

2016 Criteria Assessment Protocol Workgroup
Critical Review of the Existing Tidal Waters Chlorophyll *a* Assessment Procedures and
Recommendations on Alternatives

1ST DRAFT : September 29, 2016

2ND DRAFT, FINAL DRAFT: October 27, 2016

Final Copy: November 2, 2016

Chair: Peter Tango

Staff: Melissa Merritt

Background

In July 2016, the Chesapeake Bay Program (CBP) charged the CBP Scientific Technical Assessment and Reporting Team (STAR) to reconvene the Criteria Assessment Protocol Workgroup (CAP WG) with the tasks of 1) conducting a critical review of the existing chlorophyll *a* assessment procedures and 2) providing consensus recommendations on any alternatives to consider in revision of the procedures. The recommendations are provided only in the context of assessing the existing seasonal mean chlorophyll *a* criteria. Alternative criteria would suggest a need to reconsider the recommendations in light of the different criteria (e.g. monthly means, instantaneous values).

Existing chlorophyll *a* criteria and criteria assessment procedures were developed through the work of the CAP WG and the CBP community then published by U.S. EPA through a series of documents between 2003 and 2010 and promulgated through jurisdictions water quality standards. The following information provides a history of the documentation.

Supporting Documentation for the Review

In 2003, the U.S. EPA published the *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll *a* for the Chesapeake Bay and Its Tidal Tributaries (Regional Criteria Guidance)* in cooperation with and on behalf of the six watershed states – New York, Pennsylvania, Maryland, Delaware, Virginia, and West Virginia- and the District of Columbia. In 2004, EPA published the first addendum to the 2003 Regional Criteria Guidance (U.S. EPA 2004) which included guidance on determining where numerical chlorophyll *a* criteria should apply to local Chesapeake Bay and Tidal tributary waters. Subsequently, Delaware, Maryland, Virginia and the District of Columbia promulgated narrative chlorophyll *a* criteria in their standards regulations. Virginia promulgated numerical segment- and seasonal-specific chlorophyll *a* criteria for the James River. The District of Columbia promulgated numerical chlorophyll *a* criteria for its reach of the tidal Potomac River and its remaining tidal waters, having previously adopted numerical chlorophyll *a* criteria for protection of the Anacostia River.

In July 2007, EPA published a second addendum to the 2003 Regional Criteria Guidance (U.S. EPA 2007a). The second addendum documented the revised, refined and new criteria assessment methods for the published Chesapeake Bay criteria including chlorophyll *a*. In November 2007, U.S. EPA published a third addendum documenting numerical chlorophyll *a* criteria and reference concentrations (U.S. EPA 2007b). U.S. EPA (2007b) included a chapter on

recommended procedures for consideration in assessing attainment of numerical chlorophyll *a* criteria.

In September 2008, U.S. EPA published the fourth addendum, *2008 Technical Support for Criteria Assessment Protocols Addendum* (U.S. EPA 2008). Chapter V and Appendix G in U.S. EPA (2008) documented the updated approach and protocols for assessing numerical chlorophyll *a* criteria. The protocols were detailed as step-by-step assessment procedures while providing examples of application of the assessment methods and their outputs. The approach reflected the use of 1) 3 season-specific assessments with a seasonal mean criterion (in Virginia it is spring and summer, in the District of Columbia it is summer only) to assess water quality standards attainment of a segment, 2) surface water quality measurements (i.e., not depth integrated), 3) inverse distance weighting interpolation of the data for computing a surface of means that would be used to compute spatial violation rates of the applicable criterion, 4) presentation of the 3 years assessment in a cumulative frequency distribution curve in a 2-dimensional space-time assessment procedure and 4) having a decision rule that used a reference curve in the 2-dimensional space-time assessment procedure to decide on attainment status. Finally, U.S. EPA published a fifth addendum *2010 Technical Support for Criteria Assessment Protocols Addendum* (U.S. EPA 2010) where Chapter IV addressed revisions to the 2008 published chlorophyll *a* criteria assessment methodology. The 2010 revisions focused on the log normal nature of chlorophyll data. In U.S. EPA (2010) the revisions addressed use and implications of using log transformed chlorophyll *a* data in computing seasonal means during the assessment procedures.

Between 2010 and 2016, the James River Science Advisory Panel (JRSAP) conducted a re-evaluation of the James River chlorophyll *a* criteria in the context of new science, however, the JRSAP was not responsible for re-evaluating the methods. Coincident with the completion of the JRSAPs work, VADEQ initiated internal review of the assessment methodology and proposed directions for revisions. Two white papers and a Power Point presentation were produced. The complete list of supporting documentation provided to the CAP WG before the first meeting is provided here below:

- U.S. EPA (U.S. Environmental Protection Agency). 2003. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries (Regional Criteria Guidance)*. April 2003. EPA 903-R-03-002. Region III Chesapeake Bay Program Office, Annapolis, MD.
- U.S. EPA (U.S. Environmental Protection Agency). 2004. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2004 Addendum*. October 2004. EPA 903-R-04-005. Region III Chesapeake Bay Program Office, Annapolis, MD.
- U.S. EPA (U.S. Environmental Protection Agency). 2007a. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2007 Addendum*. July 2007. EPA 903-R-07-003. Region III Chesapeake Bay Program Office, Annapolis, MD.

