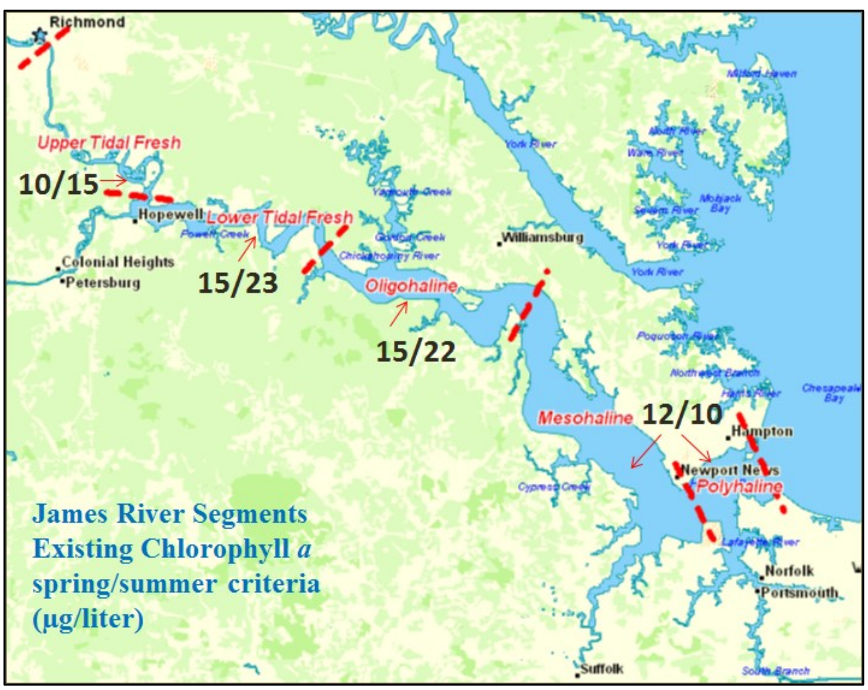
# Biological Reference Curves for Assessing the James River Chlorophyll *a* Criteria

**Final Report**
September 8, 2014

Prepared by
Claire Buchanan, PhD
Interstate Commission on the Potomac River Basin

for

Chesapeake Environmental Communications
and
Virginia Department of Environmental Quality



ICPRB Report 14-3

Interstate Commission on the Potomac River Basin
30 West Gude Dr., Suite 450
Rockville, MD 20850
301-984-1908

## Disclaimer

The opinions expressed in this report are those of the author and should not be construed as representing the opinions or policies of the United States government, or the signatories or Commissioners to the Interstate Commission on the Potomac River Basin.

## Acknowledgements

This project was supported by funds from the Virginia Department of Environmental Quality (VADEQ), by way of Chesapeake Environmental Communications (CEC), and from CWA§106 grant 98339412-0. The Interstate Commission on the Potomac River Basin, an interstate compact river basin commission of the United States government and the compact's signatories: Maryland, Virginia, Pennsylvania, West Virginia, and the District of Columbia.

# Biological Reference Curves for Assessing the James River Chlorophyll *a* Criteria

## Executive Summary

Crucial to the application of any criteria is a reference for measuring criteria attainment. A simple, symmetrical 10% hyperbolic curve is currently the default reference for Virginia's James River chlorophyll *a* criteria. The curve is intended to represent the spatial and temporal exceedances of the criteria in healthy phytoplankton populations. This study compares the 10% hyperbolic curve to "biological reference curves" derived from actual healthy populations living in high quality waters that support Chesapeake Bay open water designated uses. The study then explores some of the statistical properties of the biological curves.

Criteria approximating the upper percentiles ($90^{th}$ – $95^{th}$) of chlorophyll *a* concentrations in high quality waters produce biological reference curves that track the default 10% hyperbolic curve. Four of the ten season- and segment-specific chlorophyll *a* criteria for the James River approximate the $90^{th}$ – $95^{th}$ percentiles. In these cases, the default 10% hyperbolic curve is a good reference for measuring criteria attainment. The other James River criteria are slightly higher than the $95^{th}$%ile or lower than the $90^{th}$%ile, and their biological reference curves depart somewhat from the default curve. Overall, James River criteria are generally protective of high quality environmental conditions in Chesapeake Bay open water designated uses.

When Virginia's current 3-year assessment procedures are applied to the Chesapeake data set of high quality tidal waters, small statistical biases and artifacts inherent to the method can be seen. For example, instances of forced non-compliance may occur due to a statistical artifact ("bottleneck") inherent in the reference curves of assessment units having relatively few interpolator cells. The existing method could be slightly modified or clarified to avoid these issues.

The report's findings suggest the following changes be made to the existing Virginia Department of Environmental Quality procedures for assessing chlorophyll *a* criteria:

1) Continue to use simple, symmetrical hyperbolic curves as default reference curves in chlorophyll *a* criteria assessments, but adjust these curves so that the percent of allowable exceedances in each assessment unit reflects the biological reference curve derived with the James River criteria.

2) Increase spatial and temporal coverage with the use of data from other sampling technologies (e.g., DATAFLOW, satellite imaging, and continuous monitoring buoys) and improve the accuracy of interpolated chlorophyll assessment layers.

3) Given that the existing WQS require seasonal means in Virginia chlorophyll *a* assessments, use a longer assessment period *and* develop reference curves derived from seasonal means. This minimizes the biases introduced into the assessments by too few data layers and seasonal averaging.

4) Develop confidence intervals for the hyperbolic reference curves that account for statistical biases and artifacts inherent to the assessment method. Establish rules for assessing attainment rates that fall outside the reference curves but inside these confidence intervals.

# Introduction

The Commonwealth of Virginia has developed numeric chlorophyll *a* criteria for the tidal James River to use in CWA§303 assessments (VADEQ 2004). The state also adopted a Cumulative Frequency Distribution (CFD) methodology to measure attainment of its chlorophyll *a* criteria (USEPA 2007). Essential to this CFD method are reference curves which delineate allowable exceedances of the criteria in space and time for a given assessment period. CFD curves constructed from actual monitoring data (attainment curves) are then compared to these reference curves to determine whether or not the criteria were met.

A simple, symmetrical 10% hyperbolic curve can be used as a default reference curve for chlorophyll *a*. However, reference curves based on phytoplankton populations inhabiting high quality conditions in the upper, sunlit layer of the water column (reference conditions) would be preferred. "While there is mathematical and statistical logic underpinning the [10% hyperbolic] chlorophyll reference curve, it is important to remember that it is based on parametric models and simplifying assumptions. It is recommended that validation exercises be performed to insure that the general shape of CFD curves generated from data collected in near reference conditions is approximated by the proposed curve (USEPA 2007)." At a minimum, the reference conditions should have water column transparency deep enough to support unstressed photosynthesis in phytoplankton and nutrient concentrations low enough to limit the formation of nuisance algal blooms. Individual water samples exhibiting these conditions are found fairly often in some open water habitats of Chesapeake Bay. The samples represent discrete parcels of water flowing through the estuary. Since phytoplankton cells are short-lived and their populations adapt quickly to their surrounding environment, populations found in reference conditions are believed to represent healthy, desirable communities.

With some care, reference water samples found in the Chesapeake Bay Program (CBP) tidal water quality monitoring database can be used to create biological reference curves for Virginia's chlorophyll *a* criteria. The objective of this analysis is to investigate the properties of these reference curves, and to compare them to the default 10% hyperbolic curve.

# Methods

## Data preparation

In April 2014, D. Jasinski (Chesapeake Environmental Communications) downloaded from the CBP website (www.chesapeakebay.net, Data Library) data for Chesapeake tidal water samples collected above the pycnocline or in the surface layer between 1984 and 2013. Parameters selected for downloading included: salinity, chlorophyll *a*, dissolved inorganic nitrogen (DIN), ortho-phosphate ($PO_4$), Secchi depth, total nitrogen (TN), total phosphorus (TP), dissolved organic carbon (DOC), total organic carbon (TOC), total suspended solids (TSS), dissolved oxygen (DO), water temperature, and pheophytin. Station information included station name, water body name, CBP segment (2003), latitude, and longitude. For each station-date event, multiple measurements of a parameter in the above-pycnocline layer were averaged.

Station total depths obtained from the CBP 2012 Water Quality Users Guide or from M. Mallonee, the Water Quality Data Manager at CBPO, were incorporated. Stations with total depths less than 2 meters were then removed (they are not considered open water environments). Sampling events were assigned to one of five seasons based on date: spring (March – May), June, summer (July – September), autumn (October – November), and winter (December – February). Sampling events were assigned to one of four salinity zones based on the average above-pycnocline salinity measured on the sampling date: tidal fresh (<0.5 ‰), oligohaline (>0.5 - 5.0 ‰), mesohaline (>5.0 - 18.0 ‰), and polyhaline (>18.0 ‰). Sampling events with no Secchi depth or chlorophyll *a* measurement were removed. Events with Secchi depth equal to zero were also removed. Five sampling events with suspicious, inconsistent values for chlorophyll *a* and pheophytin or DIN and $PO_4$ were removed. If three or more sampling events occurred in a 7-day window, all but one of the records was removed to avoid over-weighting measurements from a particular location and sampling time. The QA/QC'ed data set at this point contained 64,200 records. These sample events are considered representative of Chesapeake tidal, open water environments between 1984 and 2013.

Each sampling event was grouped by season and salinity zone and classified into one of four water quality categories using the classification thresholds for Secchi depth, DIN and $PO_4$ developed by Buchanan *et al.* (2005). The classification thresholds and the four water quality categories are described below. Gaps in the nutrient data prevented definite classification of 11,672 records and they were removed, leaving 52,528 records.

## Reference conditions and populations

Chesapeake Bay Program (CBP) partners have qualitatively described in various inter-agency agreements and reports those Bay environments and designated uses they are striving to recover. The environments have nitrogen and phosphorus concentrations low enough to limit the formation of nuisance algal blooms, water column transparency clear enough to promote healthy growth of vascular plant (underwater grasses), and dissolved oxygen levels adequate for fish and bottom-dwelling communities (e.g., CBP 1987, 2000; USEPA 2003). These restoration goals are considered attainable under the present circumstance of a Bay watershed heavily influenced by humans. No longer attainable are the pre-Colonial water quality conditions, when water transparency was much deeper and the dominant primary producers were not planktonic algae (phytoplankton) but rather benthic algae (e.g., Cooper and Brush 1993) and more expansive beds of underwater grasses (e.g., Miller 1986).

Building on earlier research and data analysis results, Buchanan *et al.* (2005) developed quantitative thresholds to classify existing water quality conditions in Chesapeake Bay open waters (**Table 1**) and create distinct water quality categories relevant to phytoplankton (**Table 2**). The data are grouped into habitats defined by season and salinity zone (see above) to minimize the recognized influences of season and salinity on phytoplankton. The nutrient thresholds in **Table 1** are based on nutrient bioassays performed by Fisher and Gustafson (2003). They separate bloom-limiting and excess nutrient concentrations. The Secchi depth thresholds are from an application of the Relative Status Method to data from the 1985 – 1990 (spring and summer) and 1985 – 1999 (autumn and winter) monitoring periods as described in Buchanan *et al.* (2005) and Olson (2009). They generally separate adequate and inadequate water clarity conditions.

Phytoplankton communities in waters meeting all three thresholds (Better/Best water quality category in **Table 2**) are presumed to be the healthiest in the Bay at this time. They have consistently low and less variable total biomass, chlorophyll *a* and pheophytin (another photopigment). Their ratios of chlorophyll to biomass (Chla:C) are also low and less variable, indicating underwater light levels are high enough to avoid stressing cellular photosynthesis pathways. Their populations have relatively stable proportions of taxonomic groups, larger average cell sizes, and low biomasses of key bloom-forming taxa. Finally, median values for total biomass of the phytoplankton size fractions important to grazers (2

– 200 µm) are the same or higher than those in the degraded categories in 12 of the 16 season-salinity habitats, suggesting that the ongoing nutrient reductions will not "starve" grazers in the future.

Phytoplankton populations in the Mixed Better Light (MBL) water quality category (**Table 2**) prove to be good surrogates for those in the Better/Best category. Secchi depths in the MBL category meet their classification criteria but one or both nutrients fail their classification criteria. Phytoplankton photochemical, biomass, and taxonomic metrics in the Better/Best and MBL categories are indistinguishable in most cases (Buchanan *et al.* 2005, Lacouture *et al.* 2006, Johnson and Buchanan 2013). This is true even when samples in the MBL category have excess nitrogen *and* excess phosphorus concentrations. **Figures 1a - e** illustrates the chlorophyll *a* properties of phytoplankton in the Better/Best and MBL categories as compared to those in the degraded categories. Degraded categories have Secchi depths that fail their classification criteria and nutrient concentrations that fail one or both of their classification criteria.

Samples representing the Better/Best category in the 1984 – 2013 timeframe were rare in the tidal fresh and oligohaline and seasonally rare in the mesohaline and polyhaline (**Table 3**). For this reason, MBL populations are used in combination with the Better/Best populations in this analysis to represent reference conditions and develop the biological reference curves. Including the MBL category as a reference water quality condition increases sample numbers in each season and salinity zone and avoids giving unfair latitude to the reference classifications in tidal fresh and oligohaline salinity zones. However, when numbers of Better/Best category samples increase in response to ongoing nutrient and sediment load reductions to tidal waters, chlorophyll *a* concentrations in this category will best represent stable, desirable phytoplankton populations in a recovered Chesapeake Bay.

All open water designated uses appear to be supported in the conditions meeting reference classification thresholds for phytoplankton. Dissolved oxygen concentrations associated with Better/Best and MBL samples meet the Chesapeake Bay 30-day mean criteria for open waters often, i.e., $\geq$ 5.5 mg/liter for 0 – 0.5 ‰ salinities and $\geq$ 5.0 mg/liter for > 5.0 ‰ salinities. In spring, the success rate is more than 99.4% in samples from all salinity zones. In summer, success rates are lowest in mesohaline salinities (82.6%) and highest in oligohaline salinities (96.1%). Dissolved oxygen is $\geq$ 3.0 mg/liter in 99.2% of all samples, the exceptions being 43 mesohaline and 3 polyhaline summer samples. Water clarity criteria do not exist for open water environments of Chesapeake Bay. However, the classification thresholds used to delineate adequate Secchi depth for reference phytoplankton populations (**Table 1**) are approximately the same or higher than the original water clarity requirements for submersed aquatic vegetation, which were 0.8 m in tidal fresh and oligohaline salinities and 1.0 m in mesohaline and polyhaline salinities (Batiuk *et al.* 1992).

Virginia's chlorophyll criteria pertain only to the spring and summer seasons. Therefore, only the spring and summer reference data (9,415 records) were used in the biological reference curve analyses. The occurrence of MBL samples in tidal fresh and oligohaline waters and Better/Best and MBL samples in mesohaline and polyhaline waters is fairly evenly distributed and single stations or small groups of stations do not dominate in reference conditions.

