Progress Update for the Synthesis Project: *Quantifying the impacts of past and future climate and eutrophication on the dynamics of dissolved oxygen in the shallow waters of Chesapeake Bay*

Jeremy Testa, Damian C. Brady, Wei Liu
University of Maryland Center for Environmental Science
Chesapeake Biological Laboratory
April 6, 2021

Research Activity and Progress: Since our last update in September 2020, we have continued our statistical analysis (i.e., harmonic analysis, Classification and Regression Trees, and Generalized Additive Models) to determine climate and biogeochemical controls on daily shallow water hypoxia metrics. In this summary, we will characterize our previously reported work and outline new work related to future warming and nutrient loading effects on metabolism and oxygen variability. We are currently describing this analysis in a draft peer-reviewed manuscript that we expect to submit in June 2021. The results and conclusions of this progress update and therefore tentative.

(1) **Analysis of diel oxygen data:** We have continued to perform a wide variety of analyses on shallow water dissolved oxygen data to quantify hypoxia duration, tidal impacts on oxygen variability, and climatic and biological controls on oxygen variability. Below, we highlight three of these ongoing efforts.

(a) *Spatial analysis of controlling variables for oxygen dynamics*:

Input Data: We have developed CART models that statistically identify sets of oxygen conditions associated with different combinations of controlling variables, including physical variables (temperature, turbidity, precipitation, photosynthetically active radiation (PAR), and wind speed) and biological variables (e.g., chlorophyll-a; Fig. 1). We utilized a diverse array of available datasets to generate time-series of these input variables, including the North American Land Data Assimilation System (NLDAS), the European Centre for Medium-Range Weather Forecasts (ERA5), and the National Oceanic and Atmospheric Administration (NOAA) tide stations. We also used an assortment of regional station-data (Jug Bay NERR, UMCES for PAR data, Thomas Point Lighthouse weather data) to validate regional-scale estimates to support station-specific analysis. We also obtained all available high-frequency (15 min) water quality data from 181 stations across Chesapeake Bay (129 stations monitored by the Maryland Department of Natural Resources and 52 stations monitored by the Virginia Department of Environmental Quality and the Chesapeake Bay National Estuarine Research Reserve). These data include DO (mg O_2 L^{-1} and % saturation), station depth (m), water temperature (°C), salinity, chlorophyll-a (mg Chl m^{-3}), and turbidity (NTU). For the high-frequency data, invalid data points were removed based on Continuous Monitoring and Water Quality Mapping Error Codes provided by MD DNR.

Power Spectral Density Analysis: We used power spectral density (PSD) analysis to describe the power distribution (i.e., amplitude of DO fluctuations) as a function of frequency (e.g., semi-diurnal and diurnal signals). These analyses determine if DO variability at the diurnal (primarily

solar and biological) or semidiurnal (tidal) is higher for a given station. The PSD was calculated for the 15-minute DO time series at all ConMon stations. Unsurprisingly, we found that the diurnal variability was higher than semi-diurnal variability at the majority of the stations we analyzed, with a mean diurnal:semidiurnal PSD ratio of 2 across stations (Fig.1). Diurnal variability was weaker at locations with high turbidity, which were primarily located in upstream, relatively low salinity reaches (Fig. 1). We performed a CART analysis of the diurnal:semidiurnal PSD ratio with respect to station-specific physical and environmental conditions (Fig. 2) and found that the ratio was higher (more diurnal variability) under conditions of low turbidity and low wind speed. Thus, reductions in sediment loads to Bay tributaries, and presumed reductions in turbidity, might lead to a stronger diurnal variability as more light availability allows for more photosynthetic oxygen production.

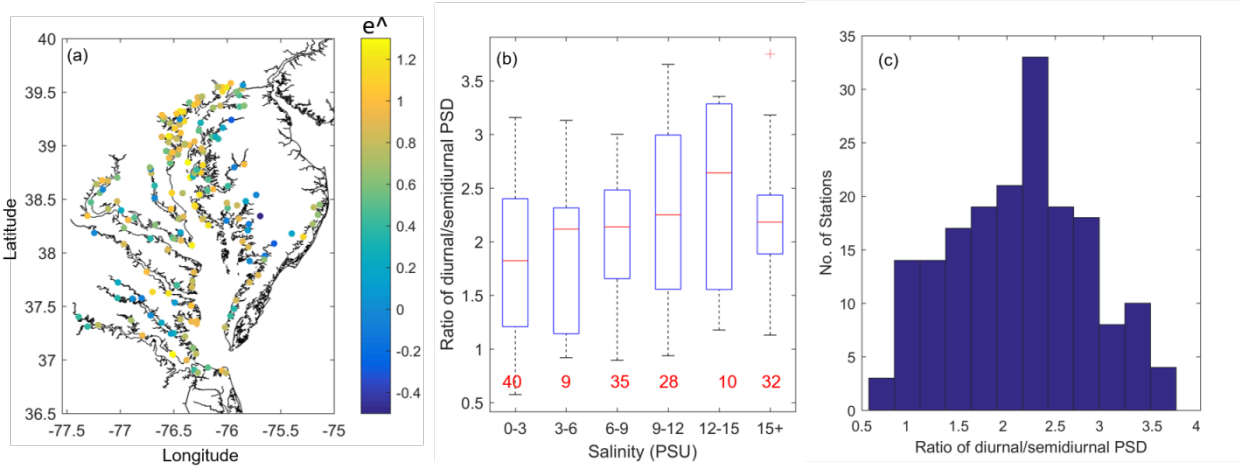

Figure 1: (a) Spatial distribution of the ratio between diurnal and semidiurnal PSDs for DO concentration on a natural log scale; (b) its correlation with mean salinity at each station, red numbers are the number of stations in each bar; and (c) the histogram of the ratio.

