Empirical Relationships Linking Algal Blooms with Threats to Aquatic Life Designated Uses in the James River Estuary

A Report from the Science Advisory Panel for the James River Chlorophyll Criteria Study

Submitted to the Virginia Department of Environmental Quality (VADEQ)

April 14, 2016

Corresponding Author: Paul A. Bukaveckas[1] (pabukaveckas@vcu.edu)

Science Advisory Panel: Clifton F. Bell[2], Brian Benham[3], Claire Buchanan[4], Todd Egerton[5], Greg Garman[1], Will Hunley[6], Rebecca LePrell[7], Kenneth Moore[8], Marjorie Mulholland[5], Kimberly S. Reece[8], Peter Tango[9], and Harry V. Wang[8].

Affiliations: (1) Virginia Commonwealth University, (2) Brown & Caldwell, (3) Virginia Tech, (4) Interstate Commission on the Potomac River Basin, (5) Old Dominion University, (6) Hampton Roads Sanitation District, (7) Virginia Department of Health, (8) Virginia Institute of Marine Science, (9) United States Geological Survey.

SUMMARY

The Commonwealth of Virginia identified a need for additional scientific study to ensure that chlorophyll-a (CHLa) criteria for the James River Estuary were appropriately protective of aquatic life designated uses. The Virginia Department of Environmental Quality (VADEQ) initiated a review of the numeric CHLa criteria for the James and established a Science Advisory Panel (SAP) to analyze the best scientific information currently available and provide recommendations as to whether the CHLa criteria were scientifically defensible.

Determining whether CHLa criteria are protective required an assessment of threats posed by algal blooms to aquatic life designated uses and the degree to which these threats would be abated if the criteria were attained. Protectiveness may be judged by various approaches which can be broadly delineated as reference-based (i.e., attaining conditions as would be expected in the absence of the stressor, or in a least-degraded state), and effects-based (mitigating observed deleterious effects of the stressor). The criteria recommended by VADEQ (2004) were derived from a review of reference- and effects-based results as well as model projections of attainable spring and summer CHLa mean concentrations. The SAP used the effects-based approach and a decade of additional data and information to reevaluate the criteria protectiveness. The use of an effects-based approach provided a means to consider multiple lines of evidence representing the various mechanisms by which algal blooms adversely affect aquatic life designated uses. These included water quality conditions (pH, dissolved oxygen, water clarity), phytoplankton community attributes (diversity, evenness, multimetric indices) and occurrence of harmful algae. For each of these metrics, the frequency of exceeding specified thresholds (e.g., nighttime dissolved oxygen < 5 mg L^{-1}) was related to co-occurring CHLa conditions in each season and segment. The frequency of occurrence for CHLa conditions in various ranges (e.g., 0-10, 10-20, etc. µg L^{-1}) was analyzed in relation to season- and segment- specific mean values. Combining these two relationships yielded estimates of the expected occurrence of threshold exceedance over a range of mean CHLa, which provided a basis for assessing protectiveness at attainment of the current criteria. This approach allowed us to quantify the expected benefits of attaining CHLa criteria in a form directly relevant to stakeholders.

Frequencies of threshold exceedance were positively related to mean CHLa thereby supporting the use of CHLa criteria to ameliorate threats to aquatic life designated uses posed by algal blooms. In all cases where protectiveness could be assessed (8 of 10 season-segment combinations), existing criteria fell

below the lower threshold of the non-protective range. This finding indicates that attainment of the existing criteria would protect aquatic life designated uses as indicated by reduced frequency of threshold exceedance. On this basis, the existing criteria were deemed to be defensible. In most cases, the criteria fell within a range below the lower threshold for non-protectiveness and above the threshold for lowest risk. Lowering the criteria may result in further improvements in water quality and phytoplankton condition, however, in most cases, these improvements would be small.

INTRODUCTION

Over-abundance of phytoplankton due to anthropogenic nutrient enrichment is one of the primary causes of poor water quality in the nation's estuaries, including Chesapeake Bay (Howarth et al. 2000; Kemp et al. 2005). Nutrients from point and non-point sources stimulate phytoplankton production and lead to secondary impacts including reductions in water clarity and submersed aquatic vegetation, depletion of dissolved oxygen and alterations in food web structure (Cooper and Brush 1993; Kemp et al. 2001). Nutrient enrichment is also associated with increasing incidence of harmful algal blooms (HABs); blooms that produce biochemicals which are harmful, and sometimes toxic, to humans and aquatic life (O'Neil et al. 2012; Paerl and Otten 2013). Though the general mechanisms linking nutrient inputs to adverse effects on aquatic systems are well known, establishing quantitative thresholds for undesirable levels of algae remains a challenge (Royer et al. 2008; Harding et al. 2014).

Linking algal abundance to adverse impacts on aquatic life designated uses is dependent upon establishing thresholds that are protective against deleterious effects on aquatic life. Numeric criteria may be established using generic indicators of algal abundance such as CHLa, or by the use of indicators that are specific to the presence of harmful algae (e.g., cell densities of toxin-producing species). CHLa concentrations are responsive to elevated nutrient inputs and are widely used for monitoring and regulatory purposes because concentrations are easily and reliably measured. Thus development of CHLa criteria may be a useful means for protecting designated uses against harmful effects of algal blooms. In addition, CHLa concentrations are amenable to deterministic modeling in which physical-chemical variables such as water residence time, underwater light conditions and nutrient availability are used to predict seasonal and spatial patterns in algal abundance. These models provide the basis for assessing attainability of proposed criteria by forecasting changes in CHLa concentrations in response to nutrient management scenarios and climatic variability. The USEPA (2003) encourages states to adopt numerical CHLa criteria for application to tidal waters in which algal-related designated use impairments are likely to persist even after attainment of the applicable dissolved oxygen and water clarity criteria. Such areas include, but are not limited to, waters that do not experience oxygen depletion for hydrodynamic reasons (e.g., shallow, well-mixed estuaries) and those in which reduced water clarity results primarily from suspended sediment rather than algae.

The 2007 CHLa Addendum to Ambient Water Quality Criteria (USEPA 2007) described approaches to developing criteria and their applications to Chesapeake Bay and its tributaries. The criteria were designed to protect the designated use to "support the survival, growth and propagation of balanced, indigenous populations of ecologically, recreationally and commercially important fish and shellfish species inhabiting open-water habitats" (USEPA 2003). The approaches were based on a consideration of historical reference conditions, effects on water quality conditions and protection from harmful algal blooms. The USEPA (2007) reference-based assessment used early historical data to characterize pre-impairment CHLa concentrations and a 90%-tile threshold approach to define a desired threshold that would protect against excessive CHLa. Key to this approach was the availability of historical, system-wide CHLa data obtained using consistent methodology, and demonstrating that CHLa had increased subsequent to the proposed reference period. Chesapeake Bay (CB) data from the 1960's and 1970's were used to define historical reference conditions, though it was recognized that these do not represent pristine conditions. A key point was demonstrating that CB had experienced increases in CHLa subsequent to this reference period, thereby allowing these data to be used as a benchmark for a less impaired state. A review of James River CHLa measurements extending from 1985 to 2015 did not reveal a period of reduced CHLa in the early historical record. It is likely that tidal tributaries such as the James exhibited eutrophication symptoms earlier than the mainstem CB thereby reducing the likelihood of finding historical data representative of pre-impairment conditions.

An alternative means for defining reference conditions is to use a space-for-time substitution approach whereby current conditions in unimpaired systems are used to infer historical reference conditions at the impaired site. This assumes that inter-system differences are minimal, and that the focal system will exhibit conditions similar to those occurring in other systems in the absence of the stressor. This approach, using CHLa concentrations associated with reference phytoplankton communities in Chesapeake Bay, was in part the approach used by the VADEQ to establish the existing numeric CHLa criteria for the James (Table 1; VADEQ 2004, 2005). The criteria are assessed on a seasonal (Spring, Summer) and segment-specific basis using pooled data of individual CHLa measurements obtained from monthly, fixed-station monitoring. The VADEQ developed CHLa criteria intended to meet a statewide use designation calling for a "balanced, indigenous population of aquatic life in all waters" (VADEQ 2004). The criteria were designed to protect against the over-abundance of nuisance or potentially harmful algal species. Criteria to protect balanced communities were based on Bay-wide data that included historical CHLa concentrations characteristic of less nutrient-enriched conditions and phytoplankton community composition characteristic of reference conditions (Buchanan et al. 2005; Lacouture et al. 2006). At the time, there was little information available to quantify the effects of algal blooms on designated uses of the James River. For example, it was unknown which toxins were present and whether toxin levels posed a risk to aquatic life. The available lines of evidence "did not clearly point to specific and defensible criteria levels" (Commonwealth of Virginia, 2010).

Segment	Spring	Summer
	CHLa (µg/L)	
TF_{up}	10	15
TF_{low}	15	23
OH	15	22
MH	12	10
PH	12	10

Table 1. Numeric CHLa criteria for James River Estuary.

The Science Advisory Panel sought to test the existing criteria using multiple lines of evidence that take into account the various mechanisms by which algal blooms adversely affect aquatic life designated uses. This effects-based analysis included water quality indicators (low dissolved oxygen, elevated pH and reduced transparency), metrics related to harmful algal blooms (e.g., cell densities of harmful taxa, presence of algal toxins) and phytoplankton community characteristics (e.g., diversity, evenness, and multimetric indices). Segment- and season- specific empirical relationships were derived relating frequencies of exceedance for these metrics as a function of CHLa. These provided a basis for communicating anticipated benefits to stakeholders as improvements in dissolved oxygen, clarity, occurrence of harmful algae, etc. (Van Dolah et al. 2015). Existing criteria were evaluated in relation to the expected frequency of occurrence for threshold exceedance to determine whether these were protective of aquatic life designated uses. In the following sub-sections, we review the rationale for selecting the various metrics and describe the data that were used to assess their relationship to CHLa.

Water Quality

Algal blooms may cause deleterious water quality conditions for aquatic life due to low dissolved oxygen (DO), elevated pH and reduced water clarity. CHLa criteria may be used to protect against adverse effects arising from poor water quality conditions in systems where empirical relationships show greater risk of poor water quality with increasing CHLa. Long-term monitoring data for the James do not show chronic oxygen depletion (e.g., seasonal hypoxia) in bottom waters. However, transient hypoxia may occur due to nighttime respiration during periods of elevated phytoplankton biomass. Continuous, fixed-station monitoring data collected from each of the five segments of the James were analyzed to determine whether the incidence of nighttime hypoxia, and corresponding daytime pH maxima, was related to CHLa. Empirical relationships relating the risk of exceeding DO and pH thresholds as a function of CHLa were evaluated by season and segment to determine whether existing criteria were protective.

Light requirements of submerged aquatic vegetation (SAV) determine the depth to which plants can grow and place an upper limit on their distribution and abundance (Kemp et al. 2004). Deterioration of water clarity is believed to be the principal cause of wide-spread decline in SAV throughout Chesapeake Bay (Orth and Moore 1983). A water clarity-based approach to defining CHLa criteria takes into account minimum light requirements for SAV and algal contributions to light attenuation (Gallegos 2001; Shields et al. 2012). Determination of the latter is complicated by the fact that there are multiple factors affecting underwater light, and that algae contribute a variable fraction of overall attenuation. Factors contributing to light attenuation include colored or chromophoric dissolved organic matter (CDOM), suspended algae (CHLa) and non-algal suspended particulate matter (SPM). Of these, contributions from non-algal SPM are often large, due to tidal re-suspension of benthic particulate matter, and highly variable, due to episodic inputs of watershed-derived particulates. Phytoplankton biomass estimates obtained from long-term monitoring at three stations in the James were analyzed to determine the proportional contribution of phytoplankton to suspended particulate matter. These data were used to derive empirical relationships relating the risk of exceeding water clarity thresholds as a function of CHLa by season and segment.

Phytoplankton

Phytoplankton community metrics are widely used in the assessment of coastal water quality conditions and as indicators of biotic responses to stressors (Martinez-Crego et al. 2010). Multimetric indices based on phytoplankton community structure (e.g., evenness, richness, diversity) are analogous to those developed for macroinvertebrate and fish communities. Multimetric indices are reported to be more robust than their component variables and have been developed for a number of estuaries, including Chesapeake Bay (Buchanan et al. 2005; Lacouture et al. 2006; Marshall et al. 2006). The Phytoplankton Index of Biotic Integrity (PIBI) for Chesapeake Bay uses 5-9 phytoplankton metrics to distinguish least degraded (reference) and degraded conditions in eight season- and salinity-based habitats. The PIBI characterizes the status of phytoplankton communities relative to nutrient and light conditions. We used long-term data collected by the Chesapeake Bay Program (CBP) and supplementary data collected in conjunction with this study to evaluate phytoplankton community metrics (evenness, taxa richness and PIBI) in relation to CHLa. Empirical relationships linking these metrics to CHLa were evaluated by season and segment.

Because harmful algal blooms (HABs) are often associated with high CHLa, a logical goal for numeric criteria is the prevention of these outbreaks. A challenge to developing HAB-based criteria is that blooms of non-harmful species can also result in high CHLa thereby resulting in poor correlations between CHLa and cell densities of harmful taxa, or related parameters, such as toxin production. For harmful taxa which comprise a consistently large fraction of the phytoplankton community (e.g., dinoflagellates in the lower James), correlations with CHLa will be stronger, thus providing a basis for CHLa-based criteria. A second challenge to linking harmful algae to effects on aquatic life is variable toxicity within and among species. Assemblages include toxic as well as non-toxic strains, and toxin production is variable even among toxic strains (Turner and Tester 1997; Landsberg et al. 2008; Mulholland et al. 2009; Kudela and Gobler 2012; Rastogi et al. 2014). Lastly, a key challenge to developing HAB-based CHLa criteria lies in identifying appropriate thresholds that are protective of aquatic life designated uses. Using previously published studies and data arising from this study, we established thresholds of concern for the dominant harmful taxa occurring in the James (the dinoflagellate *Cochlodinium* and the cyanobacterium *Microcystis*) and for the cyanotoxin microcystin. These thresholds were used to derive empirical relationships relating the risk of exceedance to CHLa.

METHODS

Data analysis performed by the Science Advisory Panel followed the typical components of stressor-response assessment which include exposure analysis (characterizing variability in the mean and range of CHLa by season and segment), effects analysis (relating water quality and phytoplankton community metrics to CHLa), and risk characterization (quantifying expected rates of threshold exceedance in relation to current criteria). In this section we provide a short background on algal blooms in the James River estuary, describe sources of data, and explain methods used for data analysis.

Site Description

The James River Estuary (Figure 1) is the southernmost and third largest sub-estuary of Chesapeake Bay based on surface area, discharge and nutrient delivery. The estuary is delineated into five segments including Upper and Lower Tidal Fresh (which together comprise 16% of total surface area), Oligohaline (21% by area), Mesohaline (50% by area) and Polyhaline (13% by area). The James experiences algal blooms throughout its length, though the nature, timing and frequency of these events differ between the upper (freshwater) and lower (saline) segments. The Upper Tidal Fresh segment extends from the Fall Line to Hopewell, VA and is characterized by a narrow, deep channel. The low ratio of photic depth to total depth creates poor light conditions resulting in low CHLa (typically $< 10\ \mu g\ L^{-1}$) despite elevated nutrient concentrations associated with riverine and point source inputs (Bukaveckas et al. 2011). The Lower Tidal Fresh segment extends from Hopewell, VA to the confluence with the Chickahominy River and is characterized by persistent (May-October) elevated CHLa due to shallower depths which provide favorable conditions of light and water residence time (Bukaveckas et al. 2011; Bukaveckas and Isenberg 2013; Wood and Bukaveckas 2014). Phytoplankton communities of the Tidal Fresh segments are dominated by diatoms and chlorophytes, but a small contribution from cyanobacteria (typically 5-10% of biomass) results in widespread occurrence of the cyanotoxin microcystin in water and tissues of fish and shellfish (Wood et al. 2014). The lower estuary (inclusive of the Oligo-, Meso- and Poly- haline segments) experiences algal blooms that are ephemeral in nature and unpredictable in their timing, location and duration (Morse et al. 2011). Blooms in the lower James are comprised of dinoflagellates (Marshall et al. 2005) that have been shown to have harmful effects on diverse aquatic biota (Kim et al. 1999; Gobler et al. 2008; Tang and Gobler 2009). The blooms are transported by currents such that sites of bloom initiation may be geographically distinct from areas where blooms develop and cause detrimental effects on water quality and living resources.

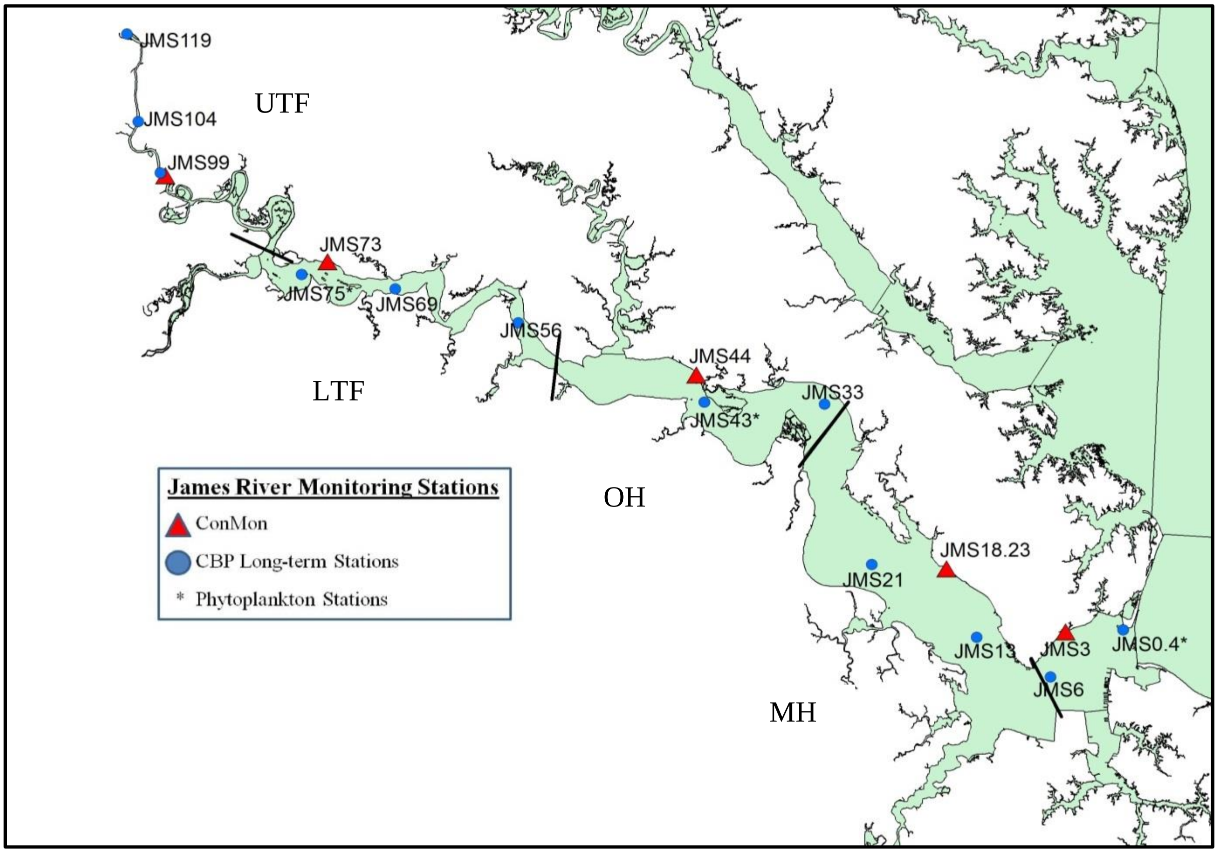


Figure 1. Map of the James River Estuary showing mainstem monthly monitoring locations (CBP) and nearshore continuous monitoring (ConMon) stations. Station numbers denote distance from the confluence with Chesapeake Bay. Solid lines delineate segments: Upper & Lower Tidal Fresh (UTF, LTF), Oligohaline (OH), Mesohaline (MH) and Polyhaline.

James River data used for this analysis were obtained from multiple sources including: (1) monthly, fixed station monitoring conducted by VADEQ for the CBP (1986-2013), (2) weekly monitoring of stations in the Tidal Fresh segments by VCU (2010-2014), (3) continuous monitoring at fixed stations by VIMS (2006-2008) and VIMS-HRSD (2012-2013; MH only) and (4) spatial mapping of water quality ('Dataflow' cruises) in the Lower James during 2005-2013 by VIMS and HRSD (Table 2). In the following sections we provide information on how these data were collected and analyzed.

