*Limnol. Oceanogr.*, 59(2), 2014, 373–384

doi:10.4319/lo.2014.59.2.0373

# Nutrient loading and meteorological conditions explain interannual variability of hypoxia in Chesapeake Bay

Yuntao Zhou,[1,2,*] Donald Scavia,[2,3,4] and Anna M. Michalak [1]

[1] Department of Global Ecology, Carnegie Institution for Science, Stanford, California
[2] Department of Civil & Environmental Engineering, University of Michigan, Ann Arbor, Michigan
[3] School of Natural Resources and Environment, University of Michigan, Ann Arbor, Michigan
[4] Graham Sustainability Institute, University of Michigan, Ann Arbor, Michigan

*Abstract*

We use geostatistical universal kriging and conditional realizations to provide the first quantitative estimates, with robust estimates of uncertainties, of the seasonal and interannual variability in hypoxic volume in Chesapeake Bay, covering early April to late October for 1985 to 2010, and explore factors controlling that variability. Results show that the time when the hypoxic volume reaches its maximum has moved from late to early July over the examined period, but that there is no trend in the seasonal-maximum hypoxic volume itself. No significant trend was found in the timing of onset of hypoxia, but the end of the hypoxic period has moved from October to September. Including nutrient loading from the Rappahannock River in addition to the Susquehanna and Potomac Rivers is found to be beneficial for explaining the interannual variability of hypoxia. Overall, January to May total nitrogen loads from these three rivers, April to August southwesterly and northeasterly winds, and April and May precipitation explain $> 85\%$ of the seasonally averaged interannual variability in hypoxic volumes. Southwesterly winds affect hypoxia by increasing vertical stratification, while precipitation likely acts as a surrogate for nonpoint sources of nitrogen downstream from monitoring stations. The relative contribution of nutrient loading to the overall interannual variability suggests that 28–35% reductions in monitored nutrient loads may not be sufficient to achieve a corresponding reduction in hypoxic conditions as had been suggested in previous studies, at least in the short term.

Hypoxia (operationally defined as dissolved oxygen [DO] concentrations $< 2$ mg $L^{-1}$) threatens many large bodies of water around the world, including the Baltic Sea (Sandberg 1994), the Black Sea (Daskalov 2003), the Yangtze River estuary (Li et al. 2002; Chen et al. 2007), the Gulf of Mexico (Rabalais et al. 2001; Obenour et al. 2013), Long Island Sound (Parker and Oreilly 1991; Welsh and Eller 1991), Chesapeake Bay (Officer et al. 1984; Breitburg 1990; Sanford et al. 1990), and Lake Erie (Burns et al. 2005; Carrick et al. 2005; Zhou et al. 2013). Hypoxia has also developed in the eastern Pacific Ocean, the South Atlantic west of Africa, the Arabian Sea, and the Bay of Bengal (Diaz and Rosenberg 2008). In the United States, increased nutrient inputs have led to substantial changes in two-thirds of all coastal systems (National Research Council 2000).

Chesapeake Bay, the largest and the most productive estuary on the East Coast of the United States, has significant bottom-water oxygen depletion due to agricultural and industrial development and population growth along its shores and in its headwaters (Cerco and Cole 1993). While first reported in Chesapeake Bay in the 1930s (Newcombe and Horne 1938; Officer et al. 1984), hypoxia became more common and more widespread in the late 1950s and early 1960s due to increased anthropogenic nutrient influx (Cronin and Vann 2003).

Because nutrient loads and summertime stratification are two primary factors leading to hypoxia in coastal systems (CENR 2003), and because increases in nutrient loads have been linked to expansions in the hypoxic volume in Chesapeake Bay (Flemer et al. 1983; Hagy et al. 2004; Liu and Scavia 2010), management efforts to improve water quality have focused on reducing that loading (U.S. EPA 2002). However, linear correlations between spring tributary nutrient loading and summertime hypoxia are relatively low (Hagy et al. 2004; Scully 2010*b*), and more complex models have increased the explanatory power only marginally (Cerco and Cole 1993; Kemp et al. 2005; Evans and Scavia 2011). It is therefore important to both quantify the uncertainty in the hypoxic volume estimates and explore additional potential drivers in new ways.

One such additional driver is weather patterns (e.g., precipitation, wind) that could play an important role in affecting both nutrient loading and stratification. For example, wind direction has been shown to affect the concentration of DO in the subpycnocline layer of Chesapeake Bay (Malone et al. 1986; Sanford et al. 1990; Lee et al. 2013), and Feng et al. (2012) suggested that wind direction, rather than wind speed, has a strong effect on hypoxia. Previous work also showed that westerly wind over Chesapeake Bay was well correlated with the summer hypoxic volume between 1950 and 2007 (Scully 2010*a*,*b*). However, Murphy et al. (2011) did not detect a significant effect of wind for 1985 to 2009. This discrepancy points to the need for a detailed investigation of how wind influences hypoxia, and the extent to which that effect can be incorporated into management-based models.

To explore the effects of nutrient loads and weather patterns in more detail, it is important to work from robust estimates of the hypoxic volume, including the uncertainty

* Corresponding author: ytzhou@stanford.edu

in those estimates. To map the detailed hypoxic conditions throughout the bay, Bahner (2006), Murphy et al. (2011), and Lee et al. (2013) provided hypoxic volume estimates from June to August starting from 1985, when the Chesapeake Bay Program began continuous water quality monitoring. However, the methods used in these previous analyses, namely inverse distance weighting, two-dimensional geostatistical ordinary kriging, and the Data Interpolating Variational Analysis (DIVA) software, respectively, cannot quantify the uncertainty associated with their estimates of hypoxic volume, and make little or no use of ancillary information about location and bathymetry. Because Chesapeake Bay is extremely narrow with highly variable depth, it creates some unique challenges for spatial interpolation.

Here we apply three-dimensional geostatistical universal kriging (UK; also known as kriging with an external drift) and conditional realizations, a set of approaches that makes it possible to incorporate information on auxiliary variables and to quantify the uncertainty associated with hypoxic volume estimates. The goals of this work are to (1) quantify hypoxic volumes and associated uncertainties for early April to late October for 1985 to 2010, (2) assess the seasonality and annual trends of hypoxia over these 26 yr, and (3) explore the relative roles of nutrient loading and weather patterns in explaining the interannual variability of hypoxia.

## Methods

*Data for estimating the spatial distribution of DO*—Chesapeake Bay includes many sub-estuaries and tributaries (Fig. 1), but the analysis presented here is restricted to the main stem of the bay. The in situ DO data were collected by the Chesapeake Bay Program and their collaborators and were obtained from the Chesapeake Bay Program Water Quality Database (Chesapeake Bay Program 2011). This database includes measurements starting in July 1984, with stations sampled once each month during late fall and winter months and twice each month from early April to late October. Each sampling cruise generally covered periods of less than a week. There are about 40 fixed monitoring stations in the main stem at which water quality parameters are measured at vertical intervals of 1–2 m through the water column (Fig. 1).

In addition to the in situ DO measurements, auxiliary variables with full spatial coverage included in the DO analysis are longitude, latitude, bathymetry, and measurement depth (i.e., the depth at which each measurement is collected). The bathymetry data, part of the National Oceanic and Atmospheric Administration (NOAA) Coastal Geospatial Data Project, were averaged to 1 km × 1 km horizontal resolution from their initial 1 m × 1 m resolution. All locations in this work were georeferenced using the coordinate system North American Datum of 1983 (NAD83), Universal Transverse Mercator (UTM) Zone 18 North.