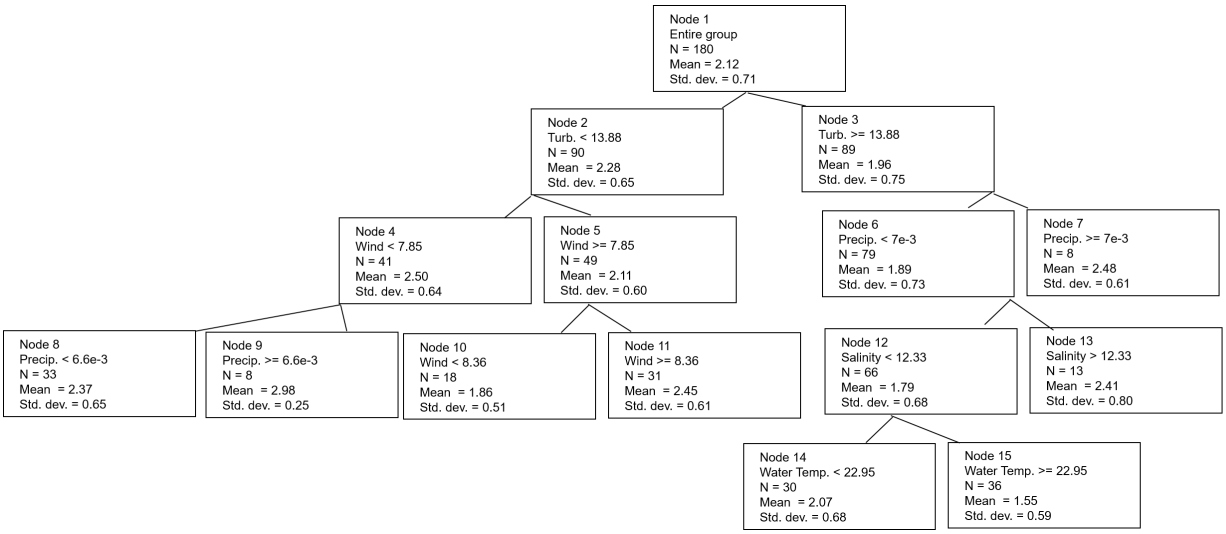

Figure 2: CART tree describing station-specific analysis of the diurnal:semidiurnal PSD, where each box includes the splitting variable, the threshold value used to group stations within each box, and the mean and standard deviation of the diurnal:semidiurnal PSD within each box.

Tidal Phase Analysis: To quantitatively identify and isolate tidal effects on DO variation, we separated tidal components from non-tidal components in DO time series using the following equation:

$$C_{DO}(t) = \overline{C_{DO}} + \sum_{n=1}^{N} A_n \cos(\omega_n t - \theta_n) + R(t)$$

where $C_{DO}(t)$ is the concentration of DO at time t; $\overline{C_{DO}}$ is the mean DO concentration at each station; $\sum_{n=1}^{N} A_n \cos(\omega_n t - \theta_n)$ is sinusoidal component of the DO time series, including tidal-induced DO variation for a total of N tidal constituents; A_n is the amplitude of DO variation due to n^{th} tidal constituent with a frequency of n and a phase of n; and $R(t)$ is the non-sinusoidal residual component at time t, representing the non-tidal contribution to DO variability. We applied a least squares method to solve this equation, where we included 35 tidal constituents and their frequencies. We used a 30 day window to solve this equation for the entire time series of DO at each station. We compared the results from this harmonic analysis with other approaches to identify harmonic behavior, including nonlinear curve-fitting, and found consistency in results. We have tested the harmonic analysis with different sizes of windows including 1, 7, 15, 30 and 90 days. All the results have steadily separated the residual from the tide-dominated variability and a 30-day window has been chosen because it balanced the capture of longer-term variability while being short relative to overall length of the record for each station (~7 months over three years).

This method has been applied to times series of both DO and % DO at each station to obtain a residual time-series ($R(t)$) at each station. CART analysis used the $R(t)$ time series to quantify key external and internal variable effects on DO. We show two examples of the harmonic analysis on time series of DO with a 30-day window in Fig. 3. One example is at Little Monie Creek, a station with relatively low DO and where the semidiurnal variation of DO is significant. In contrast, the same approach applied at the Susquehanna Flats station shows extremely high DO where the diurnal DO signal is dominant. A decomposition of the derived components of the DO time-series reveals that $R(t)$ is much more variable at Susquehanna Flats (Fig. 3).

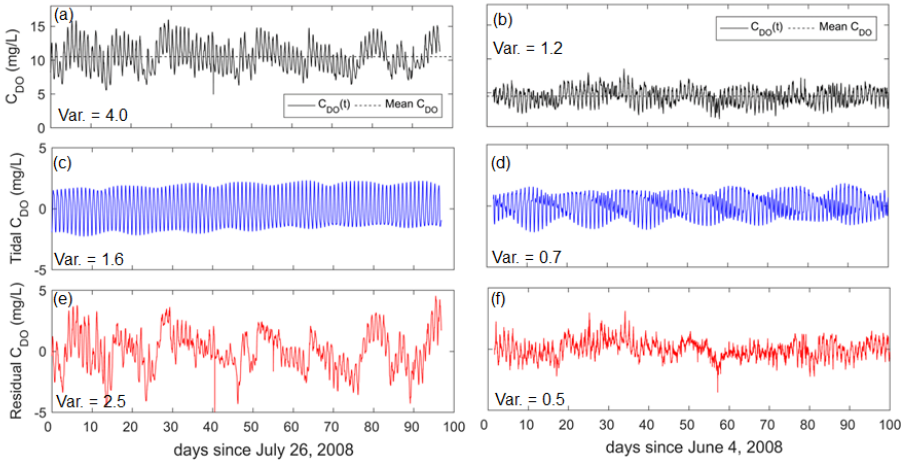

Figure 3: Examples of harmonic analysis on time series of DO with a 30-day window at Susquehanna Flats (left panels) and Little Monie Creek (right panels). Black solid lines are the original DO time series while dashed lines are the mean DO concentration (a&b). Blue lines are the sinusoidal conmponents (c&d) while red lines are the residual non-sinusoidal residuals (e&f).

CART Analysis of Controls on DO: We applied Classification and Regression Trees (CART) to the residual DO time series using the suite of potential internal and external biological and physical forcing variables. CART is a machine-learning approach that identifies threshold values of controlling variables that split the DO data into distinct homogeneous groups with high or low meanR(t). We applied CART to 181 stations, using both tide-corrected (residual DO time series) and measured (raw) DO and % DO, as well as the computed hours of hypoxia (DO <4.8 mg/L). We aggregated all DO metrics and forcing data to daily means and we analyzed DO records over the April-October period for years where we had DO data, in order to ensure seasonal consistency in our analysis. These CART models provide a single framework to determine the most significant controlling factors (e.g., PAR, temperature, precipitation) at a given station. The analysis generates two threshold values for each station, including the variable that first splits the DO data (FSV) and the *most important predictor* (MIP), which has the largest effect of splitting DO data into groups/nodes with significant differences.