Paramater	Metric & Threshold	Period	Segment	Data & Source
CHLa	Seasonal and segment-specific means compared to current criteria	2009-2014	ALL	Weekly fixed-station monitoring (TF1, TF2; VCU); weekly spatially continuous monitoring (OH,MH,PH; HRSD)
Dissolved Oxygen (DO)	Daily minima (10%-tile) < 5 mg/L	2006-2008 2012-2013	ALL MH only	Continuous fixed station monitoring (VIMS)
pH	Daily maxima (90%-tile) > 9	2006-2008 2012-2013	ALL MH only	Continuous fixed station monitoring (VIMS)
Water Clarity	Algal contributions to TSS > 20%	1985-2013	LTF, OH, PH	Monthly phytoplankton counts & TSS (CBP)
Phytoplankton	Community multimetric indices (PIBI>2.67)	1985-2013	LTF, OH, PH	Monthly phytoplankton counts (CBP)
	Community diversity & eveness	2011-2013	OH, MH, PH	Phytoplankton sampling during dataflow cruises (HRSD, ODU)
	HAB (Microcystin > 0.8 µg/L)	2011-2014	UTF, LTF	Weekly monitoring of Microcystin (VCU)
	HAB (*Microcystis* > 20k cells/ml)	1985-2014	UTF, LTF	Monthly (1985-2013) and weekly (2011-2014) phytoplankton counts (ODU)
	HAB (*Cochlodinium* > 1,000 cells/ml)	2011-2014	OH, MH, PH	Phytoplankton sampling during dataflow cruises (HRSD, ODU)

Table 2. Metrics, thresholds and data sources used to develop empirical relationships linking CHLa with threats to designated uses (UTF and LTF denote Upper and Lower Tidal Fresh segments, respectively).

Chlorophyll-a

The Upper and Lower Tidal Fresh segments were sampled weekly by VCU during 2009-2014 at 12 stations (5 of these were also CBP sites). CHLa was determined by fluorometric analysis of extracted pigments (Wood and Bukaveckas 2014). In the lower James, spatially-continuous CHLa measurements were obtained by VIMS and HRSD using a boat equipped with a YSI 6600 multiparameter sonde and a Garmin GPSMAP 168 Sounder. Established cruise tracks, which include nearshore and mid-channel habitats, were surveyed every 1-2 weeks during Spring and Summer assessment periods (mean = 10 cruises per season). CHLa was measured by *in situ* fluorescence and converted to extracted equivalents based on annual regressions. The regressions were derived by extracting CHLa at a subset of sites which spanned the range of observed CHLa. Post-processing of cruise data involved a sub-sampling procedure in which the four measurements closest to the center of each grid cell (CBP interpolator; Figure 2) were selected. The sub-sampling procedure was used to avoid over-representation of sites where the boat was at reduced speed. Data from the weekly cruises were averaged to obtain a segment- and season- specific mean, or pooled to obtain a frequency distribution representative of the segment and season.

Annual means and frequency distributions of CHLa were derived by segment for the Spring (March-May) and Summer (July-September) assessment periods using data from weekly monitoring in the Tidal Fresh segments and spatially continuous monitoring in the lower James. These data were used to test relationships for various metrics (e.g., microcystin concentrations, *Cochlodinium* cell densities) with mean CHLa by season and segment. Comparisons of mean CHLa to current criteria were made for illustrative purposes; the results are not intended to represent regulatory non-attainment.

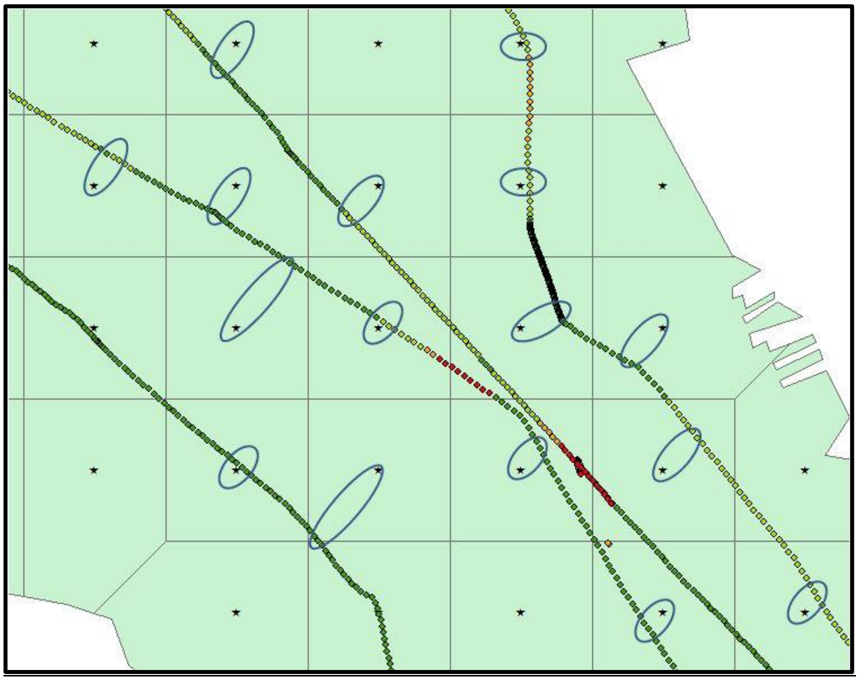


Figure 2. Example of Dataflow cruise tracks in relation to grid cell structure. The 4 measurement points located closest to the centroid of each cell were used to derive season- and segment- specific means and frequency distributions.

Two additional sources of CHLa data were used to assess relationships with candidate metrics. CHLa measurements from long-term monthly monitoring by the VADEQ for the CBP were used to test for relationships with PIBI scores and algal contributions to suspended particulate matter. CHLa was determined by pigment extraction and spectrophotometric analysis. CHLa data from continuous fixed station monitoring was used to test for relationships with daytime pH maxima and nighttime DO minima. Fixed station continuous monitoring of CHLa (15 min intervals) was conducted by VIMS during 2006-2008 at 5 nearshore stations (one per salinity segment, inclusive of Upper and Lower Tidal Fresh). Measurements of *in situ* fluorescence were converted to extracted equivalents by means of annualized regressions derived from extracted samples collected at each location and analyzed fluorometrically.

pH, Dissolved Oxygen & Water Clarity

Continuous monitoring data were collected during 2006-2008 by VIMS at sites representing each of the 5 segments. Multiparameter sondes (YSI 6600) were deployed in nearshore areas at a depth 1.5 m below mean low water. Sondes were calibrated at 2-3 week intervals. The incidence of daily DO minima less than 5 mg L^{-1} and daily pH maxima exceeding 9.0 was evaluated in relation to CHLa. Dissolved oxygen concentrations in the James are regulated by various criteria that are seasonally variable (e.g. to protect Spring migratory fish spawning) and habitat-specific (criteria for open water, deep water and deep channel). These criteria vary in their magnitude and exposure (e.g., for open water: 30-d mean > 5 mg L^{-1}, 7-d mean > 4 mg L^{-1} and instantaneous minimum > 4.3 mg L^{-1}). Our objective was not to duplicate assessments of DO impairment for the James, but to examine how the risk of low oxygen conditions varied in relation to CHLa, particularly during transient (nighttime) oxygen depletion. We analyzed variation in daily mean and minima of DO concentrations in relation to CHLa. Minima were calculated as 10%-tile values from continuous (15 min) measurements recorded at fixed stations. A corresponding analysis was performed for daily pH maxima (90%-tile values > 9.0). The 9.0 pH level was selected because it corresponds to the upper end of the range of Virginia's water quality criteria. Virginia's water quality standards do not specify an explicit duration component for the 9.0 pH criterion. However, the pH criteria were originally based on a combination of chronic/longer-term effect studies. Our use of the upper and lowest 10% of measurements during each day corresponds to an exposure period of ~2.4 h. As for dissolved oxygen, our intent here was not to characterize impairment but to assess potential risk to aquatic life in relation to CHLa based on the frequency of threshold exceedance. The number of observations used in this analysis ranged from 192 to 299 per segment and season (Table 3) where each observation corresponds to a daily pH maximum or DO minimum paired with the daily mean CHLa.

Segment	Season	pH & DO	Clarity	PIBI	Richness & Evenness	Microcystin	*Microcystis*	*Cochlodinium*
					Number of Observations			
TF-up	Spring	196	0	0	10	0	0	0
	Summer	274	0	0	32	97	0	0
TF-low	Spring	192	49	43	24	0	273	0
	Summer	254	50	49	31	229	163	0
OH	Spring	195	48	46	0	0	0	0
	Summer	276	53	49	0	0	0	0
MH	Spring	299	0	0	191	0	0	0
	Summer	254	0	0	171	0	0	171
PH	Spring	276	81	92	67	0	0	0
	Summer	275	93	107	81	0	0	81

Table 3. Number of observations used to derive relationships between water quality and phytoplankton metrics with CHLa by segment and season. 'Observations' are individual samples/measurements except for DO and pH which are daily minima and maxima, respectively, derived from 15 min measurements.

Algal contributions to water clarity were evaluated by estimating the proportion of suspended particulate matter comprised of algae. Phytoplankton biomass was determined from samples collected during monthly monitoring at stations located in the Lower Tidal Fresh (JMS75), Oligohaline (JMS43) and Polyhaline (JMS0.4) segments. Cell counts were converted to taxon-specific biovolumes based on shape and average cell dimensions, biovolumes were converted to C mass assuming taxon-specific carbon:biovolume ratios, and C mass was converted to biomass assuming that C accounted for 50% of cell mass. Total phytoplankton biomass was divided by the corresponding TSS concentration to calculate algal contribution to suspended matter. The probability of algal contributions to TSS exceeding 20% of TSS was derived to determine the frequency of threshold exceedance as a function of CHLa. The number of observations for this analysis ranged from ~50 (in each season) for the Lower Tidal Fresh and Oligohaline segments to 80-90 in the Polyhaline segment (Table 3).

Phytoplankton

PIBI values were derived for phytoplankton samples collected during 1985 to 2013 as part of the Chesapeake Bay long-term monitoring program (Buchanan et al. 2005; Lacouture et al. 2006; Johnson and Buchanan 2014). Data were obtained from three phytoplankton monitoring stations located within the Lower Tidal Fresh (JMS75), Oligohaline (JMS43) and Polyhaline (JMS0.4) segments. Mean PIBI scores and associated mean CHLa values were derived by segment and season within ranges of CHLa (0-10, 10-20, etc. µg L^{-1}). The frequency of occurrence for PIBI scores indicative of least degraded communities (PIBI $\geq$ 2.67) was determined for each range of CHLa. The number of observations for this analysis ranged from ~50 (in each season) for the Lower Tidal Fresh and Oligohaline segments to 80-90 per season in the Polyhaline segment (Table 3). A similar analysis was performed using a larger dataset of CBP samples collected from VA waters to compare against James-specific results. Lastly, supplemental data on phytoplankton communities were collected for this study during 2011-2013 in conjunction with weekly sampling in the Tidal Fresh segments and during Dataflow cruises in the Meso- and Poly- haline segments. Counting procedures for these samples differed from those used by the Bay Program which precluded using the PIBI metric. Analysis of these data focused on metrics of community structure (evenness, species richness) and the abundance of harmful taxa. Pielou's evenness index (J') ranges from 0 (least even) to 1 (most even) and was derived as:

$$J' = H'/H'max$$

with H' being the Shannon diversity index (-Σ pi Log(2)pi) (where pi = proportion of species i based on biomass) and H'max equal to Log(2)S (where S = species richness). The number of samples used for this analysis was 10-30 (in each season) for the Tidal Fresh segments, 67-81 in the Polyhaline segment, and 171-191 in the Mesohaline (Table 3).

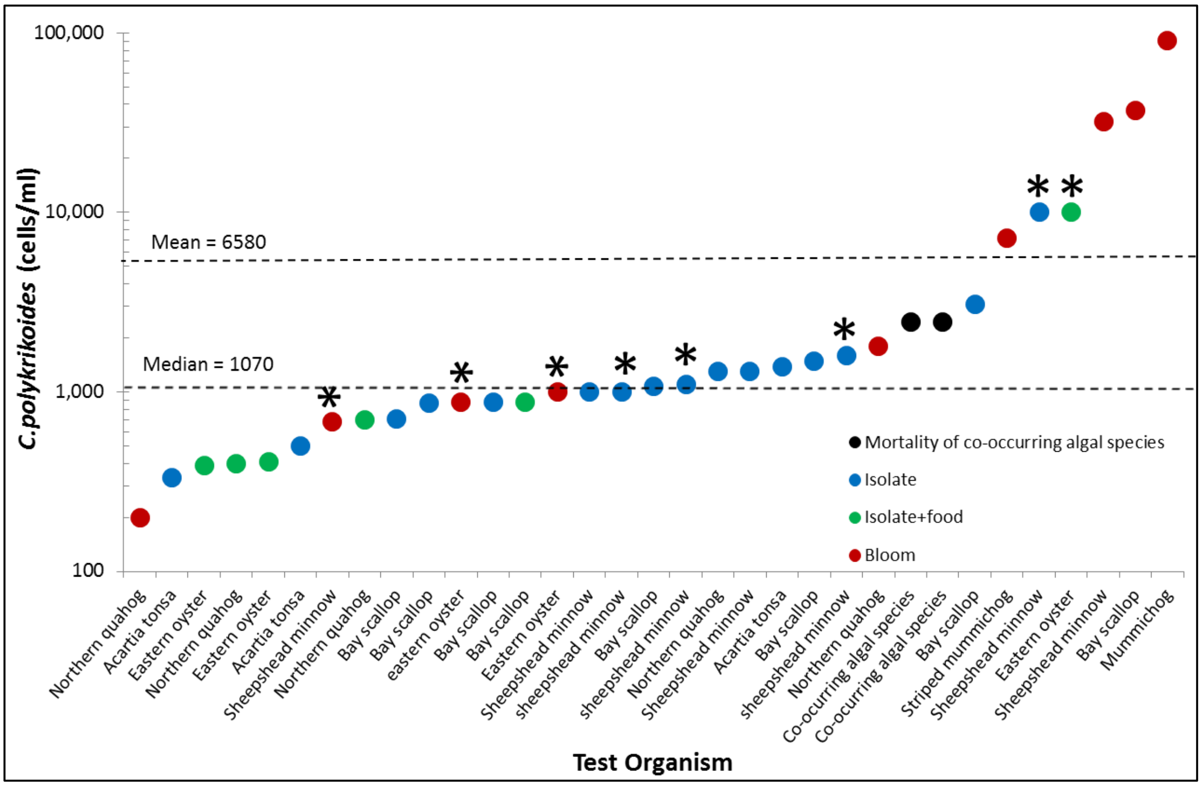


Figure 3. Compilation of results from toxicity assays showing minimum cell densities of *Cochlodinium polykrikoides* at which >20% mortality of the test organism was observed. Asterisks denote trials with James River strains.

Among harmful taxa, we focused on *Microcystis aeruginosa*, a cyanobacteria occurring in the freshwater region, and the dinoflagellate *Cochlodinium polykrikoides*, the dominant summer bloom-forming harmful taxa occurring in the saline region of the James. We analyzed variation in the abundance of these taxa in relation to CHLa. The objective was to determine whether there was an increase in the likelihood that their abundance exceeded thresholds of concern as a function of CHLa. For *Cochlodinium*, thresholds of concern were based on a literature review of previously published toxicity trials, along with toxicity assays conducted for this study (using strains from the James). This compilation showed that toxic effects (defined here as >20% mortality relative to controls) occurred over a wide range of minimum cell densities, but that the bulk of these fell within 1,000 to 10,000 cells/ml (Figure 3). We selected a cell density at the lower end of this range (1,000 cells mL^{-1}) to determine the probability of threshold exceedance as a function of CHLa. Data used for this analysis included 171 (Mesohaline) and 81 (Polyhaline) measurements of *Cochlodinium* density obtained in Summer (Table 3).

For cyanobacteria, we focused on *Microcystis* and the cyanotoxin microcystin. Variation in *Microcystis* cell densities was analyzed in relation to CHLa using long-term data from VA tidal freshwater plankton monitoring stations sampled for the CBP. A total of 273 (Spring) and 163 (Summer) observations were available for this analysis (Table 3). The same relationships were also explored in a dataset generated for this study, which included weekly sampling during 2011-2014 at several stations in the Upper and Lower Tidal Fresh segments of the James. We quantified the frequency with which cell densities exceeded threshold values as a function of CHLa. A cell density threshold of 78,600 cells ml^{-1} was selected on the basis of being the lowest cell density causing a mortality of at least 20% in bioassay experiments performed using *Ceriodaphnia dubia* and *Cyprinodon variegatus* (Table 4). While the intent of the CHLa criteria was to protect aquatic life designated uses, we also assessed exceedance rates for *Microcystis* thresholds established by the Virginia Department of Health for public safety (20,000 and 100,000 cells ml^{-1}).

Test organism	Live Cells *Microcystis*	Lysate cells/ml
C. variegatus	NA	1,570,000
C. dubia	78,600	157,000
C. dubia	160,000	160,000

Table 4. Results from toxicity assays showing the lowest cell densities of *Microcystis* that cause at least 20% mortality in test organisms (data from K. Reece).

Data on microcystin concentrations in the James Estuary were limited to the period of this study (2011-2014). During this period, microcystin was monitored weekly at two stations in the Upper Tidal Fresh segment (JMS85, JMS99) and four stations in the Lower Tidal Fresh segment (JMS75, JMS69, JMS56 and VCU Rice Pier). Samples were analyzed using the high sensitivity ADDA ELISA Kit (detection limit 0.05 µg L^{-1}; Abraxis; Warminster, PA). The assay measures numerous forms of free microcystin using polyclonal antibodies with concentrations reported in MC-LR equivalents. To release microcystin from cells, water samples were thawed and refrozen two times (as recommended by the manufacturer), and then microwaved and sonicated to improve exaction efficiency (Wood et al. 2014). We evaluated the frequency of exceeding threshold microcystin concentrations as a function of CHLa. A total of 97 (Upper Tidal Fresh) and 229 (Lower Tidal Fresh) observations were used for this analysis (all Summer; Table 3).

A review of previously published studies revealed that mortality effects occur at microcystin concentrations of 100's to 1,000's µg L^{-1} (Smith et al. 2008), which are orders of magnitude higher than those observed in the James (0.1-5 µg L^{-1}; Wood et al. 2014). Because lethal effects are unlikely to occur in this system, we assessed the likelihood of ecological effects by performing experiments to measure the effects of microcystin on consumer grazing. We focused on the wedge clam (*Rangia cuneata*) which is the dominant benthic suspension feeder in the tidal fresh James and accounts for >80% of benthic macroinvertebrate biomass based on annual CBP surveys. Our prior work has documented reductions in *Rangia* clearance rates during periods when microcystin concentrations were elevated (Wood et al. 2014). We conducted controlled experiments to measure the effects of dissolved microcystin on *Rangia* clearance rates and identify thresholds at which deleterious effects were observed. Our method for measuring clearance rates was described in Wood et al. (2014) and follows Wong et al. (2010). The relationship between *Rangia* clearance rates and microcystin was used to derive a "CR_{50}" – the toxin concentration at which clearance rates were reduced by 50%. Two replicate experiments yielded microcystin CR_{50} values of 0.40 µg L^{-1} (R^2 = 0.94) and 1.15 µg L^{-1} (R^2 = 0.61). A threshold of 0.8 µg L^{-1}

was used for deriving probabilities of exceedance as a function of CHLa and to assess threats to aquatic life designated uses. The presence of microcystin is also considered a threat to human health via recreational contact and ingestion of water or fish and shellfish (Poste et al. 2011). However, microcystin concentrations in the James did not exceed the recommended VDH contact standard (6 µg L^{-1}) during the period of monitoring (2011-2014) and an analysis by VADEQ has found that microcystin concentrations in fish and shellfish from the James did not exceed the WHO guidelines for tolerable daily intake (VADEQ 2014). Therefore we did not further consider human health risks related to the presence of microcystin.

Threshold Exceedance Analysis

To determine whether existing CHLa criteria for the James are protective against the impacts of algal blooms we assessed the likelihood of deleterious effects based on expected rates of threshold exceedance. Expected rates of threshold exceedance were derived by segment and season for each metric that exhibited statistically significant relationships to CHLa. Statistical significance was determined from a comparison of means (e.g., *Cochlodinium* densities) between populations comprised of samples collected when CHLa concentrations were above vs. below the current criteria. Our rationale was that statistically significant differences were indicative of deleterious effects associated with algal blooms, whereas metrics not displaying significant differences were uninformative with respect to protectiveness (e.g., CHLa criteria could be used to protect against excessive *Cochlodinium* densities only if *Cochlodinium* densities were related to CHLa).