*Data for analyzing drivers of interannual variability of hypoxic volume*—Nutrient load, precipitation, and wind

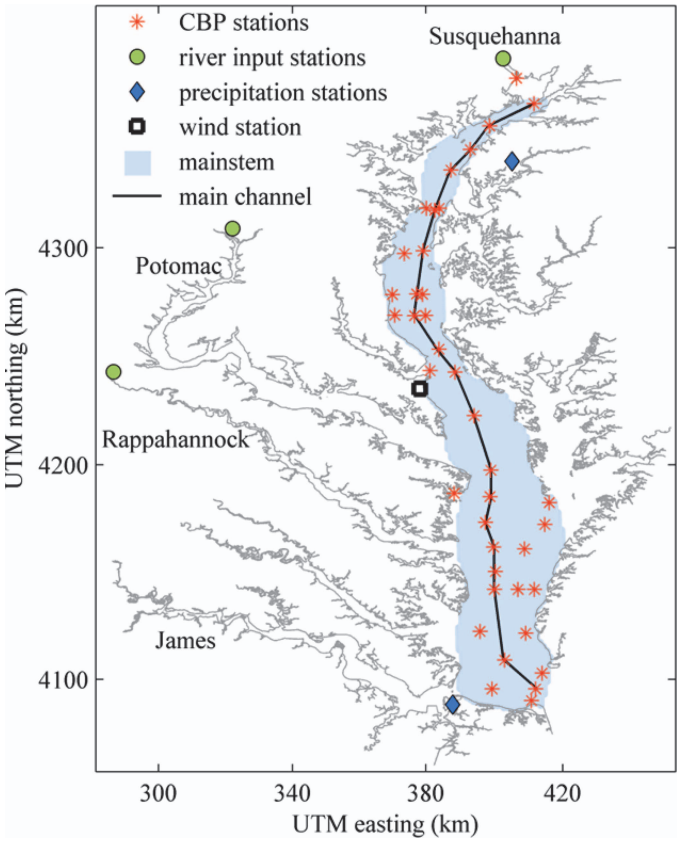


Fig. 1. Locations of the Chesapeake Bay Program (CBP) monitoring stations in the main stem of Chesapeake Bay, river input monitoring stations, precipitation monitoring stations, and the wind monitoring station. The light blue shaded area represents the main stem of the bay where the analysis was conducted. The solid black line represents the main channel of the bay.

data were used to analyze the interannual variability in the estimated hypoxic volumes. As nitrogen is the limiting nutrient for phytoplankton growth in summer (Fisher et al. 1999), we focus the nutrient loading analysis on total nitrogen (TN).

Monthly average TN loads data have been collected by the U.S. Geological Survey (USGS) Chesapeake Bay River Input Monitoring Program (USGS 2010) since 1981 for the main tributaries. We used TN loads from January to May because they have been shown to be a significant driver of hypoxia in other models (Hagy et al. 2004; Scavia et al. 2006; Murphy et al. 2011). TN loads covering the earlier portion of the spring (January–March; January–April) were also tested, but the January–May loads provided the best explanatory power. The Susquehanna, Potomac, James, and Rappahannock Rivers provide the largest nutrient loads to Chesapeake Bay. However, the James River is located at the mouth of the bay and its nutrient loading likely has little effect on conditions in the bay. As a result, loading from the James River was not considered in this work. Because the Susquehanna, Potomac, and Rappahannock Rivers all flow directly into the bay and together account for $> 90\%$ of the TN of all tributary inputs, we used TN loads from these three tributaries,

testing the explanatory power of loading from individual tributaries as well as their combined loads.

Precipitation and wind data were obtained from the NOAA's National Climatic Data Center (NCDC 2012) and the National Weather Service (NWS 2012). We used the average of observations from the Chestertown Naval Air Station (NAS), and the Maryland and Norfolk NAS stations, which both provide long-term observations (Fig. 1). Preliminary analysis showed that using total April–May precipitation yielded a model with the strongest explanatory power relative to periods that included earlier months in the year (i.e., January–May, February–May, March–May). Wind data are obtained at Patuxent River NAS, located close to the middle of the bay (Fig. 1). We considered both wind direction (including the duration of wind from cardinal or intercardinal directions) and average wind speed. We used April–August wind speed and direction because the dominant wind direction changes after August.

*Overall approach*—The approach used to quantify hypoxic volume is similar to the geostatistical methods described in Zhou et al. (2013), but expanded from two to three dimensions. The DO distribution and hypoxic volume estimates were obtained from early April to late October at half-monthly intervals for 1985 to 2010. The data for each half-month period were analyzed individually across years, without assuming any temporal covariance among different periods or different years. This was based on preliminary analysis that showed no significant temporal covariance between estimates once an average seasonal cycle was removed. The spatial resolution of the estimated DO distribution is 1 km (latitude) × 1 km (longitude) × 1 m (depth).

*Model selection for explaining the spatial distribution of DO*—Although a linear model with more parameters will always capture more of the observed variability, some of these variables may serve only to reproduce spurious correlations (Faraway 2005), thereby confounding the analysis. Various statistical approaches have been developed to identify a subset of explanatory variables that best represents the observed variability, while limiting the risk of including extraneous variables. Among these approaches, the Bayesian Information Criterion (BIC) is generally the recommended tool when explanatory power and inference are of principal interest (Ward 2008). BIC is based on the Bayes factor or the posterior probability of a model, and considers both the goodness of fit and the number of variables in a candidate model (Schwarz 1978; Anderson et al. 1998). Similarly to Zhou et al. (2013), BIC was evaluated for every possible combination of auxiliary variables using a form of BIC that accounts for correlated residuals (Mueller et al. 2010), and the set of variables with the lowest BIC defined the best model.

The auxiliary variables considered as potential explanatory factors for DO concentration were longitude (UTM easting), latitude (UTM northing), bathymetry, and measurement depth, as described in the previous section. All variables were found to be significant in explaining DO for the majority of the half-month periods. For consistency across time periods, all four auxiliary variables were therefore used for all time periods in the final analysis.

*UK for estimating the spatial distribution of DO*—In applying UK, all 26 yr of DO data ($\mathbf{z}$, $n \times 1$) for a given half-month period were organized as follows:

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_{1985} \\ \mathbf{z}_{1986} \\ \vdots \\ \mathbf{z}_{2010} \end{bmatrix} \tag{1}$$

where $\mathbf{z}_i$ ($i$ = 1985, …, 2010) is an $n_i \times 1$ vector, and $n_i$ is the number of measurements within a given half-month period (e.g., early July) during year $i$ (i.e., $n = \sum_{i=1985}^{2010} n_i$). Within the UK framework, the DO distribution was modeled as the sum of a deterministic term and a zero-mean spatially correlated residual term. The deterministic term represents the portion of the DO distribution that can be explained by categorical variables (ones and zeros) binning the data by year, and by the auxiliary variables described in the last section. The residual term represents the remaining portion of the observed variability. The categorical variables represent the spatially constant yearly offsets (i.e., intercepts) corresponding to each year. Therefore, the measurement data ($\mathbf{z}$) can also be written as:

$$\mathbf{z} = \mathbf{X}_\mathbf{z} + \mathbf{z}_{\text{res}} \tag{2}$$

where $\mathbf{X}_\mathbf{z}$ is a known $n \times (26 + 4)$ matrix of 26 categorical variables (one per year over the examined period) and four auxiliary variables that explain a portion of the DO variability, $\boldsymbol{\beta}$ is a $(26 + 4) \times 1$ vector of unknown regression coefficients corresponding to these variables, and $\mathbf{z}_{\text{res}}$ is an $n \times 1$ vector of residuals. Note that the $\boldsymbol{\beta}$ values are different for each half-month period. Overall, $\mathbf{X}_\mathbf{z}$ is expressed as:

$$\mathbf{X}_\mathbf{z} = \begin{bmatrix} \mathbf{1}_{1985} & \cdots & \mathbf{0} & \mathbf{X}_{1985} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{1}_{2010} & \mathbf{X}_{2010} \end{bmatrix} \tag{3}$$

where $\mathbf{1}_i$ ($i$ = 1985, …, 2010) are $n_i \times 1$ vectors of ones and $\mathbf{X}_i$ ($i$ = 1985, …, 2010) are $n_i \times 4$ matrices of auxiliary variables.

An $n \times n$ covariance matrix ($\mathbf{Q}_{\mathbf{zz}}$) is used to represent the spatial covariance structure of the residual term ($\mathbf{z}_{\text{res}}$), as:

$$\mathbf{Q}_{\mathbf{zz}} = \begin{bmatrix} \mathbf{Q}_{1985} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{Q}_{2010} \end{bmatrix} \tag{4}$$

where $\mathbf{Q}_i$ ($i$ = 1985, …, 2010) are covariance matrices for the residuals within a given half-month period (e.g., early July) for each year. The $\mathbf{Q}_i$ matrices for all years use the same covariance parameters:

$$Q_i(h) = \begin{cases} \sigma^2 + \sigma_Q^2; & h_{ew} = 0,\ h_{ns} = 0,\ h_v = 0 \\ \sigma^2 \exp\left(-\sqrt{\left(\frac{h_{ew}}{l_{ew}}\right)^2 + \left(\frac{h_{ns}}{l_{ns}}\right)^2 + \left(\frac{h_v}{l_v}\right)^2}\right); & \text{otherwise} \end{cases} \quad (5)$$

where $\sigma^2$ is the variance in $(\text{mg L}^{-1})^2$ of the portion of the residual DO variability that is spatially correlated and $\sigma_{\mathbf{Q}}^2$ is the measurement error, also in $(\text{mg L}^{-1})^2$. $h_{ew}$, $h_{ns}$, and $h_v$ are the separation distances in kilometers between measurement locations along east–west, north–south, and vertical directions, respectively. $3l_{ew}$, $3l_{ns}$, and $3l_v$ are the practical correlation ranges along the east–west, north–south, and vertical directions, respectively.

