapeake (Li et al. 2007). Nonetheless, a reduction in vertical mixing of lower Bay waters, where the pycnocline is not always strong (e.g., right side of Fig. 11), may be relevant to the DO concentrations in the mid-Bay hypoxic regions owing to landward transport of DO in bottom waters along the Bay's axis (Kemp et al. 1992). The vertical gradients of DO and density are strongly correlated in the lower Bay (Fig. 11), and it is also the case that June stratification strength has increased significantly in the lower Bay regions in recent

years (Table 3) so that vertical mixing of DO may now be reduced in regions of the Bay where it was not previously limited. Although the bottom waters in June in the lower Bay are not hypoxic, the presence of lower bottom-water DO concentrations due to restricted vertical mixing in recent years would contribute to less up-Bay transport of bottom-water oxygen to mid- and upper Bay regions.

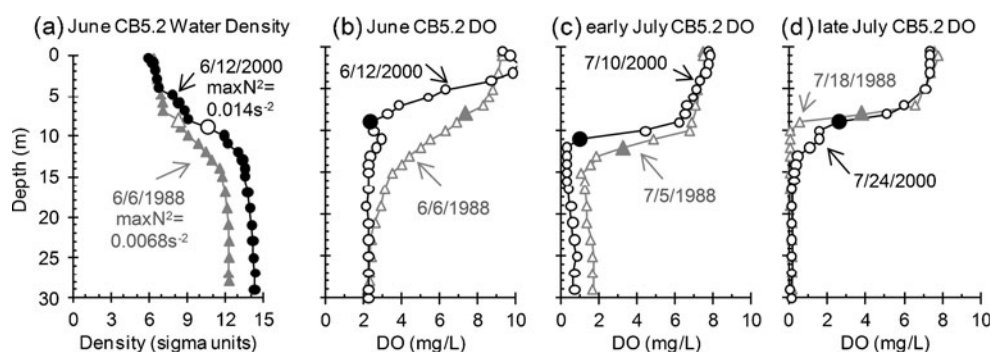
As an alternative or additional cause of increased hypoxia in the deep channel region of the Bay in recent years in June, there may be less lateral replenishment of oxygen from shallower regions. Lateral transport of oxygen from shallower regions to deeper mid-channel regions can occur via mixing from wind events (Scully 2010b), and recent shifts in such wind events have also been implicated as a cause for hypoxia trends (Scully 2010a). Research has also shown that increased vertical stratification can itself be a cause for reduced lateral mixing (Lerczak and Geyer 2004). Therefore, the stronger stratification may be causing a reduction in ventilation of hypoxic waters via reduced lateral transport between shallow flanks and deep channel waters.

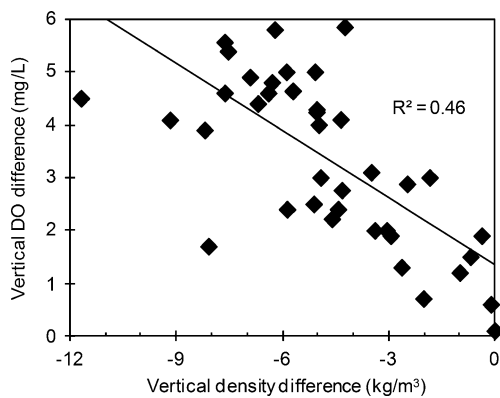
## Potential Causes of June Stratification Trend

Given the observed statistical link between long-term trends in hypoxia and stratification, we have considered potential causes of the observed stratification trends, including the possible role of several climate-related factors on June stratification strength. We stress that these analyses are intended only to identify correlations and trends, and further research will be needed to resolve the relationships and mechanisms more conclusively.

In some North American streams affected by snowmelt, there is evidence that the peak spring flow has shifted earlier in the season (e.g., Hodgkins and Dudley 2006; Regonda et al. 2005). If this was occurring in the Bay watershed, it could play a role in increasing early summer stratification. In eastern North America, however, significant temporal shifts of spring flows have only been observed in watersheds well north of Chesapeake Bay drainage (Hodgkins and Dudley 2006), and our own

**Fig. 10** Example vertical profiles at CB5.2 of June density (a) and DO in June (b), early July (c), and late July (d). The pycnocline location is indicated on each graph by contrasting symbols





**Fig. 11** At CB6.4 in lower Bay, difference in surface to bottom DO versus density for all June sampling days from 1984 to 2008 ( $n=40$ )

analyses of Susquehanna and Potomac flow data indicate no such similar temporal shift for the Chesapeake system (SM, Fig. S9). This is not surprising given that persistent snow packs are not a major source of water in the Bay's watershed. We also see no evidence of an increase in average seasonal river flow into the Bay, as demonstrated by the statistically significant increasing trend in the June stratification regression residuals after accounting for river flow (Table 2,  $p=0.004$ ).