In applying this approach across all stations in Maryland and Virginia, it is clear that certain variables are important controls on DO in spatially coherent regions of the Bay (Fig. 4). A spatial view of these CART-generated MIPs and first split variables indicates that there is strong spatial variation in what variables are most closely associated with changes in shallow water oxygen. In some cases, there are clear spatial clusters in the key controlling variables. For example, wind is an important predictor of oxygen on the western shore of the mid Bay (green circles; Fig. 4), and PAR is relevant in stations that flank the mainstem Bay (yellow squares). Finally, chlorophyll-a is generally an important discriminator in upstream regions (black diamonds, Fig. 4).

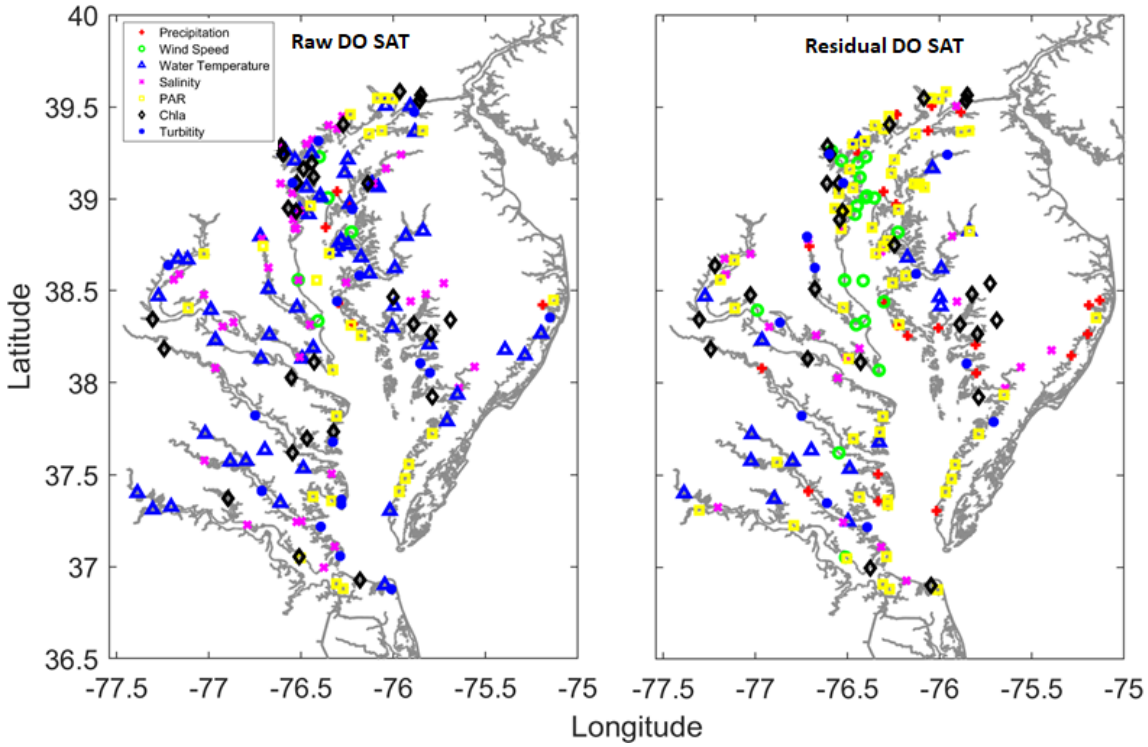

Figure 4: Spatial distribution of MIP for raw and residual DO Saturation. For residual DO saturation, there are a wind-as-MIP cluster in upper bay area and PAR-as-MIP sites all around the bay. By comparing two maps, water temperature-as-MIP sites become PAR sites after filtering tides.

We computed histograms of the by-station controlling variables (MIP) for both the raw and residual % DO to identify which variables were most frequently key controls and to identify if the variable has a positive or negative effect on DO (Fig. 5). For the purposes of this analysis, we targeted % DO saturation since it inherently accounts for the influence of temperature and salinity on DO. Water Temperature was the most important discriminator between high and low raw % DO (and had a negative effect), while PAR, precipitation, and chlorophyll-a were the two most important discriminators between high and low residual % DO saturation (Fig. 5). For residual % DO saturation (right panel in Fig. 5), significant negative effects were found between DO and three predictors: precipitation, wind speed and turbidity (Fig. 5d), while primarily positive effects were found for chlorophyll-a and PAR (Fig. 5d). In this context, we define negative effects as those where high conditions of controlling variables lead to low residual % DO, and vice versa.

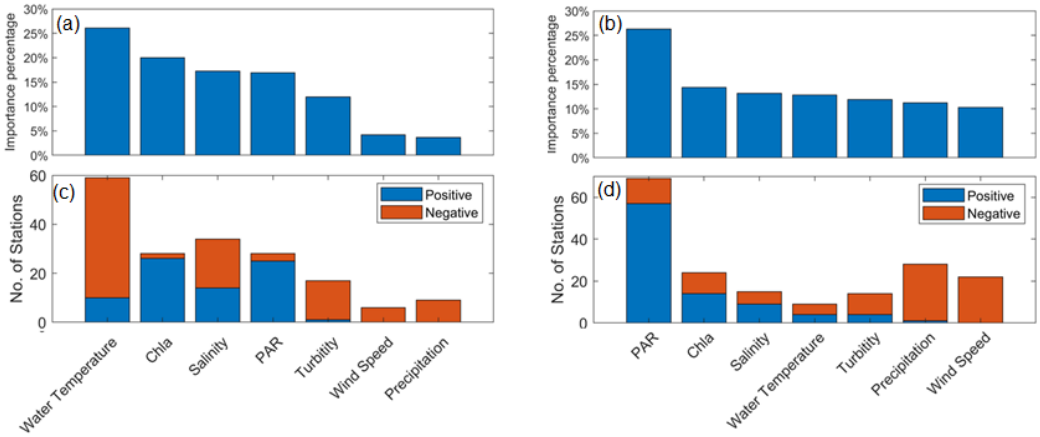

Figure 5: CART analysis on raw (left panels) and residual (right panels) DOSAT. a&b are the importance percentage for each predictor; c&d show the number of stations that have each predictor as the FSV and its positive (blue) or negative (orange) correlation with DOSAT. Chlorophyll-a and turbidity are in a log scale.