Threshold exceedance rates were derived for each metric, season, and segment based on (a) the probability of exceeding thresholds of concern within a given CHLa range, and (b) the probability of occurrence for CHLa values in that range (see Supplemental Information for sample calculations). The probability of threshold exceedance was determined by binning observations (e.g., *Cochlodinium* densities) within CHLa ranges (0-10, 10-20, etc. µg L^{-1}) to determine the proportion of samples in that range exceeding the threshold. Similarly, we determined the probability of occurrence of each CHLa range among all observations for a given season, segment and year. We restricted the analysis of CHLa distribution frequencies to the intensive sampling period (weekly monitoring of Tidal Fresh segments for 2009-2014 and Dataflow results from the lower James for 2005-2013) to ensure adequate sample sizes among the CHLa ranges (bins). From these analyses we obtained a single probability distribution for each metric by season and segment (e.g., frequency of *Cochlodinium* threshold exceedance in Summer Polyhaline), and a series of CHLa-based probability distributions, which reflected the number of years for which intensive data were available (6-9 years depending on segment).

A combined probability was derived based on the observed frequency of metric threshold exceedance within a given CHLa interval (e.g., pH > 9 when CHLa = 10-20 µg L^{-1}), and the frequency of occurrence for that CHLa interval in a given year (e.g., the proportion of CHLa measurements in the range of 10-20 µg L^{-1}). The product of these two probability distributions (when summed across bins) yielded

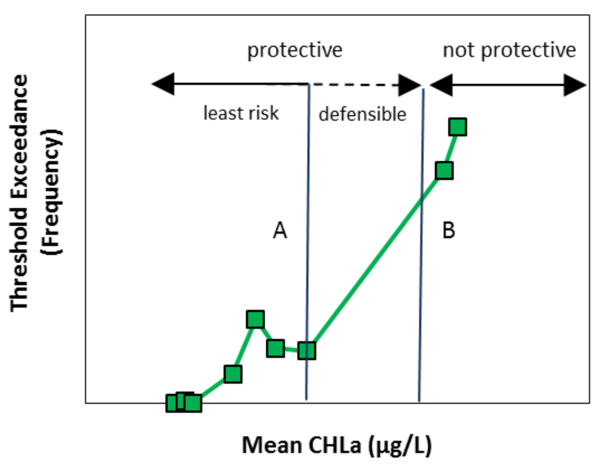


Figure 4. The relationship between the frequency of threshold exceedance and mean CHLa was used to assess the protectiveness of CHLa criteria.

an estimate of the expected frequency of threshold exceedance for a given year. Repeating this procedure across years for which intensive sampling data were available, allowed us to depict the relationship between expected rates of threshold exceedance and mean CHLa for each metric (Figure 4). Protectiveness was judged on the basis of anticipated improvements in water quality with decreasing mean CHLa, as indicated by lower rates of threshold exceedance. Existing criteria were judged to be 'not protective' if falling within the CHLa range where elevated threshold exceedance values were observed (e.g., criteria above Line B in Figure 4). Criteria were considered 'protective' if they fell within the CHLa range below the elevated exceedance values. Within the protective range, we further distinguished a range of 'least risk' for which attainment would result in expected exceedance rates at the low end of their observed distribution (e.g., criteria falling at or below line A in Figure 4). In the region between A and B, there is greater uncertainty. However, criteria in this range were considered 'defensible' in that they occupy a CHLa range below that where elevated exceedance values were observed. By this method, we established least risk, defensible and non-protective ranges for each metric, which served as the basis for assessing the protectiveness of existing criteria. In the following section we present the empirical relationships used determine the expected rates of threshold exceedance for each of the metrics by season and segment.

RESULTS

Chlorophyll-a

Highest mean CHLa concentrations were observed in the Lower Tidal Fresh segment, particularly during the summers of 2009 (mean = 98.1 ±6.9 μg L^{-1}) and 2010 (85.7 ±6.1 μg L^{-1}; Figure 5). During 2011 to 2014, Summer mean values in this segment ranged from (26.6 to 43.4 μg L^{-1}). In all years, Summer means exceeded the criteria for the Lower Tidal Fresh segment (23 μg L^{-1}). Mean CHLa values were lower in Spring (range of means = 9.3 to 16.1 μg L^{-1}) and exceeded the criteria (15 μg L^{-1}) in 2 of the 4 years for which weekly data were available. Patterns of inter-annual variation in the Upper Tidal Fresh segment were similar to the Lower Tidal Fresh, though at lower CHLa concentrations. Summer means ranged from 6.0 to 39.8 μg L^{-1} and exceeded current criteria (15 μg L^{-1}) in 2 of 6 years (2009 and 2010). Spring average values were below the current criteria (10 μg L^{-1}) in all 4 years (range of means = 2.6 to 9.1 μg L^{-1}). In the Oligohaline segment, Spring average CHLa ranged from 3.0 to 18.2 μg L^{-1} and exceeded the current criteria (15 μg L^{-1}) in 1 of 6 years. Summer average values ranged from 2.7 to 10.8 μg L^{-1} and did not exceed the 22 μg L^{-1} criteria. In the Mesohaline segment, Spring CHLa ranged from 5.3 to 22.1 μg L^{-1} and exceeded the current criteria (12 μg L^{-1}) in 3 of 9 years. Summer average values ranged from 5.6 to 16.4 μg L^{-1} and exceeded current criteria (10 μg L^{-1}) in 3 of 9 years. In the Polyhaline segment, the range of mean CHLa was similar to that observed in the Mesohaline. Spring CHLa ranged from 4.3 to 14.5 μg L^{-1} and exceeded current criteria (12 μg L^{-1}) in 1 of 9 years. Summer average values ranged from 6.6 to 16.1 μg L^{-1} and exceeded current criteria (10 μg L^{-1}) in 5 of 9 years. Corresponding patterns of exceedance were observed between the Meso- and Poly- haline segments in Spring (2005) and Summer (2005, 2008 and 2010). There was little correspondence between the upper and lower segments of the James, except for 2010 (Summer only) when all segments exceeded their respective criteria (no data for Oligohaline in this year). Overall, exceedance of current criteria was observed most often during Summer in the Lower Tidal Fresh (6/6 years) and Polyhaline (5/9 years) segments, and least often during Spring in the Upper Tidal Fresh (0/4 years), Polyhaline (1/9 years) and Oligohaline (1/6 years) segments.

Frequency distributions were used to characterize the variability in CHLa relative to the mean (Figures 6 and 7). In all segments and seasons there was a trend toward greater frequency of high CHLa and decreasing occurrence of low CHLa as mean CHLa increased. Using CHLa values exceeding 100 µg L^{-1} as an example, these accounted for 36% and 30% of observations from the Lower Tidal Fresh segment during the summers of 2009 and 2010 (respectively) when highest mean CHLa concentrations occurred. Summer CHLa values in this range were also observed in the Upper Tidal Fresh segment in 2009 (1%) and 2010 (15%), corresponding to highest mean values in this segment. When mean CHLa concentrations were low (e.g., Spring 2013 and 2014), there were few (<5%; Lower Tidal Fresh) or no (Upper Tidal Fresh) measurements above 20 µg L^{-1}. Similar patterns were observed in the segments of the Lower James, though the occurrence of high CHLa was less common. CHLa concentrations exceeding 100 µg L^{-1} accounted for ~5% of observations when mean CHLa exceeded 20 µg L^{-1} (Spring Mesohaline in 2006 and 2013). When mean CHLa was in the range from 10 to 20 µg L^{-1}, measurements exceeding 100 µg L^{-1} accounted for 1-3% of observations. When mean CHLa was less than 10 µg L^{-1}, measurements exceeding 100 µg L^{-1} were rare or absent, and measurements less than 10 µg L^{-1} accounted for 85-95% of observations. For all seasons and segments, mean CHLa was found to be a strong predictor of the increasing proportion of observations at the highest end of the CHLa range.

pH, Dissolved Oxygen & Water Clarity

Positive relationships between pH and CHLa were observed in each of the segments based on data from continuous, fixed-station monitoring (Figure 8). In the Upper Tidal Fresh, mean pH was between 7.5 and 8.0 over the range of observed CHLa. In Summer, pH was significantly higher when CHLa exceeded current criteria relative to pH when CHLa was below current criteria. There was no significant difference in Spring. There were no occurrences of daily maxima exceeding a pH of 9.0 in either season in this segment. In the Lower Tidal Fresh segment, mean pH varied over a larger range (7.2 to 8.8) with lowest values occurring during low CHLa conditions (<30 µg L^{-1}). Highest pH occurred when CHLa exceeded 60 µg L^{-1} (Spring = 8.75 ± 0.04; Summer = 8.62 ± 0.02). In this CHLa range, the incidence of daily maxima exceeding a pH of 9.0 was 60% (Spring) and 43% (Summer). At lower CHLa, the frequency of exceedance was 8% and 13% (Spring and Summer, respectively; CHLa = 30-60 µg L^{-1}). No instances of exceedance occurred when CHLa was less than 30 µg L^{-1}. A statistical comparison of pH during conditions meeting or exceeding the criteria for the Lower Tidal Fresh segment could not be performed due to the low number of measurements for the former. Positive relationships between pH and CHLa were observed in the Oligo-, Meso- and Poly- haline segments. Low pH conditions (<8) were uncommon in these segments (except Summer Oligohaline) and therefore the range of mean pH was smaller (8.0 to 8.6) relative to the Tidal Fresh. Highest pH occurred in the Oligohaline (8.69 ± 0.07) and Mesohaline (8.93 ± 0.04) in Spring when CHLa exceeded 60 µg L^{-1}. In the Mesohaline, pH was statistically significantly higher when CHLa exceeded current criteria (insufficient data for Oligohaline analysis). The frequency of daily pH maxima exceeding 9.0 was highest in Spring when CHLa exceeded 60 µg L^{-1} (Oligohaline = 47%, Mesohaline = 70%). There were no instances of daily pH maxima exceeding 9.0 in the Polyhaline, and these instances were less than 20% in the Oligo- and Meso- haline segments when CHLa was less than 60 µg L^{-1}. In Summer, the occurrence of pH exceedance was less than 10% (Oligohaline) or absent (Mesohaline and Polyhaline).

Positive relationships were observed between DO and CHLa in each of the segments (Figure 9). Average DO values typically ranged between 6 and 12 mg L^{-1} over the range of observed CHLa, with lower average DO in Summer relative to Spring. Daily minimum oxygen values (10%-tile) less than 5.0 mg L^{-1} were commonly observed during Summer in the Polyhaline and Mesohaline segments, but were less than 5% (Upper Tidal Fresh) or absent (Oligohaline, Lower Tidal Fresh) in other segments. In the Mesohaline, there was ~20% probability of occurrence of minimum DO < 5 over a range of CHLa values up to 60 µg L^{-1}, and a higher likelihood (~40%) when CHLa exceeded 60 µg L^{-1}. In the Polyhaline, the likelihood of minimum DO < 5 mg L^{-1} was ~20% and did not vary in relation to CHLa. Few instances (<1%) of DO <5 mg L^{-1} were observed in Spring in any of the segments.

The analysis comparing estimates of algal biomass (from counts) to TSS showed that algal contributions to suspended particulate matter were low in the Oligo- and Polyhaline segments, and higher in the Lower Tidal Fresh segment (Figure 10). In the Oligohaline, phytoplankton contributed 4.8% (Spring) and 7.0% (Summer) of suspended matter. In Spring, there was no significant difference ($p = 0.54$) in algal contributions to particulate matter between populations of samples meeting and exceeding current criteria (insufficient data to assess for Summer). In both seasons, the incidence of algae contributing more than 20% of suspended matter was less than 20% across the range of observed CHLa. In the Polyhaline, algal contributions to suspended matter during Summer were low (mean = 9.3%) and the incidence of contributions to suspended matter exceeding 20% was less than 12%. There was no significant difference in algal contributions to particulate matter between populations of samples meeting and exceeding Summer criteria ($p = 0.74$), whereas differences in Spring were significant ($p < 0.0001$). In Spring, algal contributions to suspended matter increased from 6.2 ± 0.7% at CHLa less than 10 µg L^{-1} to 23.5 ± 3.9% at CHLa ranging from 20 to 30 µg L^{-1}. The frequency of occurrence for algae contributing greater than 20% of suspended matter was 42% (20-30 µg CHLa L^{-1}) and 33% (30-60 µg CHLa L^{-1}), but did not exceed 12% when CHLa was less than 20 µg L^{-1}. Algae contributed a greater proportion of suspended matter in the Lower Tidal Fresh segment, particularly during high CHLa in Summer. Highest algal contributions to suspended matter occurred when CHLa exceeded 20 µg L^{-1} (range of means = 27.0 to 32.6%); at lower CHLa, average contributions were less than 12%. The incidence of algae contributing more than 20% of suspended matter was correspondingly higher when CHLa exceeded 20 µg L^{-1} (42 to 67%), relative to CHLa conditions below 20 µg L^{-1} (<20% occurrence). For the Tidal Fresh segment, there was a statistically significant difference in algal contributions to suspended matter in Summer ($p = 0.009$), but not Spring ($p = 0.89$).

Phytoplankton

Analysis of long-term samples from the James showed that mean PIBI scores for the Tidal Fresh Segment declined with increasing CHLa (Figure 11). At the lowest CHLa range (0-10 µg L^{-1}), mean PIBI scores were 3.24 ± 0.13 (Spring) and 3.00 ± 0.23 (Summer) with the majority of samples (80% and 67%, Spring and Summer, respectively) exceeding the threshold for “least degraded” (≥2.67). When CHLa concentrations were in the range of 10-20 µg L^{-1}, mean PIBI scores were lower (Spring = 2.58 ± 0.27; Summer = 2.25 ± 0.53), and the proportion considered “least degraded” declined to 44% (Spring) and 14% (Summer). At CHLa concentrations exceeding 20 µg L^{-1}, mean PIBI scores were consistently low (<2), particularly in Summer, and no samples were designated as “least degraded” (among N=42). A comparison of samples for CHLa concentrations above vs. below the current criteria revealed a statistically significant difference in PIBI for the Lower Tidal Fresh segment in both Spring ($p < 0.0001$) and Summer ($p < 0.0001$). Mean PIBI scores for phytoplankton communities of the Oligohaline segment also declined in relation to CHLa. At the lowest range (0-10 µg CHLa L^{-1}), mean PIBI scores were 2.70 ± 0.18 (Spring) and 2.38 ± 0.12 (Summer). Lower values were observed for CHLa ranges of 10-20 µg L^{-1} (Summer = 1.46 ± 0.11) and 20-30 µg L^{-1} (Spring = 1.82 ± 0.10). Differences between samples above and below current CHLa criteria were statistically significant in Spring ($p = 0.004$), but not Summer ($p = 0.63$). In Spring, the proportion of samples indicating least degraded communities declined from 52% (0-10 µg CHLa L^{-1}), and 47% (10-20 µg CHLa L^{-1}), to 0% (20-30 µg CHLa L^{-1}). In the Polyhaline segment, PIBI scores were similar across the range of observed CHLa in both Spring and Summer. There was no statistically significant difference between samples above and below the current criteria in Spring ($p = 0.79$) or Summer ($p = 0.42$), and the proportion of samples indicating least degraded communities did not decline with increasing CHLa.

A larger dataset which included PIBI scores for Virginia tidal waters outside of the James was examined to assess the responsiveness of the PIBI index to CHLa for a wider range of water quality conditions (data for high salinity CBP stations shown in Figure 12). PIBI scores were binned by CHLa (0-10, etc.) and then further grouped by water quality conditions as: Reference (REF), Near-Reference (MBL), Moderately Degraded (MPL) and Degraded. Water quality designations were based on water clarity (Secchi depth) and nutrient (DIN and PO_4) conditions; REF conditions were indicated by high light (adequate for photosynthesis) and low nutrients. The full range of PIBI scores were observed among communities associated with low CHLa (0-10 µg L^{-1}) in both Spring and Summer. Highest PIBI values were found in the Reference and Near-Reference conditions; lowest PIBI values were found in Degraded conditions. At higher CHLa (>10 µg L^{-1}), Reference and Near-Reference water quality conditions were rare and PIBI scores varied over a smaller, lower range. Mean PIBI scores across all habitat designations declined from 3.23 ± 0.07 to 2.48 ± 0.07 (Spring) and from 3.28 ± 0.06 to 2.24 ± 0.11 (Summer) between the lowest (<5 µg L^{-1}) and highest (>20 µg L^{-1}) CHLa intervals (respectively). A corresponding decline was observed in the proportion of samples representing "least degraded" communities, which decreased from 65% to 29% in Spring and from 77% to 29% in Summer (<5 and >20 µg CHLa L^{-1}, respectively). In REF waters, the overall frequency of PIBI scores indicative of desirable phytoplankton communities (i.e., PIBI≥2.67) was 84% in Spring and 77% in Summer. In DEG waters, the overall frequency of desirable index scores was much lower (19% in both Spring and Summer).

Analysis of 2011-2013 phytoplankton samples from Tidal Fresh segments showed that with higher CHLa, algal species richness increased, while evenness was unchanged (Figure 13). In the Upper Tidal Fresh segment, species richness was significantly higher (Spring p = 0.029; Summer p = 0.038) among samples exceeding current CHLa criteria (Spring = 12.0, Summer = 17.0) relative to samples meeting the current criteria (Spring = 5.9, Summer =7.1). A similar trend was observed in the Lower Tidal Fresh segment during Summer (means = 19.7 and 26.2 for samples meeting and exceeding current criteria, respectively; p = 0.019), whereas no significant difference in species richness was observed in Spring. There were no significant differences in species evenness between samples meeting or exceeding current criteria in the Upper or Lower Tidal Fresh segments during Spring or Summer. In the Mesohaline and Polyhaline segments, increasing CHLa was associated with decreasing species evenness and little change in species richness. In the Mesohaline segment, Spring phytoplankton communities exhibited greatest evenness (mean = 0.72 ± 0.03) when CHLa was lowest (0-10 µg L^{-1}), and declined to 0.43 ± 0.05 among samples in the CHLa range of 10-20 µg L^{-1}. Further increases in CHLa were associated with a progressive decline in evenness (e.g., mean = 0.17 ± 0.02 for CHLa = 60-120 µg L^{-1}). Summer Mesohaline communities exhibited high evenness at CHLa concentrations up to 30 µg L^{-1} (range of means = 0.63 to 0.69), but declined thereafter (to 0.41 ± 0.06 at CHLa = 60-120 µg L^{-1}). Differences in evenness were statistically significant in both Spring and Summer ($p \leq 0.001$). Lowest species richness was observed among samples within the lowest CHLa grouping (Spring = 6.93 ± 0.44, Summer = 5.66 ± 0.30). In Summer, greater species richness was observed at higher CHLa (range of means = 8.00 to 10.88). Differences in species richness between samples meeting and exceeding current criteria were statistically significant in Summer ($p < 0.001$), but not Spring. A similar pattern of declining evenness with little change in species richness was observed in the Polyhaline segment. Highest evenness was observed at low CHLa (Spring = 0.77 ± 0.02, Summer = 0.81 ± 0.02), which decreased to 0.35 ± 0.07 (Spring) and 0.42 ± 0.10 (Summer) in the CHLa range of 20-30 µg L^{-1}. Differences in evenness between samples meeting and exceeding current criteria were statistically significant in both Spring and Summer ($p \leq 0.001$), whereas differences in richness were not. Insufficient samples were collected from the Oligohaline to assess trends in evenness and richness in relation to CHLa.

The occurrence of harmful algae was assessed from cell densities of *Microcystis* and *Cochlodinium*, as well as concentrations of the cyanotoxin microcystin (Figure 14). Summer microcystin concentrations in the Upper and Lower Tidal Fresh segments were positively related to CHLa. When CHLa concentrations were less than 30 µg L^{-1}, microcystin was typically low (<0.4 µg L^{-1}). Highest microcystin concentrations (mean = 1.24 ± 0.37 µg L^{-1}) were observed in the Lower Tidal Fresh segment when CHLa exceeded 60 µg L^{-1}. Microcystin was not observed at CHLa concentrations below 20 µg L^{-1}, and increased to 50% when CHLa exceeded 60 µg L^{-1}. Differences in microcystin concentrations between samples meeting or exceeding current CHLa criteria were statistically significant in both the Upper and Lower Tidal Fresh segments ($p < 0.001$). *Microcystis* cell densities were positively related to CHLa in both the James-specific (2011-2014) and CBP long-term monitoring (1985-2013; including all VA tidal waters) datasets. In the James, Summer *Microcystis* cell densities in the Lower Tidal Fresh segment were significantly lower (262 ± 171 cells ml^{-1}) when CHLa conditions met the current criteria, relative to average densities when the criteria were exceeded (1,395 ± 171 cells ml^{-1}; $p = 0.002$). A similar pattern was observed in the larger CBP database with significantly higher densities occurring when the James criteria were exceeded (Spring = 837 ± 428 cells ml^{-1}; Summer = 5,627 ± 1471 cells ml^{-1}), relative to average densities when criteria were met (Spring = 53 ± 47 cells ml^{-1}; Summer = 1,101 ± 662 cells ml^{-1}; $p = 0.0002$ and $p = 0.002$ for Spring and Summer, respectively). During the 2011-2014 weekly monitoring, there were no instances of *Microcystis* cell densities exceeding the lowest of the three thresholds of concern (20,000 cells ml^{-1}). In the long-term monitoring record, there were 5 instances of *Microcystis* cell densities exceeding 20,000 cells ml^{-1}, 3 of which occurred in the James (all prior to 2000). There were no instances of *Microcystis* cell densities exceeding the 78,600 or 100,000 cells ml^{-1} thresholds in either dataset. Given the rarity of exceedance, we did not further consider threats to designated uses based on *Microcystis* cell densities.