These three covariance parameters were estimated through Restricted Maximum Likelihood (REML) for each half-month period. REML maximizes the likelihood of available observations after marginalizing with respect to the unknown regression coefficients ($\boldsymbol{\beta}$; Corbeil and Searle 1976; Kitanidis and Lane 1985; Zhou and Michalak 2009). Because of the stratification of the water column in Chesapeake Bay from late spring to late summer, the DO residual (i.e., $\mathbf{z}_{\text{res}}$) profiles still have discontinuities near the pycnocline, complicating parameter estimation. To circumvent this problem, covariance parameters in the horizontal and vertical directions were obtained separately and then combined. These covariance parameters are different for each half-month period. Taking early July as an example, Eq. 5 becomes ($Q_i(h)$, unit: $[\text{mg L}^{-1}]^2$):

$$Q_i(h) = \begin{cases} 3.4 + 0.6; & h_{ew} = 0,\ h_{ns} = 0,\ h_v = 0 \\ 3.4\exp\left(-\sqrt{\left(\frac{h_{ew}}{26}\right)^2 + \left(\frac{h_{ns}}{98}\right)^2 + \left(\frac{h_v}{0.013}\right)^2}\right); & \text{otherwise} \end{cases} \quad (6)$$

This equation suggests that the DO in early July is spatially correlated within $3 \times 26$ km ($3l_{ew}$) in east–west direction, within $3 \times 98$ km ($3l_{ns}$) in north–south direction, and within $3 \times 0.013$ km ($3l_v$) in vertical direction. As expected, the variability of DO concentration is most correlated (i.e., the correlation length is longest) along the north–south direction because the surface freshwater flows from north to south, and the bottom sea water flows from south to north.

Once the measurement vector $\mathbf{z}$, the model of the trend $\mathbf{X}_\mathbf{z}$, and the covariance $\mathbf{Q}_{\mathbf{zz}}$ are defined, DO is estimated throughout the main stem at a 1 km (latitude) $\times$ 1 km (longitude) $\times$ 1 m (depth) resolution for each half-month across all years, using standard UK equations as also implemented in Zhou et al. (2013).

*Conditional realizations for quantifying hypoxic volumes and associated uncertainties*—UK provides a direct and quantitative assessment of the DO concentration, including an assessment of its uncertainty, but cannot be used directly to estimate the uncertainty associated with hypoxic volumes (i.e., the total volume for which DO concentration is below a 2 mg $\text{L}^{-1}$ threshold). To provide probabilistic estimates of hypoxic volume, conditional realizations of DO (Gutjahr et al. 1994; Kitanidis 1995; Zhou et al. 2013) were generated by sampling the uncertainty covariance of the DO distribution. Conditional realizations were only generated for locations deeper than 2 m, because the pycnocline depth is normally around 10 m (M. Olson and G. Shenk unpubl.) and hypoxia is not expected to occur at these shallower depths. The hypoxic volume was calculated for each conditional realization by summing the volumes where the predicted DO concentration is below 2 mg $\text{L}^{-1}$. A thousand conditional realizations were generated for each half-month period.

*Model selection for explaining interannual variability of hypoxia*—After obtaining the estimated hypoxic volume, BIC was again used to test the significance of nutrient loads and meteorological variables in explaining the interannual variability in hypoxic volume. The years 2007 and 2008 were not included in the interannual analysis, due to sparse sampling during those years that focused on the lower portion of the main stem, which is less prone to hypoxia, as will be discussed in detail in later sections. Candidate variables were as described in the data section. Due to the collinearity among the examined candidate variables, BIC was applied iteratively with different subsets of variables, as described in the data and results sections.

## Results

We examined the estimated frequency of occurrence of hypoxia by location from 1985 to 2010, followed by an analysis of the seasonal variability of hypoxic volumes, and an analysis of the drivers of its interannual variability.

*Spatial distribution of hypoxic conditions*—The fraction of years for which hypoxic conditions occur at a given location for each half-month period was estimated using the UK best estimates of DO from 1985 to 2010, providing the first spatially explicit frequency analysis of the occurrence of hypoxia in Chesapeake Bay. As expected, hypoxia was most frequent 10–15 m below the water surface (i.e., along the main or central channel of the main stem; Figs. 2, 3). This is consistent with the fact that this deepest channel has the strongest stratification, and has an observed pycnocline depth of about 10 m (M. Olson and G. Shenk unpubl.). In addition, because this deep channel is narrow and therefore has a small volume, it can become oxygen depleted quickly. The near-shore areas and the lower bay become hypoxic less frequently because they are not sufficiently deep to develop strong stratification.

In all years, hypoxia first appears in the upper bay (Figs. 2, 3), driven by nitrogen input from the Susquehanna River at the head of the bay, which accounts for about 64% of the total tributary TN loads in spring (Murphy et al. 2011). This can occur as early as April in some years. The timing is also driven by the development of stratification, resulting from snowmelt and increased precipitation,

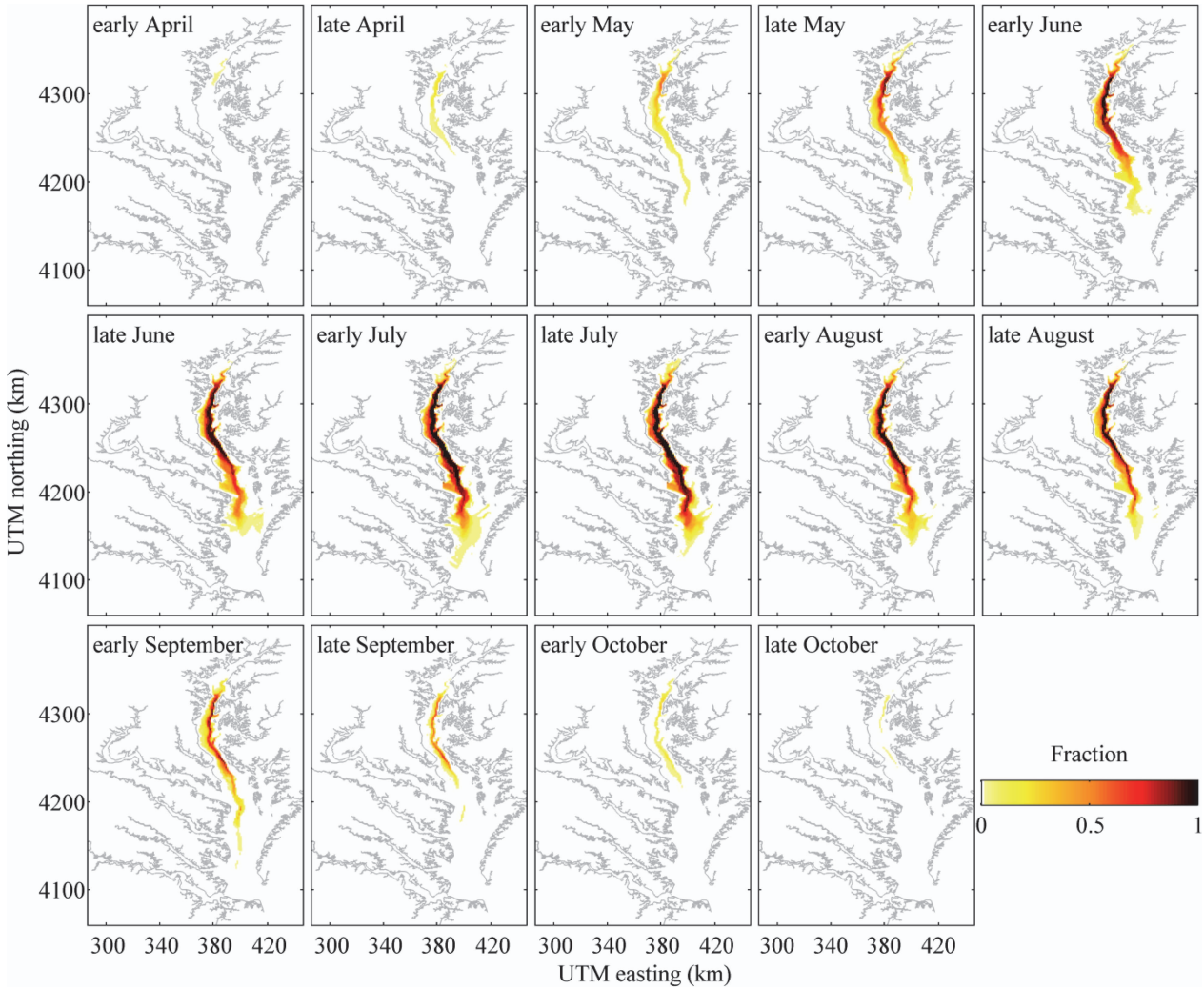


Fig. 2. Estimated frequency of hypoxia for early April to late October. Frequency is expressed as the fraction of years from 1985 to 2010 when the dissolved oxygen (DO) concentration is estimated using UK to have been below 2 mg $L^{-1}$ at a given location for at least one point in the water column.

leading to increased freshwater inflow and stronger vertical density gradients. By June, hypoxia begins to spread toward the middle and lower zones, typically peaking in July in both spatial extent and the likelihood of hypoxia at a given location. The hypoxic zone starts to recede in late August, beginning in the lower bay and receding northward, driven by cooling surface waters that break up the stable stratification and reoxygenation of bottom waters. By October, hypoxia has disappeared entirely in most years.