It has been suggested that increased coastal water temperatures could lead to higher stratification (Rabalais et al. 2009), as observed already in the Mediterranean Sea (Coma et al. 2008). In this context, rising water temperatures have indeed been reported in the Chesapeake Bay area, as presumably related to global climate change (Najjar et al. 2010; Preston 2004). Our analysis, however, showed that the changes in water temperature during this period are insufficient to account for the observed increasing pycnocline strength (SM, Fig. S10), likely because the increases in temperature have been observed for both surface and bottom waters of the Bay (Preston 2004).

An increase in salinity at the mouth of the Bay could contribute to stronger stratification in the estuary, and the Gulf Stream Index has been shown to be correlated with salinity in the Bay (Lee and Lwiza 2008). Our analysis of this indicator has, however, revealed no consistent temporal trend during the 25 years of interest and no apparent correlation with June stratification (results not shown).

Scully (2010a) has considered long-term wind patterns in relation to the average July hypoxic volume 1950–2007 record (extended from Hagy et al. 2004) and found significant correlations between July hypoxia and the frequency in occurrence of southeasterly and westerly winds. In particular, regression model results demonstrated that May–July westerly wind frequency together with total nitrogen load can account for an observed change in average July hypoxic volume during the period from 1950 to 2007 (Scully 2010a). The long-term wind record near the

Bay indicates a shift from southeasterly towards more westerly winds in the early 1980s which is explained with variations in two large-scale climatic features: the Bermuda high index and winter index of the North Atlantic Oscillation (Scully 2010a). We found that southeast wind frequency can explain some of the variability in June stratification, but that neither southeast nor westerly wind frequency (west not shown) can explain the temporal increase from 1985 to 2009 in stratification strength (Table 4). The shift in southeast and west wind frequencies (Scully 2010a) occurred in the early 1980s which is before the period we are examining for the stratification increase. Thus, although wind plays a role in explaining inter-annual variations in stratification during the past 25 years, it does not appear to be the variable causing the increase in June stratification strength.

Relative sea level, which has been rising steadily along the east coast of the USA for decades (Barbosa and Silva 2009; Zervas 2009), may result in enhanced stratification strength, as suggested in a report on possible climate change impacts on the Chesapeake (Boesch et al. 2007). We propose that more ocean water (i.e., higher sea levels) will cause a gradual increase in average salinity at the mouth of the Bay. This increase in salinity will tend to increase the estuary's longitudinal salinity gradient. Indeed, a significant positive trend in Chesapeake Bay salinity from 1949 to 2006 has been observed and correlated with sea level rise (Hilton et al. 2008). In addition, hydrodynamic model studies (Hilton et al. 2008) suggest that higher sea levels can cause salt water to intrude further into the Bay and that the increase in salinity from sea level rise scenarios may be the greatest during the times when the mean salinity of the Bay is lowest (Hilton et al. 2008).

The mean salinity of the Bay is generally lowest in April, May, and June (analysis of results from Murphy et al. 2010), at which time the stratification strength is highest (Fig. 7). We therefore expect that the increase in salinity associated with sea level rise would be greatest in these months and that it would also be more confined to bottom waters, thereby increasing its impact on stratification. In fact, we find Bay MSL to be a significant predictor of June pycnocline strength (Table 4). Further analysis shows that even after removing the temporal, flow-related, and wind-related trends from both June stratification and MSL, the residuals of pycnocline strength are positively correlated to MSL ( $p=0.1$ ; SM, Fig. S13). Although this test suggests a causal link between MSL and stratification, the result is not conclusive, and further studies are needed to investigate the impacts of sea level rise on stratification and hypoxia.