We also applied a similar CART approach for hours of hypoxia (< 4.8 mg O_2 L^{-1}) at each station. We found that a large fraction of the 181 stations had some degree of moderate hypoxic conditions in the record. The CART approach indicated that water temperature was the dominant MIP, while salinity, chlorophyll-a, and turbidity were secondary controls on hours of hypoxia.

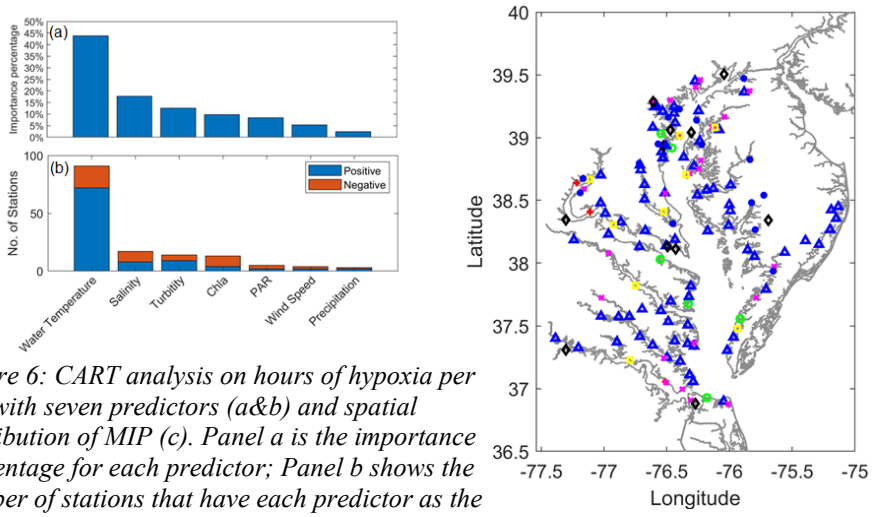

Figure 6: CART analysis on hours of hypoxia per day with seven predictors (a&b) and spatial distribution of MIP (c). Panel a is the importance percentage for each predictor; Panel b shows the number of stations that have each predictor as the First Split Variable (FSV) and its positive (blue) or negative (orange) correlation with DO saturation.

As expected, higher water temperature leads to longer lasting hypoxia in the majority of stations where temperature was MIP, while the direction of effect on DO was split between positive and negative for salinity, chlorophyll-a, and turbidity. We later revisit how extremely high chlorophyll-a leads to hypoxia at some locations, but can lead to elevated oxygen at others.

To investigate the mechanisms by which these spatial clusters operate, we examined the details of DO time-series in the context of temporal variations in the controlling variables across individual sites (Fig. 7). For example, in examining a time series at sites where wind speed is the MIP (Fig.4, right panel), we observed that increases in residual % DO saturation occur under strong southerly winds above a specific threshold (blue bars in top panel of Fig.7). These southerlies are also associated with increases in chlorophyll-a (red line, bottom panel of Fig. 8, right). It appears that high wind imports phytoplankton or encourages its growth in this region, leading to high oxygen concentrations. In contrast, under strong northerly winds, residual % DO saturation declines in step with reduced chlorophyll-a, indicating that northerly winds may be pushing turbid, low-chlorophyll waters south (Fig. 8). These linked-processes that associate oxygen with physical forcing are important for understanding oxygen concentrations in both deep and shallow waters. In-depth process based investigations based on filtering by machine learning techniques are helping to disentangle the co-variation of physical and biological controls on high frequency oxygen variations and to understand more site-specific realizations of coupled biological-physical forcing.

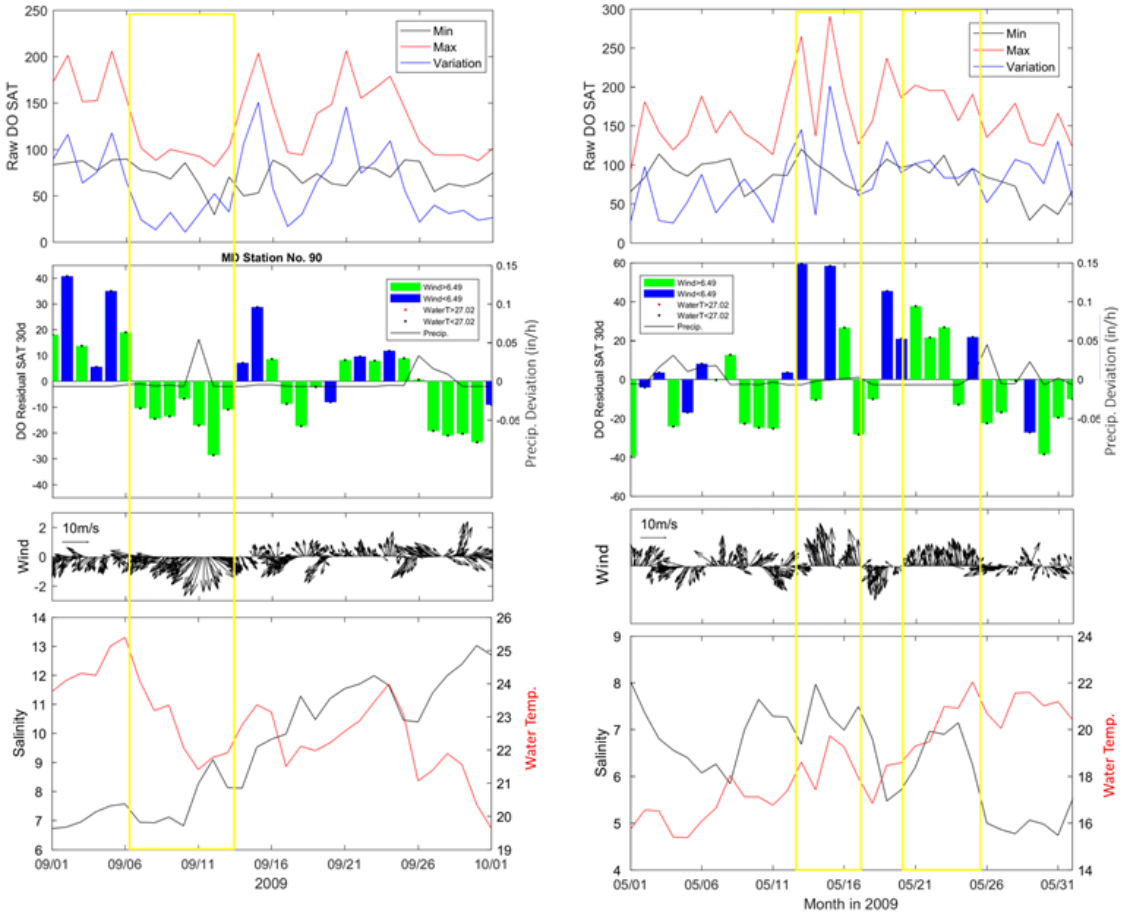

Figure 7: Time series of raw DO Saturation, DO residual, wind speed, salinity and water temperature at one wind-dominant MD station No. 90. Left panels show a significant north wind event while right significant south wind events.