- U.S. EPA (U.S. Environmental Protection Agency). 2007b. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries –Chlorophyll a Addendum*. October 2007. EPA 903-R-07-005. Region III Chesapeake Bay Program Office, Annapolis, MD.
- U.S. EPA (U.S. Environmental Protection Agency). 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*. September 2008. EPA 903-R-08-001. Region III Chesapeake Bay Program Office, Annapolis, MD.
- U.S. EPA (U.S. Environmental Protection Agency). 2010. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2010 Technical Support for Criteria Assessment Protocols Addendum*. May 2010. EPA 903-R-10-002. Region III Chesapeake Bay Program Office, Annapolis, MD.
- VADEQ. 2016a. VADEQ Critical Review of the Assessment Methodology for James River Chlorophyll Virginia Department of Environmental Quality. White paper. http://www.chesapeakebay.net/channel_files/24252/critical_review_of_jr_chlorophyll_assessment_methodology.pdf
- VADEQ. 2016b. Proposed Assessment Methodology for James River Chlorophyll Criteria. White paper. http://www.chesapeakebay.net/channel_files/24252/2016proposed_jr_chlorophyll_procedure.pdf
- VADEQ. Robertson, T. 2016. Power Point (PPT). August 9 2016 CAP WG. http://www.chesapeakebay.net/channel_files/24252/alternatives_for_chlprotocol_072816.pdf

Criteria Assessment Protocol Workgroup Review

The Chesapeake Bay Program Criteria Assessment Protocol Workgroup (CAP WG) reviewed each of five protocol elements as they relate to chlorophyll *a* criteria assessments for the tidal water James River segments. Three meetings were held between July and September 2016. The first meeting was held on August 9th as a face-to-face meeting at the USGS conference room in Baltimore, MD. Two conference calls on September 7th and September 21st were used to focus the recommendations. Comments were solicited on draft recommendations based on review and input from workgroup members has been incorporated into the final recommendations. All watershed jurisdictions were invited to participate. State, Federal, NGO and industry representatives attended the meetings. The meetings outlined five issue areas within the published assessment protocol for critical review and developing recommendations on any alternatives and options to the published protocol. Those issue areas were:

- I. DATA: SAMPLING AND ASSESSMENT PERIOD
 - 1) The sampling strategy.
 - 2) The assessment period.

- II. ASSESSMENT: INTERPOLATION, CFD and IMPAIRMENT DECISION RULE
 - 3) The data interpolation step.
 - 4) The use of the cumulative frequency distribution for assessing water quality standards attainment.
 - 5) The decision rule for declaring the status of the Chesapeake Bay segment as in or out of compliance with the water quality standards, i.e., in or out of attainment.

I. DATA: SAMPLING AND ASSESSMENT PERIOD

1) Use surface sampling or depth integrated sampling to support the chlorophyll *a* criteria assessment.

Issue: For the purpose of assessing chlorophyll *a* criteria, the water quality sampling question focused on whether to continue a surface water sampling program or move to a depth integrated sampling program?

- **Recommendation:** *No change. Continue the surface sampling basis of the chlorophyll *a* measurements that support the water quality criteria assessment. However, continue research on best practices and opportunities with development of better scientific support and understanding to apply depth-integrated assessments.*

Background

At issue is how well surface chlorophyll *a* represents the total integrated water column chlorophyll *a*—especially important in the spring when chlorophyll *a* accumulates both above and below the pycnocline until the hypoxia onset. In U.S. EPA (2007), it was reported that at mainstem water quality monitoring stations in the Bay, the annual average surface chlorophyll *a* ($\mu\text{g chl}a \cdot \text{liter}^{-1}$) correlated strongly with the annual, average, integrated water column chlorophyll *a* ($\text{mg chl}a \cdot \text{m}^{-2}$). A spring season relationship was presented showing a strong correlation (March-June 1985-2004 data) suggesting the use of surface measurements was a good reflection of and indicator for water column accumulations of phytoplankton biomass (Figure 1). However, tributary specific assessments of the relationship were not explored.

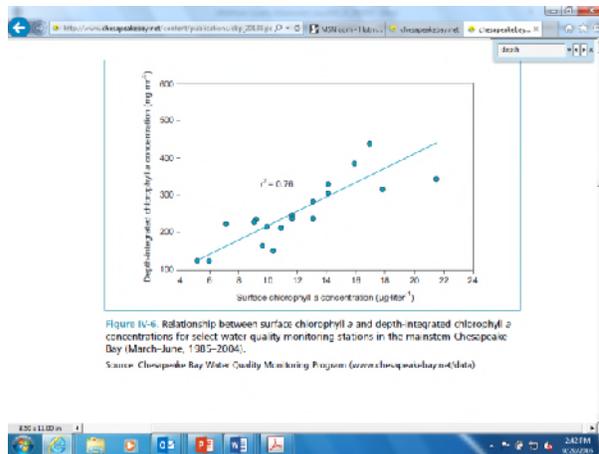


Figure 1. Surface CHLA as a predictor of water column integrated CHLA.