## Criteria attainment curves

Chlorophyll *a* criteria assessment procedures are described in Chapter 5 of USEPA (2008). CFD curves for Virginia James River assessments are currently generated from multiple data sources (routine shipboard sampling, DATAFLOW, calibration data) as follows:

1. Compile and QA/QC data set of chlorophyll *a* values for the 3-year assessment period.

2. Group data by date and segment.
3. Apply the CBP interpolation program and populate an assessment layer for each segment and sampling date with estimated chlorophyll *a* values (an assessment layer for chlorophyll *a* is the grid of surface water quality model cells in a segment).
4. For each interpolation cell, calculate a season-year arithmetic mean, or simple average, across all dates (**Figure 2**).
5. For each cell, determine if the season-year average violates the criteria.
6. Calculate the percent of all cells violating the criteria in the segment.
7. Determine the cumulative probability of the space violation rate (Weibull formula).
8. Construct a CFD.
9. If any point of this CFD crosses the reference curve, the segment is deemed "impaired."

## Reference curves

The simple, symmetric 10% hyperbolic curve used as a default reference curve for chlorophyll *a* is calculated as follows:

$$(x+b) * (y+b) = a$$

where x is %space in violation, y is %time in violation, b = 0.0429945, and a = $b^2 + b$ (CBP 2007). One data analysis objective is to compare this default curve to various biological reference curves.

Biological reference curves should technically be developed in the manner described above in "criteria attainment curves" (CBP 2007). However, reference-quality conditions would have to occur throughout the spring and summer, across entire assessment units (segments) or salinity zones, over multiple years in order to create the necessary assessment layers. Actual reference conditions are sporadic and not widespread in the Bay at this time and sufficient samples to create these layers cannot be found. Two assumptions are made that overcome the lack of coverage and sample density, maintain some degree of year-to-year natural variability found in reference-quality samples, and allow development of biological reference curves:

**Assumption 1** Reference samples collected in a particular habitat *in a given year* are assumed to represent the spatial distribution of chlorophyll *a* concentrations *in a single assessment layer* for that habitat type, regardless of which year the data were collected. For example, chlorophyll *a* concentrations in reference samples collected in mesohaline waters during summer 1992 are assumed to represent the spatial distribution of reference chlorophyll *a* concentrations across a single assessment layer for the summer mesohaline habitat. Likewise, chlorophyll *a* concentrations in reference samples collected in the segment JMSMH in summer 1992 are assumed to represent the spatial distribution of reference chlorophyll *a* concentrations across a single assessment layer for the JMSMH segment.

**Assumption 2** The proportion of chlorophyll *a* concentrations failing a criterion in a particular habitat *in a given year* represents the percent of failures in *a single assessment layer* for that habitat. By extension, the failure rates of a criterion for multiple years are assumed to represent the temporal distribution of failure rates for the corresponding number of assessment layers. For example, 12 annual failure rates calculated from summer mesohaline data correspond to the failure rates of 12 assessment layers for that habitat.

CFDs representing biological reference curves can be constructed when these assumptions are applied to the pool of reference samples and the season-salinity-year results can be used as individual assessment layers. Up to 30 assessment layers can be assembled from the 1984 – 2013 monitoring data for the 8 season- salinity habitats (**Table 4**)

Three approaches were used to develop and investigate the properties of different biological reference curves. The steps involved in creating the different curves are outlined here.

**Method 1** Treat all reference quality sampling events for a given salinity zone, season and year as representing an individual assessment layer and comparable to a layer created in step 3 of "criteria attainment curves" above. The percent of samples failing the season- and salinity-specific criteria of that salinity zone/season/year combination represents the violation rate of a single assessment layer. All years between 1984 and 2013 collected from Chesapeake Bay open water environments and containing more than two reference quality samples are utilized. For each season-salinity habitat:

1. Determine the percent of samples failing the chlorophyll *a* criteria in each assessment layer.
2. Sort the assessment layers from largest to smallest % failure rate (% space violation rate).
3. Determine the cumulative probability of failure (Weibull formula).
4. Construct a CFD curve.

A schematic of this method is shown in **Figure 3**.

**Method 2** Data are treated as in Method 1 but only assessment layers with ten or more reference quality sampling events are considered. Analysis using Method 2 was only done for the summer mesohaline habitat which has the largest numbers of reference quality sampling events per year for all 30 years of the CBP monitoring program. Method 2 was designed to investigate the influence of number of assessment layers on the shape and position of a biological reference curve. For this season-salinity habitat:

1. Assessment layers are randomly selected, with replacement, from the actual pool of 30 assessment layers.
2. Criteria violation rates of these randomly selected layers are used to constructed 50 combinations of 24 layers, 50 combinations of 18 layers, 100 combinations of 15 layers, 100 combinations of 12 layers, 100 combinations of 9 layers, 100 combinations of 6 layers and 100 combinations of 4 layers.
3. Sort each individual combination from largest to smallest failure rate.
4. Calculate the average (mean) failure rate from the 50 or 100 values associated with each rank (% time violation).
5. Determine the cumulative probability of the average failure rate, as well as that of the $10^{th}$%ile and $90^{th}$%ile of the distribution around the average (Weibull formula).
6. Construct a CFD curve.

**Method 3** Data are treated as in Method 1. Analysis using Method 3 was only done for summer mesohaline habitat. Method 3 was designed to investigate the influence of annual averaging (e.g., **Figure 2**) on the shape and position of the CFD curves. The average failure rate (% space violation) of a given rank (% time violation) is assumed to represent the failure rate for one summer month of one year in an assessment period. For this season-salinity habitat:

1. Return to the 100 randomly selected combinations of 15, 12, and 9 assessment layers produced in Method 2 step 2.
2. For each unique combination of 15 assessment layers, calculate an annual average from 3 randomly selected layers, calculate another annual average from 3 more randomly selected layers, and so forth; continue until 5 annual averages have been calculated for that particular combination of 15 assessment layers; repeat for all 100 combinations. The result is 100 series of

5 annual averages.  These are intended to represent possible annual average in a 5-year assessment period.

3. Repeat step 2 on the 100 random selected combinations of 12 assessment layers and construct 100 series of 4 annual averages. These are intended to represent possible annual averages in a 4-year assessment period.
4. Repeat step 2 on the 100 random selected combinations of 9 assessment layers and construct 100 series of 3 annual averages. These are intended to represent possible annual averages in a 3-year assessment period.
5. Sort the annual averages in the 3-, 4-, and 5-year series from largest to smallest failure rate.
6. Calculate the average (mean) failure rate from the 100 values associated with each rank in an assessment period.
7. Determine the cumulative probability of the average failure rate, as well as that of the $10^{th}$%ile and $90^{th}$%ile of the distribution around the average (Weibull formula).
8. Construct a CFD curve for each of the assessment periods.

### Chlorophyll *a* criteria tested

Three sets of chlorophyll *a* criteria were tested (**Table 5)**.  Two are based on the $90^{th}$ and $95^{th}$ percentiles of all reference quality samples (i.e., Better/Best and MBL water quality categories combined) in each season-salinity habitat.  The third set is the existing James River chlorophyll *a* criteria.  The $90^{th}$ile and $95^{th}$ile were chosen as test criteria because the US Environmental Protection Agency recommends an approximately 10% allowable criteria exceedance if a default CFD reference curve is used (USEPA 2003).

The above-pycnocline salinity measured at the time of sampling decides which $90^{th}$ile and $95^{th}$ile criterion applies to a given sampling event, regardless of where the station is located.  This approach differs from CBP procedures which use segment-specific rather than salinity-specific criteria.  The James River segment JMSMH, for example, is designated mesohaline and chlorophyll *a* data collected in that segment are assessed only with the mesohaline criteria of 12 (spring) and 10 (summer). However, the segment's routine monitoring stations (LE5.2, LE5.3) experience salinities between 0.2‰ to 26.4‰.  For the purpose of developing biological reference curves, the James River segment-specific criteria were used as if they were salinity-specific criteria.  The James River polyhaline segment criteria of 12 (spring) and 10 (summer) were applied to all samples in the reference data set associated with >18‰ salinity and not simply to CBP segments designated as polyhaline (PH).  Similarly, James River segment-based criteria for mesohaline, oligohaline, and tidal fresh were applied according to the salinity measured at time of sampling and not simply to CBP segments designated as mesohaline (MH), oligohaline (OH) and tidal fresh (TF), respectively.

## Results

**1. What nutrient concentrations occur in reference water quality conditions?**

Reference-quality waters are defined for this analysis as the Better/Best and MBL water quality categories. The Better/Best category has water clarity adequate for unstressed phytoplankton photosynthesis and nutrient concentrations known to be low enough to limit bloom formation in open water environments (i.e., $\leq$ 0.07 mg DIN/liter, $\leq$ 0.007 mg $PO_4$/liter).  The MBL category, which is used as a surrogate for reference quality conditions to increase sample numbers for the analysis, has adequate water clarity but one or both nutrients are above bloom-limiting concentrations.  MBL is considered reference-quality because its phytoplankton communities are essentially indistinguishable from those in the Better/Best category.

Analysis of nutrient concentrations in the MBL category shows the following:

- when one of the two nutrients is limiting in the MBL category, it is most often $PO_4$ (**Table 6**);
- the exceptions are summer mesohaline and polyhaline which are mostly limited by DIN;
- when both nutrients are present in excess concentrations in the MBL category, $PO_4$ is rarely greater than 10x (4.3% of all samples) and usually less than 5x (80.6% of all samples) the bloom-limiting concentration of 0.007 mg $PO_4$/liter; and
- when both nutrients are present in excess concentrations in the MBL category, DIN is less than 5x the bloom-limiting concentration of 0.07 mg/liter in about 50% of all samples and greater than 10x of the threshold in about 25% of all samples.

The fact that neither DIN nor $PO_4$ is hugely greater than its bloom-limitation threshold when both are present in excess amounts suggests additional phytoplankton growth in these particular MBL samples would have been limited soon by one of the nutrients.

**2. Can biological reference curves be developed from chlorophyll a concentrations observed in reference water quality conditions?** Yes.

**Figures 4a - h** show the biological reference curves produced with Method 1 when the season- and salinity-specific 90$^{th}$%ile and the 95$^{th}$%ile criteria (**Table 5**) are applied to all available reference-quality samples from Chesapeake open water environments. Assessment layers with as few as 3 samples per layer were used in this analysis, so some bias in the CFD curves due to small sample sizes is expected. The CFDs produced with the 90$^{th}$%ile and 95$^{th}$%ile criteria generally follow the default 10% hyperbolic curve. Percentiles *between* the 90$^{th}$%ile and 95$^{th}$%ile produce CFDs more closely overlay the 10% hyperbolic curve (not shown). Limiting the curves to assessment layers with 10 or more samples did not greatly change the general shape or position of most curves, although the bias created by the layers with small sample sizes became more evident.

**3. Do James River chlorophyll criteria produce reference curves that follow the default 10% hyperbolic curve?** Sometimes.

**Figures 4a – h** also show the biological reference curves produced when James River criteria are applied using Method 1. When James River criteria range between the 90$^{th}$%ile and 95$^{th}$%ile criteria, they produced CFDs that closely follow the default 10% hyperbolic curve, i.e., spring tidal fresh (upper, lower), summer tidal fresh (lower), and summer oligohaline. James River criteria for the spring and summer polyhaline are higher than their corresponding 95$^{th}$%ile criteria and their CFD curves fall noticeably below the 10% hyperbolic curve. This suggests the James River polyhaline criteria may be somewhat under-protective of reference conditions. James River criteria for summer tidal fresh (upper), spring oligohaline, and spring and summer mesohaline are to varying degrees lower than their corresponding 90$^{th}$%ile criteria and their CFD curves are noticeably above the 10% hyperbolic curve. These criteria may be somewhat over-protective of reference conditions.

The James River criteria correspond to the following percentiles of the reference chlorophyll data in the corresponding season-salinity zone (asterisk * indicates the CFD curve approximates the 10% hyperbolic curve):

Spring TF upper 89.1$^{st}$%ile*

| | | |
|---|---|---|
| | TF lower | 95.5$^{th}$%ile* |
| | OH | 73.8$^{th}$%ile |
| | MH | 82.8$^{th}$%ile |
| | PH | 97.3$^{rd}$%ile |
| Summer | TF upper | 86.6$^{th}$%ile |
| | TF lower | 94.0$^{th}$%ile* |
| | OH | 94.4$^{th}$%ile* |
| | MH | 81.0$^{nd}$%ile |
| | PH | 97.8$^{th}$%ile |

Hyperbolic curves can be formulated to specifically overlay biological reference curves derived with the existing James River chlorophyll criteria. These hyperbolic curves have percents of allowable exceedances ranging from 26.2% to 2.2% instead of the default allowance of 10%.

**4. Does the number of individual assessment layers affect the shape and position of the biological reference curve?** Yes, when there are fewer than 9 assessment layers.

**Figures 5a – c** show biological reference curves developed from 24, 15, 12, and 9 individual assessment layers and **Figures 6a – c** show biological reference curves developed from 6 and 4 individual layers. Reference data for the summer mesohaline habitat were used in the analysis because this habitat had the most data. (We assume the other season-salinity habitats will behave the same.) All curves were developed using Method 2. The James River (10 µg/liter), 90$^{th}$%ile (11.89 µg/liter) and 95$^{th}$%ile (14.22 µg/liter) criteria were each tested. When 9 or more individual assessment layers are used to create the CFDs, the number of assessment layers does not appear to influence the overall shape and position of the curves relative to the observed 30-layer curve. When fewer than 9 assessment layers are used (i.e., 6 or 4), the points on the CFD curve in the middle and lower right corner begin to pull up and away from the observed 30-layer curve into the non-compliance zone. This is likely an artifact reflecting the statistical limitations of using few assessment layers to construct the CFD curves.

**5. Do annual averages (each calculated from 3 assessment layers) produce biological reference curves comparable to those constructed from individual assessment layers?** To varying degrees, no.

**Figure 7** uses a hypothetical example to illustrate the effect of annual averaging on a CFD curve. In the example, nine assessment layers – one for each month-year – are generated for a 3-year assessment window (e.g., March-Yr1, April-Yr1, May-Yr1, March-Yr2, … May-Yr3) and are used to calculate three annual means. Both CFD curves are plotted in the figure's graph. Distributions of chlorophyll *a* values are typically log-normal meaning the mean value of several measurements is higher than most of the measurements. This is the case in the hypothetical example. Year 3 has the lowest annual mean (14.2%) so it is positioned at %time = 75% on the 3-point, annual mean CFD curve. Two of the three monthly values making up Year 3's average are lower than 14.2%. They are in fact the two lowest monthly values and as such are plotted at %time = 90% and 80% on the monthly, 9-point CFD curve. Their low values coupled with their high positions on the y-axis locate them closer to the reference curve (default 10% hyperbolic curve) than their corresponding annual mean on the 3-point, annual mean CFD curve. Similarly, Year 2 has the highest annual mean (37%) and is positioned at %time = 25% on the 3-point, annual mean CFD curve. Only one of the three monthly values making up Year 2's average is higher than 37% while the other two values are lower. So, Year 2's mean value ends up closer to the reference curve than the monthly, 9-point CFD curve.