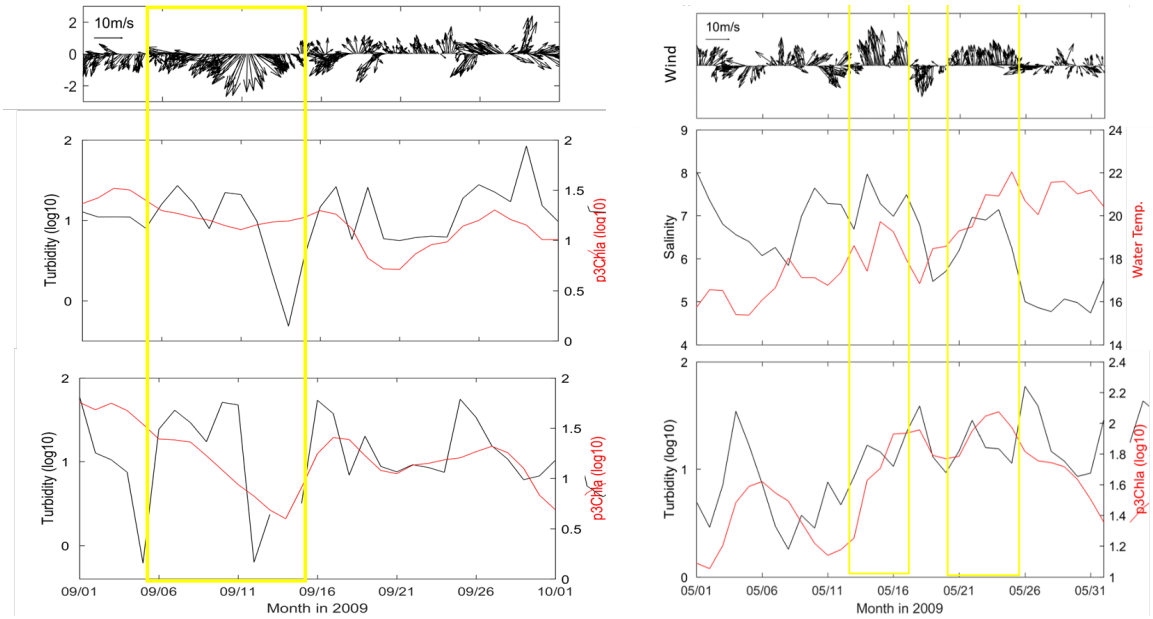

Figure 8: Time series of wind speed and direction (top), salinity and temperature (middle), and chlorophyll-a, and turbidity (bottom) at a wind speed MIP station as in Figure 7. Left panels show a significant north wind event while right significant south wind events.

Indices of Ecosystem Metabolism: We have also derived estimates of ecosystem gross primary production (P_g), respiration (R_t), and net ecosystem metabolism (NEM) from a subset of the stations in the Chesapeake Bay ConMon dataset. These rates of ecosystem metabolism represent changes in the eutrophication process in shallow waters with changes in nutrient loading and biological communities. For example, metabolic rates derived from a single location in the Back River declined significantly during the last 3 decades consistent with nutrient and chlorophyll-*a* declines related to major WWTP upgrades (Fig. 9; Testa et al. in prep). P_g and R_t approached 1000 mmol O_2 m^{-2} d^{-1} (32 g O_2 m^{-2} d^{-1}) in June and July of the mid-1990s prior to major nitrogen load reductions from WWTPs, approached 700 mmol O_2 m^{-2} d^{-1} (22 g O_2 m^{-2} d^{-1}) in 2014-2017 following BNR, and were typically below 500 mmol O_2 m^{-2} d^{-1} (15 g O_2 m^{-2} d^{-1}) in 2018 following ENR. Declines in P_g and R_t were apparent in all summer months, and histograms of all measured rates indicated a clear shift in the magnitude of the peak rates associated with nutrient load reductions (Fig. 9). NEM also declined steadily during the record, where NEM was positive (i.e. net autotrophy) in the mid-1990s and was reduced to near metabolic balance ($P_g = R_t$) during 2018 (Fig. 9). These data suggest that nutrient load reductions will reduce the diurnal variability in oxygen concentrations associated with ecosystem metabolism