Further, U.S. EPA (2003) and Harding et al. (2013) found a threshold relationship between surface chlorophyll *a* measurements and blooms of the potentially toxic cyanobacteria *Microcystis*. Tango and Butler (2008) then found a positive relationship in Chesapeake Bay tidal waters between surface collected *Microcystis aeruginosa* counts and the hepatotoxin microcystin. Linking the three elements of the relationship, it has been suggested that harmful algal bloom based criteria using chlorophyll *a* thresholds could be used to protect against harmful algal bloom related water quality impairments. Ultimately, surface and depth-averaged water column chlorophyll *a* computed from vertical profiles of water quality in Chesapeake Bay and its tributaries were used in deriving recommendations for chlorophyll *a* criteria (Harding et al. 2013). During the CAP WG review of the chlorophyll *a* criteria assessment protocol, consideration was given to both measurement approaches going forward in recommending which to use in a sampling procedure.

CAP WG Discussion

The CAPWG reviewed whether there was sufficient new science to support a change from using surface measurements that support the existing chlorophyll *a* assessment methods, or, to recommend moving to a vertical profiling of chlorophyll *a* sampling basis to support a volume-based assessment protocol. The CAP WG recognizes portions of the plankton community migrate through the water column and therefore representativeness of the water column dynamics is a challenge with surface only measurements. Katherine C. Filippino (ODU) provided a presentation during the 9/7/16 CAP WG meeting covering the most current work on understanding the relationship between surface spatial and vertical depth integrated chlorophyll *a* distributions. The work showed that variability in chlorophyll *a* with depth was not well characterized for the James River. Further, the basis for chlorophyll *a* criteria re-evaluation by the recent James River Science Advisory Panel effort (2011-2015) was evaluated as a function of surface water chlorophyll relationships. The CAP WG supported a theme where criteria derivation and assessment should have the same basis. Recognizing the method of chlorophyll *a*

criteria derivation, chlorophyll *a* criteria re-evaluation, USEPA (2007) relationship illustration between surface and the more work intensive depth integrated chlorophyll *a* sampling results, the CAP WG reinforced its recommendation for surface-only monitoring approach at this time as suitable to support the assessment of the chlorophyll *a* criteria. The CAP WG further recommends that work continue on understanding the relationship between surface and depth integrated measures of chlorophyll *a*. This additional work is needed in order to support future improvements in characterizing water quality status of chlorophyll and the scientific basis to support updates to the assessment protocol.

For the future, chlorophyll *a* distribution in the water column and assessment seasons will deserve further research at the tributary level for any adjustments to the protocol. Published literature speaks to the mechanistic relationships between biological and physical forcings on hypoxia that include the relationship to phytoplankton biomass accumulation. Chlorophyll *a* remains the measure of that phytoplankton biomass. Hypoxia is a dissolved oxygen impairment with its applicable designated uses. Testa and Kemp (2014) provide this mechanistic understanding that provides insights for this future consideration in evolving criteria and protocols such that:

*“Indeed, O₂ depletion rates in Chesapeake Bay were significantly related to chlorophyll *a* deposition rates derived from sediment traps over the course of several years (Boynton and Kemp 2000), and biomarker studies indicate that sediment organic matter is most labile during the spring bloom (Zimmerman and Canuel 2001). Thus, there is strong evidence for the association of winter–spring O₂ declines with the bottom water biomass of phytoplankton over a large section of Chesapeake Bay.”*

However, Testa and Kemp (2014) go on to discuss the weakening of the relationship between winter-spring bottom water chlorophyll to summer hypoxia and present evidence that winter-spring season bottom chlorophyll conditions relate to the early portion of the hypoxia phenomenon in Chesapeake Bay while summer plankton biomass impacts the duration of mid-late summer hypoxia conditions. Therefore, there is evidence to build upon that firm relationships exist between bottom water chlorophyll and dissolved oxygen conditions that point to considering a different depth of measurement that relates to dissolved oxygen impairments. Additionally, that winter season, as well as spring and summer seasons, is part of this relationship. Further work is needed for translating this recent research into future criteria and assessment protocols with mechanistic relationships evident at the tributary scale.

2) Assessment period. Maintain a 3 year assessment period or use a different length assessment window?

Issue: The existing chlorophyll *a* criteria assessment protocol defines the water quality standards attainment assessment period as 3 consecutive years of monitoring results for the applicable seasons and its respective criteria (VA: spring and summer. DC: summer)

- **Recommendation:** Adopt longer assessment periods. 5-6 years is recommended.

Background

In defining what it means for the criteria to be met and the standard to be attained, stressor magnitude, duration, return frequency, spatial extent and temporal assessment period must be accounted for. The recommended temporal assessment period for measuring attainment of all three Chesapeake Bay criteria (i.e., for dissolved oxygen, water clarity and chlorophyll *a*) within the respective designated use habitats and at the spatial scale of the Chesapeake Bay segments (spatial extent) was for using the most recent three consecutive years of applicable tidal water quality monitoring data (U.S. EPA 2003a). However, in circumstances where three consecutive years of data are not available it was recommended that a minimum of three years within the most recent five years should be used (U.S. EPA 2003a).

Application of a three-year period in the Chesapeake Bay chlorophyll *a* water quality standards attainment assessment procedure has been consistent with the water quality status assessment period used for over a decade by the Chesapeake Bay Program partners prior to the establishment of the Chesapeake Bay water quality criteria and their subsequent adoption into water quality standards (e.g., Alden and Perry 1997, U.S. EPA 2003a). A three-year period includes some natural year-to-year variability largely due to climatic events, and it also addresses residual effects of one year's conditions on succeeding years. Two years has been considered insufficient time to assess central tendency, and four or more years has been viewed as delaying a management response to problems that may be detected. Longer periods were considered more appropriate for detecting trends than for characterizing current water quality conditions (U.S. EPA 2003a). A comparison of criteria attainment across one-, three- and five-year assessment periods confirmed the selection of three years as the appropriate temporal averaging period. Attainment levels were highly variable using single-year periods. The five year period smoothed much of the variability and resulted in little difference between one assessment period and the next (U.S. EPA 2003a).