This pattern of bias is found in the summer mesohaline biological reference curves developed with 3-, 4- and 5-year annual means (Method 3). When compared to the CFD curves made from the original 9, 12 and 15 assessment layers, the annual mean curves are biased outward toward non-compliance in the upper left "bottleneck" area and biased inward toward compliance in the lower right side of the curve (**Figure 8**). Annual averaging essentially compresses the natural variability seen over time in the data. The bias is an artifact of averaging the individual assessment layers. Its affect diminishes for criteria values closer to the $90^{th}$%ile – $95^{th}$%ile, whose biological reference curves approach the 10% hyperbolic curve (e.g., compare the James River criteria with the $95^{th}$%ile criteria results in summer mesohaline habitats).

**6. Can criteria exceedances rate be affected by the number of interpolator cells used to generate an assessment layer?** Yes.

This can be demonstrated with an example. For a 3-year assessment period, a reference curve "bottleneck" occurs at %space (x-axis) = 1.355% and %time (y-axis) = 75% (**Figure 7**). Assume a segment has 100 interpolator cells and, in the year having the lowest annual average, one cell in one month of the year fails and no cells fail in the other two months.

March of Low Year - 1% (1 of 100 cells fail)
April of Low Year - 0% (0 of 100 cells fail)
May of Low Year – 0% (0 of 100 cells fail)

This produces an annual average failure rate of 0.333% from a total of 300 evaluated cells. While the lowest possible failure rate is 0%, the second lowest possible rate for a 100-cell segment and three monthly interpolations would be 1/300 or 0.333%. This failure rate will be in compliance (not exceed the reference curve) at %time = 75%, which is the y-axis position of the year having the lowest annual average in a 3-year assessment window.

Now, let's say that the segment had 24 interpolator cells and once again one cell in one month of the year fails and no cells fail in the other two months.

March of Low Year – 4.167% (1 of 24 cells fail)
April of Low Year - 0% (0 of 24 cells fail)
May of Low Year – 0% (0 of 24 cells fail)

This produces an annual average failure rate of 1.389% from a total of 72 evaluated cells. This failure rate is positioned to the right of the 3-year assessment reference curve point at %time = 75% and thus is slightly out of compliance. For a 3-year assessment period, if a segment has fewer than 72 interpolator cells evaluated in a season (fewer than 24 cells per month) then the resulting CFD curve is destined to fail at the bottleneck. Likewise, if two cells fail in one month and no cells fail in the other two months, then the season-year having the lowest annual average will fail at the bottleneck if the segment has fewer than 144 cells per season (48 cells per month).

The minimum number of interpolator cells per segment required for the year with the lowest annual average to simply pass through the bottleneck changes as the assessment period lengthens. For a 4-year assessment period, the bottleneck occurs at %time (y-axis) = 80% and corresponds to %space in violation (x-axis) = 1.020%. A minimum of 98 interpolator cells per segment per season (or 33 cells per segment per month) are required to pass the reference curve bottleneck if one cell in any month fails.

For a 5-year assessment period, the bottleneck occurs at %time values = 83.3% and corresponds to %space = 0.818%. A minimum of 123 interpolator cells per segment per season (or 41 cells per segment per month) are required to pass the reference curve bottleneck if one cell in any month fails.

**7. Is there a relationship between mean chlorophyll a and the frequency of criteria exceedance regardless of how the data are grouped?** Yes.

The 1984 – 2013 CBP monitoring data for chlorophyll *a* from all water quality conditions in open water environments were used to address this question. Data were grouped three ways: season – station (a), season – CBP segment – year (b), and season – salinity zone – year (c). Each sample was scored if it exceeded 10, 20, or 30 µg/liter. The means and frequencies of exceedance of the 10, 20, and 30 µg/liter thresholds were then determined for each group. Relationships between the mean and the frequency of exceeding the three thresholds were tight and often nearly identical regardless of how the data were grouped. **Figure 9** shows the relationships across all Chesapeake Bay tidal waters. Coefficient of determination ($r^2$) values ranged between 0.85 and 0.97. **Figure 10** shows the (b) and (c) relationships for just the James River.

## Discussion

Reference water quality conditions as defined for this study have water clarity adequate for unstressed photosynthesis, as indicated by low, stable Chla:C ratios. Reference condition concentrations of two key nutrients, DIN and $PO_4$, can exceed the bioassay thresholds known to limit algal bloom formation (Fisher and Gustafson 2003). However, when this occurs one or the other of these two nutrients is usually limiting, or the quantities of excess $PO_4$ are comparatively low. It is important to recall that these reference conditions represent the best available at this time. Ongoing nutrient and sediment load reductions to tidal waters should begin to increase numbers of samples meeting all three classification criteria. Chlorophyll *a* concentrations in that category (Better/Best) will most accurately represent stable, desirable phytoplankton populations in a recovered Chesapeake Bay.

Biological reference curves for chlorophyll *a* can clearly be developed from phytoplankton populations inhabiting reference quality conditions in open water environments. Criteria values between the 90$^{th}$ and 95$^{th}$ percentiles of chlorophyll *a* concentrations in present-day reference conditions generate CFD curves that closely approximate the default 10% hyperbolic reference curve for the two seasons and four salinity zones assessed by Virginia. This indicates the 10% hyperbolic curve is a reasonable representation of the natural spatial and temporal extent of algal blooms in Chesapeake Bay under reference conditions. Another CFD analysis based on 1960s Chesapeake chlorophyll *a* data shows a similar closeness to the 10% hyperbolic curve (Curve 3 in Figure 4.1, Appendix A, USEPA 2007). The analysis used the observed means and 90$^{th}$%iles of chlorophyll *a* in the different habitats and made certain assumptions about spatial and temporal variances.

In general, all of the existing James River chlorophyll *a* criteria are protective of high quality habitat conditions. James River criteria values approximating the 90$^{th}$ to 95$^{th}$ percentile values in reference populations will produce biological reference curves that closely follow the default 10% hyperbolic curve. James River criteria values lower than the reference 90$^{th}$%ile may be slightly over-protective of reference conditions; criteria values larger than the 95$^{th}$%ile may be slightly under-protective of reference conditions. This is not to say that the James River criteria which differ from the 90$^{th}$ – 95$^{th}$ percentile range are *not* protective against impairment. It only means they are more or less protective of the reference water quality conditions that support the healthiest phytoplankton populations presently found in Bay tidal waters.

Several avoidable biases and artifacts were found in the biological reference curves.

- When just six or four assessment layers are used to construct biological reference curves, points on the middle and lower right portions of the CFD curve tend to shift toward the non-compliance zone. This is in comparison to curves constructed with nine or more layers. It may be advisable for Virginia to use nine or more assessment layers to build attainment CFD curves if a 10% hyperbolic curve serves as the default reference.
- The shape of the biological reference curve changes when it is constructed from 3, 4, or 5 annual averages instead of the 9, 12, or 15 corresponding monthly assessment layers. Points in the upper left region of the 3, 4, and 5 point CFD curves shift toward the non-compliance zone whereas points in the lower right region shift towards the compliance zone.
- The reference curve bottleneck (upper left region of the CFD curve) can force out of compliance segments having few interpolator grids, even if their criteria failure rates are very low. This artifact of the interpolator approach may be avoided if individual (monthly) assessment layers are used in Virginia assessments rather than annual averages.

It is apparent in the results that monitoring data with more intensive spatial and temporal coverage will improve the CFD approach. CFDs constructed from higher spatial and temporal density data will be less subject to bias and artifact. Regardless, chlorophyll *a* criteria somewhere between the overall $90^{th}$ and $95^{th}$ percentiles of each season and salinity zone habitat will most closely parallel the default 10% hyperbolic curve. More stringent selection of the reference quality samples from which the reference curves are developed could further align the biological reference curves and the 10% hyperbolic curve.

Strong relationships occur between the mean chlorophyll *a* concentration and the frequency of exceeding chlorophyll *a* threshold concentration, regardless of how the data are analyzed. The same mean concentration of chlorophyll *a* at a station, in a CBP segment, or in an entire salinity zone moving longitudinally with wind and flow over time appears to have about the same probability of exceeding a specific threshold. Documenting the relationships may appear tangential to the development of biological reference curves. However, a concern was that salinities higher or lower than a segment's designated salinity would affect the segment's exceedance frequencies. Tidal waters are in no way bound to the salinity designations of CBP segments. Severe droughts shift oligohaline waters well into segments designated as tidal fresh and large storms push low salinity waters into segments designated as polyhaline. Another concern was the James River's shallow bathymetry and its possible effect on exceedance frequencies.

Finding the relationships has several implications. First, the strong similarities between relationships in different data groupings indicate it was acceptable to use the James River criteria as salinity-specific in developing and characterizing biological reference curves. When used as salinity-specific criteria, James River criteria are applied according to the salinity observed at the time of sampling. This is how the $90^{th}$%ile, 95%ile, and James River criteria were used to develop the biological reference curves. When used as segment-specific criteria, the criterion applied to a sample is decided by the salinity designation of the sampling station's segment. Second, the relationships might assist in future chlorophyll *a* assessments. Assessment units with questionable exceedance frequencies can be check against the expected frequencies calculated from established relationships between the mean and the frequency of exceeding specific chlorophyll *a* criteria. Third, and perhaps most importantly, the relationships demonstrate that exceedance frequencies are not strongly controlled by salinity or locational features.

This further support the idea that chlorophyll *a* concentrations and bloom frequencies are most strongly controlled by the water quality conditions surrounding the phytoplankton population.

## Recommendations

The report's findings suggest the following changes to the existing Virginia Department of Environmental Quality procedures for assessing chlorophyll *a* criteria:

1) Continue to use simple, symmetrical hyperbolic curves as default reference curves in chlorophyll *a* criteria assessments, but adjust these curves so that the percent of allowable exceedances in each assessment unit reflects the biological reference curve derived with the James River criteria.

2) Increase spatial and temporal coverage with the use of data from other sampling technologies (e.g., DATAFLOW, satellite imaging, and continuous monitoring buoys) and improve the accuracy of interpolated chlorophyll assessment layers.

3) Given that the existing WQS require seasonal means in Virginia chlorophyll a assessments, use a longer assessment period and develop reference curves derived from seasonal means. This minimizes the biases introduced into the assessments by too few data layers and seasonal averaging.

4) Develop confidence intervals for the hyperbolic reference curves that account for statistical biases and artifacts inherent to the assessment method. Establish rules for assessing attainment rates that fall outside the reference curves but inside these confidence intervals.

## Literature cited

Batiuk, R. A., P. Heasley, R. Orth, K. Moore, J. C. Stevenson, W. Dennison, L. Staver, V. Carter, N. B. Rybicki, R. Hickman, S. Kollar, S. Bieber and P. Bergstrom. 1992. Chesapeake Bay Submerged Aquatic Vegetation Habitat Requirements and Restoration Targets: A Technical Synthesis. CBP/TRS 83/92. U.S. EPA Chesapeake Bay Program, Annapolis, Maryland.

Buchanan, C., R. V. Lacouture, H. G. Marshall, M. Olson, and J. M. Johnson. 2005. Phytoplankton Reference Communities for Chesapeake Bay and its Tidal Tributaries. Estuaries 28(1):138-159.

Chesapeake Bay Program (CBP). 1987. 1987 Chesapeake Bay Agreement. Chesapeake Bay Program Office, Annapolis, MD. Available online at www.chesapeakebay.net.

Chesapeake Bay Program (CBP). 2000. Chesapeake 2000 Agreement. Chesapeake Bay Program Office, Annapolis, MD. Available online at www.chesapeakebay.net.

Cooper, S. R. and Brush, G. S. 1993. A 2,500-year history of anoxia and eutrophication in Chesapeake Bay. Estuaries 16(3):617-626.

Fisher, T. R. and A. B. Gustafson. 2003. Nutrient-addition bioassays in Chesapeake Bay to assess resources limiting algal growth. Progress report: August 1990 – December 2002. Prepared for Maryland

Department of Natural Resources, Chesapeake Bay Water Quality Monitoring Program, by University of Maryland Horn Point Laboratory, Cambridge, MD.

Johnson, J. M. and C. Buchanan. 2014. Revisiting the Chesapeake Bay Phytoplankton Index of Biotic Integrity. Environmental Monitoring and Assessment 186:1431-1451.

Lacouture, R. V., J. M. Johnson, C. Buchanan, H. G. Marshall. 2006. Phytoplankton Index of Biotic Integrity for Chesapeake Bay and its tidal tributaries. Estuaries 29(4):598-616.

Miller, H. M. 1986. Transforming a “Splendid and Delightsome Land”: Colonists and Ecological Change in the Chesapeake 1607 - 1820. Journal of the Washington Academy of Sciences 76(3):173-187.

Olson, M. 2009. Relative Status Indicator. Development and Evolution of a Relative Measure of Condition for Assessing the Status of Water Quality and Biological Parameters Tracked in the USEPA Chesapeake Bay Program Long Term Monitoring Programs, Final Report. Prepared for the Interstate Commission on the Potomac River Basin. ICPRB Report 09-4. Available online at www.potomacriver.org.

Robertson, T. 2014. Chlorophyll *a* Assessment: a General Overview. Slide 14 of presentation made to the James River Chlorophyll Criteria Scientific Advisory Panel on 11 April, 2014.

U.S. Environmental Protection Agency (USEPA). 2003. Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll *a* for the Chesapeake Bay and its Tidal Tributaries. EPA 903-R-03-002. Region III Chesapeake Bay Program Office, Annapolis, MD. Available online at www.chesapeakebay.net.

U.S. Environmental Protection Agency (USEPA). 2007. Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll *a* for the Chesapeake Bay and its Tidal Tributaries. 2007 Addendum. EPA 903-R-07-003. Region III Chesapeake Bay Program Office, Annapolis, MD. Available online at www.chesapeakebay.net.

U.S. Environmental Protection Agency (USEPA). 2008. Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll *a* for the Chesapeake Bay and its Tidal Tributaries. 2008 Technical Support for Criteria Assessment Protocols Addendum. EPA 903-R-08-001. Region III Chesapeake Bay Program Office, Annapolis, MD. Available online at www.chesapeakebay.net.

Virginia Department of Environmental Quality (VADEQ). 2004. Chlorophyll *a* Numerical Criteria for the Tidal James River, November 30, 2004 (revised 1/12/2005). Technical Report prepared by Virginia Department of Environmental Quality.