Cochlodinium exhibited a pattern of increasing cell densities in relation to CHLa in the Meso- and Polyhaline segments. In the Mesohaline segment, mean *Cochlodinium* densities were low (<500 cells ml^{-1}) at CHLa ranges up to 30 µg L^{-1}. Mean densities increased to 1,856 (± 617) cells ml^{-1} and 14,042 (± 3,667) cells ml^{-1} among samples in the CHLa ranges of 30-60 and >60 µg L^{-1} (respectively). Exceedances of the 1,000 cells ml^{-1} threshold were not observed when CHLa was less than 10 µg L^{-1}, and there were no instances of exceedance when CHLa was below the current criteria. The frequency of exceedance increased to 47% and 77% when CHLa was in the range of 30-60 and >60 µg L^{-1} (respectively). In the Polyhaline segment, *Cochlodinium* cell densities did not exceed 1000 cells ml^{-1} at CHLa less than 10 µg L^{-1}, but exceeded 1,000 cells ml^{-1} more than 10% of the time among samples in the CHLa ranges of 20-60 µg L^{-1}. At highest CHLa (>60 µg L^{-1}), cell densities in the Polyhaline segment (16,037 ± 2,724 cells ml^{-1}) were similar to those observed in the Mesohaline segment. The incidence of Cochlodinium exceeding 1,000 cells ml^{-1} in the Polyhaline segment was greater than 50% when CHLa exceeded 20 µg L^{-1}. *Cochlodinium* densities in samples from the Meso- and Polyhaline segments were significantly greater when CHLa concentrations were above the summer criteria (10 µg L^{-1}) vs. below ($p < 0.001$). *Cochlodinium* was not observed in either segment during the Spring assessment period.

Threshold Exceedance Analysis

Expected rates of threshold exceedance in relation to mean CHLa were analyzed by season and segment for each metric that exhibited statistically significant differences above and below the current criteria. These were derived from their frequency of threshold exceedance within specified CHLa ranges (Figures 8, 9, 10, 11, 13 and 14) and the frequency of occurrence for the specified CHLa ranges (Figures 6 and 7). The resulting combined probability reflects both the likelihood of exceeding a threshold at a given CHLa, and the likelihood of exceeding that CHLa at a given mean CHLa (Figure 15). The number of metrics that exhibited statistically significant differences above and below the current criteria ranged from 1 to 4 per season-segment combination (Table 5). In two cases (Upper Tidal Fresh – Spring, Oligohaline –

Summer) there were no metrics showing statistically significant differences between pooled observations above and below the current criteria. Patterns in the relationships between expected threshold exceedance with mean CHLa were used to infer the limits of lowest risk, defensible and non-protective ranges (Table 5). Standard errors associated with these mean values were used to assign confidence intervals to thresholds. Standard errors were ~10% for the tidal fresh segments and ~5% in the saline segments (see also Figure 5). The smaller standard errors for the lower segments are due to the larger number of observation generated by the Dataflow sampling.

				CHLa Ranges (µg/L)		
Segment	Season	Criteria	Metric	Least Risk	Defensible	Non-Protecti
UTF	Spring	10	None	NA	NA	NA
	Summer	15	Microcystin	≤12±2	12-21	≥21±2
LTF	Spring	15	PIBI	≤10±1	10-16	≥16±2
			pH	≤16±2	10-19	≥19±2
	Summer	23	PIBI	≤27±2	27-31	≥31±2
			Clarity	≤32±2	32-43	≥43±4
			pH	≤32±2	32-43	≥43±4
			Microcystin	≤32±2	32-43	≥43±4
OH	Spring	15	pH	≤7±0.1	7-18	≥18±0.7
			PIBI	≤7±0.1	7-18	≥18±0.7
	Summer	22	None	NA	NA	NA
MH	Spring	12	pH	≤13±0.3	13-21	≥21±0.6
	Summer	10	DO	≤8±0.3	8-13	≥13±0.2
			Cochlodinium	≤8±0.3	8-13	≥13±0.2
PH	Spring	12	Clarity	≤7±0.1	7-11	≥11±0.2
	Summer	10	*Cochlodinium*	≤8±0.2	8-12	≥12±0.2

Table 5. Indicator metrics and associated CHLa thresholds (±SE) by segment and season.

In the Upper Tidal Fresh segment, the expected occurrence of microcystin concentrations exceeding the 0.8 µg L^{-1} threshold was near zero when mean CHLa concentrations were at or below 12 µg L^{-1}. At mean CHLa concentrations of 21 and 40 µg L^{-1} (as observed in 2009 and 2010), the expected occurrence of microcystin exceeding 0.8 µg L^{-1} was 2.0 and 9.3% (respectively). The current criterion for this season and segment (15 µg L^{-1}) is considered defensible as it falls below the lower limit for the non-protective range (21 µg L^{-1}). At attainment of the existing criterion, the expected occurrence of threshold exceedance would be less than 2%.

In the Lower Tidal Fresh segment, the expected occurrence of PIBI scores indicative of least degraded communities was ~63% when Spring mean CHLa concentrations were ≤10 µg L^{-1}. The occurrence of least degraded communities declined to ~40% when mean CHLa was >16 µg L^{-1}. The current criterion for this season-segment (15 µg L^{-1}) fell within the range that would be considered defensible. In Summer, a similar pattern of decreasing incidence of least degraded PIBI scores with increasing CHLa was observed, though the occurrence of least degraded communities was overall lower. The expected occurrence of PIBI scores indicative of least degraded communities was 12% at the lowest observed mean CHLa (27 µg L^{-1}) and declined to 5% expected occurrence at a mean CHLa of 31 µg L^{-1}. For pH, the expected rate of threshold exceedance was low (<2%) up to a mean CHLa of 32 µg L^{-1}. Therefore the current criterion (23 µg L^{-1}) fell within the range of least risk for daytime pH exceedances. Expected exceedances of the microcystin threshold followed the same pattern with lower rates (8-10%) in the range of mean CHLa from 27 to 32 µg L^{-1} and higher rates (26-47%) when CHLa exceeded 43 µg L^{-1} indicating that the current criterion was in the range of least risk. The same pattern was observed for exceedances of the water clarity threshold, but with an overall higher frequency of occurrence. For mean CHLa in the range from 27 to 43 µg L^{-1}, the expected rate of threshold exceedance was 33 to 44%. At higher mean CHLa (86 and 98 µg L^{-1}), the expected exceedance increased to 65%.

In the Oligohaline segment, there were no metrics demonstrating statistically significant differences in relation to CHLa during Summer, thus precluding an analysis of expected frequency exceedance. In Spring, the expected occurrence of least degraded phytoplankton communities was 52% when mean CHLa was in the range from 3 to 7 µg L^{-1}. The expected occurrence of least degraded communities declined to 48% and 43% when mean CHLa was 11 and 18 µg L^{-1} (respectively). A similar pattern was observed for pH threshold exceedance which was near zero at mean CHLa of 3 to 7 µg L^{-1} and increased to 3% and 7% at mean CHLa of 11 to 18 µg L^{-1} (respectively). The current criterion for this segment (15 µg L^{-1}) was considered defensible based on these two metrics.

In the Mesohaline segment, the expected pH exceedance during Spring was 0-2% over a range of mean CHLa from 5 to 13 µg L^{-1}. At higher mean CHLa (21-23 µg L^{-1}), the expected exceedance was 6-7%. Attainment of the current criterion (12 µg L^{-1}) would result in an expected exceedance rate ~1%, and within the range of lowest risk for daytime pH exceedance. The Summer Mesohaline exhibited expected DO exceedances of 14 to 15% over a range of mean CHLa from 6 to 8 µg L^{-1}. There was a small increase in the expected rate of DO exceedance (16-18%) when mean CHLa was >13 µg L^{-1}. The current criterion (10 µg L^{-1}) was considered defensible, though attainment of mean CHLa values within this range would have only a small effect on the frequency of exceeding the DO threshold. Expected exceedance of the *Cochlodinium* threshold (>1,000 cells mL^{-1}) ranged from 1 to 3% over a range of mean CHLa from 6 to 8 µg L^{-1}. The expected frequency of exceeding the *Cochlodinium* threshold was 7-9% when mean CHLa was >13 µg L^{-1}. The current criterion (10 µg L^{-1}) fell within the defensible range based on this metric.

In the Polyhaline segment, expected exceedances of the clarity threshold were 3-5% over a range of Spring CHLa from 4 to 7 µg L^{-1}. The expected exceedance increased to 11-12% when mean CHLa was 11-14 µg L^{-1}. The upper limit of the defensible range (11.4 ± 0.4 µg L^{-1}) was similar to the current criterion (12 µg L^{-1}). In Summer, the expected frequency of exceeding the *Cochlodinium* threshold was <4% when mean CHLa was in the range of 7-8 µg L^{-1}. The expected frequency of exceedance increased to 12-19% when CHLa ranged from 12 to 16 µg L^{-1}. The current criterion (10 µg L^{-1}) fell within the defensible range (8-12 µg L^{-1}).

Lastly, we considered whether inferences regarding protective and non-protective ranges of CHLa were sensitive to the selection of threshold values used to determine frequencies of threshold exceedance. We evaluated a range of thresholds for three metrics (Microcystin, pH and *Cochlodinium*), including values above and below those used in the threshold exceedance analysis. As expected, the use of more or less stringent thresholds resulted in higher and lower rates of exceedance, but patterns in relation to CHLa were generally similar among the range of thresholds (Figure 16). For Microcystin (Summer Lower Tidal Fresh), the use of a lower threshold (0.4 µg MC L^{-1}) resulted in a higher frequency of threshold exceedance (26-74%) relative to the threshold used to assess protectiveness (0.8 µg MC L^{-1}; range = 8-47%), but with similar patterns in relation to CHLa. A comparable result was obtained for daytime pH exceedances in the Lower Tidal Fresh (Summer). At the protective CHLa threshold (32 µg L^{-1}), pH exceedance values were less than 5% for all pH thresholds (8.8, 9.0, and 9.2). At the non-protective threshold (42 µg CHLa L^{-1}), the frequency of pH exceedance was 20% (pH > 9.0) and 31% (pH > 8.8). At the higher pH threshold (>9.2), expected rates of exceedance were less than 10% over the entire range of observed CHLa (27-98 µg L^{-1}). Corresponding patterns of threshold exceedance were observed across the range of pH thresholds in both the Oligohaline and Mesohaline segments, indicating that delineation of the protective CHLa range was not influenced by the choice of threshold. Analysis of *Cochlodinium* exceedance values for the Mesohaline and Polyhaline segments yielded a similar conclusion, though it was noted that the use of a less stringent threshold (2,000 vs. 1,000 cells mL^{-1}) would have yielded a higher upper limit for the lowest risk CHLa threshold (12 vs. 8 µg L^{-1}, respectively). Overall, these results suggest that inferences regarding protective and non-protective CHLa ranges are robust with respect to selection of metric thresholds used to determine frequencies of threshold exceedance.

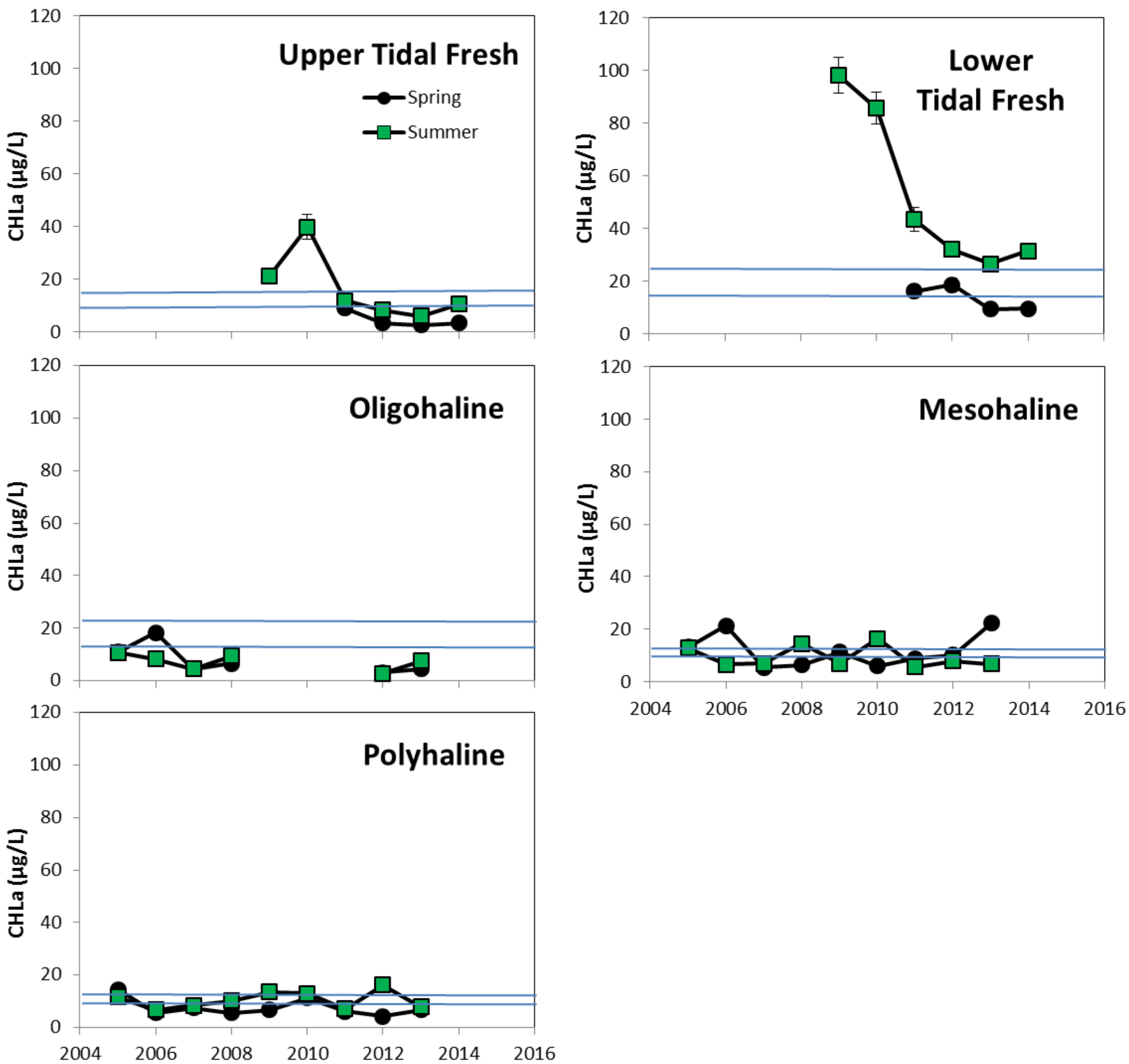


Figure 5. Inter-annual variation in Spring (March-May) and Summer (July-August) CHLa concentrations (mean ± SE) in the five segments of the James River Estuary. Results for the Tidal Fresh segments are based on weekly sampling at 6 stations (upper) and 4 stations (lower). Results for the saline segments are from weekly Dataflow cruises. Horizontal lines denote current CHLa criteria.

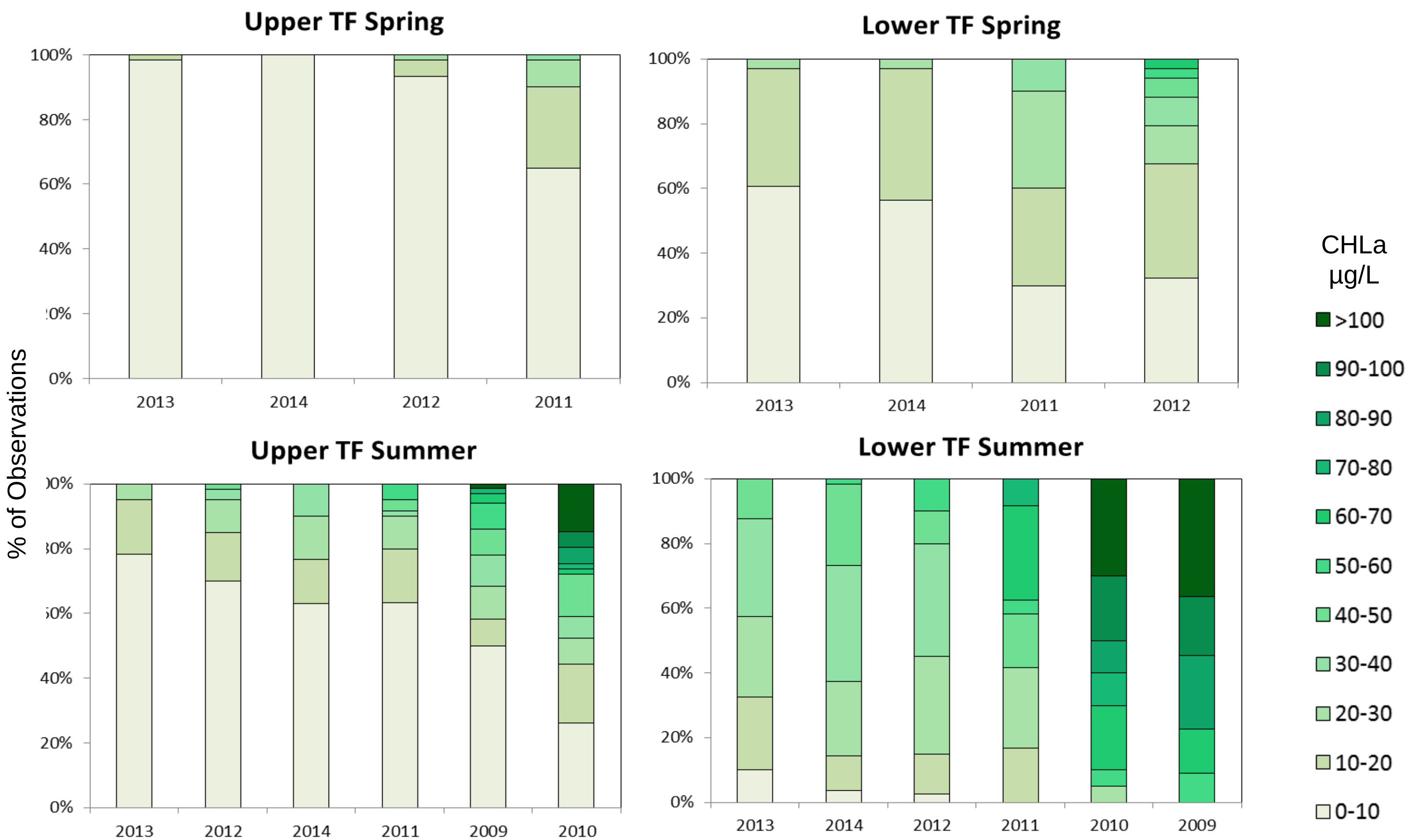


Figure 6. Ranges of CHLa concentrations observed in the Upper and Lower Tidal Fresh segments of the James River Estuary during Spring and Summer. Annual data are shown in order of increasing mean CHLa.