Additional support for the 3 year assessment period is similarly provided by other U.S. EPA guidance (U.S.EPA 2003b). U.S. EPA (2003b) states "EPA guidance recommends use of a 1 in 3 year maximum allowable excursion recurrence frequency – the number of times conditions in a water are worse than those specified by the concentration and duration components of a freshwater life criterion for a toxic chemical". A key basis for this recommendation was a 1989 literature survey of over 150 studies looking at recovery rates of freshwater ecosystems from various kinds of natural disturbances and anthropogenic stressors. USEPA (2003b) states that the vast majority of macroinvertebrate and fish metric endpoints recovered in 2 years or less. Therefore, 3 years was again considered a sufficient assessment period given the majority of recovery endpoints highlighted in the literature were 2 years or less.

CAP WG Discussion

Within the CAP WG review, suggestions for longer assessment periods were supported by 1) results of simulation tests that conducted repeated chlorophyll *a* 3-year attainment assessments by applying the existing fixed station sampling protocol on known chlorophyll *a* distributions, 2) assessment periods applied to other permit assessments in Virginia, and 3) new literature on longer recovery rates of important ecosystem metrics in estuarine and coastal waterways. First,

the protocol assessment simulation tests (Perry 2015 *in* VADEQ 2016) showed that 3 years of the seasonal average criteria assessment from the fixed station monitoring protocol and interpolation of that data for mapping violation rates provided too few data points for an effective evaluation of the health of the system. Low sample site density and low (biweekly) sampling intensity during the season showed that assessments were highly biased towards finding non-attainment when the data set was known to be in attainment (Perry 2015 *in* VADEQ 2016). Therefore, it was demonstrated that there can be a high probability of considering a segment as out of attainment even when it is known to be in attainment when applying the existing long term Chesapeake Bay Program fixed station biweekly to monthly sampling and 3-year assessment protocol with chlorophyll *a*. Additional data provides more power to effectively assess the condition and therefore, longer assessment periods were recommended to improve temporal density of the seasonal mean assessments to support a more accurate assessment of status.

Second, the CAP WG noted that most of Virginia's wastewater permits are based upon 5 year cycles. It was further noted that 6 year assessments are common for many water quality parameters in Virginia. The CAP WG discussed that other states use similarly long assessment periods. While the CAP WG did not provide a particular summary of other State's assessment periods, references are added here. For example, chlorophyll *a* assessments for Georgia lakes are assessed on a 5-year basis

https://epd.georgia.gov/sites/epd.georgia.gov/files/related_files/site_page/303d_Listing_Methodology_Y2014.pdf.

Coincidentally, a common assessment methodology used in surface water quality assessments by the State of Arkansas also bases their standards attainment decisions on a five year period (Scott and Haggard 2015). However, at the far end of the spectrum, Scott and Haggard further evaluated a suite of chlorophyll *a* criteria assessment period options from 5-11 years for use in Beaver Lake, Arkansas, where the concerns are for protecting its drinking water designated use. They suggest one alternative to assessing changes in average condition in a 5-year windows may be to use a window as large as 10 years. The 10-year window suggestion was based on literature identifying decadal-scale climate patterns in chlorophyll *a* behavior in lakes (e.g. Hampton et al. 2008).

The CAP WG discussions highlighted that if there is continued use of the CFD assessment approach for determining chlorophyll *a* standards attainment, that by any means it would occur, more data points were desirable for decision support. Including more data in the use of the CFD approach more effectively tracked the hyperbolic shaped reference curves. Various efforts have been used to look at the shape of CFD curves (e.g. Dissolved oxygen deep channel instantaneous minimum curves), the most recent assessment on shape of curves was conducted during the recent critical evaluation of the chlorophyll attainment assessment method by VADEQ (see Buchanan 2015 *in* VADEQ 2016). The collective experience of the CAP WG in its practice of using the CFD for Chesapeake Bay water quality standards analyses, and in reviewing the most recent reference curve analysis, suggested that 9 to 10 data points were considered a desirable data density to support attainment decisions with this approach. Under the present assessment framework that evaluates a standard based on seasonal means, we would therefore be looking at a 9-10 year assessment window. The CAP WG did not suggest using a 9-10 year assessment

window as suggested as a possibility with scientific basis by Scott and Haggard 2015), however, it used this line of reasoning regarding data needed to support an accurate assessment when using the CFD approach to evaluate water quality standards attainment to support a recommendation for longer assessment periods. Assessment periods of 5-6 years were suggested as a compromise and in line with other lines of evidence such as the literature on ecosystem recovery rates.

Additionally regarding the issue of data density, the CAP WG recognized that there are multiple factors in play affecting this recommendation, particularly data density needs and feasibility of the assessment. With respect to data density, Elgin Perry (pers. comm. 2016) noted that in the original planning for using the CFD for chlorophyll *a* criteria assessment of the James River, that he and others anticipated that spring and summer seasons would be evaluated in a combined assessment on the same CFD. This approach would raise the number of data points from 1 to 2 per year and from 3 to 6 total within a 3-year period under a joint spring-summer season assessment. However, considering a combined spring-summer season type of adjustment to the protocol would appear to require a change in the chlorophyll *a* standards. The CAP WG was not addressing issues that were intended to change the existing standards.