**Table 1**. Classification thresholds used to delineate adequate water clarity (Secchi depth) for phytoplankton and concentrations of dissolved inorganic nitrogen (DIN) and ortho-phosphate ($PO_4$) that limit the formation of nuisance algal blooms. Seasons: spring (March – May); June*; summer (July – September); autumn (October – November); winter (December – February). Salinity zones: TF, tidal fresh ($\leq$0.5 ‰); OH, oligohaline (>0.5 - 5.0 ‰); MH, mesohaline (>5.0 - 18.0 ‰); PH, polyhaline (>18.0 ‰). From Buchanan *et al.* (2005). * June was not included in the original classification scheme. It was subsequently added and assigned the summer thresholds.

| | Spring | June | Summer | Autumn | Winter |
|---|---|---|---|---|---|
| Secchi depth (m) | | | | | |
| TF | >0.9 | >0.8 | >0.8 | >0.9 | >0.6 |
| OH | >0.7 | >0.6 | >0.6 | >0.5 | >0.6 |
| MH | >1.8 | >1.45 | >1.45 | >2.0 | >1.8 |
| PH | >2.15 | >1.85 | >1.85 | >2.5 | >2.3 |
| DIN (mg/liter) | $\leq$0.07 (all seasons and salinity zones) | | | | |
| $PO_4$ (mg/liter) | $\leq$0.007 (all seasons and salinity zones) | | | | |

**Table 2**. Water quality categories. See Table 1 for classification thresholds.

| Category name | Description |
|---|---|
| Better/Best | meets all thresholds for Secchi, DIN, & $PO_4$ |
| Mixed Better Light (MBL) | meets Secchi threshold, fails DIN and/or $PO_4$ threshold |
| Mixed Poor Light (MPL) | fails Secchi threshold, meets DIN and/or $PO_4$ threshold |
| Poor/Worst | fails all thresholds for Secchi, DIN, & $PO_4$ |

**Table3**. Number of water quality sampling events in each season-salinity habitat category of the Chesapeake open water environment (i.e., $\geq$ 2 meter depth) after data preparation. All sampling events have Secchi depth and chlorophyll *a* measurements, can be associated with a salinity zone, and can be definitely classified into one of the four water quality categories. Total number of sampling events is 52,528. See Tables 1 and 2 headings for details.

| Salinity zone | Water quality category | Spring | June | Summer | Autumn | Winter |
|---|---|---|---|---|---|---|
| TF | Better/Best | 0 | 0 | 1 | 1 | 1 |
| | MBL | 350 | 168 | 523 | 229 | 433 |
| | MPL | 723 | 240 | 619 | 191 | 121 |
| | Poor/Worst | 2630 | 825 | 1781 | 1018 | 1397 |
| OH | Better/Best | 1 | 1 | 20 | 14 | 0 |
| | MBL | 275 | 120 | 469 | 389 | 182 |
| | MPL | 385 | 100 | 475 | 177 | 118 |
| | Poor/Worst | 800 | 291 | 663 | 346 | 471 |
| MH | Better/Best | 100 | 96 | 792 | 99 | 93 |
| | MBL | 1061 | 396 | 1007 | 293 | 589 |
| | MPL | 5117 | 1852 | 3950 | 1767 | 2287 |
| | Poor/Worst | 1071 | 490 | 2485 | 1464 | 796 |
| PH | Better/Best | 132 | 27 | 79 | 63 | 194 |
| | MBL | 203 | 95 | 622 | 139 | 158 |
| | MPL | 966 | 504 | 1787 | 1185 | 1333 |
| | Poor/Worst | 247 | 165 | 1158 | 865 | 253 |

**Table 4.** Years in the reference water quality data set that can represent assessment layers for a given season-salinity habitat and their corresponding sample numbers. **Bold black text** with gray highlight indicates years with greater than 10 samples. Some analyses used only these samples; others included all season-salinity habitats with more than 2 samples. Season-salinity habitats with 1 or 2 samples are considered insufficient to calculate a meaningful %failure of a criterion (red text).

| Assessment Layer (Year) | Spring TF | Spring OH | Spring MH | Spring PH | Summer TF | Summer OH | Summer MH | Summer PH |
|---|---|---|---|---|---|---|---|---|
| 1984 | 6 | 0 | 0 | 0 | 3 | **12** | **46** | **20** |
| 1985 | 6 | 0 | **61** | **28** | **20** | 4 | **50** | **59** |
| 1986 | **16** | 6 | **103** | **21** | **20** | 6 | **81** | **34** |
| 1987 | **10** | 7 | **64** | 6 | **13** | **10** | **113** | **73** |
| 1988 | **12** | 8 | **98** | 8 | **17** | **15** | **101** | **35** |
| 1989 | 7 | 6 | **84** | **27** | **16** | **18** | **53** | 8 |
| 1990 | 7 | 2 | **31** | **18** | **22** | **12** | **61** | **34** |
| 1991 | **26** | **10** | **98** | **30** | **18** | **19** | **68** | **54** |
| 1992 | **16** | 7 | **76** | **38** | **28** | **17** | **83** | **34** |
| 1993 | 7 | **22** | **32** | 7 | **29** | **29** | **75** | **23** |
| 1994 | 9 | **44** | **71** | 6 | **20** | **17** | **76** | **28** |
| 1995 | 8 | 5 | **56** | **32** | **12** | **26** | **68** | **45** |
| 1996 | **10** | **22** | **35** | 5 | 6 | **12** | **45** | 4 |
| 1997 | **18** | 9 | **33** | 2 | **16** | **16** | **57** | **29** |
| 1998 | 8 | **36** | **12** | 3 | **11** | **25** | **81** | **12** |
| 1999 | **28** | 7 | **41** | 5 | **34** | **16** | **62** | **17** |
| 2000 | **27** | 3 | **18** | **10** | **41** | **16** | **56** | 9 |
| 2001 | **14** | 1 | **38** | **22** | **21** | **20** | **95** | **31** |
| 2002 | **15** | 3 | **28** | **34** | **23** | **26** | **77** | **66** |
| 2003 | 2 | **12** | **17** | 1 | **11** | **19** | **41** | 4 |
| 2004 | 1 | **10** | **24** | 4 | **15** | **22** | **51** | 5 |
| 2005 | **13** | 6 | **17** | 0 | **19** | **18** | **44** | 9 |
| 2006 | **12** | 2 | **11** | 1 | 5 | 6 | **23** | 5 |
| 2007 | **19** | 8 | 7 | 0 | **17** | **32** | **68** | 9 |
| 2008 | **10** | **10** | **19** | 1 | **28** | **32** | **50** | 6 |
| 2009 | **15** | 6 | **42** | **14** | **28** | **20** | **81** | 3 |
| 2010 | 8 | 7 | **18** | 8 | **10** | **10** | **24** | **17** |
| 2011 | 7 | **11** | **12** | 1 | 8 | 4 | **10** | 1 |
| 2012 | 7 | 4 | **13** | 2 | **10** | 5 | **33** | **23** |
| 2013 | 6 | 2 | 2 | 1 | 3 | 5 | **26** | 4 |

**Table 5.** Chlorophyll *a* criteria investigated. The 90th and 95th percentile criteria (rounded to 2 decimals) are calculated from all reference quality samples in each season-salinity habitat (i.e., Better/Best and MBL categories combined). The James River tidal fresh is divided into two segments with different criteria for assessments: the JMSTF1 (upper) segment between Richmond and Hopewell, and the JMSTF2 (lower) segment between Hopewell and the JMSOH boundary. See text for details of how criteria are applied to chlorophyll *a* measurements.

| | TF | OH | MH | PH |
|---|---|---|---|---|
| 90th%ile criteria | | | | |
| Spring | 10.39 | 22.54 | 15.49 | 7.94 |
| Summer | 16.93 | 17.23 | 11.89 | 7.37 |
| 95th%ile criteria | | | | |
| Spring | 13.53 | 28.61 | 22.5 | 10.05 |
| Summer | 24.16 | 23.75 | 14.22 | 8.65 |
| James River criteria (upper, lower) | | | | |
| Spring | 10, 15 | 15 | 12 | 12 |
| Summer | 15, 23 | 22 | 10 | 10 |

**Table 6.** Analysis of the Mixed Better Light (MBL) category samples where one nutrient concentration is bloom-limiting and the other is not. Bloom-limiting concentrations are 0.07 mg DIN/liter and 0.007 mg $PO_4$/liter (from Fisher and Gustafson 2003).

| | Spring | | Summer | |
|---|---|---|---|---|
| Salinity zone | % of samples limited by DIN | % of samples limited by $PO_4$ | % of samples limited by DIN | % of samples limited by $PO_4$ |
| TF | 0.00% | 100.00% | 9.84% | 90.16% |
| OH | 0.00% | 100.00% | 16.42% | 83.58% |
| MH | 0.13% | 99.87% | 59.91% | 40.09% |
| PH | 21.21% | 78.79% | 92.95% | 7.05% |

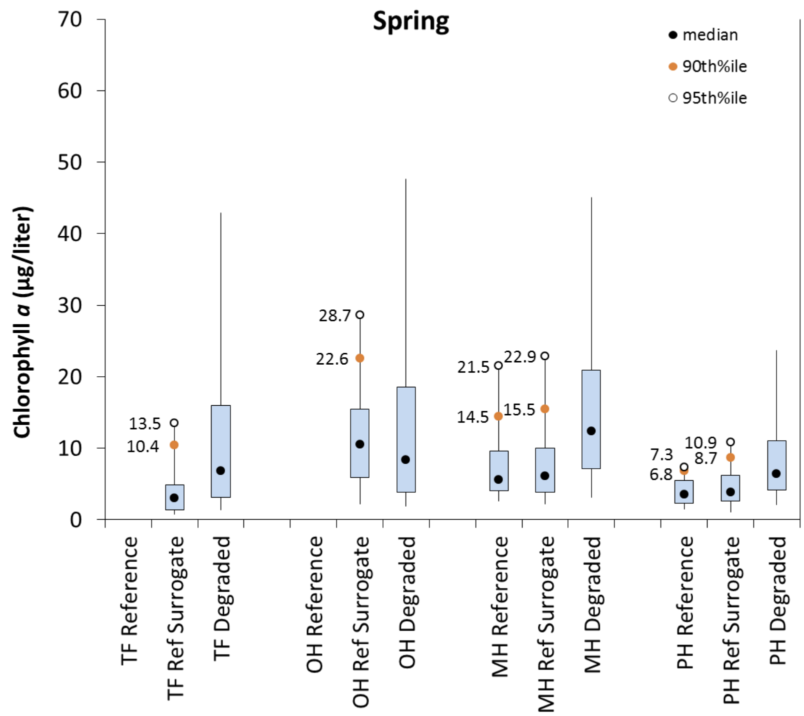



**Figure 1a.** Comparisons of spring (March – May) chlorophyll *a* concentrations in the salinity-specific Better/Best categories ("Reference"), Mixed Better Light categories ("Ref Surrogate"), and combined Mixed Poor Light and Poor/Worst categories ("Degraded") for all Chesapeake open water habitats, 1984-2013. Box, $25^{th}$ – $75^{th}$ percentile; whiskers, $5^{th}$ – $95^{th}$ percentile. Values of $90^{th}$ and $95^{th}$ percentiles are shown for Reference and Ref Surrogate categories. Habitats with < 20 samples are not shown.