In summary, data analyses made possible through decades of water quality monitoring and data archiving, together with recently facilitated access to these data, have led us to the conclusion that a combination of factors have

caused a trend of increasing hypoxic volume over time in early summer, despite stable or even declining nutrient inputs to the Bay in recent years. Statistical correlations point to both the mid-1980s shift in prevailing summer wind directions (Scully 2010a) and the gradual increase (1985 to 2009) in June stratification (this study). We suggest that the increase in stratification is perhaps due to a combination of factors, with the shift to more westerly winds in the mid-1980s causing reduced vertical mixing that in turn helped to confine to the bottom layers the saltier water that is associated with sea level rise.

Overall, it seems clear that physical mechanisms have resulted in less replenishment of DO to bottom waters during the crucial time when DO is being depleted via respiration of the spring bloom, thereby increasing hypoxic volume in early summer. Our analyses also demonstrate that hypoxia duration throughout the summer and hypoxia in the mid-to-late summer are largely controlled by spring nutrient loads, thereby suggesting that nutrient management efforts in the Chesapeake system have had beneficial effects on controlling Bay hypoxia. Furthermore, it appears that if the changes in physical conditions in the early summer had not occurred, we would likely be observing a slight decrease of hypoxic volume throughout the entire summer, instead of just the late summer.

**Acknowledgments** This study was supported by funding from the National Science Foundation under Grant no. 0618986 and the US Department of Commerce, NOAA Grant no. NAO7NOS4780191. We would like to thank EPA Chesapeake Bay Program for access to the monitoring data; Randal Burns and Eric Perlman for CBE0 testbed development; Frank Curriero for the aid with kriging methods; Dominic Di Toro, Jeremy Testa, Damian Brady, and Malcolm Scully for useful discussions; and two anonymous reviewers for helpful comments.