Longer time windows were considered by the CAP WG as a viable means of increasing data density that would provide a more effective characterization of a segment. More samples informing the assessment further result in sample mean estimates that are nearer the true mean, reducing the standard error, and therefore uncertainty, on the estimate of the mean. However, the CAP WG also recognized the spatial domain of the assessment in reducing uncertainty. VADEQ (2016a, b) both addressed station density and the representativeness of stations. VADEQ analyses of a DATAFLOW dataset suggested that some segments have subsegment chlorophyll *a* dynamics where unsampled regions behave differently than other sampled regions of the segment. Representativeness of sites to relatively homogeneous behavior of the water quality around it are desirable in the assessment network. Under a fixed station approach, VADEQ is considering adding additional sites for improving spatial characterization of a river segment. This work is tremendously important to the final sampling strategy supporting the assessment protocol, however, the additional insights on adjusting the spatial representativeness of the sampling strategy did not change the CAP WG's consideration for the longer time period of assessment to support an attainment decision.

Finally, new literature on ecosystem recovery rates is available and highlights longer recoveries from stress that U.S. EPA (2003b). Using ecosystem recovery rates as a basis for defining an appropriate assessment period (e.g., U.S. EPA 2003b), the new literature would suggest a longer assessment period may be warranted and supported. U.S. EPA (2003b) states "EPA guidance recommends use of a 1 in 3 year maximum allowable excursion recurrence frequency – number of times conditions in a water are worse than those specified by the concentration and duration components of a freshwater life criterion for a toxic chemical". As described previously, a key basis for this recommendation that defined a decision rule within a 3-year assessment period context was a 1989 literature survey of over 150 studies looking at recovery rates of freshwater ecosystems from various kinds of natural disturbances and anthropogenic stressors. The vast majority of macroinvertebrate and fish metric endpoints recovered in 2 years or less. However, a

more recent review on recovery rates specific to estuarine and coastal ecosystems was published by Borja et al. (2010) in *Estuaries and Coasts*. Borja et al. summarized results from 51 studies used to evaluate recovery patterns as a function of various stressors. Many of the studies cited by Borja et al. (2010) pertained to systems that had experienced long-term degradation from a variety of stressors, as opposed to episodic impacts associated with a single parameter. Similar to the 1989 EPA review, some studies showed near-term (months to a few years) recoveries of certain taxonomic groups. However, the lower boundaries for the majority of studies (see Table 2 in Borja et al.) is frequently 2-3 years while central tendencies are longer, and many recoveries take 6 years or more. Following the basis for EPA's support for decision rules based on a 3 year assessment, and now using the same logic with the support of this synthesis of new research by Borja et al. (2010) would suggest longer assessment periods of 5-6 years are supported and warranted.

ASSESSMENT: INTERPOLATION, CFD and IMPAIRMENT DECISION RULE

3) Interpolation

Issue: Interpolation has been used to populate a surface of values on a 1km x 1km grid to estimate the chlorophyll *a* distribution in the unsampled space of an entire segment. With fixed station sampling, when 1 or 2 sites are present in a segment, not uncommon within Chesapeake Bay and its tidal tributaries, the interpolation is not adding any information present in a simpler interpretation of the results.

- **Recommendation** - *Discontinue application of the present application of the Chesapeake Bay interpolator for interpolation. Compute the season mean based on a mean of spatial means given that sample sites are distributed in sufficient density to be considered representative of habitats within a segment that are defined as homogenous regions. Assuming fixed site sampling, weighting of the site-means by the % spatial area representation within a segment implicitly accounts for spatial variability in a manner similar to but not identical to the inverse distance weighting interpolation approach without overinterpreting the variability of the unsampled regions.*

Background

The following overview of the interpolation method supporting the water quality standards attainment assessments is largely derived from U.S.EPA (2003a). Recognize that chlorophyll *a* assessments are surface-only as described earlier. Text that refers to volumetric calculations are a reflection of the full capacity of the interpolation approach and an extension of the surface-only assessments.

For the Chesapeake Bay and its tidal tributaries, using a grid-based spatial interpolation software provides a common spatial framework and agreed upon method of spatial extrapolation. Spatial interpolation provides estimates of water-quality measures for all

locations within a spatial assessment unit. This is accomplished at any single location by linear interpolation of the data of all its nearest neighbors. This approach provides an estimate of the water quality measure at all locations within the spatial unit being considered.

Using spatial interpolation, chlorophyll *a* concentrations were estimated for all locations in a segment. Based on those estimates, the spatial distribution of chlorophyll *a* can be shown to exhibit spatial gradients that tend to occur throughout an area. Those gradients need to be accounted for in order to accurately assess the extent of criteria exceedance.