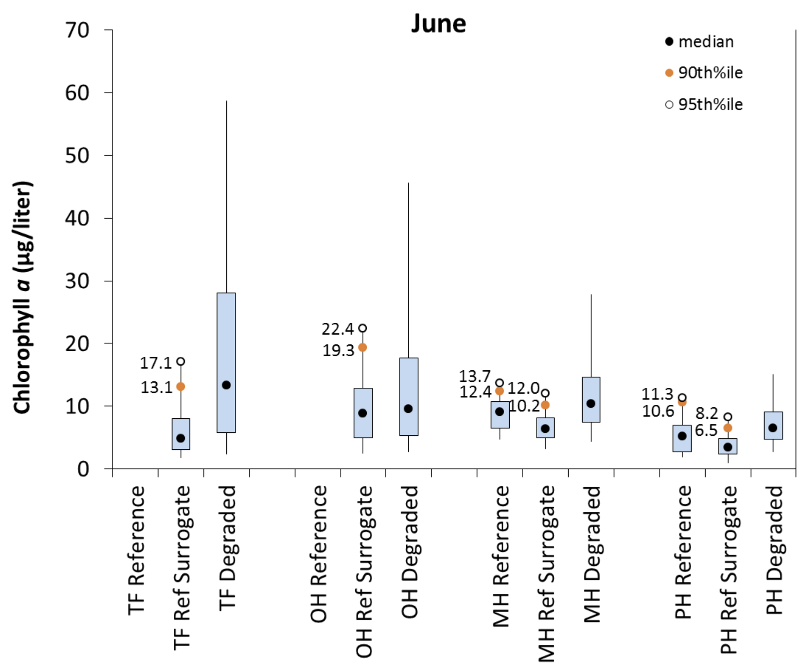



**Figure 1b.** June chlorophyll *a* concentrations.

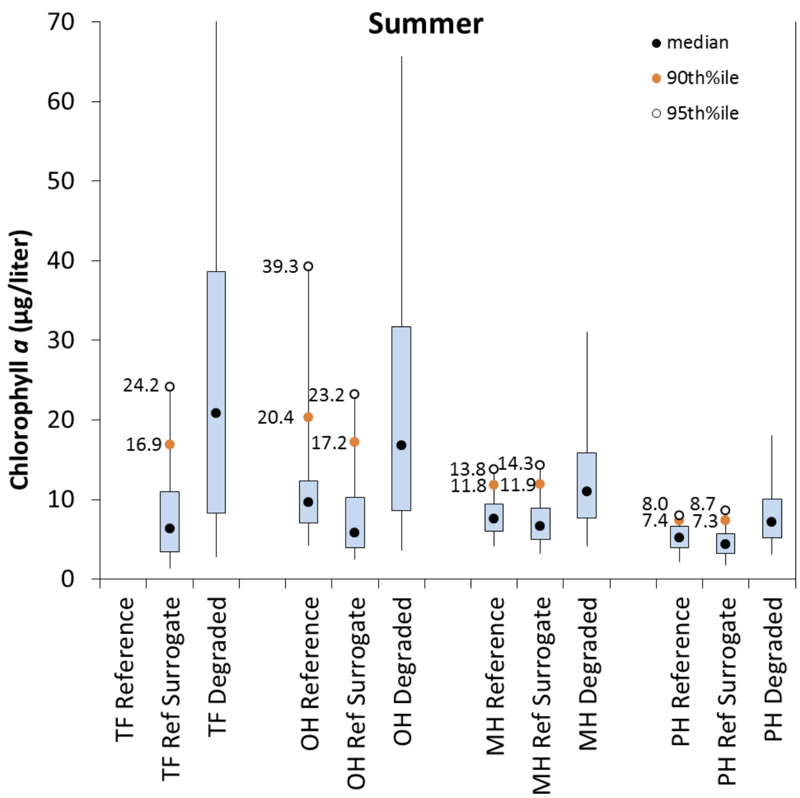



**Figure 1c.** Summer (July – September) chlorophyll *a* concentrations.

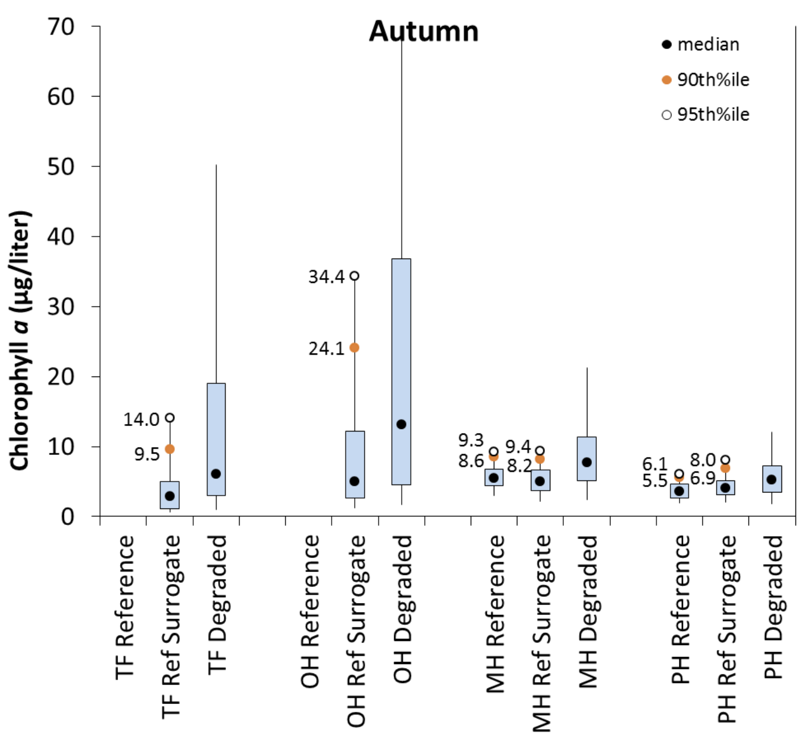



**Figure 1d.** Autumn (October – November) chlorophyll *a* concentrations.

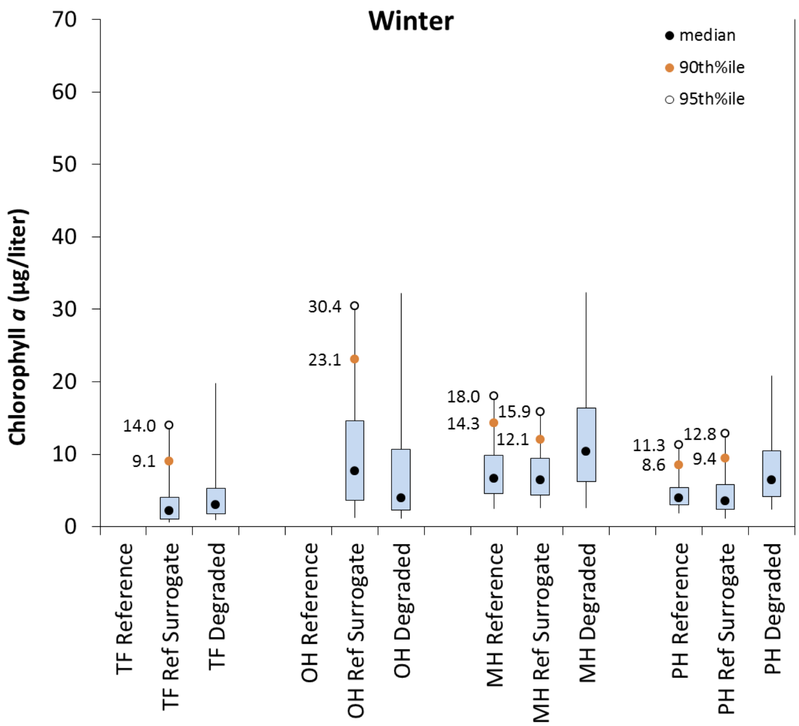



**Figure 1e.** Winter (December – February) chlorophyll *a* concentrations.

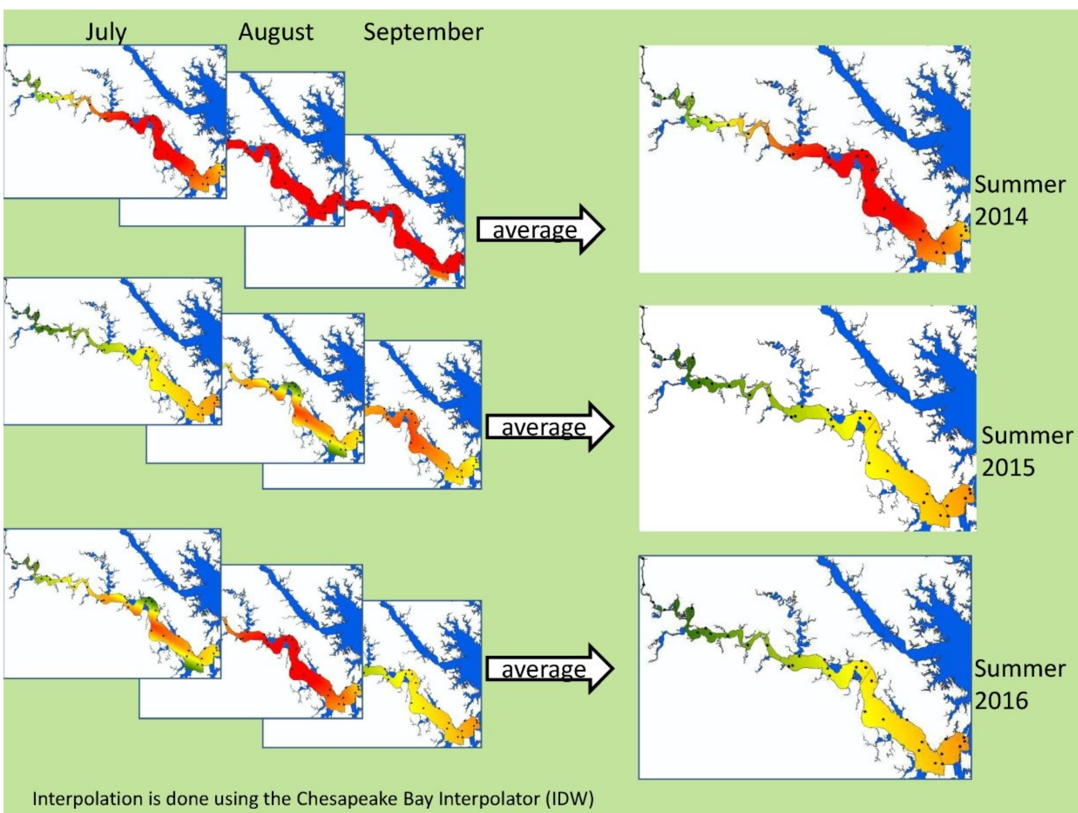
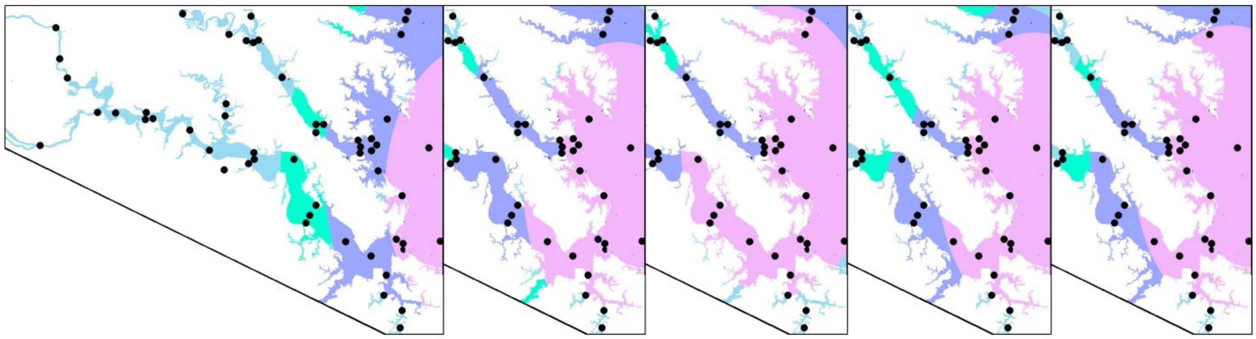
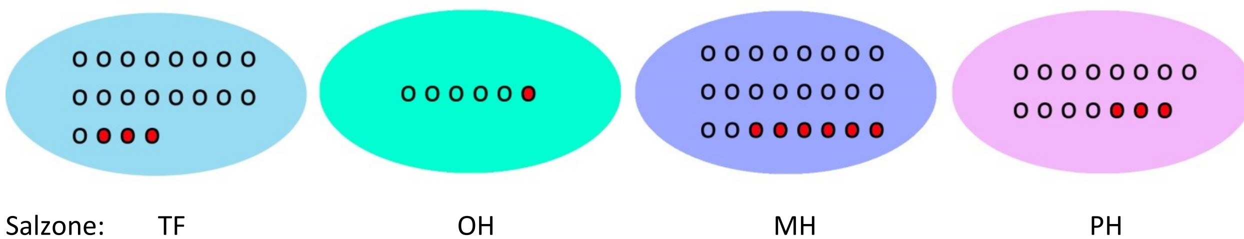



**Figure 2**. Illustration of Virginia method of annual averaging from monthly chlorophyll *a* interpolations (from T. Robertson 2014).

A) Individual cruise periods in one season-year, bay wide (all samples)



B) One season-year (all samples from reference quality conditions regardless of cruise periods)




C) Multiple season-years

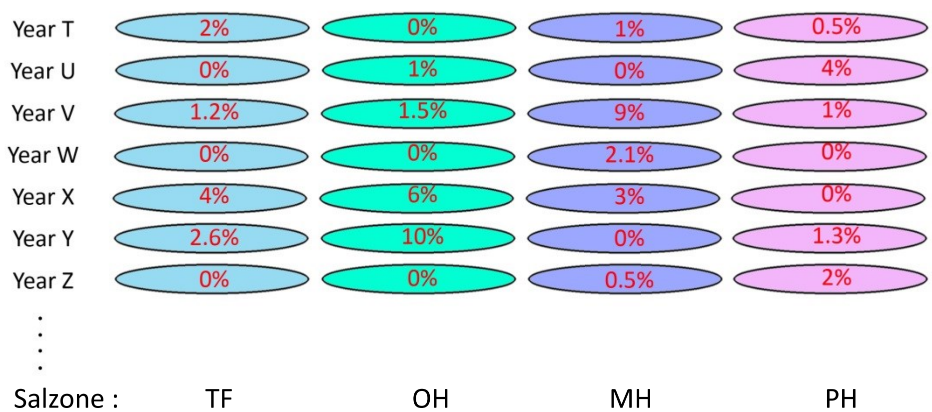



**Figure 3a-c**. Illustration of Method 1 used to develop biological reference curves for chlorophyll *a*. The method assumes that the proportion of observations violating a criterion in a given season-year is comparable to the exceedance rate in space of a single assessment layer developed with the prescribed VADEQ method. A) All sampling events in a given season-year are classified according to the salinity observed at the time of sampling (light blue = tidal fresh; green = oligohaline; blue = mesohaline; purple = polyhaline). B) All sampling events in the season-year whose water quality conditions meet the reference classification criteria are extracted. Samples are scored by the appropriate season-salinity criterion (dots = total number of reference sampling events; red dots = criterion failed). C) Each year is treated as if it were an individual assessment layer representative of the given season anywhere in the given salinity zone. Attainment rates for the individual layers can be used to construct biological reference curves from multiple layers.

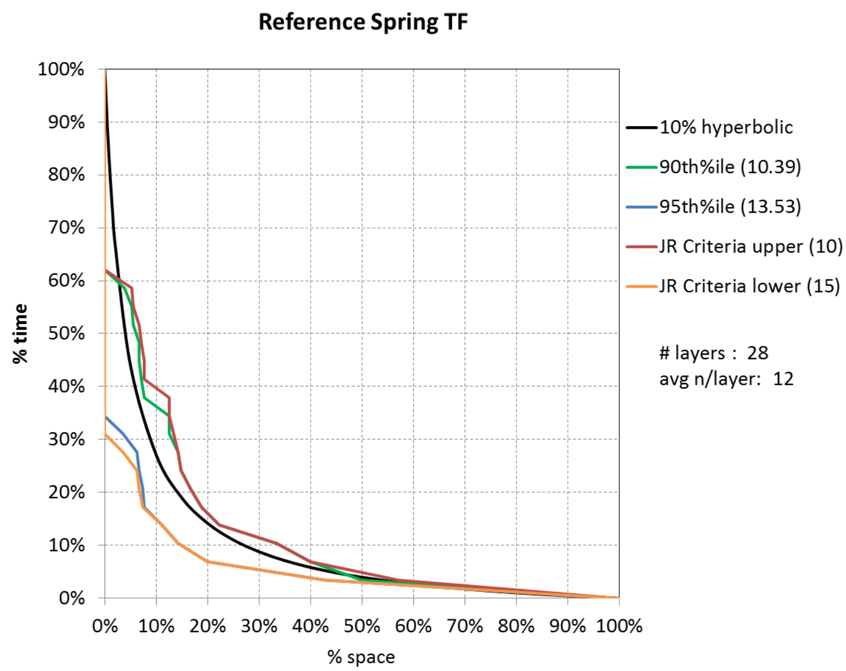



**Figure 4a**. Spring tidal fresh CFD curves for chlorophyll *a* from reference water quality conditions found in Chesapeake Bay open waters between 1984 and 2013. The number of assessment layers used to create the CFD curves and the average number of samples per layer are indicated on the right. The values of the criteria applied to the data are indicated in parentheses.