## References

- Bahner, L. 2006. User guide for the Chesapeake Bay and tidal tributary interpolator. NOAA Chesapeake Bay Office, Annapolis, MD. <http://noaa.chesapeakebay.net/Interpolator.aspx>.
- Barbosa, S.M., and M.E. Silva. 2009. Low-frequency sea-level change in Chesapeake Bay: Changing seasonality and long-term trends. *Estuarine, Coastal and Shelf Science* 83: 30–38. doi:10.1016/j.ecss.2009.03.014.
- Boesch, D.F., V.J. Coles, D.G. Kimmel, and W.D. Miller. 2007. Coastal dead zones & global climate change: Ramifications of climate change for Chesapeake Bay hypoxia. Pew Center on Global Climate Change. [http://www.pewclimate.org/regional\\_impacts](http://www.pewclimate.org/regional_impacts).
- Boicourt, W.C. 1992. Influences of circulation processes on dissolved oxygen in the Chesapeake Bay. In *Oxygen dynamics in the Chesapeake Bay: A synthesis of recent research*, ed. D.E. Smith et al., 7–59. College Park: Maryland Sea Grant Publication.
- Box, G.E.P., and D.R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society: Series B: Methodological* 26: 211–252.
- Boynton, W.R., and W.M. Kemp. 2000. Influence of river flow and nutrient loads on selected ecosystem processes. In *Estuarine science: A synthetic approach to research and practice*, ed. J.E. Hobbie, 269–298. Washington: Island.
- Breitburg, D.L. 1992. Episodic hypoxia in Chesapeake Bay: Interacting effects of recruitment, behavior, and physical disturbance. *Ecological Monographs* 62: 525–546.
- Breitburg, D.L. 2002. Effects of hypoxia and the balance between hypoxia and enrichment, on coastal fishers and fisheries. *Estuaries* 25: 767–781. doi:10.1007/BF02804904.
- CBE0 Project Team (Ball, W., D. Brady, M. Brooks, R. Burns, B. Cuker, D. DiToro, T. Gross, W.M. Kemp, L. Murray, R. Murphy, E. Perlman, M. Piasecki, J. Testa, I. Zaslavsky). 2008. Prototype system for multidisciplinary shared cyberinfrastructure: Chesapeake Bay Environmental Observatory. *Journal of Hydrologic Engineering* 13: 960–970. doi:10.1061/(ASCE)1084-0699(2008)13:10(960).
- Chesapeake Bay Program. 1993. Guide to using Chesapeake Bay Program water quality monitoring data. CBP/TRS 78/92, Annapolis, MD.
- Chesapeake Bay Program. 2008. Chesapeake Bay historical data sets. <http://archive.chesapeakebay.net/data/historicaldb/historicalmain.htm>. Accessed 8 Oct 2008.
- Chesapeake Bay Program. 2010. CBP Water Quality Database (1984–present). [http://www.chesapeakebay.net/data\\_waterquality.aspx](http://www.chesapeakebay.net/data_waterquality.aspx). Accessed 20 Jul 2010.
- Codiga, D.L., H.E. Stoffel, C.F. Deacutis, S. Kieman, and C.A. Oviatt. 2009. Narragansett Bay hypoxic event characteristics based on fixed-site monitoring network time series: Intermittency, geographic distribution, spatial synchronicity, and interannual variability. *Estuaries and Coasts* 32: 621–641. doi:10.1007/s12237-009-9165-9.
- Coma, R., M. Ribes, E. Serrano, E. Jimenez, J. Salat, and J. Pascual. 2008. Global warming-enhanced stratification and mass mortality events in the Mediterranean. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* 106: 6176–6181. doi:10.1073/pnas.0805801106.
- Conley, D.J., J. Carstensen, G. Ærtebjerg, P.B. Christensen, T. Dalsgaard, J.L.S. Hansen, and A.B. Josefson. 2007. Long-term changes and impacts of hypoxia in Danish coastal waters. *Ecological Applications* 17(5 Supplement): S165–S184.
- Conley, D.J., J. Carstensen, R. Vaquer-Sunyer, and C.M. Duarte. 2009. Ecosystem thresholds with hypoxia. *Hydrobiologia* 629: 21–29. doi:10.1007/978-90-481-3385-7\_3.
- Cressie, N.A.C. 1993. *Statistics for Spatial Data*, Revised ed. New York: Wiley.
- Cronin, W.B., and D.W. Pritchard. 1975. *Additional statistics on the dimensions of Chesapeake Bay and its tributaries: Cross-section widths and segment volumes per meter depth. Reference 75–3. Special Report 42*. Baltimore: Chesapeake Bay Institute, The Johns Hopkins University.
- Diaz, R.J., and R. Rosenberg. 2008. Spreading dead zones and consequences for marine ecosystems. *Science* 321: 926–929. doi:10.1126/science.1156401.
- Diggle, P.J., and P.J. Ribeiro. 2007. *Model-based geostatistics*. New York: Springer.
- Edwards, M., and A.J. Richardson. 2004. Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430: 881–884. doi:10.1038/nature02808.
- Fisher, T.R., A.B. Gustafson, K. Sellner, R. Lacouture, L.W. Haas, R. L. Wetzel, R. Magnien, D. Everitt, B. Michaels, and R. Karrh. 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. *Marine Biology* 133: 763–778. doi:10.1007/s002270050518.
- Goodrich, D.M., W.C. Boicourt, P. Hamilton, and D.W. Pritchard. 1987. Wind-induced destratification in Chesapeake Bay. *Journal of Physical Oceanography* 17: 2232–2240.
- Guo, X., and A. Valle-Levinson. 2008. Wind effects on the lateral structure of density-driven circulation in Chesapeake Bay.