The Chesapeake Bay Program spatial-interpolation software (or 'CBP interpolator') computes water quality concentrations throughout the Chesapeake Bay and its tidal tributaries from measurements collected at point locations or along cruise tracks (Bahner 2001). It estimates water quality concentrations at all locations in a two dimensional area or in a three-dimensional volume. The CBP interpolator is cell-based. Fixed cell locations are computed by interpolating the nearest number (*n*) of neighboring water quality measurements, where *n* is normally 4, but is adjustable. Typically an interpolation is performed for the entire Chesapeake Bay for a single monitoring event (e.g., a monthly data collection cruise). In this way all monitoring stations are used to develop a baywide picture of the spatial variation of the parameter being considered. Segment and designated use boundaries can then be superimposed over the baywide interpolation to assess the spatial variation of the parameter in any one segment's designated use(s). *The CAP WG notes, however, that interpolations supporting the water quality standards attainment assessments have been and are presently performed on a segment by segment basis. There is no cookie-cutting of baywide interpolations with a segment overlay.*

Cell size in the Chesapeake Bay was chosen to be 1 kilometer (east-west) by 1 kilometer (north-south) by 1 vertical meter, with columns of cells extending from the surface to the bottom of the water column, thus representing the three-dimensional volume as a group of equal-sized cells. The tidal tributaries are represented by various cell sizes, depending on the geometry of the tributary, since the narrow upstream portions of the tidal rivers require smaller cells to represent the river's dimensions accurately. This configuration results in a total of 51,839 cells for the mainstem Chesapeake Bay and a total of 238,669 cells for the Chesapeake Bay and its tidal tributaries.

Where volumetric assessments are involved for the water column, the CBP interpolator is tailored for use in the Chesapeake Bay in that the code is optimized to compute concentration values that closely reflect the physics of stratification. The Chesapeake Bay is very shallow despite its width and length; hence water quality varies much more vertically than horizontally. The CBP interpolator uses a vertical filter to select the vertical range of data for each calculation. For instance, to compute a model cell value at 5-meters deep, monitoring data at 5 meters are preferred. If fewer than *n* (4) monitoring data values are found at the preferred depth, the depth window is widened to search up to *d* (normally ± 2 m) meters above and below the preferred depth, with the window being widened in 0.5-meter increments until *n* monitoring values have been found for the

computation. The user is able to select the smallest n value that is acceptable. If fewer than n values are located, a missing value (normally a -9) is calculated for that cell.

A second search radius filter is used to limit the horizontal distance of monitoring data from the cell being computed. Data points outside the radius selected by the user (normally 25,000 meters) are excluded from calculation. This filter is included so that only data near a specific location are used for interpolation. In the current version of the CBP interpolator, segment and region filters have been added (Bahner 2001)

The CBP interpolator uses an inverse distance weighting (IDW) approach to interpolation and assumes a linear distribution of the data between points. Given the dynamic nature of estuaries, this is obviously a conservative assumption. However, the spatial limitations of the data make the simplest approach the most prudent. The strength of the CBP interpolator's output is directly related to the quality and spatial resolution of the input data. As sample size increases, interpolation error decreases. For more detailed documentation on the Chesapeake Bay Program interpolator and access to a downloadable version, refer to the Chesapeake Bay Program web site at <http://www.chesapeakebay.net/tools.htm>.

CAP WG Discussion

The implementation of fixed station monitoring sites in the Chesapeake Bay long term water quality monitoring network in 1984/85 were applied in a design intended to capture seasonal and annual patterns when documenting status and trends. With the advent of DATAFLOW technology to map the surface distribution of water quality conditions, high spatial density data have been used for water quality criteria assessments. The common denominator of the monitoring program, however, remains the fixed station long-term water quality monitoring program network sites. The application of DATAFLOW technology for water quality monitoring in Chesapeake Bay and its tributaries during the recent decade has provided new, extremely valuable and insightful data sets. These high spatial resolution data coupled with in-situ high temporal density continuous monitoring site data collections have provided unprecedented understanding and opportunities to assess the representativeness of the existing fixed station networks to represent the water quality of a segment.

VADEQ analyses for data from the James River used a form of clustering to assess homogeneity of chlorophyll dynamics in segments that had both DATAFLOW and fixed sites (VADEQ 2016a). The resulting analyses showed that heterogeneity of chlorophyll behavior in the tidal fresh upper James River could be characterized by subdividing the segment into 4 subregions. Further, while 3 of the subregions already have long-term fixed station monitoring sites, there is one subregion that expressed significantly different chlorophyll *a* behavior and is not represented by the Chesapeake Bay water quality monitoring program network. This is an issue that interpolation of the present fixed site data would not effectively address unless there are more sampling sites. By having a minimum of one monitoring site per subregion, computing the subregion seasonal mean provides an implicit interpolation assigning its value to the area it represents. Conducting water quality standards analysis based on the area weighted means for a season therefore continues to

support a space-time evaluation of water quality with fewer assumptions than with the application of the Chesapeake Bay interpolation software.

The cluster analysis that defines the region of representativeness for a fixed station might be viewed at odds with the finding that the simulation analysis by Elgin Perry on chlorophyll variance and station representativeness can only be extended 1-2 km around the point. However, in San Francisco Bay, Jassby et al. (1997) used high spatial density data to similarly cluster regions of similar, homogenous behavior for chlorophyll *a* and found a sampling site may represent an 8 km square area. Regarding the concern that insufficient data density with fixed station assessments can miss hot spots in space, VADEQ found that spatial density of approximately every 7 km was found to take account of spatial autocorrelation in chlorophyll *a* distributions while enhancing spatial representativeness over the existing water quality monitoring program network (September 16 CAP WG presentation by VADEQ's Arianna Johns and Tish Robertson). Therefore, within the published literature, and James River specific analyses, a subsegment size of 1-8 km² fits within the range of the scientific support for a monitoring network supportive of the non-IDW interpolated season mean standards assessment.