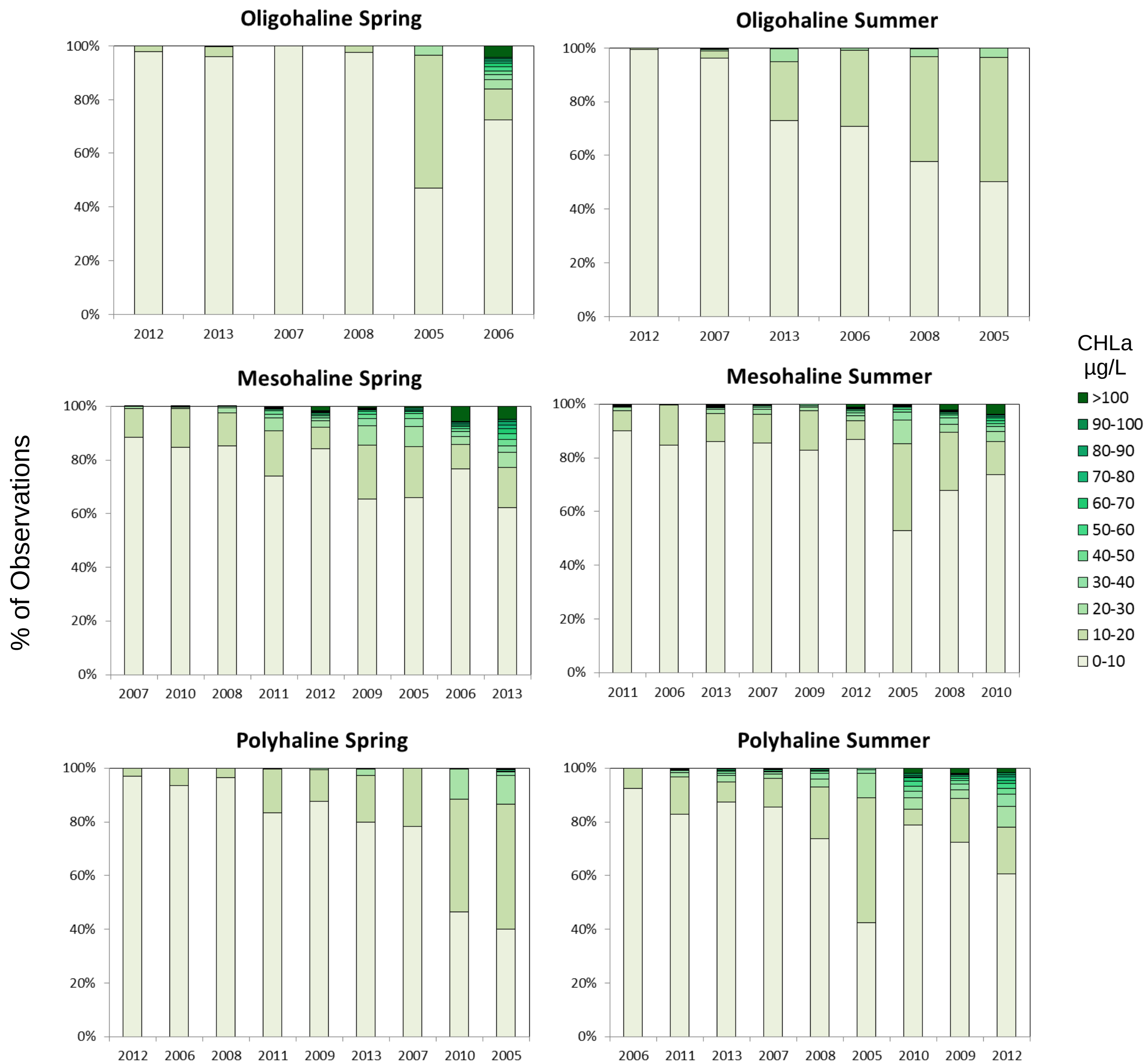


Figure 7. Ranges of CHLa concentrations observed in the Oligo-, Meso- and Poly- haline segments of the James River Estuary during Spring and Summer. Annual data are shown in order of increasing mean CHLa.

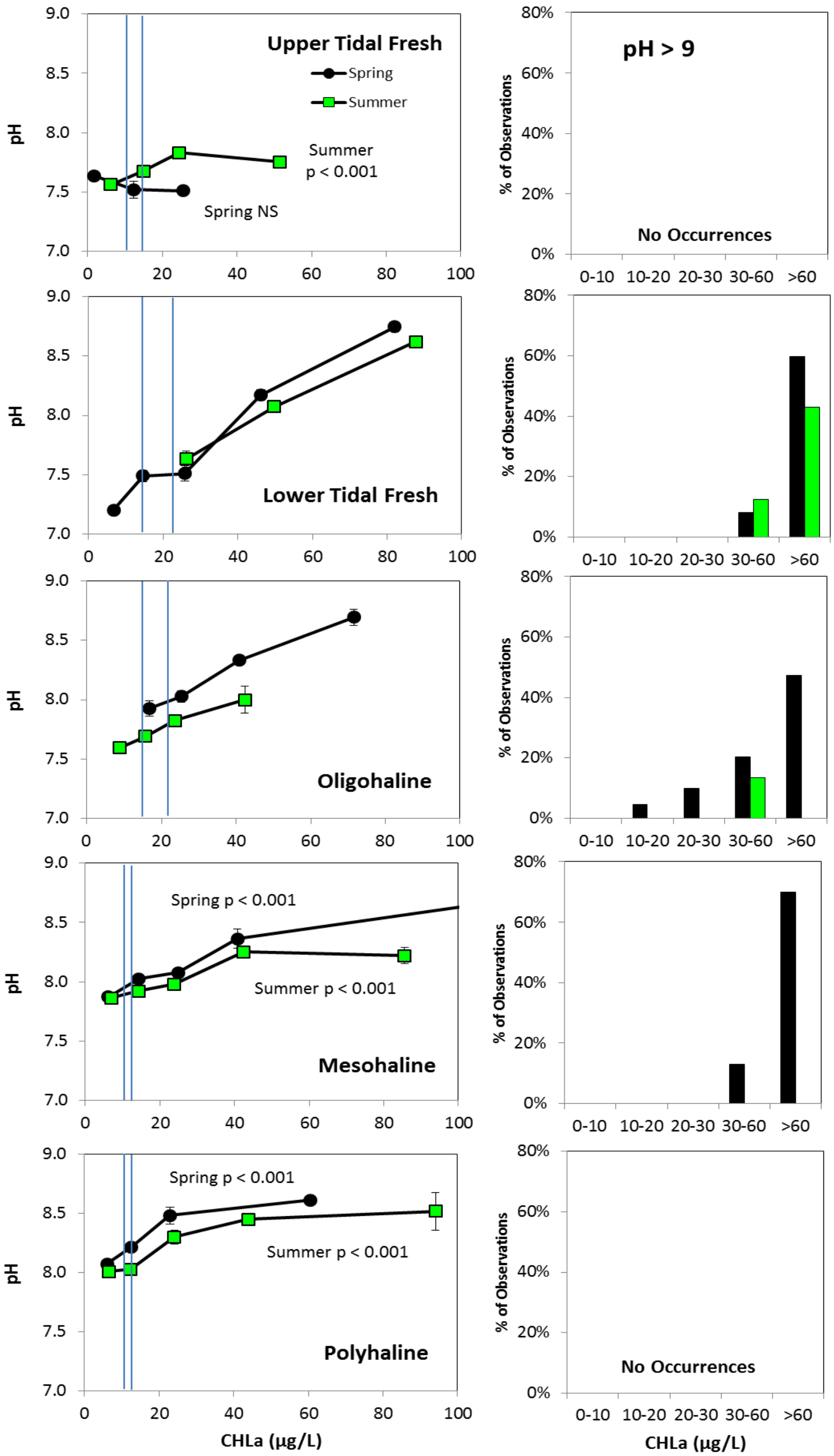


Figure 8. Relationships between pH and CHLa in the James River Estuary based on continuous monitoring data for 2006-2008. Left: means of average daily pH and CHLa (with SE) during Spring and Summer assessment periods. Right: incidence of daily pH maxima (90%-tile) exceeding 9.0. Statistical results are comparisons of mean pH above and below current CHLa criteria (shown as vertical lines for Spring and Summer). No statistics are shown for the Lower Tidal Fresh and Oligohaline segments due to insufficient samples for CHLa meeting current criteria.

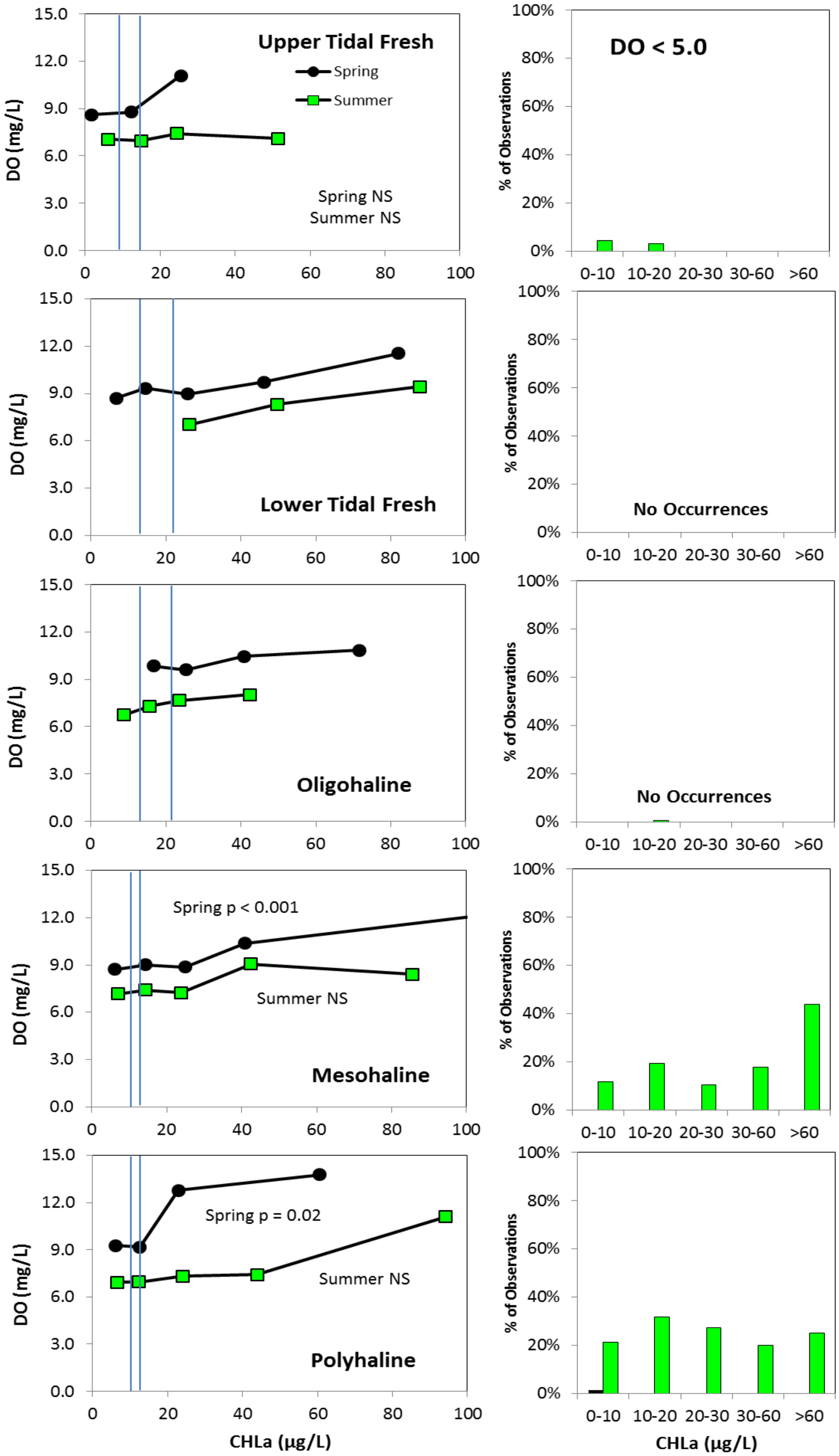


Figure 9. Relationships between DO and CHLa in the James River Estuary based on continuous monitoring data for 2006-2008. Left: means of average daily DO and CHLa (with SE) during Spring and Summer assessment periods. Right: incidence of daily DO minima (10%-tile) less than 5.0 mg/L. Statistical results are comparisons of mean DO above and below current CHLa criteria (shown as vertical lines for Spring and Summer). No statistics for the Lower Tidal Fresh and Oligohaline segments are due to insufficient samples for CHLa meeting current criteria.

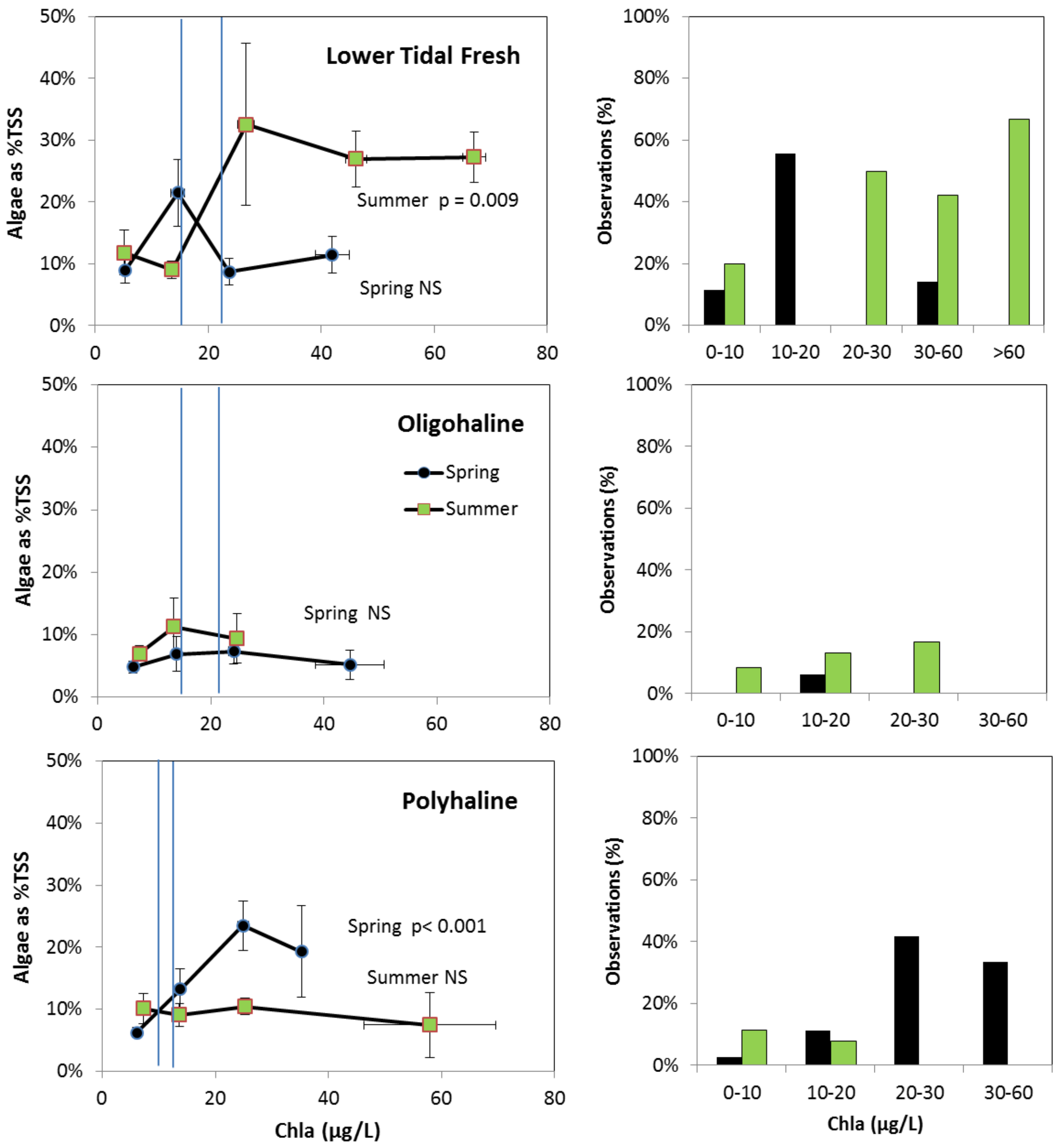


Figure 10. Algal contributions to suspended particulate matter based on measurements of algal biomass and total suspended solids obtained monthly during 1986-2010 at three phytoplankton monitoring stations in the James River Estuary. Left: mean values with SE (some not visible) for Spring and Summer. Right: frequency of occurrence for algal contributions exceeding 20 % of TSS. Statistical results are comparisons of samples above and below current CHLa criteria (shown as vertical lines for Spring and Summer).

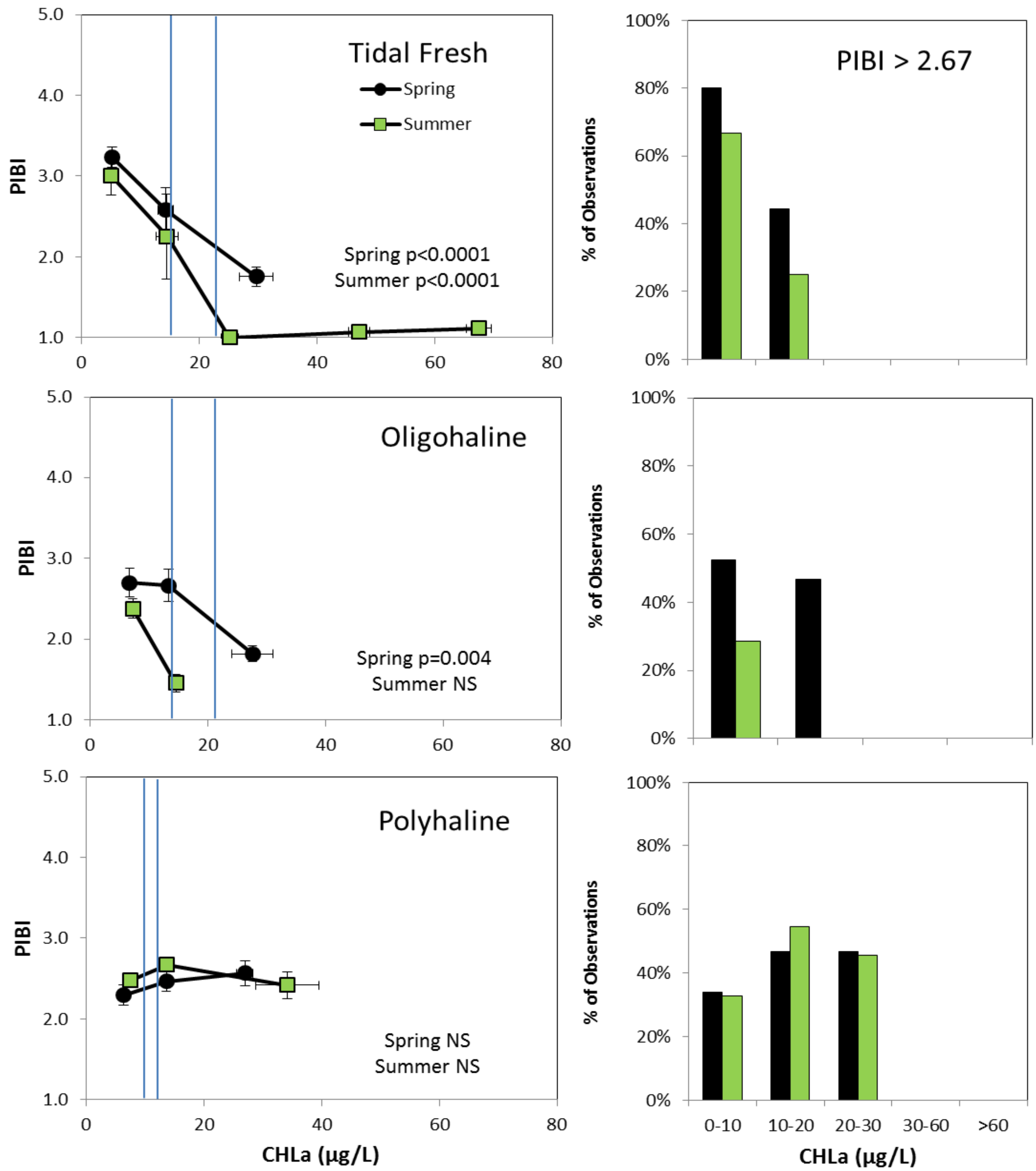


Figure 11. Relationships between PIBI and CHLa in the James River Estuary based on monthly samples collected at three stations (1986-2013). Left: mean values (with SE) during Spring and Summer assessment periods. Right: incidence of PIBI values indicative of least degraded communities (PIBI>2.67). Statistical results are comparisons of samples above and below current CHLa criteria (shown as vertical lines for Spring and Summer).