- Continental Shelf Research* 28: 2450–2471. doi:10.1016/j.csr.2008.06.008.
- Hagy, J.D., W.R. Boynton, C.W. Keefe, and K.V. Wood. 2004. Hypoxia in the Chesapeake Bay, 1950–2001: Long-term changes in relation to nutrient loading and river flows. *Estuaries* 27: 634–658. doi:10.1007/BF02907650.
- Harding, L.W. 1994. Long-term trends in the distribution of phytoplankton in Chesapeake Bay: Roles of light, nutrients and streamflow. *Marine Ecology Progress Series* 104: 267–291.
- Hilton, T.W., R.G. Najjar, L. Zhong, and M. Li. 2008. Is there a signal of sea-level rise in Chesapeake Bay salinity? *Journal of Geophysical Research* 113: 1–12. doi:10.1029/2007JC004247.
- Hodgkins, G.A., and R.W. Dudley. 2006. Changes in the timing of winter–spring streamflows in eastern North America, 1913–2002. *Geophysical Research Letters* 33: L06402. doi:10.1029/2005GL025593.
- Jolliffe, I.T. 2002. *Principal component analysis*, 2nd ed. New York: Springer.
- Kemp, W.M., and W.R. Boynton. 1984. Spatial and temporal coupling of nutrient inputs to estuarine primary production: The role of particulate transport and decomposition. *Bulletin of Marine Science* 35: 522–535.
- Kemp, W.M., P.A. Sampou, J. Garber, J. Tuttle, and W.R. Boynton. 1992. Seasonal depletion of oxygen from bottom waters of Chesapeake Bay: Roles of benthic and planktonic respiration and physical exchange processes. *Marine Ecology Progress Series* 85: 137–152.
- Kemp, W.M., W.R. Boynton, J.E. Adolf, D.F. Boesch, W.C. Boicourt, G. Brush, J.C. Cornwell, T.R. Fisher, P.M. Glibert, J.D. Hagy, L.W. Harding, E.D. Houde, D.G. Kimmel, W.D. Miller, R.I.E. Newell, M.R. Roman, E.M. Smith, and J.C. Stevenson. 2005. Eutrophication of Chesapeake Bay: Historical trends and ecological interactions. *Marine Ecology Progress Series* 303: 1–29.
- Kemp, W.M., J.M. Testa, D.J. Conley, D. Gilbert, and J.D. Hagy. 2009. Temporal responses of coastal hypoxia to nutrient loading and physical controls. *Biogeosciences* 6: 2985–3008.
- Knauss, J.A. 1997. *Introduction to physical oceanography*, 2nd ed. Upper Saddle River: Prentice Hall.
- Kromkamp, J.C., and T.V. Engeland. 2010. Changes in phytoplankton biomass in the Western Scheldt estuary during the period 1978–2006. *Estuaries and Coasts* 33: 270–285. doi:10.1007/s12237-009-9215-3.
- Langland, M.J., D.L. Moyer, and J. Blomquist. 2007. Changes in streamflow, concentrations, and loads in selected nontidal basins in the Chesapeake Bay watershed, 1985–2006. Open File Report 2007-1372. U.S. Geological Survey. <http://pubs.usgs.gov/of/2007/1372/>.
- Lee, Y.J., and K.M.M. Lwiza. 2008. Factors driving bottom salinity variability in the Chesapeake Bay. *Continental Shelf Research* 28: 1352–1362. doi:10.1016/j.csr.2008.03.016.
- Lerczak, J.A., and W.R. Geyer. 2004. Modeling the lateral circulation in straight, stratified estuaries. *Journal of Physical Oceanography* 34: 1410–1428.
- Li, M., L. Zhong, W.C. Boicourt, S. Zhang, and D.-L. Zhang. 2007. Hurricane-induced destratification and restratification in a partially-mixed estuary. *Journal of Marine Research* 65: 169–192.
- Lomas, M.W., P.M. Glibert, F. Shiah, and E.M. Smith. 2002. Microbial processes and temperature in Chesapeake Bay: Current relationships and potential impacts of regional warming. *Global Change Biology* 8: 51–70.
- Ludsin, S.A., X. Zhang, S.B. Brandt, M.R. Roman, W.C. Boicourt, D. M. Mason, and M. Costantini. 2009. Hypoxia-avoidance by planktivorous fish in Chesapeake Bay: Implications for food web interactions and fish recruitment. *Journal of Experimental Marine Biology and Ecology* 381: S121–S131. doi:10.1016/j.jembe.2009.07.016.
- Malone, T.C. 1992. Effects of water column processes on dissolved oxygen, nutrients, phytoplankton and zooplankton. In *Oxygen dynamics in the Chesapeake Bay: A synthesis of recent research*, ed. D.E. Smith et al., 61–112. College Park: Maryland Sea Grant Publication.
- Malone, T.C., L.H. Crocker, S.E. Pike, and B.W. Wendler. 1988. Influences of river flow on the dynamics of phytoplankton production in a partially stratified estuary. *Marine Ecology Progress Series* 48: 235–249.
- Miles, J.W. 1961. On the stability of heterogeneous shear flows. *Journal of Fluid Mechanics* 10: 496–508.
- Murphy, R.R., F.C. Curriero, and W.P. Ball. 2010. Comparison of spatial interpolation methods for water quality evaluation in the Chesapeake Bay. *Journal of Environmental Engineering* 136: 160–171. doi:10.1061/(ASCE)EE.1943-7870.0000121.
- Najjar, R.G., C.R. Pyke, M.B. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M.R. Mulholland, M. Paolisso, D. Secor, K. Sellner, D. Wardrop, and R. Woodm. 2010. Potential climate-change impacts on the Chesapeake Bay. *Estuarine, Coastal and Shelf Science* 86: 1–20. doi:10.1016/j.ecss.2009.09.026.
- National Climatic Data Center. 2009. Global summary of the day, Patuxent Naval Air Station. <http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd>. Accessed 12 Jul 2010.
- National Oceanic and Atmospheric Administration. 2010. Historic tide data. [http://tidesandcurrents.noaa.gov/station\\_retrieve.shtml?type=Historic+Tide+Data](http://tidesandcurrents.noaa.gov/station_retrieve.shtml?type=Historic+Tide+Data). Accessed 17 Jul 2010.
- Officer, C.B., R.B. Biggs, J.L. Taft, L.E. Cronin, M.A. Tyler, and W. R. Boynton. 1984. Chesapeake Bay anoxia: Origin, development, and significance. *Science* 223: 22–27.
- Paerl, H.W., J.L. Pinckney, J.M. Fear, and B.L. Peierls. 1998. Ecosystem responses to internal and watershed organic matter loading: Consequences for hypoxia in the eutrophying Neuse River Estuary, North Carolina, USA. *Marine Ecology Progress Series* 166: 17–25.
- Preston, B.L. 2004. Observed winter warming of the Chesapeake Bay estuary (1949–2002): Implications for ecosystem management. *Environmental Management* 34: 125–139. doi:10.1007/s00267-004-0159-x.
- R Development Core Team. 2008. The R project for statistical computing. <http://www.r-project.org/>. Accessed 24 Aug 2008.
- Rabalais, N.N., R.E. Turner, R.J. Diaz, and D. Justic. 2009. Global change and eutrophication of coastal waters. *ICES Journal of Marine Science* 66: 1528–1537. doi:10.1093/icesjms/fsp047.
- Regonda, S.K., B. Rajagopalan, M. Clark, and J. Pitlick. 2005. Seasonal cycle shifts in hydroclimatology over the western United States. *Journal of Climate* 18: 372–384.
- Ribeiro, P.J., and P.J. Diggle. 2008. geoR: A package for geostatistical analysis using the R software. <http://leg.ufpr.br/geoR/>. Accessed 18 Aug 2008.
- Ripley, B.D., and M. Lapsley. 2008. RODBC: ODBC database access. <http://cran.r-project.org/web/packages/RODBC/index.html>. Accessed 22 May 2008.
- Sampou, P., and W.M. Kemp. 1994. Factors regulating plankton community respiration in Chesapeake Bay. *Marine Ecology Progress Series* 110: 249–258.
- Scully, M.E. 2010a. The importance of climate variability to wind-driven modulation of hypoxia in Chesapeake Bay. *Journal of Physical Oceanography* 40: 1435–1440. doi:10.1175/2010JPO4321.1.
- Scully, M.E. 2010b. Wind Modulation of Dissolved Oxygen in Chesapeake Bay. *Estuaries and Coasts* 33: 1164–1175. doi:10.1007/s12237-010-9319-9.