4. CFD

Issue: Under the present fixed station sampling design, with 1-3 stations sampled biweekly, interpolated and assessed for spatial violations on a 1km² grid, the CAP WG finds the assessment of attainment difficult to support and defend.

Further, the derivation of the criteria are based on relationships with seasonal means. The CAP WG finds that the science to date has not defined an agreed upon set of spatio-temporal criteria exceedance structures for chlorophyll *a* distributions that the CFD is an effective decision tool for deciding chlorophyll *a* attainment. Not having such reference has further made the explanation of attainment trends a challenge for managers and stakeholders.

- **Recommendation:** *Use a decision framework no longer dependent on the CFD but conceptually consistent with the mutual accounting for spatial and temporal criteria exceedances.*

Background:

The following overview is derived as background from U.S.EPA (2003a) regarding the CFD:

Per the publication of U.S. EPA (2003a), the use of cumulative frequency distributions (CFDs) is recommended for assessing spatial and temporal water quality criteria exceedance in the Chesapeake Bay. CFDs offer a number of advantages over other techniques that are applied for this purpose. First, the use of CFDs is well established in both statistics and hydrologic science. CFDs have been used for much of the past century to describe variations in hydrologic assessments (Haan 1977). For example, the U.S. Geological Survey has traditionally used CFDs to describe patterns in historical

streamflow data for the purpose of evaluating the potential for floods or droughts (Helsel and Hirsch 1992).

Second, the application of the CFD for evaluating water quality criteria attainment in the Chesapeake Bay allows for the evaluation of both spatial and temporal variations in criteria exceedance. Methods currently used for the assessment of criteria attainment are based only on temporal variations because measurements are usually evaluated only at individual monitoring station locations. One of the limitations of this approach is that it is often difficult to determine whether an individual sampling location is representative, and there is always potential for bias. In a water body the size of the Chesapeake Bay, accounting for spatial variation can be very important and in that respect, the CFD approach has been considered to represent a significant improvement over methods used in the past.

A CFD is developed first by quantifying the spatial extent of criteria exceedance for every monitoring event during the assessment period. Compiling estimates of spatial exceedance through time accounts for both spatial and temporal variation in criteria exceedance. Assessments are performed within spatial units defined by the intersection of Chesapeake Bay Program segments and the refined tidal-water designated uses and temporal units of three-year periods. Thus, individual CFDs will be developed for each spatial assessment unit over three-year assessment periods.

CAP WG Discussion

The CAP WG recognizes the value of the CFD for assessment purposes and is not critical of the CFD in concept and as a tool. It further recognizes that this approach assesses against an “allowable exceedance” of the stated criterion, in part as an accounting for the uncertainty in the assessments due to data density of the fixed station water quality monitoring program. However, the CAP WG recognizes there are a variety of mismatches with the present application of the CFD for chlorophyll *a* criteria assessment.

Challenges to the use of the CFD have focused on examining the shapes of the CFD. Ideally, a bio-reference CFD would be used as the basis for making a decision about impairment status of the waterway. However, to date, there is no agreement within the community on an appropriate reference community, and the reference curves compared to date with the 10% reference curve have been based on metrics other than a collection of season means (e.g. instantaneous values and 90th %-iles of distributions of chlorophyll *a* values in communities deemed reference by the author of those assessments, see Buchanan 2015 in VADEQ 2016a).

Further, STAC (2006), the application of a 10% reference curve is “the expected distribution of exceedance given a criterion that is derived from the 90th percentile of a reference population.” Seasonal mean chlorophyll *a* criteria are not derived from the 90th percentile of a reference population. The CFD has not been shown to describe the behavior of a waterbody being assessed

with a seasonal mean criterion. Its use, therefore, is considered a mismatch between criteria derivation, application and assessment.

Building off an earlier point on data density: more effective characterization of a segment in space and time will occur with larger samples. Larger samples in space result in sample mean estimates that are nearer the true mean and reduce the standard error on the estimate of the mean. Greater frequency in time will resolve bloom frequency and distribution and likewise better estimate mean conditions. As referenced by T. Robertson in VADEQ (2016a) referring to a Perry 2015 analyses, “as long as VADEQ continues to perform fixed station monitoring using the CFD approach, the true state of the estuary must far exceed the requirements of its designated use in order to have a high probability of correctly identifying the passing condition in order to remove it from the [impaired waters] list.” This means the sensitivity of the method under the existing monitoring network design is low requiring that water quality must be exceedingly good for the method to consistently and correctly classify a segment in attainment.

The CAP WG practitioners of the CFD have further identified a significant communications challenge of the approach and its results when working with stakeholders. In statistical testing we frequently put error bars around an estimate that reflects the uncertainty associated with subsampling a population. We may state that there is or is not a significant difference based on confidence bounds of our uncertainty with our estimates. The CFD provides an “allowable exceedance” or an “allowable failure”. The translation of the 10% default reference curve was recently described in the T. Robertson August 9, 2016 presentation to the CAP WG (see Figure 2 below):

- 1 of 3 years may exceed 10% but only by a small amount, to 11.0%, after which the seasonal mean will create an exceedance and the entire assessment period is out of attainment.
- If the one year passes the criterion, then the combination of exceedances in the other 2 years must not exceed at a rate of 4.0% and 1.4% respectively.
- If the entire 3 year period does not follow this distribution, then the system is considered out of attainment.