*April to October hypoxic volumes*—Hypoxic volumes and associated uncertainties are presented in Fig. 4 and the Web Appendix (www.also.org/lo/toc/vol_59/issue_2/0373a.html). For periods when DO data are not available, hypoxic volumes were approximated by averaging values for the same half-month period in other years, and these periods are indicated using white, rather than filled, uncertainty bars. We averaged across years instead of averaging earlier and later half-months within the same year because there was no within-year temporal correlation after a long-term average seasonality was removed from the estimates. The uncertainty standard deviations for these unsampled periods were approximated as the standard deviation across the hypoxic volumes for the same half-month period in other years, and are much higher, as expected, relative to periods with available DO observations.

Prior to the analysis presented here, the most expansive analyses of hypoxic volumes in the main stem of Chesapeake Bay were those conducted using the Chesapeake Bay Interpolator (CBI) Tool (Bahner 2006) and through two-dimensional ordinary kriging (OK) of DO (along northing and vertical directions) by Murphy et al. (2011) from June to September and from May to September, respectively, starting in 1985 (Fig. 4). These earlier estimates did not cover the full seasonal cycle of hypoxia, did not take advantage of the auxiliary variables used here, did not provide uncertainties associated with the estimates, and in the case of the OK analysis, assumed that the vertical DO profile estimated in the main channel remained constant laterally across the bay and thus did not

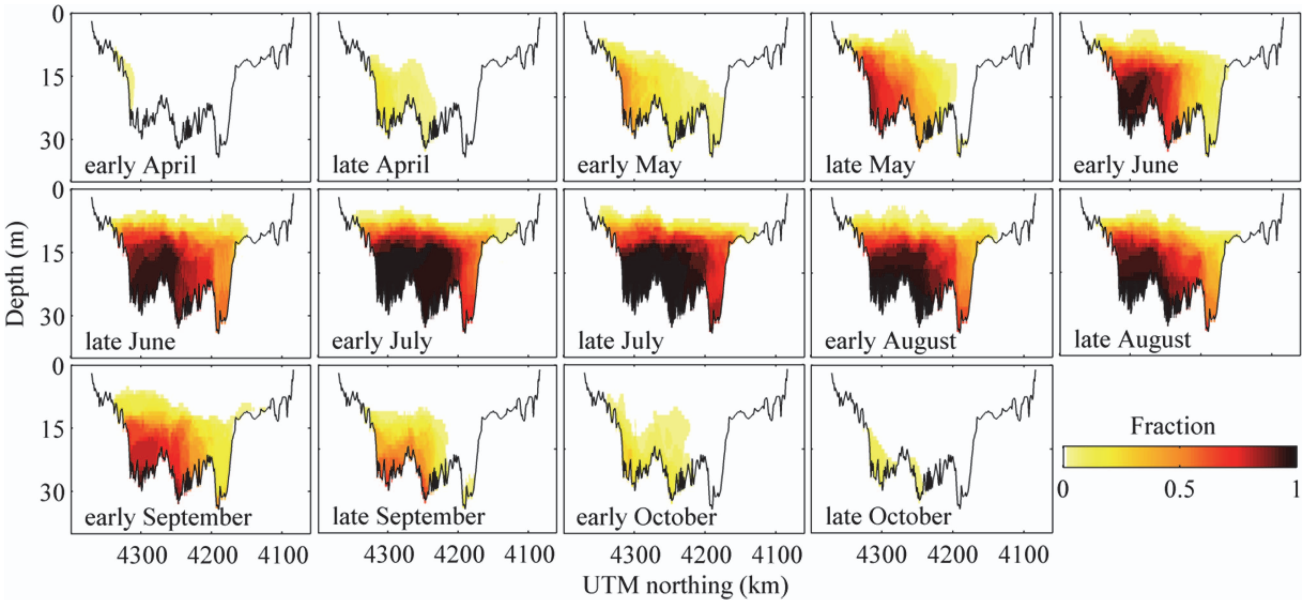


Fig. 3. Estimated frequency of hypoxia for early April to late October. Frequency is expressed as the fraction of years from 1985 to 2010 when the dissolved oxygen (DO) concentration is estimated using UK to have been below 2 mg $L^{-1}$ at a given depth along the main channel.

represent any variability in DO across the main stem (Fig. 1). Our estimates are generally consistent with these two previous analyses for overlapping months, but we find that the OK estimates (Murphy et al. 2011) were higher than our estimates due to their assumption that hypoxia in the main channel extended laterally at a given depth (Fig. 4). This effect is most significant ($p = 0.01$) during the height of the hypoxic period in July and August, when the OK estimates were higher by an average of 1.4 $km^3$ (21%) and 1.2 $km^3$ (25%), respectively. CBI estimates are also slightly higher than those presented here (Fig. 4), but the differences are not significant given the uncertainties on the estimates.

Results confirm that hypoxic volumes exhibit a strong seasonal cycle, with the maximum volumes consistently found in July or August (Fig. 4). We find that, because the summer months also have large areas with estimated DO near the 2 mg $L^{-1}$ hypoxic threshold, summer estimates also have higher absolute uncertainties (Fig. 4).

The duration of hypoxic conditions and the seasonal timing of hypoxia changed from 1985 to 2010 (Fig. 5). We used a threshold volume of 0.85 $km^3$ to explore changes in the onset and duration of hypoxia. This threshold was selected because it represents a typical uncertainty in the volume estimates, obtained by averaging the variance associated with individual periods with DO observations, and then defining a typical uncertainty as plus or minus the square root of that average variance. Whereas there is no significant trend for the initiation of hypoxic conditions ($p = 0.27$), the duration of hypoxic conditions has decreased from 5 to 4 months ($p = 0.02$), and the end of the hypoxic season has moved from October to September ($p = 0.005$). The conclusion about the end of hypoxic season is unchanged if years with missing data in September are removed from the analysis. Although hypoxic duration decreased, annual mean hypoxic volumes show no clear trend (Fig. 6), because the hypoxic volumes in September and October are much lower than those in the summer months.

Finally, the time with the peak hypoxic volume has moved from late to early July ($p = 0.07$; Fig. 5). Murphy et al. (2011) suggested that this shift might be due to sea level rises because rising sea level could result in enhanced stratification. However, we did not find a significant linear correlation with sea level.

*Interannual variability of hypoxic volumes*—Over the 26 yr period, the largest hypoxic zone occurred in early July 2003, likely resulting from heavy rainfall that spring and associated large nutrient loading (Fig. 4; Lewis et al. 2007). However, hypoxic conditions decreased sharply by late July of that year due to storm-induced mixing of the water column (Lewis et al. 2007). Such results imply that hypoxic volumes at synoptic (i.e., sub-monthly) time scales can be greatly affected by isolated meteorological events.

To avoid such short-term effects, we focus our interannual analysis primarily on seasonally averaged hypoxic conditions. For example, peak hypoxic volume in summer 1996, a year with one of the highest nitrogen loads, was lower than that in 1995 (Hagy et al. 2004), but averaging over the entire season reveals that the mean hypoxic volume from April to October for 1996 was indeed higher than that for 1995 (Fig. 4). The average hypoxic conditions are expressed here as the mean estimated hypoxic volume ($km^3$) from April to October. Preliminary analyses confirmed that a model of interannual variability targeting the seasonally averaged hypoxic volume was more powerful ($R^2 = 0.85$) than one targeting any specific half-month period ($R^2 \leq 0.43$) or one targeting the maximum observed hypoxic volume ($R^2 = 0.46$) in a given year.