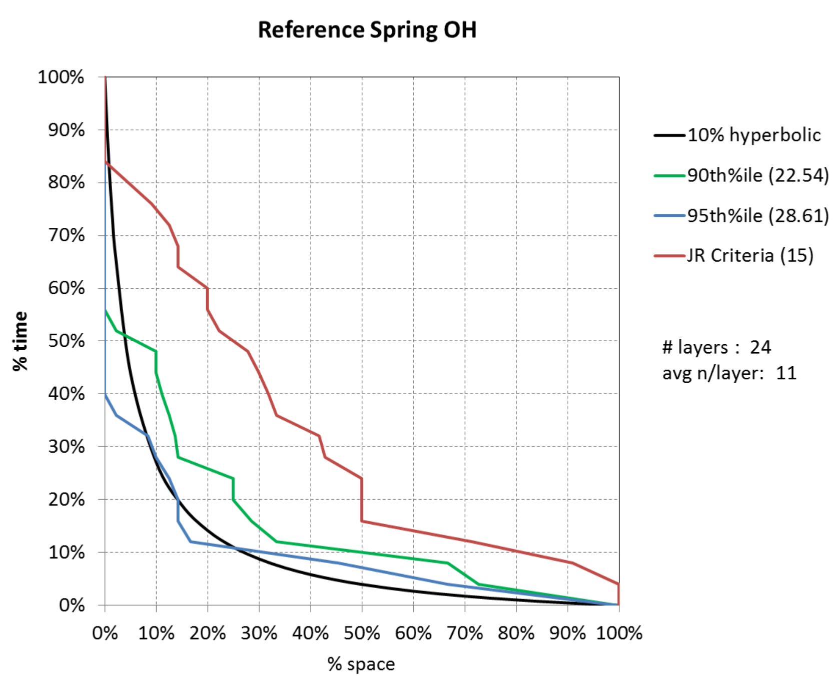



**Figure 4b.** Spring oligohaline CFD curves for chlorophyll *a* from reference water quality conditions.

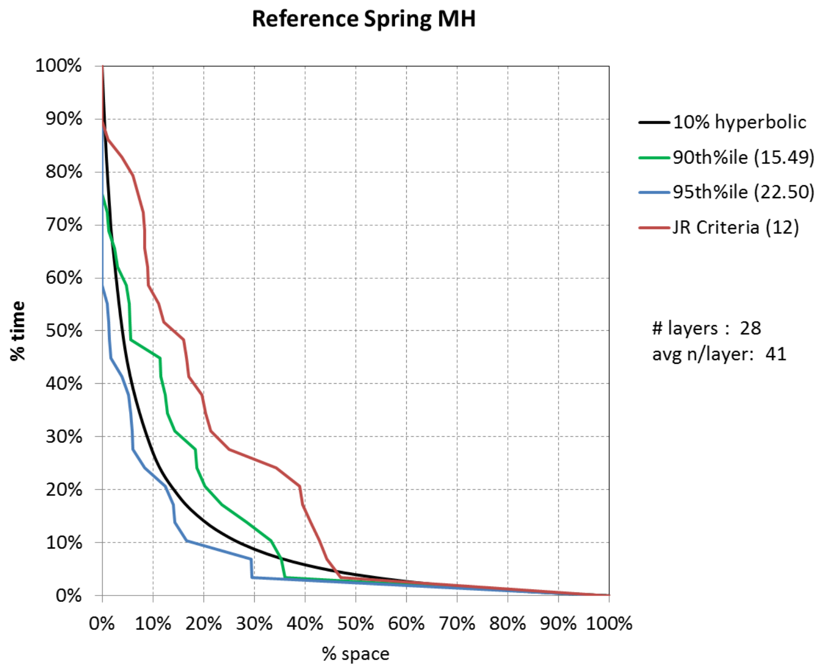



**Figure 4c.** Spring mesohaline CFD curves for chlorophyll *a* from reference water quality conditions.

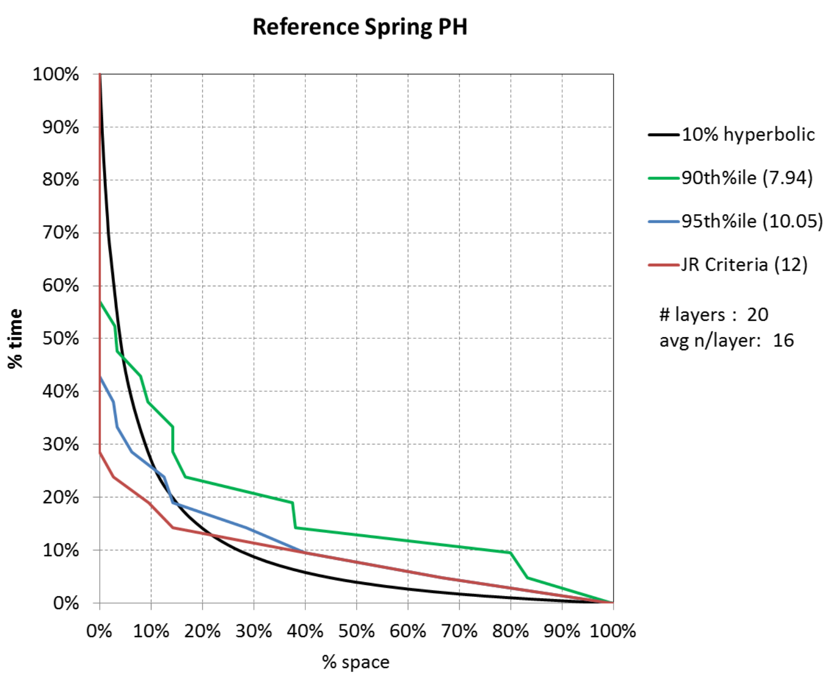



**Figure 4d.** Spring polyhaline CFD curves for chlorophyll *a* from reference water quality conditions.

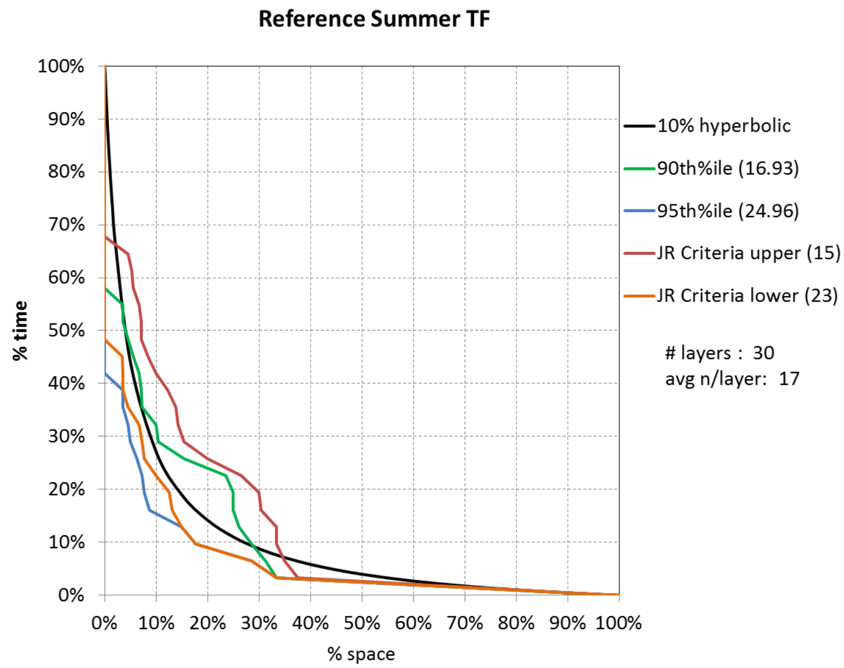



**Figure 4e.** Summer tidal fresh CFD curves for chlorophyll *a* from reference water quality conditions.

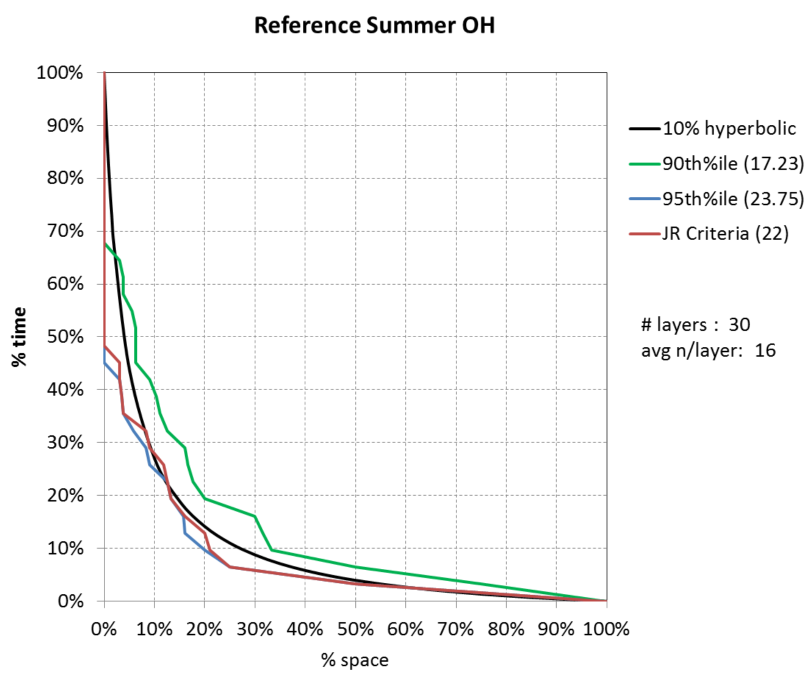



**Figure 4f.** Summer oligohaline CFD curves for chlorophyll *a* from reference water quality conditions.

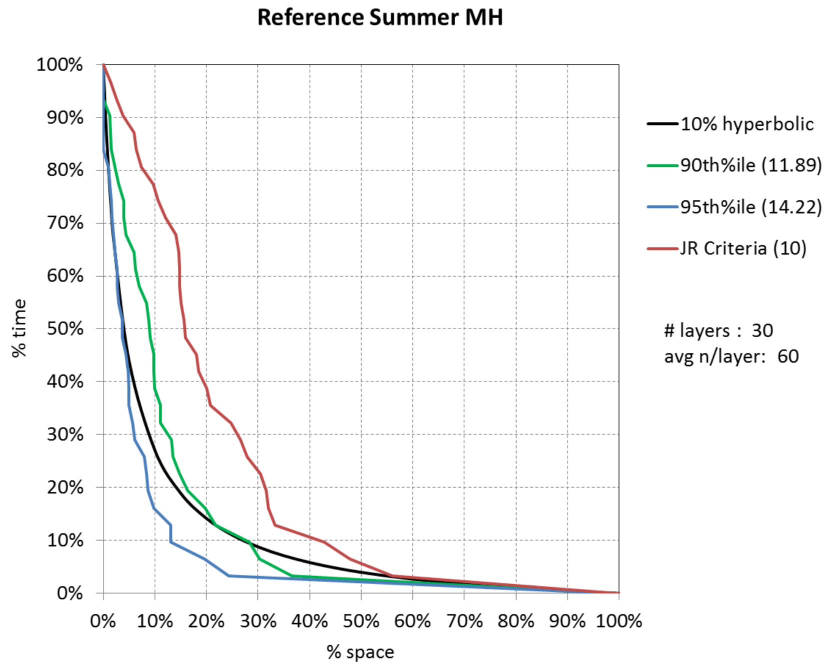



**Figure 4g.** Summer mesohaline CFD curves for chlorophyll *a* from reference water quality conditions.

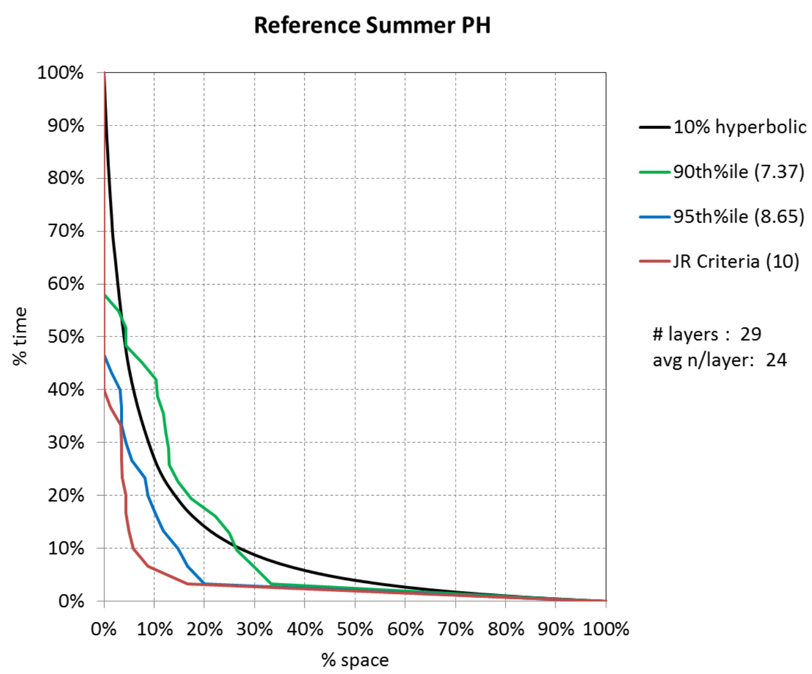



**Figure 4h.** Summer polyhaline CFD curves for chlorophyll *a* from reference water quality conditions.

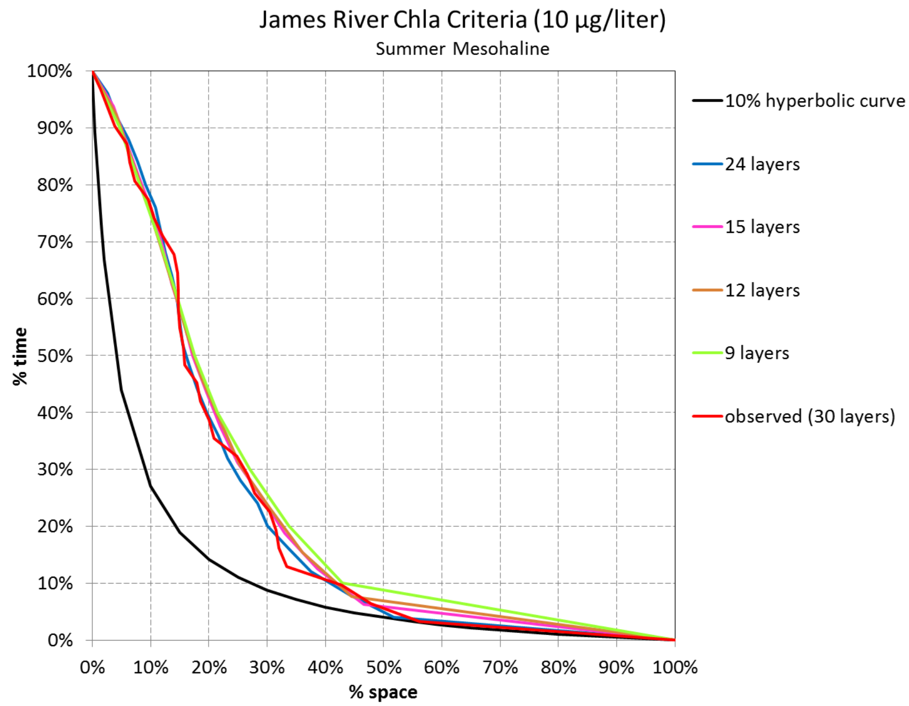



**Figure 5a**. Biological reference curves based on the James River criterion for summer mesohaline habitat and developed from 24, 15, 12, and 9 assessment layers. Layers were created using random sampling with replacement (Method 2). See text for details.