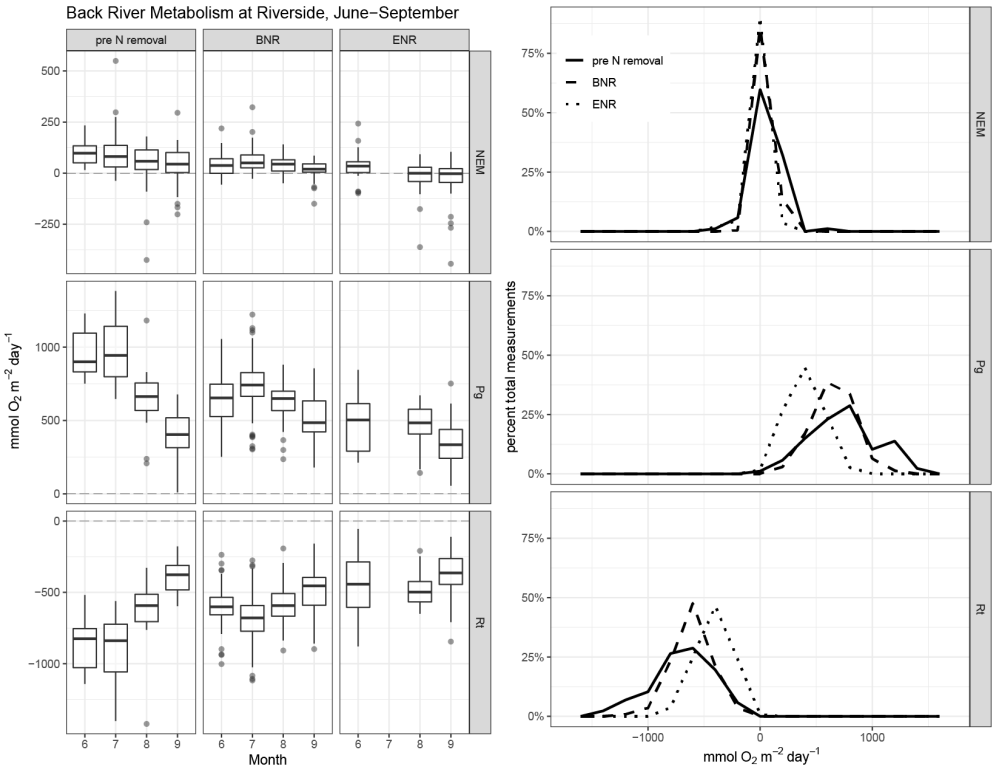

Figure 9: (left panel) Box plots of estimates of net ecosystem metabolism (NEM; top), ecosystem gross primary production (Pg; middle), and ecosystem respiration (Rt, bottom) from June to September in periods before BNR, after BNR, and after ENR. (right panel) Histograms of NEM (top), Pg (middle), and Rt (bottom) for each period of nitrogen removal treatment at the Back River WWTP. Metabolic rates derived from continuous dissolved oxygen measurements at one station in the Back River.

In a separate analysis in the Chester River estuary, we found that estimates of P_g, R_t, and NEM from observed O_2 time series were highly correlated with temperature on a seasonal basis, but between stations, the magnitude of rates varied with chlorophyll-a (Fig. 10). May-October mean values of respiration, for example, ranged from 50-250 mmol O_2 m^{-2} d^{-1} (~1.6-8 g O_2 m^{-2} d^{-1}) over 3-8 years, and were elevated within the sub tributary stations of the Corsica River (The Sill, Sycamore Point, Possum Point, Emory Creek) relative to the two stations in the main body of the Chester (Rolph's Wharf, Deep landing; Fig. 10). Median measured chlorophyll-a concentrations at the Corsica River stations were 13.8-36.3 µg/L, which were 50% higher than chlorophyll-a concentrations measured over a similar period in the adjacent Chester River (6.3-12.5 µg/L). Respiration rates at a given temperature were higher under conditions of elevated chlorophyll within stations (Fig. 10). For the two stations located within a small inlet in the lower Chester River estuary (Kent Narrows 'Inside' and 'Outside'), computed respiration rates were a factor of two larger at the inner, more protected station than the outer station (Fig. 10). As a consequence, the temperature-respiration slope was 2.5 times higher in the inner station than the outer station (Fig. 10). Despite these spatial differences, rates of Rt were significantly correlated with temperature at all sites, but with lower slopes in the main Chester channel region (Fig. 10). Regressions of Rt versus temperature reveal that mean slopes of -8 mmol O_2 m^{-2} d^{-1} $°C^{-1}$ and -

14.1 mmol O_2 m^{-2} d^{-1} °C^{-1} for the Chester River estuary and Corsica River estuary, respectively. These analyses reveal an interplay between temperature and chlorophyll-a in determining respiration rates, which serve as the primary sink of oxygen depletion in Chesapeake Bay. Future warming will thus increase respiration rates, but especially so when phytoplankton biomass is high, increasing the vulnerability to diel cycling hypoxia.

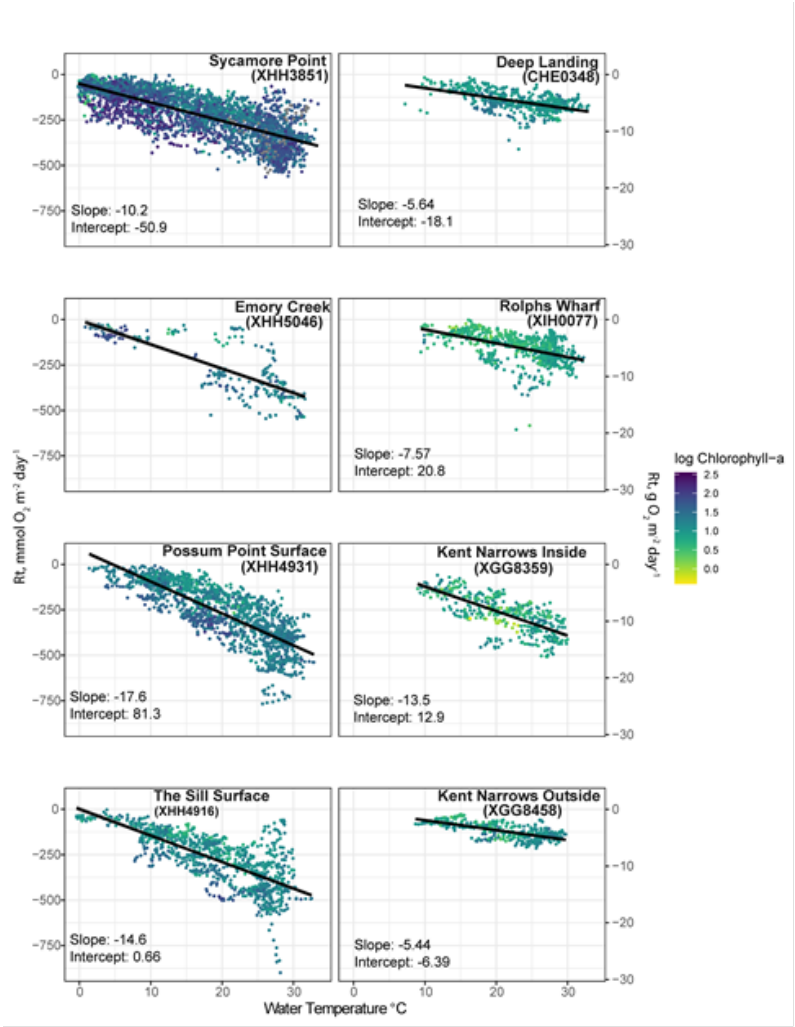

Figure 10: Relationships between daily water temperature and daily rates of ecosystem respiration derived from dissolved oxygen time series. Color of circles represents mean daily log chlorophyll-a. Rt is a rate of oxygen uptake, where increasingly negative values indicate higher respiration. For each regression, the slope and intercept (in units of mmol O_2 m^{-2} d^{-1}) are included.