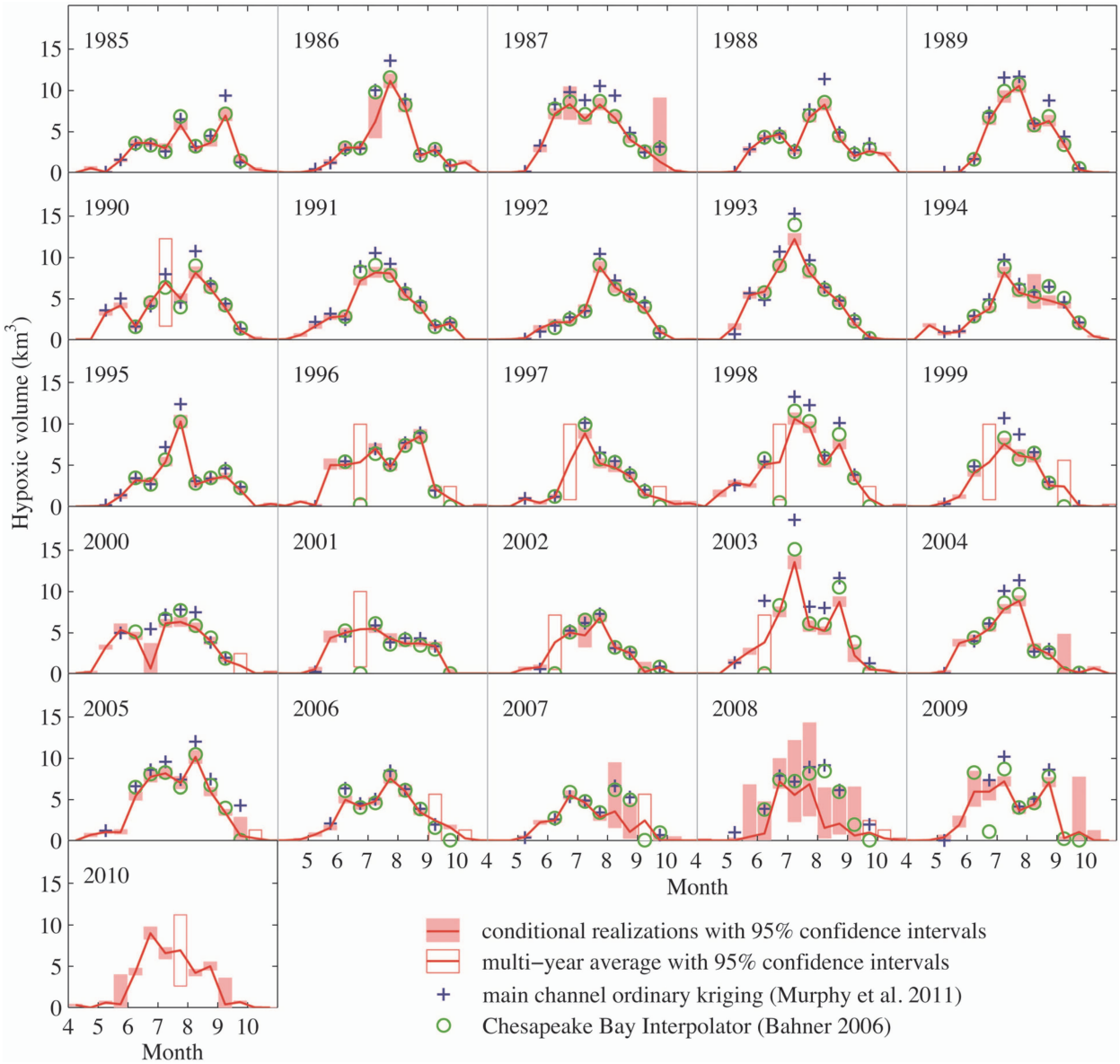


Fig. 4. Estimated hypoxic volumes, expressed as the median of 1000 conditional realizations, and their associated 95% confidence intervals, also derived from the ensemble of 1000 conditional realizations, from April to October 1985–2010 (*see* Web Appendix). The white uncertainty bars represent time periods for which DO measurements are not available, and hypoxic volumes and uncertainties were estimated for these periods as described in the text. Also included for comparison are the estimated hypoxic volumes for concurrent periods from the Chesapeake Bay Interpolator (Bahner 2006), and from two-dimensional ordinary kriging along the main channel, extrapolated across the main stem, as reported in Murphy et al. (2011).

The candidate variables described previously, including TN loads, precipitation, wind speed, and wind direction, were evaluated using BIC to identify a best subset to explain the interannual variability of mean hypoxic volume. Note that 2007 and 2008 had much sparser sampling relative to other years, and the sampling that did take place focused on the lower portions of the bay that are less affected by hypoxia, leading to high uncertainties on estimated hypoxic volumes (Fig. 4) and estimates that may be biased low. The estimated average hypoxic volumes for these two years were therefore not included in the analysis of the interannual variability.

Total January to May combined TN loads from the Susquehanna, Potomac, and Rappahannock Rivers were selected through BIC. Loads from the Susquehanna River have received the most attention in past research because they contribute the most nitrogen to the system (Hagy et al. 2004). In addition to the Susquehanna, Murphy et al. (2011) included loading from the Potomac River in their analysis of the variability of hypoxic volume. Here, we find that including the loading from the Rappahannock together with Susquehanna and Potomac in a combined loading term yields a model with marginally higher explanatory power than a modeling using the combined

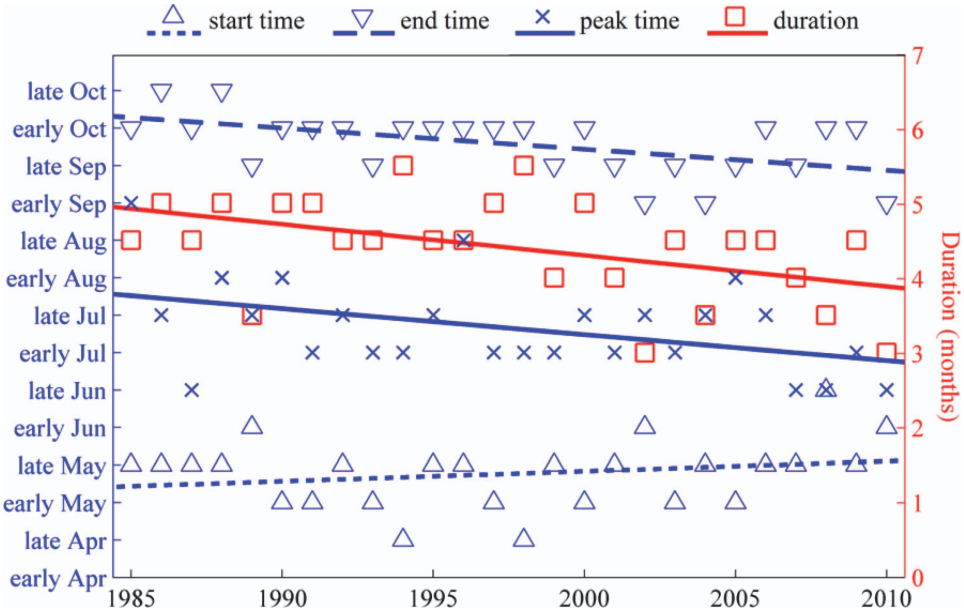


Fig. 5. Start and end time of hypoxic conditions in Chesapeake Bay, time of peak hypoxic volume, and total duration of hypoxic conditions. All quantities expressed at a half-monthly resolution.

loading from only the Susquehanna and Potomac. The relative contribution of the Potomac and Rappahannock Rivers is higher for years with larger TN loads ($p = 0.004$), indicating that loading from these rivers is particularly important in years with high hypoxic volumes. For example, in years when the total loading was in the range of 30–40 $\times$ $10^6$ kg, these two rivers contributed on average 25% of the total load of the three rivers, whereas when the total loading was above 50 $\times$ $10^6$ kg they contributed 31%. We also tested models with multiple variables representing loading from individual tributaries, but the reduction in the number of degrees of freedom yielded a higher (i.e., worse) BIC score relative to simply using the combined load. Springtime (April–May) precipitation was also selected using BIC, and we argue below that it likely represents unmonitored nitrogen loading from nonpoint sources directly to the bay and to tributaries downstream of the monitoring stations.

The explanatory power of the duration of April to August winds from the four cardinal directions (north, east, south, west), as well as the intercardinal directions (southwest [SW], northeast [NE], southeast, northwest), was also tested to reflect the potential effect of wind direction on water column stratification. Some studies have suggested that wind speeds need to be above a given threshold to have a substantive effect on stratification (Hunter et al. 2008), and we therefore also examined the correlations between wind duration from cardinal and intercardinal directions above various wind speed thresholds (6, 8, 10, 12, and 15 m $s^{-1}$) and stratification, as approximated using the average pycnocline depth across monitoring stations (Fig. 1). We found that limiting the directional analysis to winds above any of the examined thresholds decreased the correlation with stratification, and the subsequent analysis therefore used total wind duration by direction, irrespective of wind speed.