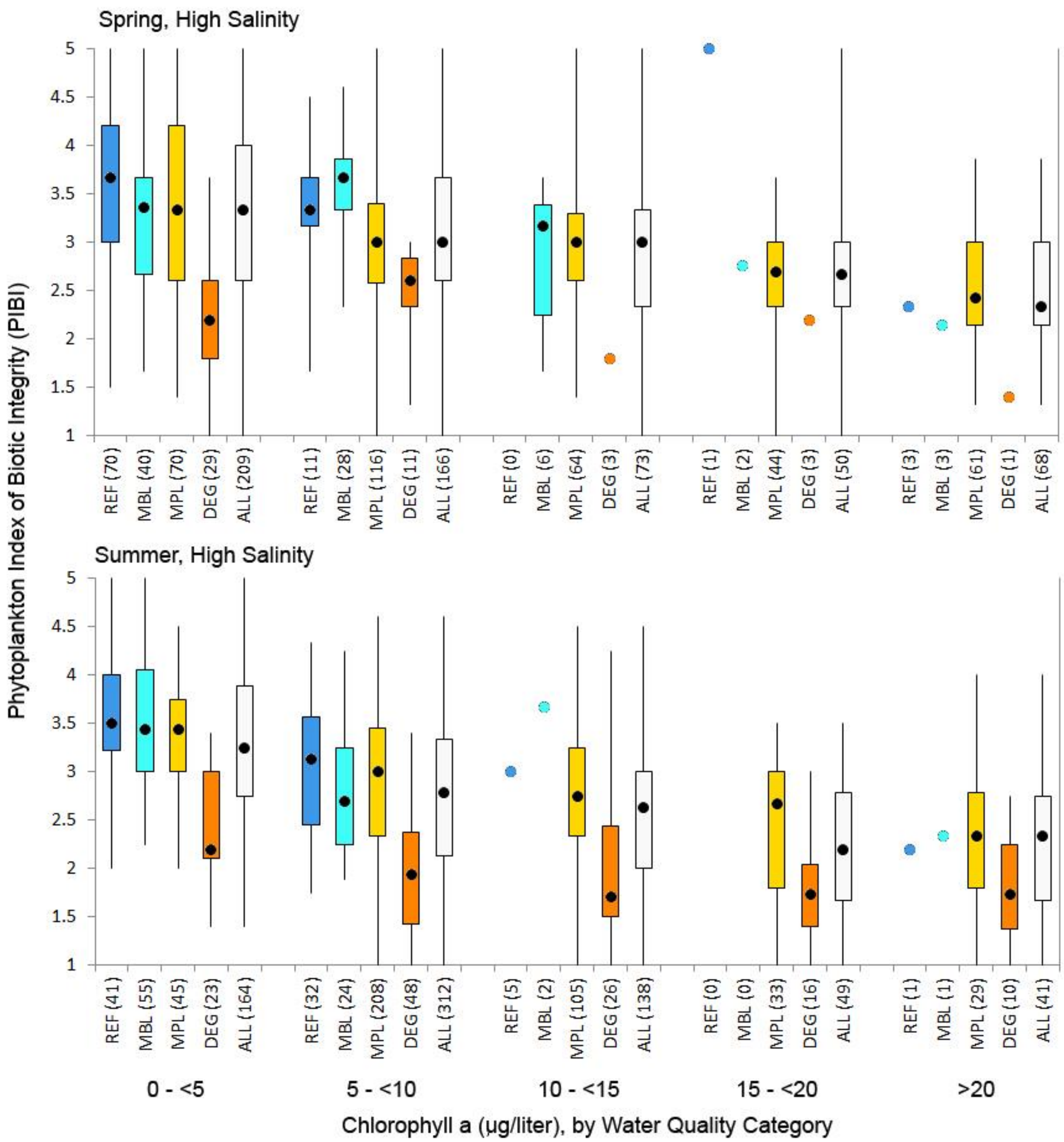


Figure 12. PIBI scores for Spring and Summer phytoplankton communities from high salinity regions (>10 ppt) in Virginia waters of Chesapeake Bay categorized by water quality conditions and CHLa. Water quality conditions were designated as Reference (REF), Near-Reference (MBL), Moderately Degraded (MPL) and Degraded (DEG) based on light and nutrient measurements at the time of sample collection.

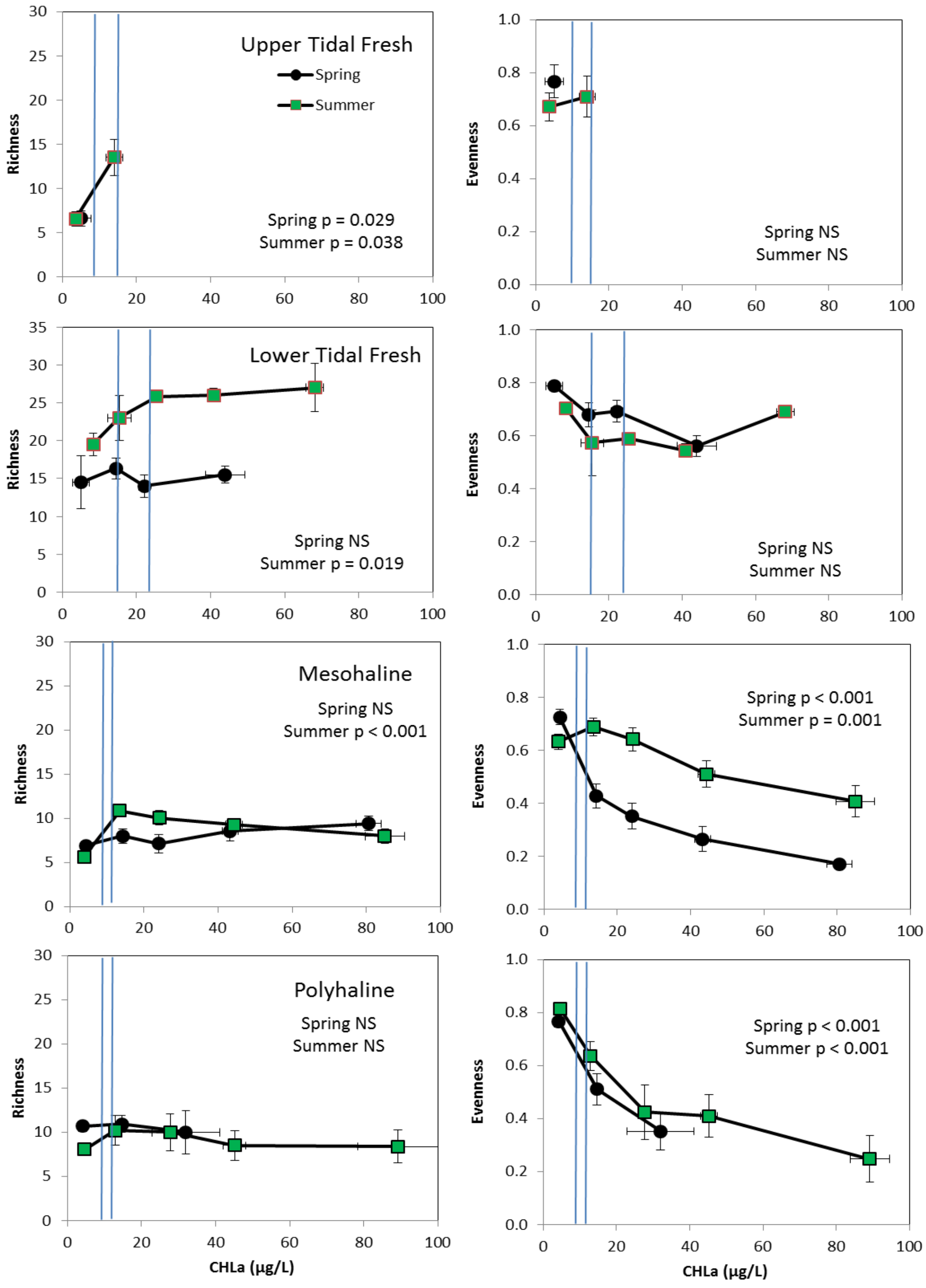


Figure 13. Phytoplankton species richness and evenness in relation to CHLa. Phytoplankton samples were collected during 2010-2013. Statistical results are comparisons of samples above and below current CHLa criteria (shown as vertical lines for Spring and Summer).

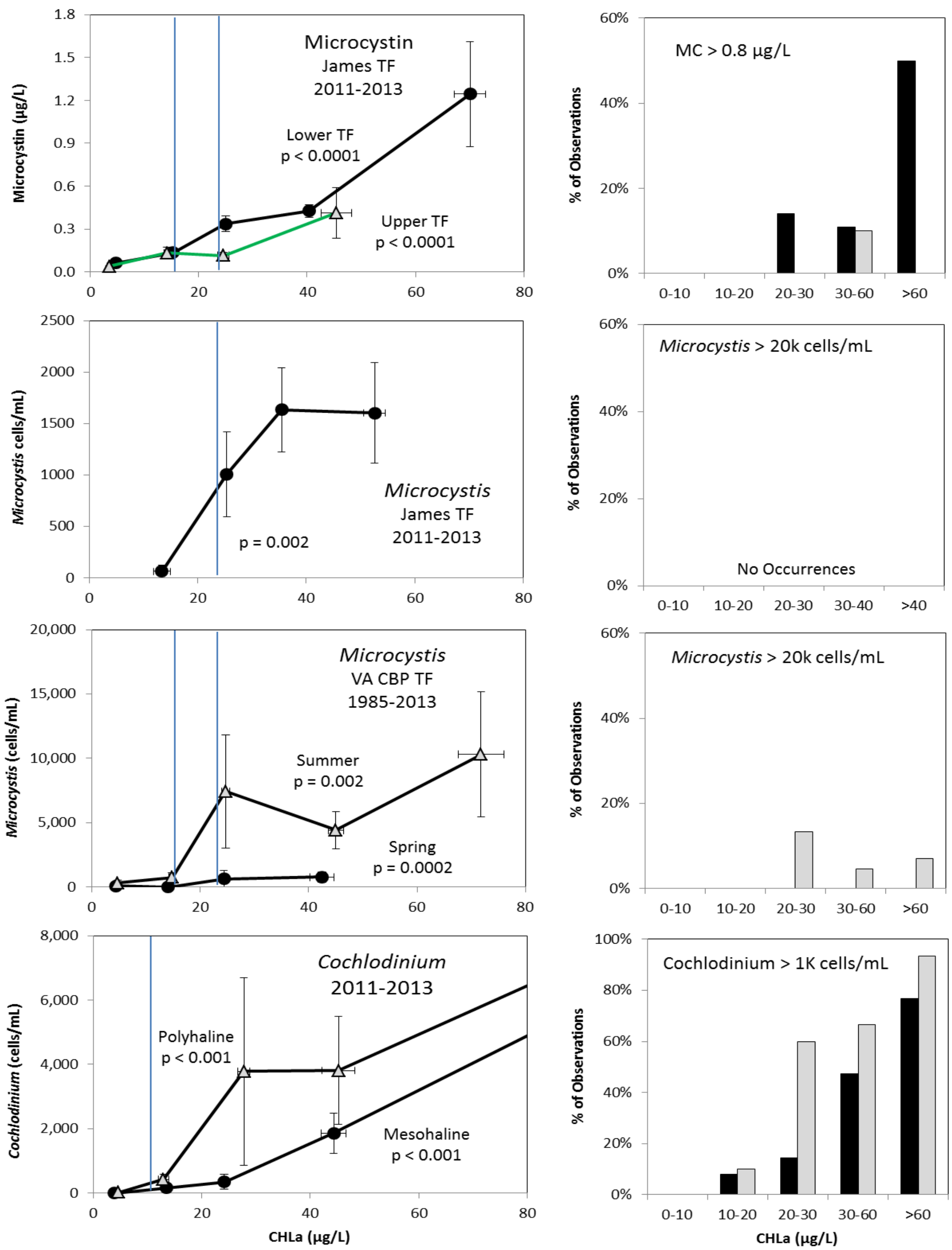


Figure 14. Occurrence of harmful algae (*Microcystis* and *Cochlodinium*) and the cyanotoxin microcystin in relation to CHLa *Microcystis* data are shown separately for the Lower Tidal Fresh segment of the James (2011-2013) and for Virginia tidal fresh waters sampled by the CBP (1985-2013). *Cochlodinium* data are for the Mesohaline and Polyhaline segments of the James during Summer. Statistical results are comparisons of samples above and below CHLa criteria (shown as vertical lines).

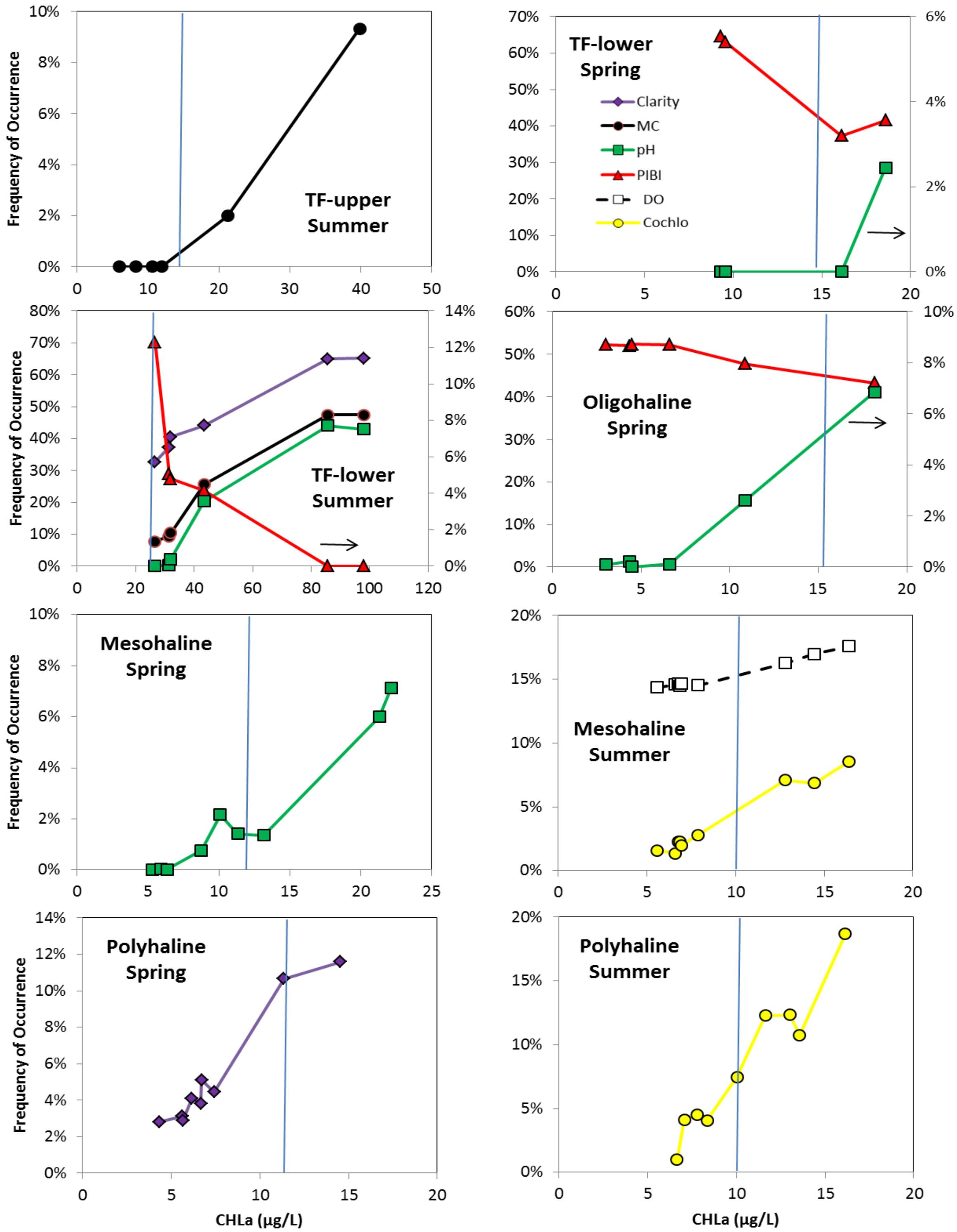


Figure 15. Expected frequency of exceeding thresholds for water clarity, microcystin (MC), pH, PIBI, DO and *Cochlodinium* (Cochlo) in relation to mean CHLa and current CHLa criteria (vertical lines). Arrows denote metrics plotted on secondary axes.

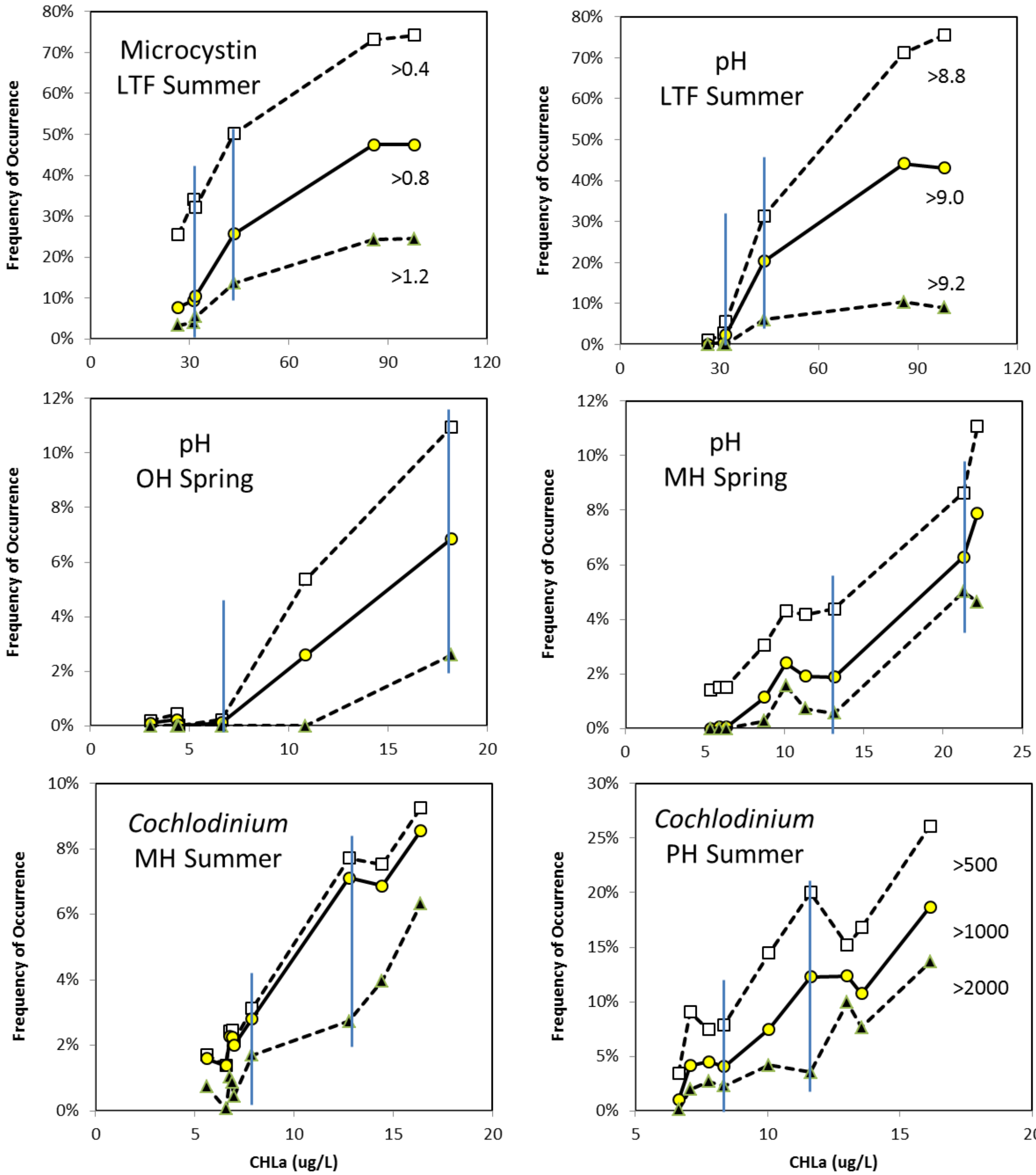


Figure 16. Frequencies of exceedance in relation to mean CHLa for a range of thresholds for Microcystin (0.4, 0.8 and 1.2 µg L^{-1}), pH (8.8, 9.0 and 9.2), and *Cochlodinium* (500, 1000, and 2000 cells ml^{-1}). Vertical lines denote upper limit of lowest risk range (at left) and lower limit of non-protective range (at right) derived from the thresholds used in the effects-based analysis (Microcystin = 0.8 µg L^{-1}; pH = 9.0 and *Cochlodinium* = 1000 cells ml^{-1}).

DISCUSSION

Determining whether CHLa criteria are protective of aquatic life designated uses requires site-specific information characterizing the spatial and temporal occurrence of algal blooms and their deleterious effects on aquatic life. With respect to the former, the James River Estuary is a well-characterized system. The VADEQ has monitored CHLa at 12 stations along the length of the estuary since the 1980's. Their monthly monitoring provides 6-9 measurements of CHLa per segment and season in each year. This long-term effort has been supplemented in recent years with additional sampling at greater spatial and temporal intensity. The weekly monitoring conducted at 12 stations in the two Tidal Fresh segments yields ~80 measurements annually for each season and segment. Spatially continuous monitoring in the Oligo-, Meso- and Poly- haline segments provides ~1000's of CHLa measurements annually per season and segment. A potential limitation of the supplementary CHLa data is that these are surface measurements, which may not reflect depth-integrated concentrations. In the freshwater estuary, positively buoyant cyanobacteria could aggregate near the surface resulting in over-estimation of depth-integrated CHLa. However, as cyanobacteria account for less than 10% of phytoplankton biomass, their effect on vertical CHLa distribution would likely be small. In the lower (saline) segments of the James, algal blooms are dominated by dinoflagellates which are known to have non-uniform vertical distributions. In Monterey Bay, *Cochlodinium* was shown to migrate vertically in the water column on a diel cycle with maximum densities near the water surface during the day, and in deeper waters at night (Kudela et al. 2008). Vertical profiling in the James Estuary indicated heterogeneous vertical CHLa concentrations and evidence of vertical migration (Morse et al. 2011). These findings suggest that surface measurements in the lower James may not be representative of the water column. A project is underway (ODU-HRSD) to collect additional vertical profiles and investigate diurnal patterns during 2015.