Effects of Warming on Future Oxygen: In an effort to use CART results describing relationships between temperature and hypoxia, we extracted the temperature threshold that emerged from CART analyses of hypoxic duration as the First Split Variable (FSV) at a subset of ConMon Stations (Fig. 11). We then computed the percentage of time in each April-October period where temperatures exceeded this threshold value under observed past conditions (Fig. 11, left panel). Then, assuming that shallow-water temperatures will rapidly equilibrate with air temperatures, we simulated a 0.5, 1, and 1.5 °C warming across all days in the time series, and computed the increase in time spent above the FSV threshold. We found that a 1.5 °C warming would increase the time spent above the threshold for hypoxia by ~10%, suggesting that warming will have a non-trivial, direct effect on hypoxia vulnerability. Much of this temperature effect emerges because oxygen solubility at current summer temperatures in Chesapeake Bay is not far above the 30-day oxygen criterion of 5 mg L^{-1}, so even modest warming will increase vulnerability to this degree of low oxygen condition.

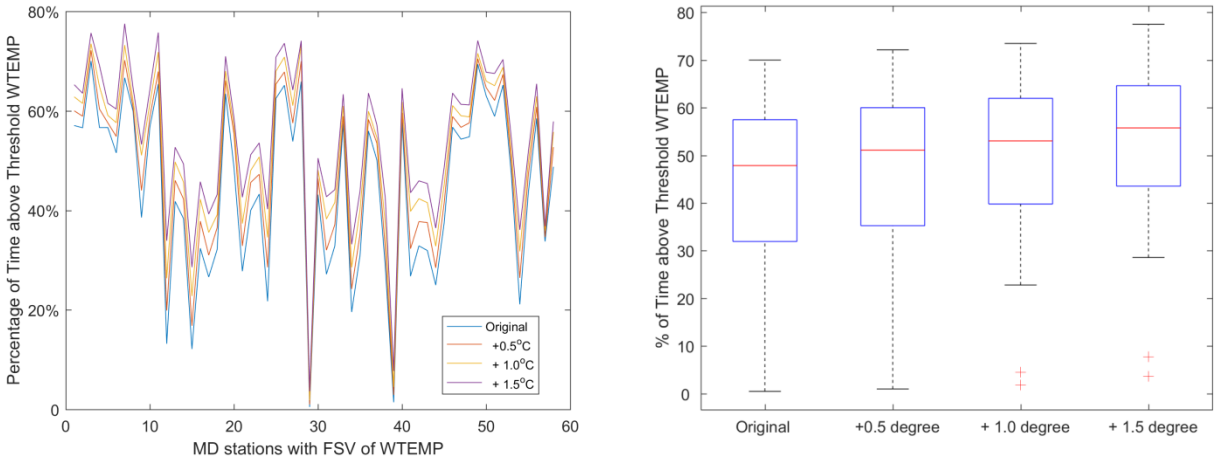

Figure 11: (left panel) Percentage of time between April to October at Maryland ConMon stations currently experiencing hypoxia that water temperatures exceeded the thresholds FSV value for hypoxia duration at each station under current conditions and under warming. (right panel) box plots representing all Maryland stations of % of time above the FSV threshold under current and warming scenarios (red line = median, upper and lower edge of box = 25th and 75th percentile, dashed lines and bars = extremes).

Effects of Eutrophication on Hypoxia: We synthesized relationships between chlorophyll-a and the duration of more severe hypoxia (hours below 2 mg O_2 L^{-1} and 3.2 mg O_2 L^{-1}) in select systems where chlorophyll-a levels were particularly high. These systems include relatively enriched areas of Chesapeake Bay's shallow waters, including the Corsica River, highly enriched regions of the Maryland Coastal Bays, and a few Western Shore tributaries (Fig. 12). In all three regions, we found positive relationships between water-column chlorophyll-a (summer mean) and the duration (hours) of hypoxia. Importantly, determining how to control chlorophyll in these shallows regions may reduce the number of hours of severe hypoxia per day. However, these relationships are highly-driven by location and year when chlorophyll-a was an average of 30 mg m^{-3} and higher (Fig. 12), indicating that there may be a threshold of extensive phytoplankton biomass that generates the organic matter and respiration rates necessary to drive nighttime oxygen concentrations below critical thresholds.

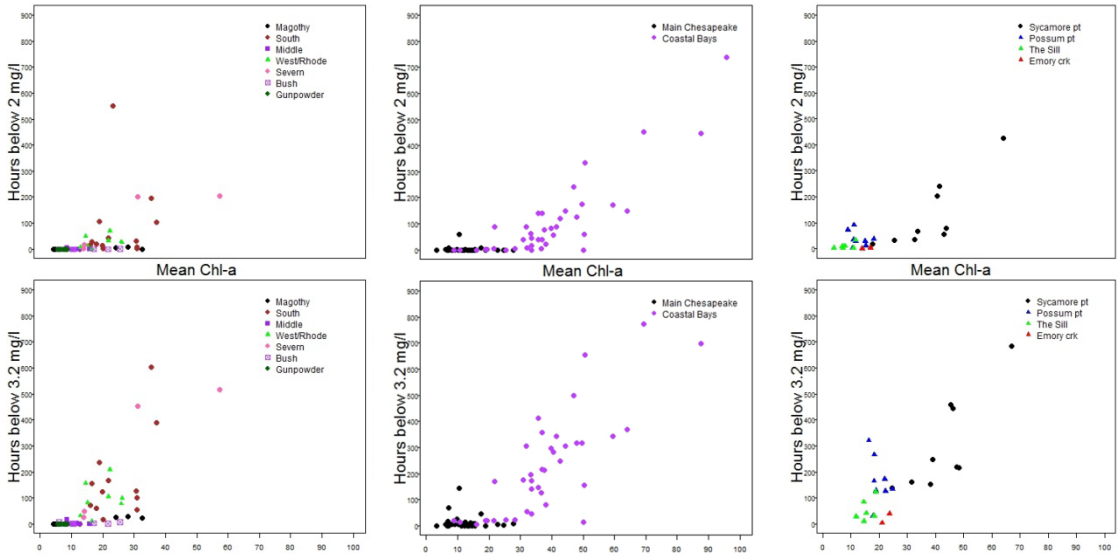

Figure 12: Correlation between mean summer chlorophyll-a concentrations and hours where oxygen is below 2 mg/L (top panels) and 3.2 mg/L(bottom panels) measured at ConMon sites on the western shore (left panels), Maryland Coastal Bays (middle panels), and Corsica River (right panels).