Further preliminary analysis showed that models using wind along intercardinal directions outperformed models using cardinal directions. In addition, among the intercardinal directions, only southwesterly (negatively correlated to hypoxic volume) and northeasterly (positively correlated to hypoxic volume) wind durations were BIC selected, and replacing these two variables by their ratio (SW wind duration/NE wind duration), termed henceforth the "dominant wind effect," was found to yield a further reduced (i.e., improved) BIC score. This ratio is above 1 when southwesterly winds are more frequent than northeasterly winds, and below 1 when northeasterly winds dominate. It is this ratio of wind durations that forms the basis of the discussion in the next section. Furthermore, although Lee et al. (2013) suggested that winter–spring NE–SW wind speed affected hypoxic volume by affecting the distribution of phytoplankton biomass, we did not find spring–summer wind speed to be as significant as wind direction.

Overall, total April–May precipitation (**P**, mm), cumulative January–May TN loading (**N**, $10^6$ kg) from the Susquehanna, Potomac, and Rappahannock, and the April–August dominant wind effect (**W**, i.e., SW wind duration/NE wind duration) were selected by BIC for the final model, and together explained 85% of the variability in the mean hypoxic volume from early April to late October for 1985 to 2010, excluding 2007 and 2008 as described earlier due to data limitations (Fig. 6):

$$\mathbf{V} = 1.94 + 0.021 \times \mathbf{N} + 0.0029 \times \mathbf{P} - 0.38 \times \mathbf{W} \qquad (7)$$

where **V** ($km^3$) represents the predicted annual mean hypoxic volume.

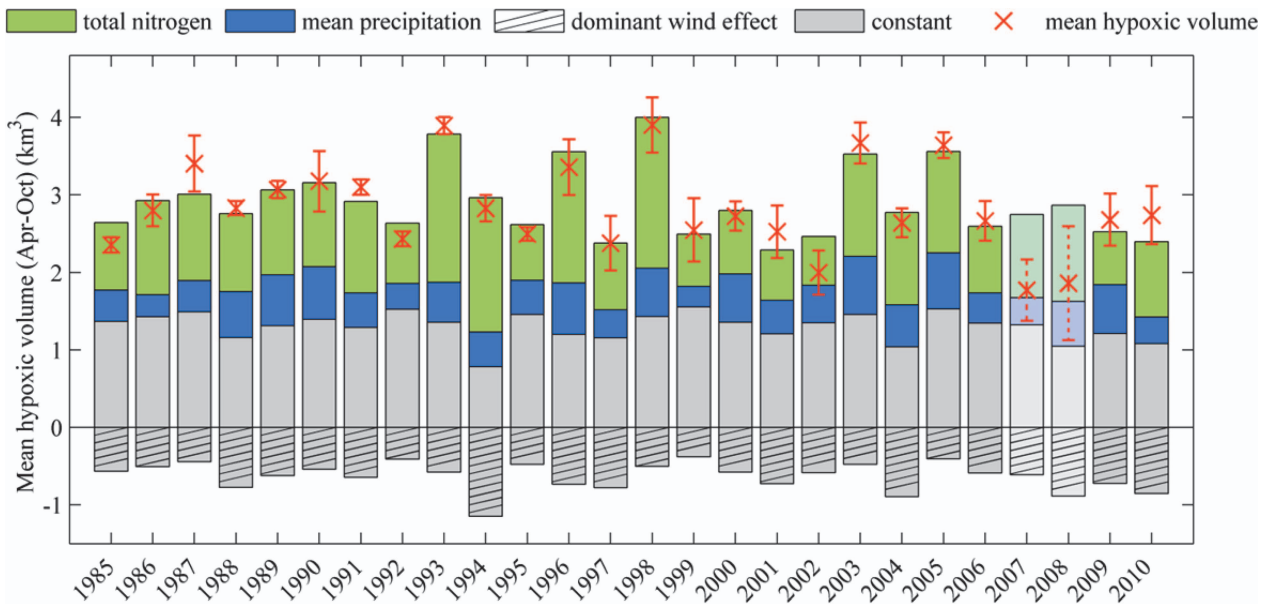


Fig. 6. Stacked bar plot of variables explaining the April to October mean hypoxic volume from 1985 to 2010 (**V** in Eq. 5). The dominant wind effect (dashed bar, **W** in Eq. 5) is calculated as the ratio of the duration of southwesterly and northeasterly winds. The gray bars represent a constant (1.94 km$^3$) that remains the same throughout the 26 yr period (Eq. 5). Note that these gray bars reach to different mean hypoxic volumes in different years simply because they start at different negative values representing the dominant wind effect for each year. Red crosses and lines represent conditional-realization-derived April to October mean hypoxic volume estimates and their uncertainties, respectively. These uncertainties are expressed as ± 2 standard deviations, where the standard deviation is estimated as $\sqrt{\sum_{1}^{14}\sigma_i^2/14^2}$, where 14 represents the total number of periods within a year, and $\sigma_i$ is the standard deviation of the uncertainty associated with each half-monthly hypoxic volume from April to October during a given year (Fig. 4). Note that data from 2007 and 2008 were not used in the analysis, and those years are therefore represented using lighter colors to distinguish them from the data used in the regression analysis.

The correlation coefficients between these explanatory variables are relatively low ($\rho_{\mathbf{N},\mathbf{P}} = 0.36$, $\rho_{\mathbf{N},\mathbf{W}} = 0.20$, $\rho_{\mathbf{P},\mathbf{W}} = 0.03$), and their contributions towards explaining the interannual variability in hypoxic volume can therefore be largely interpreted individually in Fig. 6. The contribution of nutrient loading to the total hypoxic volume varies by 1.31 km$^3$ between years, making it the strongest single factor in explaining interannual variability, whereas the contribution of the dominant wind effect varies by 0.77 km$^3$, and that of precipitation varies by 0.49 km$^3$ between years.

## Discussion

Overall, new estimates of early-April to late-October hypoxic volumes for 1985 to 2010 show that the time (i.e., date) when the hypoxic volume reaches its maximum has moved from late to early July, but that there is no trend in the seasonal-maximum hypoxic volume itself. Furthermore, no significant trend was found in the timing of onset of hypoxia, but the end of the hypoxic period has moved from October to September over the examined period. This analysis also shows that the mean hypoxic volume from April to October is better correlated with the examined factors than is the maximum or any half-monthly hypoxic volume, suggesting that the mean hypoxic volume is a better metric for elucidating mechanistic relationships between the interannual variability in hypoxia and its contributing factors. This is likely because the effects of the loads play out over the full season, and thus the mean hypoxic volume over the entire hypoxic season is a more direct indication of that cumulative load.

Whereas previous work had only considered TN loading from the Susquehanna (Hagy et al. 2004; Evans and Scavia 2011; Lee et al. 2013), or the Susquehanna and Potomac (Murphy et al. 2011), including the loading from the Rappahannock was found to be beneficial for explaining the interannual variability in mean hypoxic volume. This may be due to the fact that loading from the Potomac and Rappahannock Rivers constitutes a higher fraction of total TN loading in years with high loads. The positive correlation between TN loads and hypoxic volume is expected. On their own, the TN loads can explain the majority ($R^2 = 0.64$) of the variability in the mean hypoxic volume, and, using only the dominant wind effect and precipitation reduces the overall model's explanatory power dramatically ($R^2$ decreases from 0.85 to 0.36). This confirms that TN loading is the primary contributing factor to the interannual variability of hypoxia in Chesapeake Bay from 1985 to 2010.

Including precipitation as an explanatory factor in the model is unique to the work presented here. Precipitation

alone explains approximately one-third ($R^2 = 0.31$) of the variability in mean hypoxic volume, and including it in the model in addition to TN loading and the dominant wind effect yields a significant improvement ($p = 0.007$). Eliminating precipitation (i.e., using a model with only TN loading and the dominant wind effect) reduces the explanatory power ($R^2$ decreases from 0.85 to 0.78) of the model, although not as dramatically as for the other variables. The inferred positive drift coefficient between precipitation and hypoxic volume, even after accounting for the monitored nutrient load, suggests that precipitation acts as a surrogate for unmonitored nutrient loading into the bay, both through loads to tributaries downstream of the monitoring locations and through direct nonpoint source loading into the bay. This interpretation is plausible, because the monitoring stations on the three examined tributaries are located substantially upstream of the discharge into the bay (ranging from 15 km to $>$ 200 km). Furthermore, USGS (2010) showed that monitored nitrogen loading from tributaries represent only 53% of the total loads, whereas nonpoint sources downstream from the tributary monitoring stations were estimated to contribute 23%. The fact that the nonpoint contributions below the monitoring sites contribute about half of the monitored loads, and that the contribution of precipitation to the model proposed here (Fig. 6) is also approximately half that of monitored loads, further supports the idea that precipitation was BIC-selected because it represents this unmonitored component of the load.