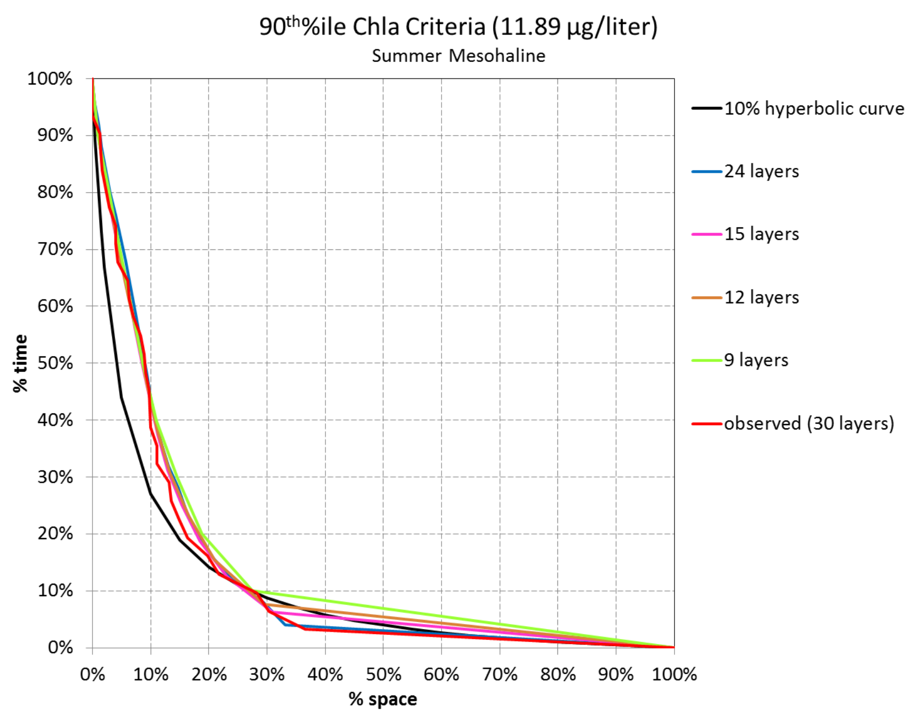



**Figure 5b**. Biological reference curves based on the $90^{th}$%ile criterion for summer mesohaline habitat and developed from 24, 15, 12, and 9 assessment layers.

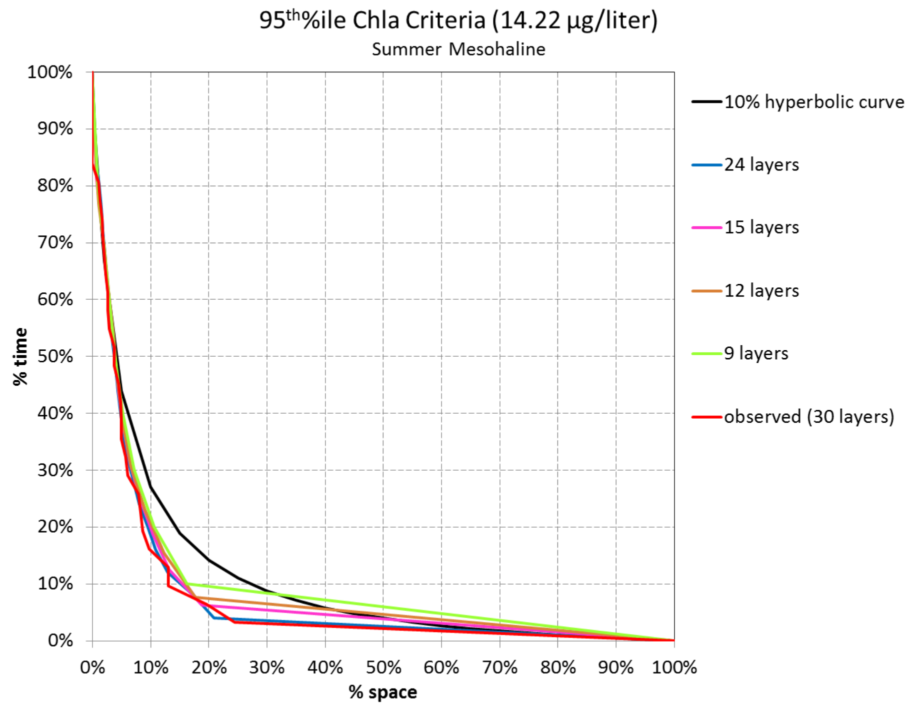



**Figure 5c**. Biological reference curves based on the 95th%ile criterion for summer mesohaline habitat and developed from 24, 15, 12, and 9 assessment layers.

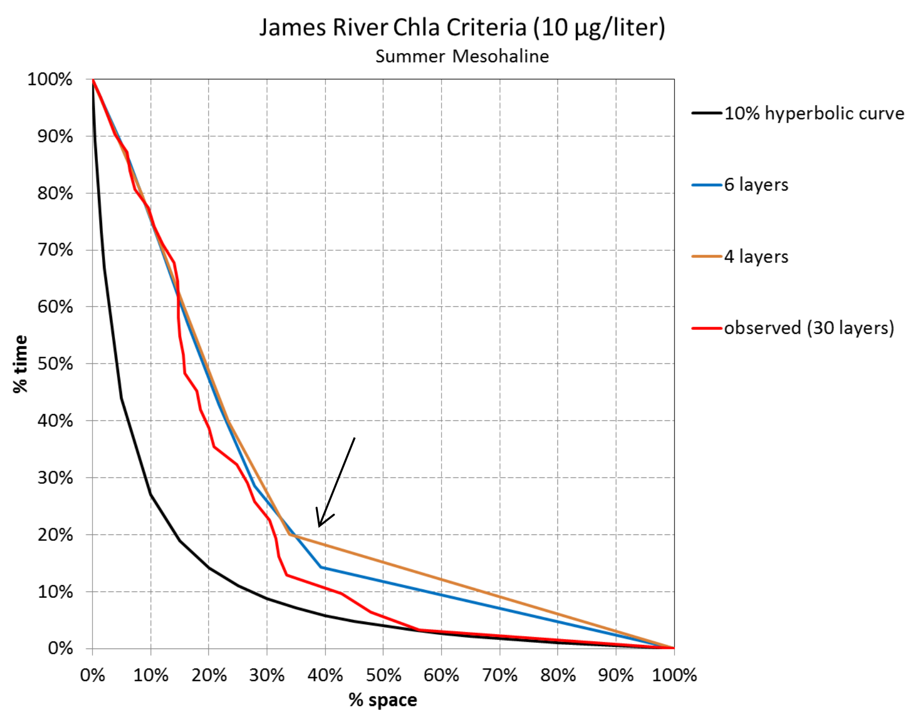



**Figure 6a**. Biological reference curves based on the James River criterion for summer mesohaline habitat and developed from 6 and 4 assessment layers. Layers were created using random sampling with replacement (Method 2). Arrow indicates bias caused by too few layers (see text for details).

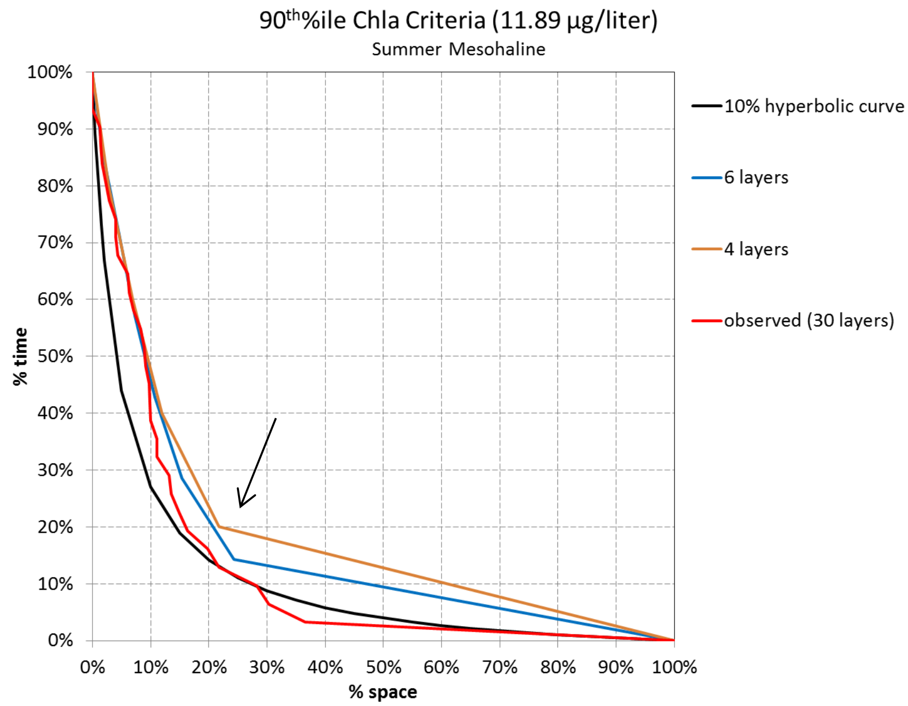



**Figure 6b**. Biological reference curves based on the 90th%ile criterion for summer mesohaline habitat and developed from 6 and 4 assessment layers.

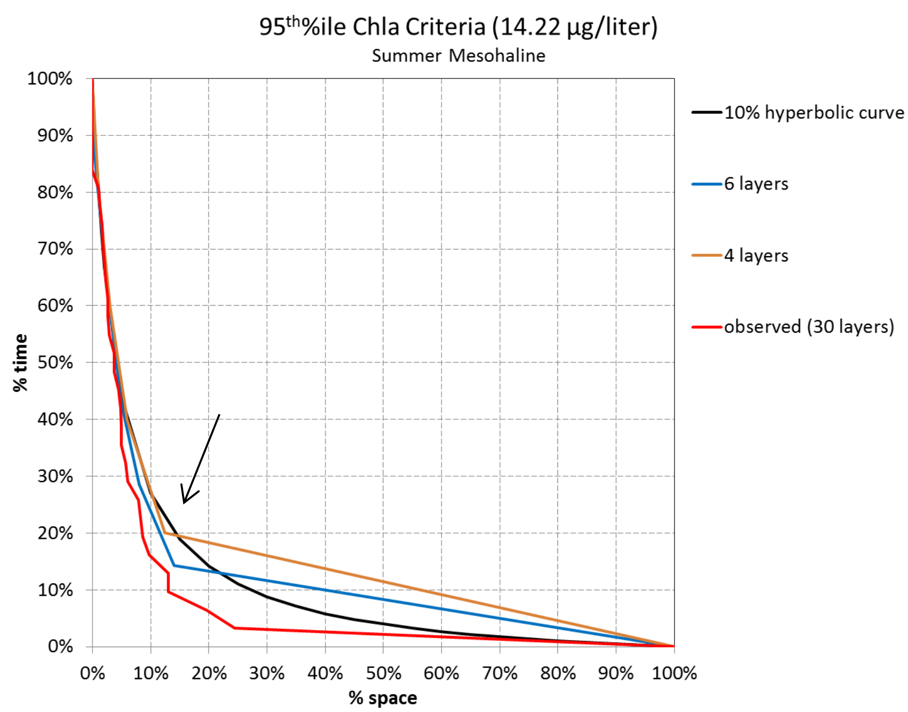



**Figure 6c**. Biological reference curves based on the 95th%ile criterion for summer mesohaline habitat and developed from 6 and 4 assessment layers.

Individual Layers

| %time | %space | Year |
|---|---|---|
| 0.0% | 100.0% | |
| 10.0% | 85.0% | 2 |
| 20.0% | 65.0% | 1 |
| 30.0% | 36.0% | 3 |
| 40.0% | 20.0% | 2 |
| 50.0% | 18.2% | 1 |
| 60.0% | 14.8% | 1 |
| 70.0% | 6.0% | 2 |
| 80.0% | 5.0% | 3 |
| 90.0% | 1.5% | 3 |
| 100.0% | 0.0% | |

Annual Means

| %time | %space | Year |
|---|---|---|
| 0.0% | 100.0% | |
| 25.0% | 37.0% | 2 |
| 50.0% | 32.7% | 1 |
| 75.0% | 14.2% | 3 |
| 100.0% | 0.0% | |