(2) **Modeling Shallow Water Hypoxia:** We combined simulations of a relatively high resolution shallow water biogeochemical model (ROMS-RCA, 200 m resolution) in the Chester River estuary with very-fine scale simulations using SCHISM-ICM in the Corsica River estuary (a Chester tributary). The Chester River ROMS-RCA simulation revealed a sensitivity of hypoxia to warming in both deep and shallow waters (data not shown), but the model suggested that (1) warming effects on Chesapeake Bay hypoxia that are communicated through bottom waters to the Chester estuary are a dominant control, and (2) that even a 200 meter resolution model using similar biogeochemical formulations to CBWQSTM cannot resolve shallow-water short-term oxygen dynamics. For example, at a station in the Chester River estuary where 15-minute O_2 data were available in 2003, hourly O_2 variations were substantial (hourly standard deviation occasionally > 2 mg O_2 L^{-1}; Fig. 13) and during two events led to O_2 departures below 4 mg O_2 L^{-1} for several hours. Model-simulated O_2 did capture episodic variations outside of the annual seasonal cycle, but the short-term variations (~ hourly) were not as large (~0.1 mg O_2 L^{-1}) as observed (Fig. 13). The implication of this underestimation of diel-variability is that modeled metabolic rates were likely lower than observed. However, warming scenarios led to clear downward departures in daily O_2 minima (data not shown), suggesting that warming will lead to not only reductions in mean O_2 but also increases in the duration of diel-cycling hypoxia.

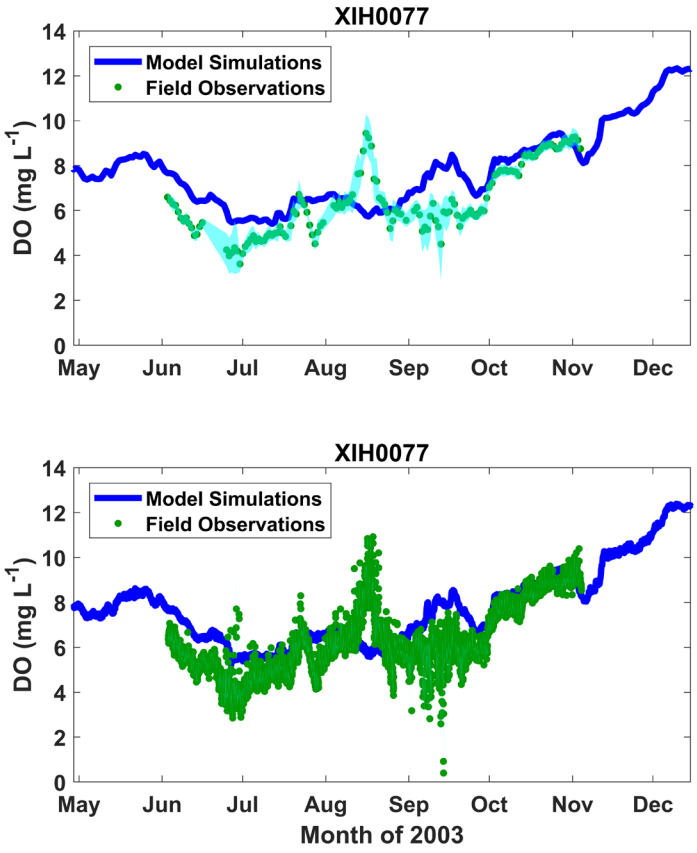

Figure 13: Comparisons of modeled (dark blue lines) and observed (green circles) oxygen concentrations at station XIH0077. The top panel includes modeled and observed oxygen values averaged over each 24 hour day (where light blue lines are 24-hour standard deviation of observations) and the bottom panel is hourly mean modeled and observed oxygen concentrations. Oxygen concentrations at XIH0077 were measured by in-situ sensors every 15-minutes from 1 meter below the surface.

Our collaboration with the CBP Modeling team, notably Richard Tian, has been identifying the key aspects of a very-fine resolution grid (SCHISM-ICM) on predicting shallow-water oxygen conditions. We presented the early results of that work at the April CBP quarterly meeting and again, we expect, at subsequent Modeling Quarterly meetings to report our progress.

(3) Tentative Conclusions and Recommendations: Here we summarize the key findings from our synthesis of shallow-water oxygen dynamics in Chesapeake Bay

1. Shallow-water models currently used in Bay management, even when implemented at extremely high resolution, do not capture diurnal variations in dissolved oxygen. Continued investigations into why these models fail should continue, particularly addressing the following questions: (1) Is natural variability in PAR adequately forced on the model at short enough (~hour) time steps?, (2) Do the metabolic rates of primary production and respiration computed within the model agree with the substantial rates derived from observations?, (3) Is wind-stress properly applied in protected shallow tributaries, given most wind products are based on larger scales?, and (4) Will fine-scale watershed model inputs be necessary to represent fine-scale effects of freshwater inputs to shallow waters and their associated circulation effects?

2. The relative importance of internal and external controls on DO are locally-dependent. Thus, persistent shallow-water monitoring and modeling is necessary to understand local implications of TMDL-associated responses of dissolved oxygen.

3. Future temperature increases will increase the vulnerability of shallow waters to failing 30-day oxygen criteria (and likely other criteria). Vulnerability to warming effects will be especially acute in locations that are highly turbid, and thus don't generate substantial oxygen during the day. Some naturally turbid waters (upstream, near tidal marshes) have always been vulnerable to low oxygen concentrations. Reduced sediment loads would potentially alleviate temperature stress in shallow waters

4. Oxygen concentrations in moderately eutrophic, low-turbidity shallow waters may actually be supplemented by high algal growth, given the associated production of oxygen during that day. Thus nutrient reduction strategies may make some shallow-water habitats more vulnerable to oxygen concentrations falling below 4.8 mg/L at night. However, under conditions of extremely high chlorophyll-a, hours of hypoxia are strongly correlated with chlorophyll-a, suggesting that elevated respiration of these algal communities enhances diel cycling hypoxia. High chlorophyll-a also increases the respiration rate for a given temperature, suggesting that warming combined with sustained eutrophication creates conditions for substantial diel cycling hypoxia.