The relationship between the dominant wind effect and hypoxic volume indicates that southwesterly winds are associated with reduced hypoxic volumes. The dominant wind effect explains only a small amount of the interannual variability in hypoxic volume on its own ($R^2 = 0.05$), but removing this factor from the overall model (i.e., using a model based only on TN loading and precipitation), decreases the explanatory power of the overall model more substantially (i.e., $R^2$ reduced from 0.85 to 0.73). This indicates that wind represents the residual variability after nutrient loading (as quantified by monitoring loading and precipitation) have been taken into account. The significance of wind is expected due to its effect on stratification. These results are also consistent with the findings by Valle-Levinson et al. (1998) and Guo and Valle-Levinson (2008), who demonstrated that the northeasterly winds can enhance stratification in the lower part of Chesapeake Bay, resulting in larger hypoxic volumes. In addition, Cho et al. (2012) confirmed that down-estuary (e.g., NE wind) local wind stress tends to enhance stratification under moderate wind speeds, and the up-estuary (e.g., SW wind) local wind stress tends to reduce stratification by reversing gravitational circulation. The current study and these earlier studies are inconsistent with the work by Scully (2010*a*), however, who found a negative correlation between July hypoxia and the southeasterly wind and a positive correlation between July hypoxia and westerly wind for a period from 1950 to 2007. It is important to note, however, that Scully (2010*a*) identified a shift in SE and westerly wind frequencies in the early 1980s, such that their results may be more representative of periods preceding those covered by the current work.

Together, the monitored nitrogen loading, precipitation, and wind direction explain 85% of the interannual variability in April–October mean hypoxic volume in Chesapeake Bay. While monitored nutrient loading explains the largest portion of the variability, nonpoint loading that is not proportional to monitored loading and wind direction together account for a substantial additional proportion of the variability. Their effect therefore needs to be explicitly considered in the development of management strategies for Chesapeake Bay.

As described previously, hypoxic volume estimates from 2007 and 2008 were not included in deriving the model for interannual variability because the sparse monitoring in those years focused on the lower portions of the bay. Results of the interannual analysis confirm that the mean hypoxic volumes were likely underestimated in those years, with the mean hypoxic volumes estimated from the DO data being significantly lower ($p \sim 0$ for 2007; $p = 0.003$ for 2008) than those predicted from the model in Eq. 7.

Beyond the three main explanatory variables, the model (Eq. 7) also includes a constant factor of 1.94 $km^3$, which accounts for approximately half of the overall hypoxic volume for most years (Fig. 6). For comparison, monitored TN loading contributes between 0.63 $km^3$ (2002) and 1.93 $km^3$ (1998). Taken at face value, the large constant factor runs counter to earlier analyses that suggested that nutrient load reductions of 28% (Hagy et al. 2004) or 35% (Cerco 1995; Scavia et al. 2006) would return the bay to 1950–1970 conditions when hypoxic volumes were 36–68% lower than they are today. However, it is also likely that the range of hypoxic volumes observed during the 1985–2010 period, and their relationships to nutrient loading during that time, may represent too small of a range of variability to assess the benefit that could be achieved by a substantial long-term reduction in nitrogen loading. Some recent work (Hagy et al. 2004) suggests an increasing sensitivity of hypoxia to nutrient loads in recent decades, which would represent a nonlinearity in the load–response curve. Such a nonlinearity, often resembling a saturating function, has also been used in other models of Chesapeake Bay (Scavia et al. 2006; Evans and Scavia 2011), the Gulf of Mexico (Scavia et al. 2003; R. E. Turner pers. comm.), and Lake Erie (D. Rucinski unpubl.). The results presented here suggest that the hypoxic volumes observed over the examined period can, in and of themselves, neither confirm nor deny the validity of such an assumption, because the loads and hypoxic volumes at which such nonlinearities would have a substantial effect are outside of the range observed during this period.

However, our results do indicate that management strategies that aim to achieve a given reduction in hypoxic conditions must account for the fact that predictions of benefits carry with them a substantial amount of uncertainty. For example, our results show that years with substantially lower nitrogen loading did not have substantially lower hypoxic volumes (Fig. 6). For example, the years 1999 to 2002 had monitored nitrogen loadings that were 37% below the 26 yr average, and the unmonitored loads as represented through the precipitation term were

28% lower, but the hypoxic volumes during those years were only 16% below average. This may be due to extenuating factors not considered here, but it nevertheless suggests that reductions in monitored nutrient loadings in the 28–35% range may not be sufficient to achieve a corresponding reduction in hypoxic conditions as had been suggested, at least in the short term. This uncertainty is further compounded by the observed effect of nonpoint loading and wind conditions.

*Acknowledgments*

We thank Rebecca Murphy for providing the two-dimensional ordinary kriging estimates of dissolved oxygen concentrations, the Chesapeake Bay Program for access to the monitoring data, and Mary Anne Evans and Caroline Wicks for helping us obtain the Chesapeake Bay Interpolator estimates. We thank Chao Li, Eva Sinha, Dan Obenour, and Jeff Ho for feedback on this work, Xuemei Qiu for help with illustrations, and two anonymous reviewers for their constructive input. This work is supported by National Science Foundation under grant 0644648. Additional support for D.S. was provided by the Graham Sustainability Institute at the University of Michigan.

## *References*

Anderson, D. R., K. P. Burnham, and G. C. White. 1998. Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. J. Appl. Stat. **25:** 263–282, doi:10.1080/02664769823250

Bahner, L. 2006. User guide for the Chesapeake Bay and tidal tributary interpolator [Internet]. Annapolis (MD): NOAA Chesapeake Bay Office [accessed 01 October 2011]. Available from http://archive.chesapeakebay.net/cims/interpolator.pdf

Breitburg, D. L. 1990. Near-shore hypoxia in the Chesapeake Bay—patterns and relationships among physical factors. Estuar. Coast. Shelf Sci. **30:** 593–609, doi:10.1016/0272-7714(90)90095-9

Burns, N. M., D. C. Rockwell, P. E. Bertram, D. M. Dolan, and J. J. H. Ciborowski. 2005. Trends in temperature, secchi depth, and dissolved oxygen depletion rates in the central basin of Lake Erie, 1983–2002. J. Great Lakes Res. **31:** 35–49, doi:10.1016/S0380-1330(05)70303-8

Carrick, H. J., J. B. Moon, and B. F. Gaylord. 2005. Phytoplankton dynamics and hypoxia in Lake Erie: A hypothesis concerning benthic-pelagic coupling in the central basin. J. Great Lakes Res. **31:** 111–124, doi:10.1016/S0380-1330(05)70308-7

CENR 2003. An assessment of coastal hypoxia and eutrophication in U.S. waters [Internet]. Washington, D.C.: National Science and Technology Council Committee on Environment and Natural Resources [accessed 20 July 2011]. Available from http://oceanservice.noaa.gov/outreach/pdfs/coastalhypoxia.pdf

Cerco, C. F., and T. Cole. 1993. 3-dimensional eutrophication model of Chesapeake Bay. J. Environ. Eng. ASCE **119:** 1006–1025, doi:10.1061/(ASCE)0733-9372(1993)119:6(1006)

———. 1995. Response of Chesapeake Bay to nutrient load reductions. J. Environ. Eng. **121:** 549–557, doi:10.1061/(ASCE)0733-9372(1995)121:8(549)

Chen, C.-C., G.-C. Gong, and F.-K. Shiah. 2007. Hypoxia in the East China Sea: One of the largest coastal low-oxygen areas in the world. Mar. Environ. Res. **64:** 399–408, doi:10.1016/j.marenvres.2007.01.007

Chesapeake Bay Program. 2011. Chesapeake Bay Program water quality database (1984–present) [Internet]. City (ST): Chesapeake Bay Program. Accessed 00 Month 2011. Available from http://www.chesapeakebay.net/data_waterquality.aspx

Cho, K.-H., H. V. Wang, J. Shen, A. Valle-Levinson, and Y.-C. Teng. 2012. A modeling study on the response of Chesapeake Bay to hurricane events of Floyd and Isabel. Ocean Modell. **49–50:** 22–46, doi:10.1016/j.ocemod.2012.02.005

Corbeil, R. R., and S. R. Searle. 1976. Restricted maximum likelihood estimation of variance components in the mixed model. Technometrics **18:** 31–38, doi:10.2307/1267913

Cronin, T. M., and C. D. Vann. 2003. The sedimentary record of climatic and anthropogenic influence on the Patuxent Estuary and Chesapeake Bay ecosystems. Estuaries **26:** 196–209, doi:10.1007/BF02695962

Daskalov, G. M. 2003. Long-term changes in fish abundance and environmental indices in the Black Sea. Mar. Ecol. Prog. Ser. **255:** 259–270, doi:10.3354/meps255259

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

Evans, M. A., and D. Scavia. 2011. Forecasting hypoxia in the Chesapeake Bay and Gulf of Mexico: Model accuracy, precision, and sensitivity to ecosystem change. Environ. Res. Lett. **6:** 015001, doi:10.1088/1748-9326/6/1/015001

Faraway, J. 2005. Linear models with R. CRC Press.