Despite the potential limitations of surface measurements, the supplementary CHLa data were a valuable addition to this study in that these allowed for a more robust assessment of how the distribution of CHLa values (as individual observations) varied in relation to the mean. Across seasons, segments and years we observed a consistent pattern whereby the highest CHLa ranges accounted for an increasing proportion of measurements in years with higher mean CHLa (see also Supplemental Information). In addition, the concurrent monitoring of water quality conditions and phytoplankton communities provided an opportunity to test for relationships between algal blooms and adverse effects by more densely populating empirical models, and over a broader range of CHLa, than would be possible from reliance on the long-term monthly data alone. Data for high CHLa conditions (above current criteria) provided a basis for identifying and quantifying deleterious effects (e.g., incidence of high pH, low DO, etc.); while similar observations at low CHLa (below current criteria) were used to determine whether the current criteria were protective against these effects (i.e., to determine the incidence of threshold exceedance at attainment of criteria).

The use of multiple datasets, including long-term monitoring and short-term studies, allowed consideration of a range of metrics and multiple lines of evidence to evaluate the protectiveness of the existing criteria. Metrics included, but were not limited to, those based on phytoplankton community structure (e.g., PIBI, diversity/evenness), as the goal was to consider the broad range of mechanisms by which algal blooms may adversely affect aquatic life designated uses. For the 8 metrics evaluated, statistically significant differences were found between mean values above and below current CHLa criteria for 8 of the 10 season-segment combinations (excluding Upper Tidal Fresh segment in Summer and Summer Oligohaline). The largest number of these (four) occurred in the Lower Tidal Fresh segment during Summer; other segments and seasons were represented by 1-2 metrics. PIBI, pH and harmful algae exceedances were observed in multiple segments and seasons, whereas exceedances for DO and algal contributions to clarity were less common. One of the metrics for which we evaluated relationships to CHLa (species evenness) was not considered in this analysis. Species evenness was found to decline with increasing CHLa in the Mesohaline and Polyhaline segments where algal blooms were dominated by a few species of dinoflagellates. The decline in evenness, if accompanied by a reduction in species richness, could be viewed as an impairment of aquatic life designated uses. However, as declines in evenness were not accompanied by a loss of species richness, we focused instead on metrics related specifically to the occurrence of harmful algae (*Microcystis*, microcystin and *Cochlodinium*).

A combined probability approach was used to derive expected frequencies of threshold exceedance as a function of mean CHLa to determine whether attainment of these criteria would result in low rates of threshold exceedance. In all cases, the likelihood of threshold exceedance increased as a function of CHLa (or, in the case of PIBI, the likelihood of occurrence for least degraded communities decreased). Positive slopes for these relationships reflect an increase in the likelihood of threshold exceedance with rising CHLa, as well as the greater probability of occurrence for high CHLa measurements with increasing mean CHLa. In addition to exhibiting positive slopes, the relationships often showed near-zero intercepts. This indicates that attainment of CHLa conditions at the low end of the observed CHLa range would result in low rates of threshold exceedance for most of these metrics. There were exceptions to this pattern. For example, the incidence of dissolved oxygen minima less than 5 mg L^{-1} ranged from 15% to 18% in the Summer Mesohaline over a span of CHLa from 5 to 15 µg L^{-1}. These cases represent instances where attainment of CHLa criteria would be expected to lessen, but not eliminate threats to aquatic life designated uses as indicated by high rates of threshold exceedance throughout the observed range of CHLa.

Relationships between expected frequencies of threshold exceedance and mean CHLa were used to delineate ranges in which criteria would be protective. In most cases, the criteria were within the defensible range, falling below the lower threshold for non-protectiveness and above the threshold for lowest risk. This finding indicates that attainment of the existing criteria would protect against deleterious effects of algal blooms as indicated by reduced frequency of threshold exceedance. This finding also suggests that lowering the criteria could result in further improvements in water quality and phytoplankton condition, however, in most cases, these improvements would be small. For example, at attainment of the current criterion for the Upper Tidal Fresh during Summer (15 µg L^{-1}), the expected frequency of exceeding the microcystin threshold would be ~2%. Lowering the criterion to the upper threshold of the lowest risk range (12 µg L^{-1}) would reduce the frequency of exceedance to ~0%. In the Oligohaline (Spring), reducing the criterion from its current value (15 µg L^{-1}) to the upper limit of the low risk range (7 µg L^{-1}) would reduce the frequency of pH exceedance from 3% to ~0% and increase the frequency of least impaired phytoplankton communities from 48% to 52%. In these and most other cases, lowering the criteria from their current values to the upper limit of the low risk range would result in small reductions (<5%) in the expected frequency of threshold exceedance. The Polyhaline segment was an exception in that the expected frequency of exceeding the clarity threshold in Spring was 11% at attainment of the current criterion (10 µg L^{-1}), but dropped to 4% at the upper limit of the low risk range (8 µg L^{-1}). The Summer assessment period for the Lower Tidal Fresh was unique in having no mean values lower than the current criterion (23 µg L^{-1}) during the period of enhanced monitoring. Therefore, there is greater uncertainty to assessing protectiveness in the absence of estimates for expected frequencies of threshold exceedance at attainment (see below). However, at the lowest observed CHLa (27 µg L^{-1}), the expected frequency of exceedance was less than 10% for 2 of the 4 metrics (microcystin, pH) and the expected occurrence of least degraded phytoplankton communities (PIBI) was at the maximum of the observed range for this season-segment. Overall, the results do not provide compelling evidence for lowering the existing CHLa criteria based on observable effects on water quality and phytoplankton communities.

The method for judging whether current criteria are protective is based on an assessment of relative risk as indicated by expected frequencies of threshold exceedance in relation to mean CHLa. An alternative approach is to select a fixed rate of exceedance (e.g., 10%) that is considered an unacceptably high risk to aquatic life designated uses. However, there are a number of concerns with using a fixed exceedance value. Foremost among these is that frequencies of exceedance are determined by the selection of thresholds used to characterize risk to aquatic life designated uses (Figure 16). More stringent thresholds result in higher frequencies of exceedance at the same CHLa range. By contract, the relative risk of exceedance was largely unaffected by the selection of more or less stringent thresholds. Thus an assessment of protectiveness based on relative risk was more robust than the quantification of absolute risk. It should be noted that the thresholds selected for the exceedance analysis were not specifically linked to criteria used to assess impairment (e.g., dissolved oxygen criteria for the James). For some of the metrics, these criteria do not exist (e.g., microcystin, *Cochlodinium*), therefore, restricting the analysis to consideration of impairment would not have allowed for a broad assessment of potential risks to aquatic life. In cases where existing criteria were available (pH, DO), the thresholds used for the exceedance analysis were more conservative. For example, we assessed the frequency of occurrence for daily oxygen minima (10%-tile) less than 5 mg L^{-1}, rather than using one of the established criteria (e.g., e.g., instantaneous minimum < 3.2 or 4.3 mg L^{-1}, depending on water temperature) as the established criteria resulted in few instances of occurrence and therefore precluded an assessment of how the risk of low oxygen conditions varied in relation to CHLa. The intent of the frequency of exceedance analysis was to assess threats to aquatic life designated uses associated with elevated CHLa, not to specifically link CHLa to previously-established criteria for impairment.

The SAP considered the strengths and weaknesses of the effects-based approach in the context of prior and concurrent efforts to develop criteria using a reference-based approach (USEPA 2003; Buchanan 2016). A strength of the effects-based approach is that it directly links deleterious effects of algal blooms with threats to aquatic life designated uses based on documented thresholds for adverse impacts. These relationships provide a basis for communicating to stakeholders the benefits of attaining criteria. Reference-based approaches do not quantify these relationships but rather rely on the use of phytoplankton metrics as indicators of biological quality. A limitation of the effects-based approach is that assessment of protectiveness is restricted to those parameters where data were available to test for relationships with CHLa (e.g., dissolved oxygen, clarity, harmful algae, etc.). The metrics used in this analysis are broadly representative of the various direct and indirect effects by which algal blooms are known to cause harmful effects, however, we cannot discount the possibility that deleterious effects of algal blooms extend beyond the range of metrics that are available. Furthermore, it is important to consider that our assessment of protectiveness is based on historical relationships between these metrics and CHLa, and that these are subject to change. For example, if climatic changes were to favor greater dominance by cyanobacteria in the tidal fresh segments, this might result in greater microcystin yield (i.e., per unit CHLa) and more frequent threshold exceedance within a given CHLa range. Similarly, our assessment of HAB threats in the saline segments is based on the abundance and toxicity of *Cochlodinium*, which at present is the dominant bloom-forming dinoflagellate. Its replacement by taxa of greater or lower toxicity would require a revised assessment of protectiveness. Fortunately, most of the metrics used in this analysis are subject to continued monitoring which would allow periodic re-evaluation of protectiveness based on updated relationships between threshold exceedance and CHLa. While recognizing these limitations, the observed relationships between expected rates of threshold exceedance and mean CHLa represent the best currently available science for linking deleterious effects with algal blooms in this system, and for communicating the benefits of CHLa criteria attainment.

A second issue with the effects-based approach is that the analysis of expected threshold exceedances was limited to the range of mean CHLa observed during the period of enhanced monitoring. For most segments this was not a concern as the available data spanned a range of mean CHLa above and below the current criteria, which allowed us to assess protectiveness of the criteria. However, as noted above, we did not observe Summer mean CHLa levels in the Lower Tidal Fresh below the current criteria. Therefore our assessment of protectiveness, which is based on minimizing threshold exceedance to the low end of the observed range, does not consider the potential for further improvements in the range between the lowest observed CHLa (27 µg L^{-1}) and the criterion (23 µg L^{-1}). Reference-based approaches can be used for establishing restoration targets below currently observed conditions, though in the case of the James, these require use of data from outside the system.

We compared results from the effects-based assessment to a reference-based analysis using Bay-wide data (Buchanan 2016). Comparisons between the two approaches were complicated by the fact that different statistical measures were used (see Supplemental Information). Defensible ranges derived by the effects-based analysis are expressed as arithmetic means, whereas the reference-based ranges represent the span of multiple indices of central tendency (medians, arithmetic & geometric means) and multiple measures of reference conditions. In addition, the criteria themselves were derived and originally assessed as arithmetic means, but are assessed as geometric means in the current standards. To consider the implications of using different measures of central tendency, we present the effects-based results as both arithmetic and geometric means (Figure 17; upper and lower panels, respectively). When the current criteria are compared to defensible ranges expressed as arithmetic means they fall within or below the defensible ranges, with one exception (Spring Polyhaline). However, when the current criteria are compared to the defensible ranges expressed as geometric means, they fell within or below the defensible ranges in 5 of 8 season-segment combinations. Thus conclusions about protectiveness are somewhat dependent on whether the defensible ranges are expressed as arithmetic or geometric means.

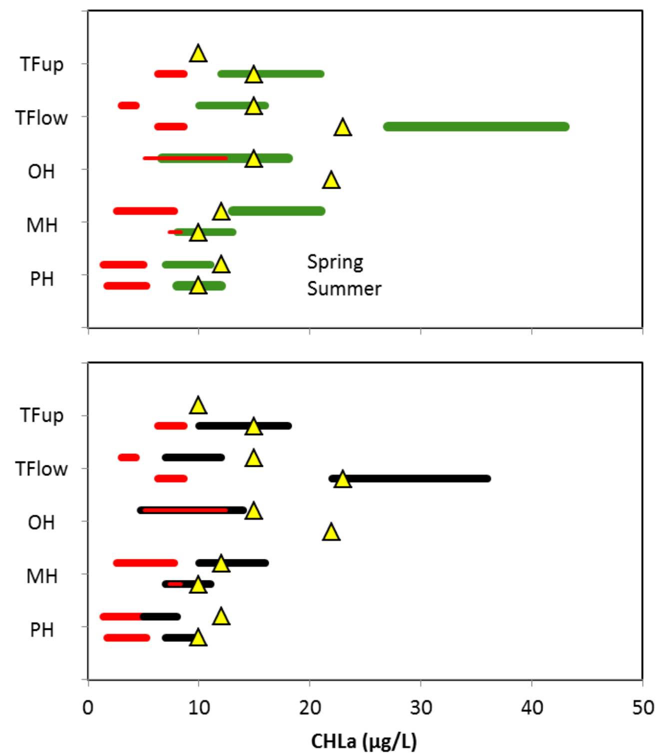


Figure 17. James River CHLa criteria (yellow symbols) shown in relation to protective ranges based on reference conditions (red bars; from Buchanan 2016) and defensible ranges derived from the effects-based analysis. Defensible ranges are shown as arithmetic means (green bars, upper panel) and as geometric means (black bars, lower panel).

In contrast, the reference-based analysis showed that current criteria were not protective in all cases. The reference-based ranges were consistently below those derived by the effects-based approach, with the exception of the Spring-Oligohaline and Mesohaline-Summer, where there was some overlap. The results from these comparisons show that the level of CHLa reduction needed to achieve reference phytoplankton communities is higher than the level of reduction needed to ameliorate observed deleterious effects from algal blooms. Attainment of reference conditions may present a difficult challenge in the short-term because they are dependent on nutrient limited and light-saturated conditions. This level of protectiveness may represent a desirable long-term outcome after addressing the immediate need to mitigate observed harmful effects. Decisions on the use of the reference community as a long-term restoration goal would benefit from monitoring of the degree to which light conditions in the James River estuary respond to reductions in external nutrient and sediment loads.

In summary, this report describes a novel, effects-based approach for examining the protectiveness of James-specific CHLa criteria. The analysis of phytoplankton and water quality data for the James River Estuary showed that elevated CHLa conditions were associated with a range of deleterious effects. Effects were observed in every segment, though in some cases (Upper Tidal Fresh and Oligohaline), not in both seasons of the assessment period. These results support the use of CHLa criteria in the James as a means of protecting aquatic life designated uses. The method used to judge whether the current criteria are protective was based on threats to aquatic life, as represented by relationships between frequencies of threshold exceedance and CHLa. The results of the effects-based analysis suggest that the current criteria are defensible in that they fall below the non-protective range. In most cases, the criteria fall above the upper threshold for low risk indicating that lowering the values of the criteria may result in further improvements in water quality and phytoplankton condition. However in most cases, anticipated reductions in frequency of exceedance at attainment of the low risk threshold were small. The criteria were found to be less protective when interpreted as geometric means, indicating that conclusions regarding protectiveness are somewhat sensitive to the methodology by which attainment of the criteria is determined. The approach builds upon prior and concurrent efforts to establish CHLa thresholds using reference-based and other approaches, and offers the advantages of explicitly linking elevated CHLa to measurable effects and quantifying the benefits of attaining CHLa criteria.

REFERENCES

Buchanan, C., R. Lacouture, H. G. Marshall, M. Olson, and J. M. Johnson. 2005. Phytoplankton reference communities for Chesapeake Bay and its tidal tributaries. Estuaries 28: 138-159.

Buchanan, C. 2016. From programmatic goals to criteria for phytoplankton chlorophyll-a. Interstate Commission on the Potomac River Basin. ICPRB Report 16-1.

Bukaveckas, P. A., L. E. Barry, M. J. Beckwith, V. David, and B. Lederer. 2011. Factors determining the location of the chlorophyll maximum and the fate of algal production within the tidal freshwater James River. Estuaries and Coasts 34: 569-582.

Bukaveckas, P. A., and W. N. Isenberg. 2013. Loading, transformation and retention of nitrogen and phosphorus in the tidal freshwater James River (Virginia). Estuaries and Coasts 36: 1219-1236.

Commonwealth of Virginia. 2010. Chesapeake Bay TMDL Phase I Watershed Implementation Plan. 141 p.

Cooper, S.R. and G.S. Brush. 1993. A 2,500-year history of anoxia and eutrophication in Chesapeake Bay. Estuaries 16:617–626.

Dennison, W.C., R.J. Orth, K.A. Moore, J.C. Stevenson, V. Carter, S. Kollar, P. Bergstrom, and R. Batiuk. 1993. Assessing water quality with submersed aquatic vegetation. Habitat requirements as barometers of Chesapeake Bay health. Bioscience 43: 86–94.

Gallegos, C. L. 2001. Calculating optical water quality targets to restore and protect submersed aquatic vegetation: Overcoming problems in partitioning the diffuse attenuation coefficient for photosynthetically active radiation. Estuaries 24: 381-397.

Gobler, C.J., D.L. Berry, O.R. Anderson, A. Burson, F. Koch, B.S. Rodgers, L.K. Moore et al. 2008. Characterization, dynamics, and ecological impacts of harmful *Cochlodinium polykrikoides* blooms on eastern Long Island, NY, USA. Harmful Algae 7: 293-307.

Harding, L. W., R. A. Batiuk, T. R. Fisher, C. L. Gallegos, T. C. Malone, W. D. Miller, M. R. Mulholland, H. W. Paerl, E. S. Perry, and P. Tango. 2013. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. Estuaries and Coasts DOI 10.1007/s12237-013-9656-6.

Howarth, R.W., D. Anderson, J. Cloern, C. Elfring, C. Hopkinson, B. Lapointe, T. Malone, N. Marcus, K. McGlahery, A. Sharpley, and D. Walker. 2000. Nutrient pollution of coastal rivers, bays and seas. Ecology 7:15.

Johnson, J. M. and C. Buchanan. 2014. Revisiting the Chesapeake Bay phytoplankton index of biotic integrity. Environmental Monitoring and Assessment 186: 1431-1451.

Kemp, W.M., M.T. Brooks, and R.R. Hood. 2001. Nutrient enrichment, habitat variability and trophic transfer efficiency in simple models of pelagic ecosystems. Marine Ecology Progress Series 223:73–87.

Kemp, W.M., R. Batiuk, R. Bartleson, P. Bergstrom, V. Carter, G. Gallegos, W. Hunley, L. Karrh, E. Koch, J. Landwehr, K. Moore, L. Murray, M. Naylor, N. Rybicki, J.C. Stevenson, and D. Wilcox. 2004. Habitat requirements for submerged aquatic vegetation in Chesapeake Bay: Water quality, light regime, and physical-chemical factors. Estuaries 27:263–377.

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, L.W. Harding, Jr., 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.

Kim, C.S., S.G. Lee, C.K. Lee, H.G. Kim, and J. Jung. 1999. Reactive oxygen species as causative agents in the ichthyotoxicity of the red tide dinoflagellate *Cochlodinium polykrikoides*. Journal of Plankton Research 21: 2105-2115.

Kudela, R. M., and C. J. Gobler. 2012. Harmful dinoflagellate blooms caused by Cochlodinium sp.: Global expansion and ecological strategies facilitating bloom formation. Harmful Algae 14: 71-86.

Kudela, R. M., and others. 2008. Linking the physiology and ecology of *Cochlodinium* to better understand harmful algal bloom events: a comparative approach. Harmful Algae 7: 278-292.

Lacouture, R., J. M. Johnson, C. Buchanan, and H. G. Marshall. 2006. Phytoplankton index of biotic integrity for Chesapeake Bay and its tidal tributaries. Estuaries and Coasts 29: 598-616.

Landsberg, J. H., L. J. Flewelling, and J. Naar. 2008. Karenia brevis red tides, brevetoxins in the food web, and impacts on natural resources: Decadal advancements. Harmful Algae 8: 598-607.

Marshall, H. G., L. Burchardt, and R. Lacouture. 2005. A review of phytoplankton composition within Chesapeake Bay and its tidal estuaries. Journal of Plankton Research 27: 1083-1102.

Marshall, H. G., R. Lacouture, C. Buchanan, and J. M. Johnson. 2006. Phytoplankton assemblages associated with water quality and salinity regions in Chesapeake Bay, USA. Estuarine, Coastal and Shelf Science 69: 10-18.

Martinez-Crego, B., T. Alcoverro, and J. Romero. 2010. Biotic indices for assessing the status of coastal waters: a review of strengths and weaknesses. Journal of Environmental Monitoring 12: 1013-1028.