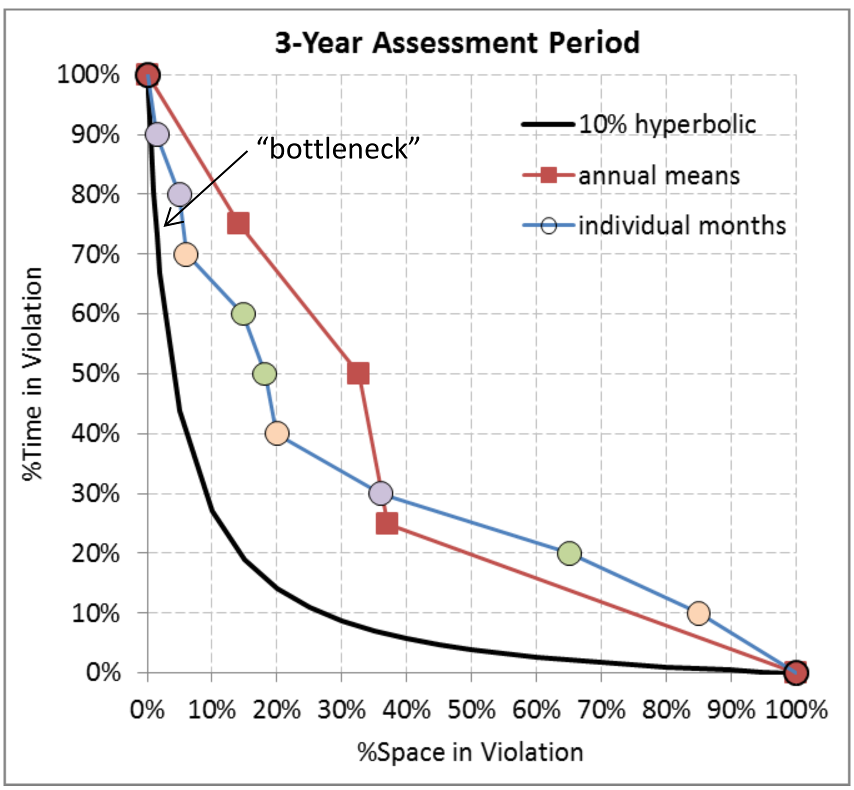



**Figure 7.** Hypothetical example of CFD curve biases created when annual (seasonal) means are used.

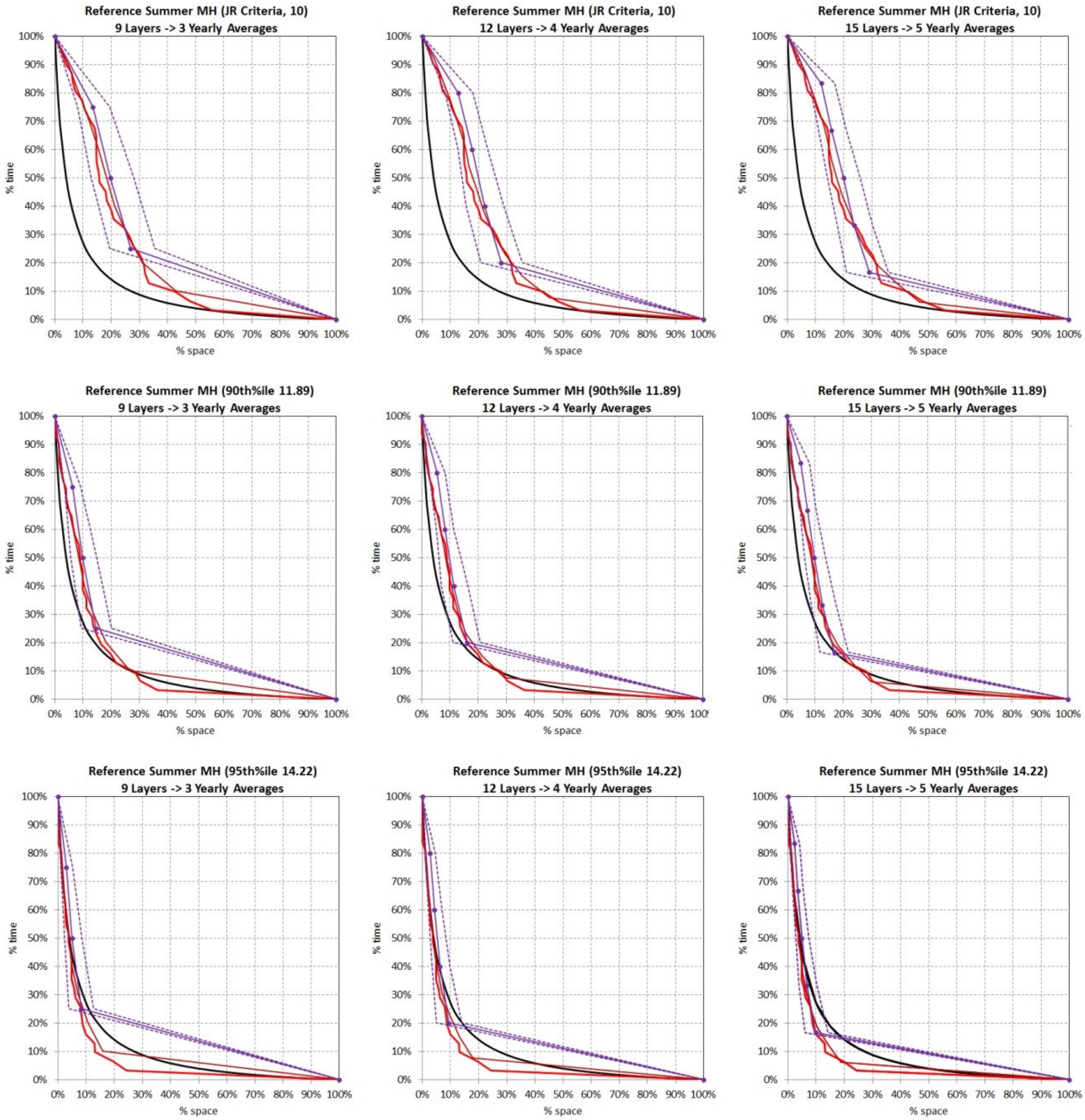
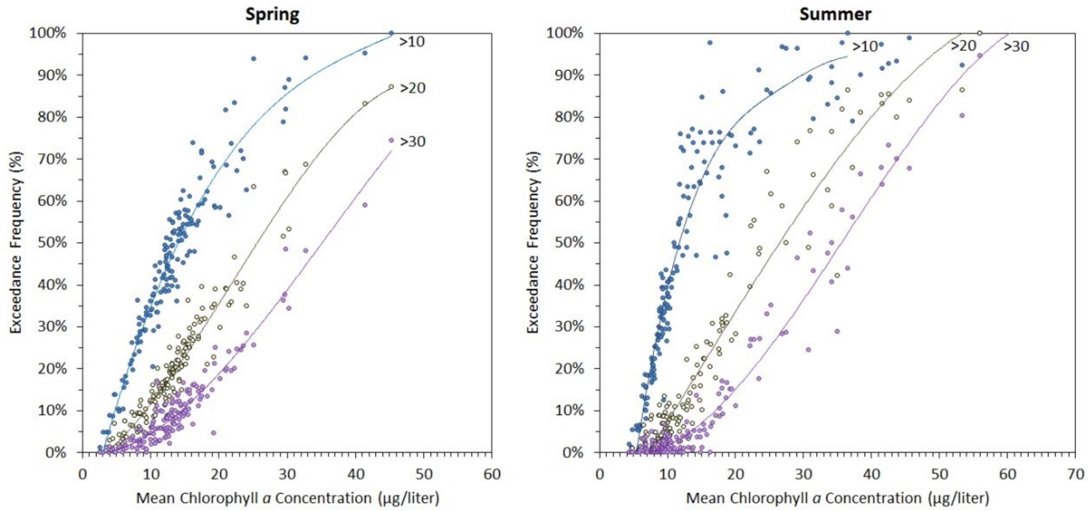
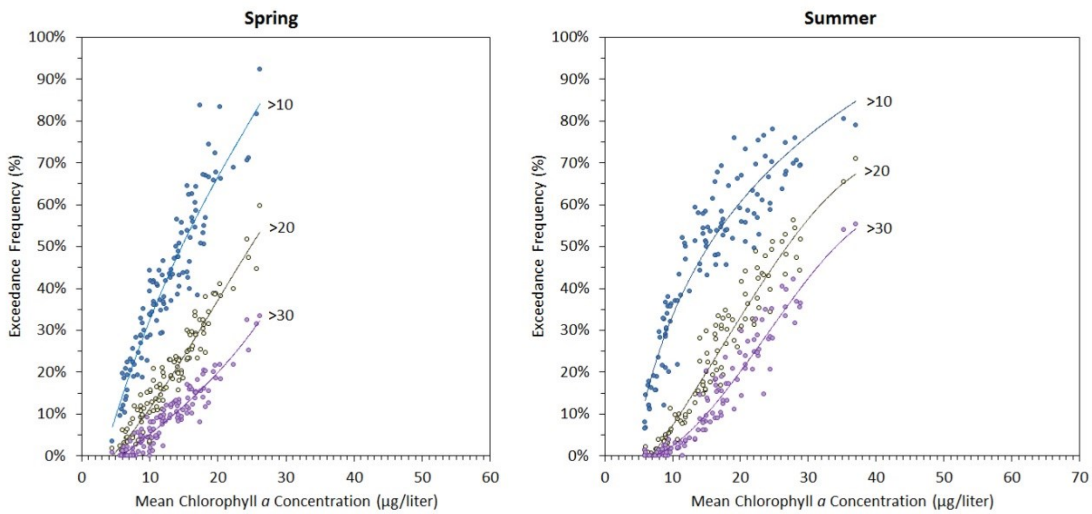



**Figure 8.** Biological reference curves for summer mesohaline built from 3, 4, and 5 annual averages. The 10% hyperbolic curve, the complete 30-point CFD curve, and the CFDs built from the underlying 9, 12, 15 individual assessment layers are shown for comparison. Assessment layers for annual averages were created using Method 3. See text for details. Red solid line: CFD curve of all 30 summer mesohaline assessment layers (representing different summer months in this example). Orange solid line: CFD curve base on the 9, 12, and 15 assessment layers used to calculate the 3, 4, and 5 yearly averages for a season, respectively. Purple solid line: CFD curve built from the yearly averages for a season (see Method 3 for details). Purple dashed line: $10^{th}$%ile and $90^{th}$%ile around the CFD curve.

A)




B)




C)

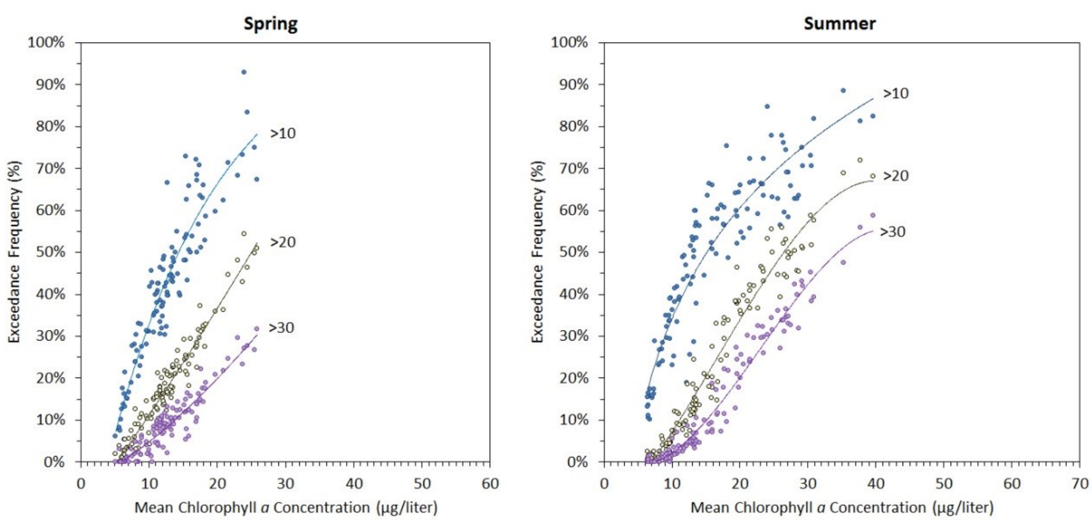
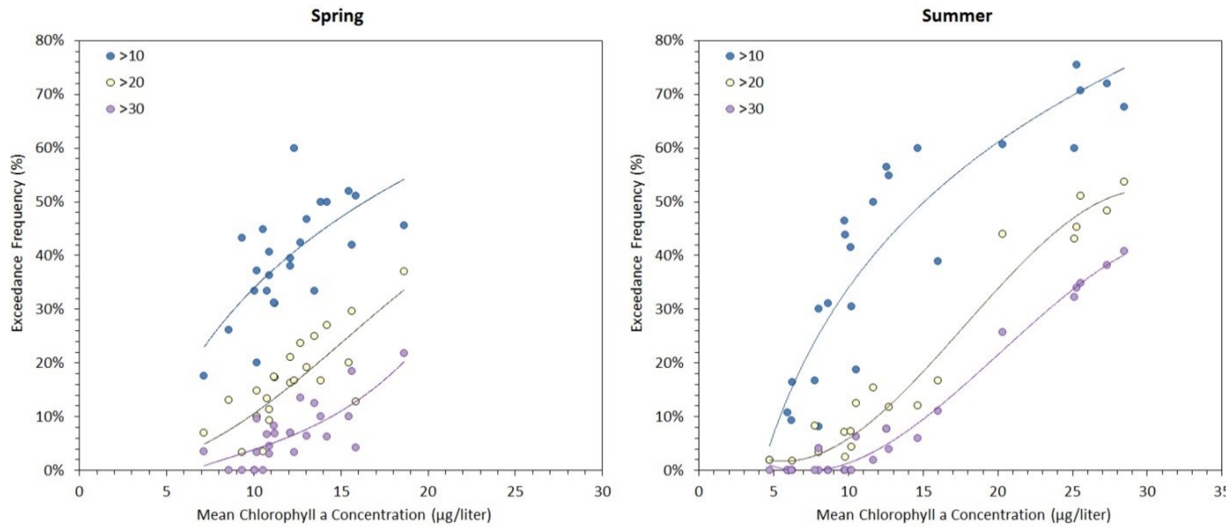



**Figure 9**. Relationships in entire Chesapeake Bay between mean chlorophyll *a* and frequency of exceeding 10, 20, and 30 µg/liter. Seasonal data grouped by station (A), CBP segment and year (B), and salinity zone and year (C). Points with n<30 not shown.

A)




B)

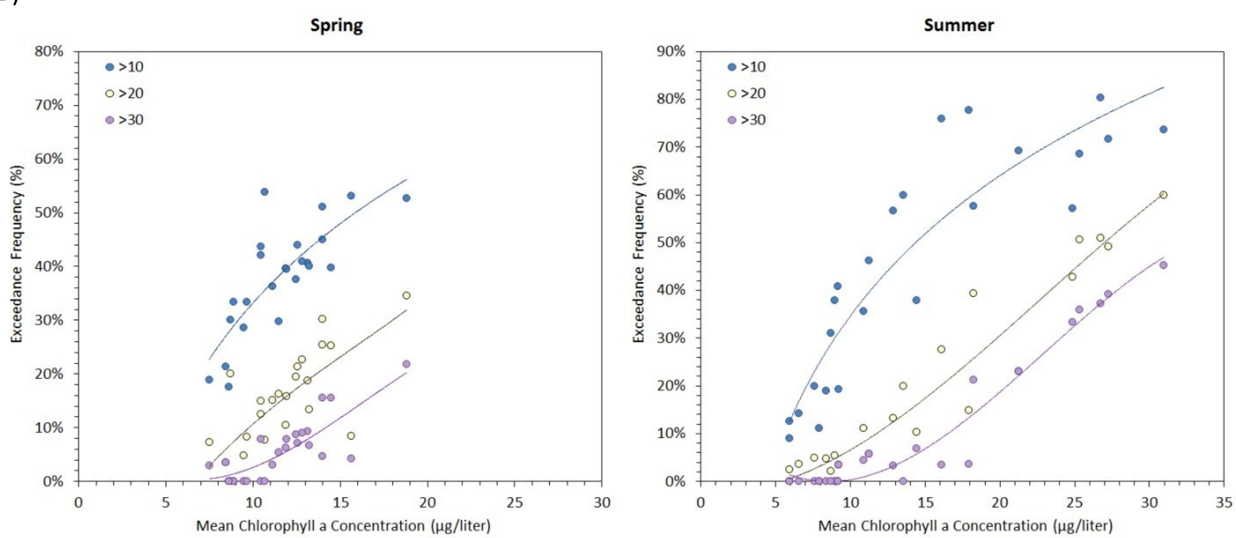



**Figure 10.** Relationships in the James River between mean chlorophyll *a* and frequency of exceeding 10, 20, and 30 µg/liter. Points with n<24 not shown. Data were grouped by season and the James River CBP segment JMSTF, JMSOH, JMSMH, and JMSPH (A) or by season and salinity zone (B). Data were then further divided into six time periods (1985 – 1989, 1990 – 1994, 1995 – 1999, 2000 – 2004, 2005 – 2009, and 2010 – 2013) instead of by year in order to ensure sufficient numbers of samples per point.