Feng, Y., S. F. Dimarco, and G. A. Jackson. 2012. Relative role of wind forcing and riverine nutrient input on the extent of hypoxia in the northern Gulf of Mexico. Geophys. Res. Lett. **39:** L09601, doi:10.1029/2012gl051192

Fisher, T. R., and others. 1999. Spatial and temporal variation of resource limitation in Chesapeake Bay. Mar. Biol. **133:** 763–778, doi:10.1007/s002270050518

Flemer, D. A., and others. 1983. Chesapeake Bay: A profile of environmental change appendices, p. B1–B59. *In* E. G. Macalaster, D. A. Barker, and M. E. Kasper [eds.]. U.S. Environmental Protection Agency Chesapeake Bay Program Office. Annapolis, MD.

Guo, X., and A. Valle-Levinson. 2008. Wind effects on the lateral structure of density-driven circulation in Chesapeake Bay. Cont. Shelf Res. **28:** 2450–2471, doi:10.1016/j.csr.2008.06.008

Gutjahr, A., B. Bullard, S. Hatch, and L. Hughson. 1994. Joint conditional simulations and the spectral approach for flow modeling. Stoch. Hydrol. Hydraul. **8:** 79–108, doi:10.1007/BF01581391

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

Hunter, P. D., A. N. Tyler, N. J. Willby, and D. J. Gilvear. 2008. The spatial dynamics of vertical migration by *Microcystis aeruginosa* in a eutrophic shallow lake: A case study using high spatial resolution time-series airborne remote sensing. Limnol. Oceanogr. **53:** 2391–2406, doi:10.4319/lo.2008.53.6.2391

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

Kitanidis, P. K. 1995. Quasi-linear geostatistical theory for inversing. Water Resour. Res. **31:** 2411–2419, doi:10.1029/95WR01945

———, and R. W. Lane. 1985. Maximum-likelihood parameter-estimation of hydrologic spatial processes by the Gauss-Newton method. J. Hydrol. **79:** 53–71, doi:10.1016/0022-1694(85)90181-7

Lee, Y. J., W. R. Boynton, M. Li, and Y. Li. 2013. Role of late winter–spring wind influencing summer hypoxia in Chesapeake Bay. Estuaries Coasts **36:** 683–696, doi:10.1007/s12237-013-9592-5

Lewis, B. L., and others. 2007. Short-term and interannual variability of redox-sensitive chemical parameters in hypoxic/anoxic bottom waters of the Chesapeake Bay. Mar. Chem. **105:** 296–308, doi:10.1016/j.marchem.2007.03.001

Li, D. J., J. Zhang, D. J. Huang, Y. Wu, and J. Liang. 2002. Oxygen depletion off the Changjiang (Yangtze River) Estuary. Sci. China, Ser. D: Earth Sci. **45:** 1137–1146, doi:10.1360/02yd9110

Liu, Y., and D. Scavia. 2010. Analysis of the Chesapeake Bay hypoxia regime shift: Insights from two simple mechanistic models. Estuaries Coasts **33:** 629–639, doi:10.1007/s12237-009-9251-z

Malone, T. C., W. M. Kemp, H. W. Ducklow, W. R. Boynton, J. H. Tuttle, and R. B. Jonas. 1986. Lateral variation in the production and fate of phytoplankton in a partially stratified estuary. Mar. Ecol. Prog. Ser. **32:** 149–160, doi:10.3354/meps032149

Mueller, K. L., V. Yadav, P. S. Curtis, C. Vogel, and A. M. Michalak. 2010. Attributing the variability of eddy-covariance $CO_2$ flux measurements across temporal scales using geostatistical regression for a mixed northern hardwood forest. Global Biogeochem. Cycles **24:** GB3023, doi:10.1029/2009GB003642

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

National Research Council. 2000. Clean coastal waters. National Academy Press.

NCDC. 2012. National Climatic Data Center [Internet]. Asheville (NC): Publisher [accessed 01 August 2012]. Available from http://www.ncdc.noaa.gov/cdo-web/

Newcombe, C. L., and W. A. Horne. 1938. Oxygen-poor waters of the Chesapeake Bay. Science **88:** 80–81, doi:10.1126/science.88.2273.80

NWS. 2012. National Weather Service [Internet]. Silver Spring (MD): Publisher [accessed 01 August 2012]. Available from http://www.weather.gov/

Obenour, D. R., D. Scavia, N. N. Rabalais, R. E. Turner, and A. M. Michalak. 2013. Retrospective analysis of midsummer hypoxic area and volume in the northern Gulf of Mexico, 1985–2011. Environ. Sci. Technol. **47:** 9808–9815, doi:10.1021/es400983g

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, doi:10.1126/science.223.4631.22

Parker, C. A., and J. E. Oreilly. 1991. Oxygen depletion in Long-Island Sound—a historical perspective. Estuaries **14:** 248–264, doi:10.2307/1351660

Rabalais, N. N., R. E. Turner, and W. J. Wiseman, Jr. 2001. Hypoxia in the Gulf of Mexico. J. Environ. Qual. **30:** 320–329, doi:10.2134/jeq2001.302320x

Sandberg, E. 1994. Does short-term oxygen depletion affect predator-prey relationships in zoobenthos—experiments with the isopod *Saduria entomon*. Mar. Ecol. Prog. Ser. **103:** 73–80, doi:10.3354/meps103073

Sanford, L. P., K. G. Sellner, and D. L. Breitburg. 1990. Covariability of dissolved-oxygen with physical processes in the summertime Chesapeake Bay. J. Mar. Res. **48:** 567–590, doi:10.1357/002224090784984713

Scavia, D., E. L. A. Kelly, and J. D. Hagy. 2006. A simple model for forecasting the effects of nitrogen loads on Chesapeake Bay hypoxia. Estuaries Coasts **29:** 674–684.

———, N. N. Rabalais, R. E. Turner, D. Justic, and W. J. Wiseman. 2003. Predicting the response of Gulf of Mexico hypoxia to variations in Mississippi River nitrogen load. Limnol. Oceanogr. **48:** 951–956, doi:10.4319/lo.2003.48.3.0951

Schwarz, G. 1978. Estimating dimension of a model. Ann. Stat. **6:** 461–464, doi:10.1214/aos/1176344136

Scully, M. E. 2010*a*. The importance of climate variability to wind-driven modulation of hypoxia in Chesapeake Bay. J. Phys. Oceanogr. **40:** 1435–1440, doi:10.1175/2010JPO4321.1

———. 2010*b*. Wind modulation of dissolved oxygen in Chesapeake Bay. Estuaries Coasts **33:** 1164–1175, doi:10.1007/s12237-010-9319-9

———, W. R. Geyer, and J. H. Trowbridge. 2011. The influence of stratification and nonlocal turbulent production on estuarine turbulence: An assessment of turbulence closure with field observations. J. Phys. Oceanogr. **41:** 166–185, doi:10.1175/2010JPO4470.1

U. S. EPA. 2002. The state of the Chesapeake Bay. A report to the citizens of the Bay region. U.S. Environmental Protection Agency for the Chesapeake Bay Program.

USGS. 2010. Measuring nutrient and sediment loads to Chesapeake Bay [Internet]. City (ST): United States Geological Survey [accessed 00 December 2010]. Available from http://va.water.usgs.gov/chesbay/RIMP/

Valle-Levinson, A., J. L. Miller, and G. H. Wheless. 1998. Enhanced stratification in the lower Chesapeake Bay following northeasterly winds. Cont. Shelf Res. **18:** 1631–1647, doi:10.1016/S0278-4343(98)00067-3

Ward, E. J. 2008. A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. Ecol. Modell. **211:** 1–10, doi:10.1016/j.ecolmodel.2007.10.030

Welsh, B. L., and F. C. Eller. 1991. Mechanisms controlling summertime oxygen depletion in western Long-Island Sound. Estuaries **14:** 265–278, doi:10.2307/1351661

Zhou, Y., and A. M. Michalak. 2009. Characterizing attribute distributions in water sediments by geostatistical downscaling. Environ. Sci. Technol. **43:** 9267–9273, doi:10.1021/es901431y

———, D. R. Obenour, D. Scavia, T. H. Johengen, and A. M. Michalak. 2013. Spatial and temporal trends in Lake Erie hypoxia, 1987–2007. Environ. Sci. Technol. **47:** 899–905, doi:10.1021/es303401b

*Associate editor: David A. Caron*

*Received: 27 June 2013*
*Accepted: 06 November 2013*
*Amended: 12 November 2013*