Morse, R. E., J. Shen, J. L. Blanco-Garcia, W. S. Hunley, S. Fentress, M. Wiggins, and M. R. Mulholland. 2011. Environmental and physical controls on the formation and transport of blooms of the dinoflagellate *Cochlodinium polykrikoides* Margalef in the lower Chesapeake Bay and its tributaries. Estuaries and Coasts 34: 1006-1025.

Mulholland, M. R., R. E. Morse, G. E. Boneilo, P. W. Bernhardt, K. C. Filippino, L. A. Procise, J. L. Blanco-Garcia, H. G. Marshall, T. A. Egerton, W. S. Hunley, K. A. Moore, D. L. Berry, and C. J. Gobler. 2009. Understanding causes and impacts of the dinoflagellate, *Cochlodinium polykrikoides*, blooms in the Chesapeake Bay. Estuaries and Coasts 32: 734-747.

O'Neil, J. M., T.W. Davis, M.A. Burford, and C.J. Gobler. 2012. The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. Harmful Algae 14: 313–334.

Orth, R.J. and K.A. Moore. 1983. Chesapeake Bay: An unprecedented decline in submerged aquatic vegetation. Science 222:51–53.

Paerl, H. W. and T.G. Otten. 2013. Harmful cyanobacterial blooms: Causes, consequences, and controls. Microbial Ecology 65: 995–1010.

Poste, A. E., R. E. Hecky, and S. J. Guilidford. 2011. Evaluating microcystin exposure risk through fish consumption. Environmental Science and Technology 45: 5806-5811.

Rastogi, R. P., R. P. Sinha, and A. Incharoensakdi. 2014. The cyanotoxin-microcystins: current overview. Reviews in Environmental Science and Biotechnology 13: 215-249.

Royer, T. V., M. B. David, L. E. Gentry, C. A. Mitchell, K. M. Starks, T. Heatherly, and M. R. Whiles. 2008. Assessment of chlorophyll-a as a criterion for establishing nutrient standards in the streams and rivers of Illinois. Journal of Environmental Quality 37: 437-447.

Shields, E. C., K. A. Moore, and D. B. Parrish. 2012. Influences of salinity and light availability on abundance and distribution of tidal freshwater and oligohaline submersed aquatic vegetation. Estuaries and Coasts 35: 515-526.

Smith, J. L., G. L. Boyer, E. L. Mills, and K. L. Schulz. 2008. Toxicity of Microcystin-LR, a cyanobacterial toxin, to multiple life stages of the burrowing mayfly, Hexagenia, and possible implications for recruitment. Environmental Toxicology 499-506.

Tang, Y.Z. and C.J. Gobler. 2009. Characterization of the toxicity of *Cochlodinium polykrikoides* isolates from Northeast US estuaries to finfish and shellfish. Harmful Algae 8: 454-462.

Turner, J. T. and P. Tester. 1997. Toxic marine phytoplankton, zooplankton grazers and pelagic food webs. Limnology and Oceanography 42: 1203-1214.

USEPA 2003. Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for Chesapeake Bay and Its Tidal Tributaries. EPA 903-R-03-002. Region III Chesapeake Bay Program Office, Annapolis, MD.

USEPA 2007. Ambient water quality criteria for dissolved oxygen, water clarity and Chloriphyll a for the Chesapeake Bay and its tidal tributaries: 2007 chlorophyll criteria addendum. EPA 903-R-07-005 CBP/TRS 288/07. Region III Chesapeake Bay Program Office.

VADEQ 2004. Commonwealth of Virginia Department of Environmental Quality Technical Report on Chl-a Numeric Criteria for the Tidal James River. 55 pp.

VADEQ. 2005. Technical Report – Chlorophyll a Numerical Criteria for the Tidal James River. 57 p. http://www.deq.virginia.gov/Portals/0/DEQ/Water/WaterQualityStandards/JamesRiverChlorophyllStudy/James_chl_a_criteria_tech_doc_2004.pdf.

VADEQ 2014. Commonwealth of Virginia Department of Environmental Quality Technical Report on Concentrations of Microcystin in Fish and Blue Crabs Detected in the James River During 2012.

Wainer, H., M. Gessaroli, and M. Verdi. 2006. Finding what is not there through the unfortunate binning of results: the Mendel effect. Chance, 19: 49 -52.

Van Dolah, E. R., M. Paolisso, K. Sellner and A. Place. 2015. Employing a socio-ecological systems approach to engage harmful algal bloom stakeholders. Aquatic Ecology**,** DOI 10.1007/s10452-015-9562-z.

Wong, W. H., N. N. Rabalais, and R. E. Turner. 2010. Abundance and ecological significance of the clam *Rangia cuneata* (Sowerby, 1831) in the upper Barataria Estuary (Louisiana, USA). Hydrobiologia 651: 305-315.

Wood, J. D., and P. A. Bukaveckas. 2014. Increasing severity of phytoplankton nutrient limitation following reductions in point source inputs to the tidal freshwater segment of the James River Estuary. Estuaries and Coasts 37: 1188-1201.

Wood, J. D., R. B. Franklin, G. C. Garman, S. P. McInich, A. J. Porter, and P. A. Bukaveckas. 2014. Exposure to the cyanotoxin microcystin arising from inter-specific differences in feeding habits among fish and shellfish from the James River Estuary, Virginia. Environmental Science and Technology 48: 5194-5202.

SUPPLEMENTAL INFORMATION

The effects-based approach for assessing protectiveness of the existing CHLa criteria used a combined probability analysis to determine expected rates of threshold exceedances over a range of mean CHLa. We present here some additional explanation (including a sample calculation) and further consideration on various aspects of these computations.

General Approach: Constructing stressor-response curves based on instantaneous measurements is problematic for this application because the criteria are not assessed on an instantaneous basis. For example, modeling *Cochlodinium* densities as a function of instantaneous CHLa might reveal that densities tend to exceed a threshold of concern (e.g., >1000 cells/ml) at an instantaneous CHLa of 15 µg/L. It would be inappropriate to judge protectiveness by comparing this instantaneous CHLa value to the criterion, given that the latter is assessed as a mean. Thus the question arises – what is the likelihood for instantaneous CHLa values to exceed 15 µg/L at a mean CHLa equivalent to the criterion? We can assess the likelihood of occurrence for CHLa values in ranges (e.g., 15-20 µg/L, etc.) and thereby determine their expected frequency of occurrence at any mean CHLa for which we have adequate data to characterize the frequency distribution (i.e., using intensive sampling data generated by VCU and VIMS-HRSD). We can similarly calculate the probability of exceeding the *Cochlodinium* threshold for each of the CHLa bins. By combining these two probabilities we maximize the information contained in the *Cochlodinium*-CHLa relationship and the CHLa distribution to derive an integrated estimate of the expected frequency of threshold exceedance at a given mean CHLa. Provided that we have a range of mean CHLa above and below the criteria (true for all but one segment), we have a basis for assessing the protectiveness of the criteria.

Sample Calculation: The sample calculation is for the expected frequency of occurrence of Microcystin concentrations exceeding the threshold of concern (>0.8 µg MC L^{-1}) during Summer in the Lower Tidal Fresh segment (Figure 14). The expected probability of exceeding the threshold of concern was calculated based on the observed occurrence of threshold exceedances across CHLa bins (e.g., 0-10, 10-20, etc. µg L^{-1}) and the observed distribution of CHLa values in years for which intensive monitoring data were available. The combined probability may be expressed as:

$$\mathrm{p(TE_{year})} = \sum_{x=1}^{n} \mathrm{p(MCx)} * \mathrm{p(CHLx, year)}$$

Where $p(TE_{year})$ is the probability of threshold exceedance in a given year, *p(MCx)* is the probability of exceeding the Microcystin threshold in a given CHLa bin (*x*), $p(CHL_{x,year})$ is the probability of occurrence of CHLa values in that bin for a given year, and *n* is the total number of bins. The first array of probabilities *[p(MCx])* is not year-specific as this is derived by pooling Microcystin measurements across years. For this analysis, a total of 228 measurements of Microcystin were made in the Lower Tidal Fresh segment during 2011-2014. There were 25 instances of Microcystin concentrations exceeding the threshold of concern. The frequency of occurrence *[p(MCx])* increased with higher CHLa (Col D below). Multiplying these individual probabilities (for each CHLa bin) by the corresponding probability of occurrence for CHLa values in that bin ($[p(CHL_{x,year})]$; Cols E, F and G) yields the combined probability estimates ($[p(TE_{year})]$; Cols H, I and J) which are summed by year. This analysis was performed using data for 7 summers during which weekly sampling of CHLa was conducted (3 of the 7 years shown below for illustration).

[A]	[B]	[C]	[D]	[E]	[F]	[G]	[H]	[I]	[J]
				CHLa Mean & Distribution			Combined Probability		
		MC>0.8 µg/L		2013	2011	2010	2013	2011	2010
CHLa Bin µg/L	Total N Obs	N Obs	% Obs	26.6	43.4	85.7	26.6	43.4	85.7
0-10	35	0	0%	10%	0%	0%	0%	0%	0%
10-20	38	0	0%	23%	17%	0%	0%	0%	0%
20-30	57	8	14%	25%	25%	5%	4%	4%	1%
30-40	46	5	11%	30%	0%	0%	3%	0%	0%
40-50	25	2	8%	13%	17%	0%	1%	1%	0%
50-60	12	2	17%	0%	4%	5%	0%	1%	1%
60-70	11	6	55%	0%	29%	20%	0%	16%	11%
70-80	4	2	50%	0%	8%	70%	0%	4%	35%
Total	228	25					8%	26%	47%

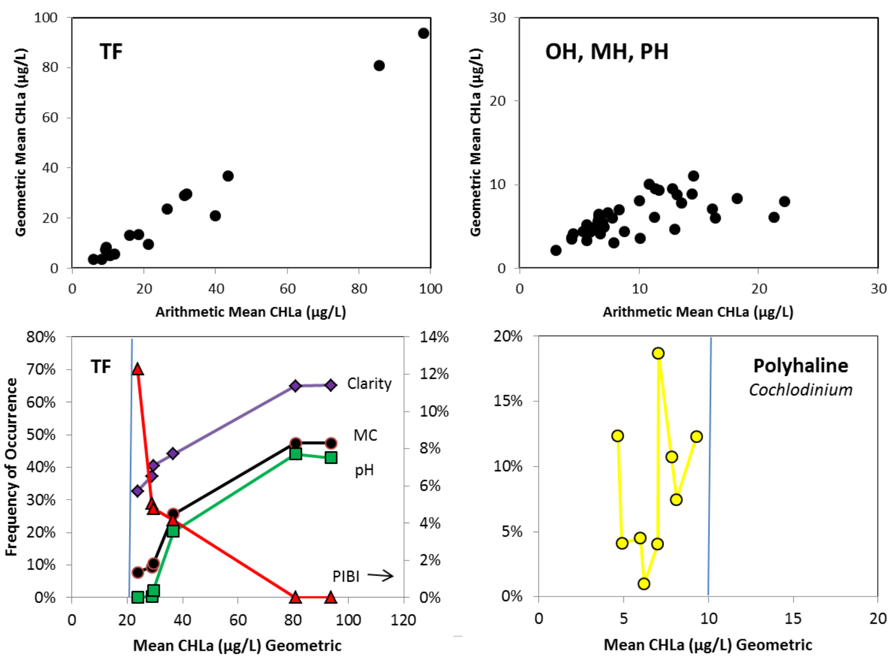
The results show that among three summers ranging from low (2013) to high (2010) mean CHLa, the expected frequency of Microcystin threshold exceedance increased from 8% to 47% (highlighted values). This increase is a result of greater frequency of threshold exceedance when CHLa is high (e.g., exceedance ~50% when CHLa > 60 µg L^{-1}) and because the frequency of occurrence of high CHLa measurements is greater during years with high mean CHLa (e.g., % of measurements > 60 µg L^{-1} = 0% and 90% in 2013 and 2010, respectively). Thus the method explicitly considers both changes in the frequency of threshold exceedance in relation to CHLa, and changes in the distribution of CHLa as a function of the mean CHLa.

Binning of Data: For the combined probability analysis, we elected to use binned data to assess variation in the probability of threshold exceedance over a range of CHLa conditions. Datasets for the various metrics (DO, Microcystin, etc.) provided sample sizes sufficient to populate the bins (ca. 50 to >200 observations distributed across 6-10 bins; Table 3). While there is some loss of information when converting a continuous distribution, in this case CHLa, to bins (rankings), the method is analogous to non-parametric statistics, which are widely used with water quality datasets. There are concerns regarding binning as this may lead to false correlations and can mask the structure of the data when too few bins are used (Wainer et al. 2006). However, we did not feel this was an issue with the present analysis as the correlations were evident in the raw data (Figures 8, 9, 10, 11 and 14), and the analysis used sufficient bins to reveal the pattern over the CHLa range of interest. It was felt that the use of bins provided a relatively simple means of determining the proportion of observations within a given CHLa range that exceeded the specified threshold. Linking the probability of threshold exceedance to CHLa provided a basis for communicating to stakeholders the benefits of incremental reductions in mean CHLa (e.g., as changes in the frequency of occurrence of deleterious water quality conditions).

Use of Models: We explored the utility of fitting models that relate variation in the probability of threshold exceedance to CHLa and using the model outputs to derive the threshold exceedance values. The use of modeled outputs may be advantageous where there is an uneven distribution of observations among the bins leading to greater uncertainty of estimates from poorly populated bins. For example, among three successive bins with high, low and high numbers of observations, the probability of threshold exceedance for the middle (poorly sampled) bin may be more reliably inferred by interpolating between the adjoining bins, rather than using the observed value. We tested a variety of model-fitting procedures on both the primary data (e.g., *Cochlodinium* densities) and probability values (proportion of samples exceeding threshold). The use of LOGIT and LOWESS showed some promise, but not in all cases. Neither linear nor non-linear models provided an adequate fit to the observed patterns in the data. As there were frequent instances of modeled values showing poor correspondence to observed values for bins with adequate sample size, we opted to use the observed values (bin-specific probabilities) in the combined probability analysis. We felt that these provided the best estimate for the probability of exceeding the threshold of concern within a given CHLa bin, and that this approach did not require assumptions about the form of the relationship.

<u>Arithmetic vs. Geometric Means</u>: Expected rates of threshold exceedance were analyzed in relation to mean CHLa as a basis for assessing protectiveness of existing criteria, which are also assessed as mean values. A complicating factor is that the defensible ranges arising from the effects-based analysis were specified as arithmetic means. James River CHLa criteria were derived and originally assessed as arithmetic means, but are now (since 2010) assessed as geometric means. This raised the concern whether a comparison of arithmetic-mean based defensible ranges to criteria assessed as geometric means was appropriate, and more broadly, the question of which measure of central tendency was more appropriate for this analysis. Here we provide some further rationale for the use of arithmetic means in assessing the protectiveness of the CHLa criteria.

The geometric mean is commonly used for CHLa criteria in other states, and is considered a superior measure to the arithmetic mean for estimating central tendency of log-normally distributed variables such as CHLa (USEPA, 2010). By "superior", it may be stated that the geometric mean is a better measure of central tendency because it is more stable (less sensitive) in the presence of extremes (e.g., infrequent, high CHLa associated with bloom events). Conversely, arithmetic means were shown to be a superior indicator of the integrated probabilities of various harmful effects in the James River, because these effects were associated with the frequency and extent of bloom events. A comparison of arithmetic and geometric means for the Tidal Fresh segments showed that these were in good agreement over the observed range of mean CHLa, whereas in the lower segments, the correspondence between the two statistics was poor, particularly at the high end of the range (i.e. arithmetic means > 15 µg L^{-1}; see figure below). As a result, patterns in the relationship between threshold exceedance and geometric means for the tidal fresh were similar to those based on arithmetic means, whereas relationships for saline segments were not. In the saline segments, relationships with arithmetic mean CHLa were in many cases not apparent when depicted in relation to the geometric mean CHLa (e.g., threshold exceedance of *Cochlodinium* in the Polyhaline show below). As the



arithmetic mean is more influenced by the tail of the distribution, and these infrequent, high CHLa values were associated with higher rates of exceedance, the arithmetic mean was found to be more useful in assessing protectiveness.

These findings suggest that the arithmetic mean is superior to the geometric mean for expressing CHLa targets when the management response of interest represents the integrated probability of harmful effects across the range of CHLa concentrations, instead of merely the central tendency of the data. We provide below some further arguments to support this position. These observations are not intended to suggest that the geometric mean CHLa is not a valid target in some applications, or that the 2010 switch to assessing the James River CHLa criteria as geometric means was in error. That decision pre-dated the more recent empirical analysis of integrated probabilities; it was reasonable to interpret the existing criteria as simple measures of central tendency at that time.

The arithmetic mean is more useful than the geometric mean for assessing cumulative or integrated effects, even for log-normally distributed variables. An arithmetic mean is analogous to an integral, in that the arithmetic mean of $y_i \ldots y_n$ is computed as:

$$\frac{1}{n}\sum_{i=n}^{n} y_i$$

As *n* tends to infinity, the arithmetic could be expressed as an integral function:

$$\frac{\int_{x_0}^{x_n} f(x)dx}{(x_n - x_0)}$$

Although the geometric mean may also be expressed as an integral function, it would be best described as a "product integral" or "multiplicative integral". Although product integrals have applications in mathematics, the integral (represented by the arithmetic mean) has more physically-based applications for considering the cumulative effects of functions over time, space, or some other independent variable. This is true even for log-normally distributed variables, as illustrated by the following example:

> *Example*: Consider a situation in which one wanted to estimate an annual load using mean streamflow. Streamflow is log-normally distributed (similar to CHLa), so use of the geometric mean would be superior for measuring the "typical" load delivered on any given day. But use of the geometric mean would underestimate the annual (i.e., cumulative or integrated over time) load, given that a large proportion of the annual load is delivered by a small number of large storm events. The arithmetic mean would provide a superior indication of the total load delivered over time. So the question of which metric is "superior" depends on the application, even for a log-normally distributed variable.

The effects-based analysis empirically links CHLa to probabilities of threshold exceedance that are integrated over time, space, and CHLa range. Much of the SAP's progress to date has been to improve on past methods by directly linking CHLa to harmful effects. CHLa is variable over time and space, and the risk of threshold exceedance is variable over the CHLa range. In order to derive practical targets, the SAP used an empirical approach that integrated these effects by deriving the probability of exceeding an effect threshold as a function of CHLa, and the distribution of CHLa in time and space.

> Let *f(x)* = the "instantaneous" probability of exceeding an effect threshold as a function of a single-sample CHLa concentration, and *g(x)* = the probability distribution of CHLa values over the temporal and/or spatial averaging period of interest. In this case, the mean exceedance rate of the effect threshold can be expressed as a joint probability integrated over the CHLa range of interest:

$$Effect\ threshold\ exceedance\ rate = \int_0^x f(x)g(x)dx$$

For some segment-seasons, it would be possible to use regression or other curve-fitting methods to model the CHLa vs. effect relation as continuous functions. However, the chlorophyll-effects relations in many of the segment-seasons are not amenable to simple line or curve-fitting techniques (see above), and the real-world departures from the fitted curves would be a source of error in the method. Instead, the SAP applied a discrete and data-based version of the integral function above, in which the mean exceedance rates were calculated as:

$$Effect\ threshold\ exceedance\ rate = \sum_1^n p_n * g_n$$

Where:

p_n= the probability of exceeding the effect threshold of a sample in CHLa bin n

g_n = the mean % space that falls in CHLa bin n over the temporal period of interest

The region of central tendency represents CHLa conditions of low probability of threshold exceedance, but relatively high frequency of occurrence. The region containing the tail of the distribution represents CHLa conditions of high probability of threshold exceedance, but relatively low frequency of occurrence. Both regions contribute to the integrated estimate of threshold exceedance. The CHLa range of interest is not limited to the tail of the distribution (blooms) or the central tendency, but extends across the entire range. As the overall risk of threshold exceedance is dependent upon the full range of CHLa conditions, the arithmetic mean is the superior metric. In this context, the arithmetic mean is also more useful than an upper-percentile metric (e.g., 95^{th} percentile) because the integrated probability of threshold exceedance is a function of the entire CHLa range, not just the uppermost conditions. It is also relevant to consider that water quality models used for management are more accurate for predicting seasonal mean conditions than short-term bloom events. This second item is important because of the practical need to establish resulting load allocation as accurately as possible within the constraints of modeling capabilities.

In summary, the arithmetic mean-based analysis described in this report provides an assessment of the risk of threshold exceedance that is integrated over time, space and CHLa range, and which was shown to be a more reliable indicator of threshold exceedance than the geometric mean. Therefore, we feel that the arithmetic mean-based defensible ranges are the appropriate benchmarks for assessing the protectiveness of the current criteria. However, we recognize the complication arising from a misalignment between the statistics used to assess the protectiveness of the criteria (arithmetic means) and those used to assess attainment of the criteria (geometric means).