- Seitz, R.D., D.M. Dauer, R.J. Llansó, and C. Long. 2009. Broad-scale effects of hypoxia on benthic community structure in Chesapeake Bay, USA. *Journal of Experimental Marine Biology and Ecology* 381: S4–S12. doi:10.1016/j.jembe.2009.07.004.
- Turner, R.E., N.N. Rabalais, and D. Justic. 2008. Gulf of Mexico hypoxia: Alternate states and a legacy. *Environmental Science & Technology* 42: 2323–2327.
- Ukrainskii, V.V., and Y.I. Popov. 2009. Climatic and hydrophysical conditions of the development of hypoxia in waters of the northwest shelf of the Black Sea. *Physical Oceanography* 19: 140–150. doi:10.1007/s11110-009-9046-6.
- U.S. Geological Survey. 2010a. Chesapeake Bay river input monitoring program. <http://va.water.usgs.gov/chesbay/RIMP/>. Accessed 17 May 2010.
- U.S. Geological Survey. 2010b. Surface-water data for the nation. <http://waterdata.usgs.gov/nwis/sw>. Accessed 1 Dec 2010.
- Zervas, C. 2009. Sea level variations of the United States 1854–2006. Technical Report NOS CO-OPS 053. National Oceanic and Atmospheric Administration, National Ocean Service, Center for Operational Oceanographic Products and Services, Silver Spring, Maryland. <http://tidesandcurrents.noaa.gov/pub/html/sltrends>.