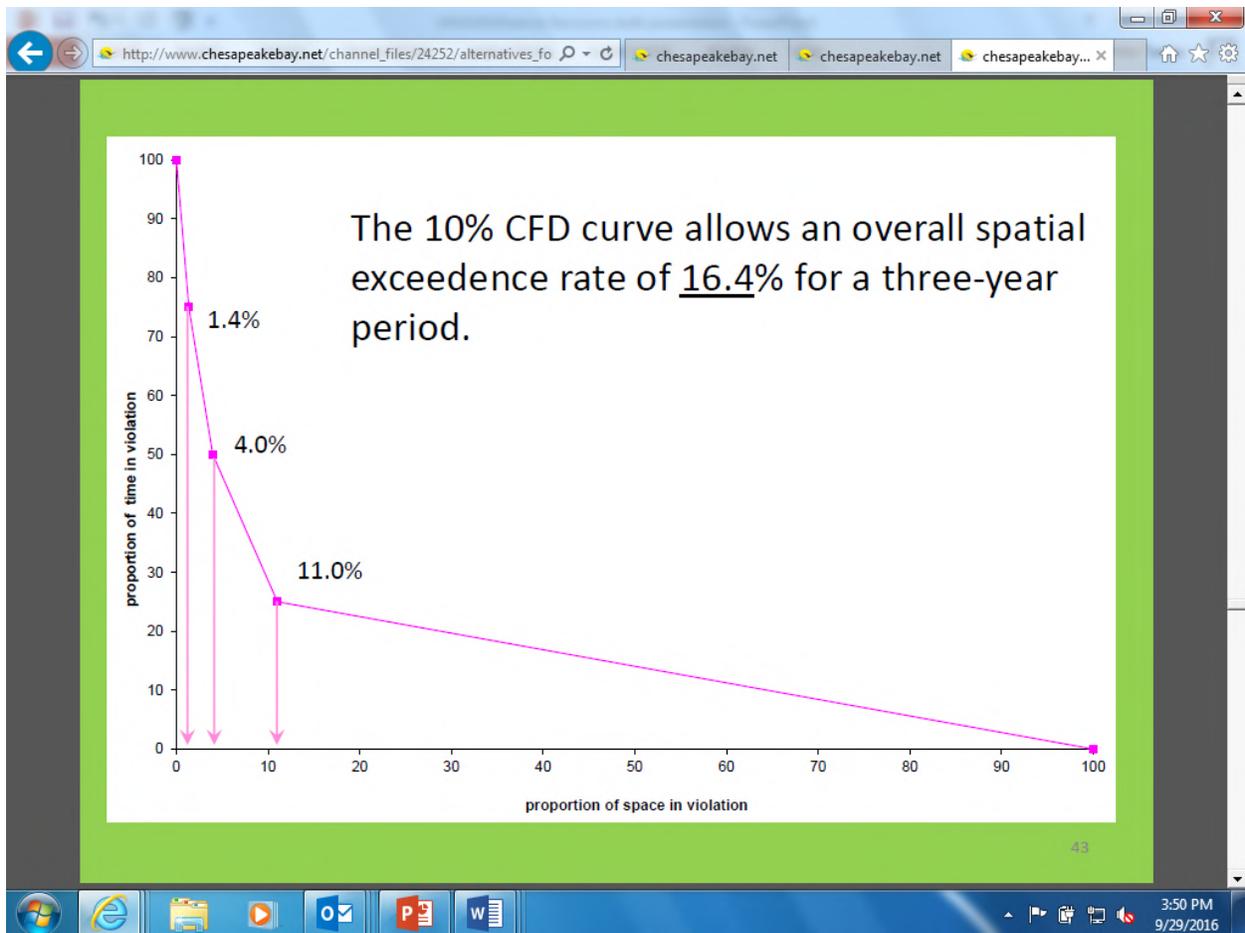


Figure 2. The assessment CFD for seasonal mean chlorophyll. Source: Robertson 2016 PPT

To date, there is no science that supports such a specific spatial and temporal distribution for a 3 year period as being in exceedance of the criteria resulting in impairment. And by using the interpolations to spatially assess exceedances, there is not yet the science available that describes the structure of the exceedances that leads to impairment. This combination of insights has made assessment by the CFD difficult to defend and communicate the nature of the chlorophyll *a* impairment.

Given the many mismatches of the application of the CFD and the present chlorophyll *a* assessment protocol, the CAP WG has suggested eliminating use of the CFD in lieu of using an alternative decision rule. The violation rates are to be based on number of years of exceedances of the season-mean criterion within the assessment period. The CAP WG recognized such an approach as a more straightforward assessment and supporting more direct communication of the results to stakeholders compared with the CFD.

5) Decision rule on attainment

Issue: The present framework relies on a 3 year assessment of seasonal means against the CFD in space and time. The CFD assessment is presently compared to a 10% reference curve. The

CAP WG reviewed many potential decision rules (Figure 3). An alternative based on the logic used by EPA in 2003 guidance (U.S. EPA 2003b) updated with recent estuarine science insights (Borja et al. 2010) is proposed.

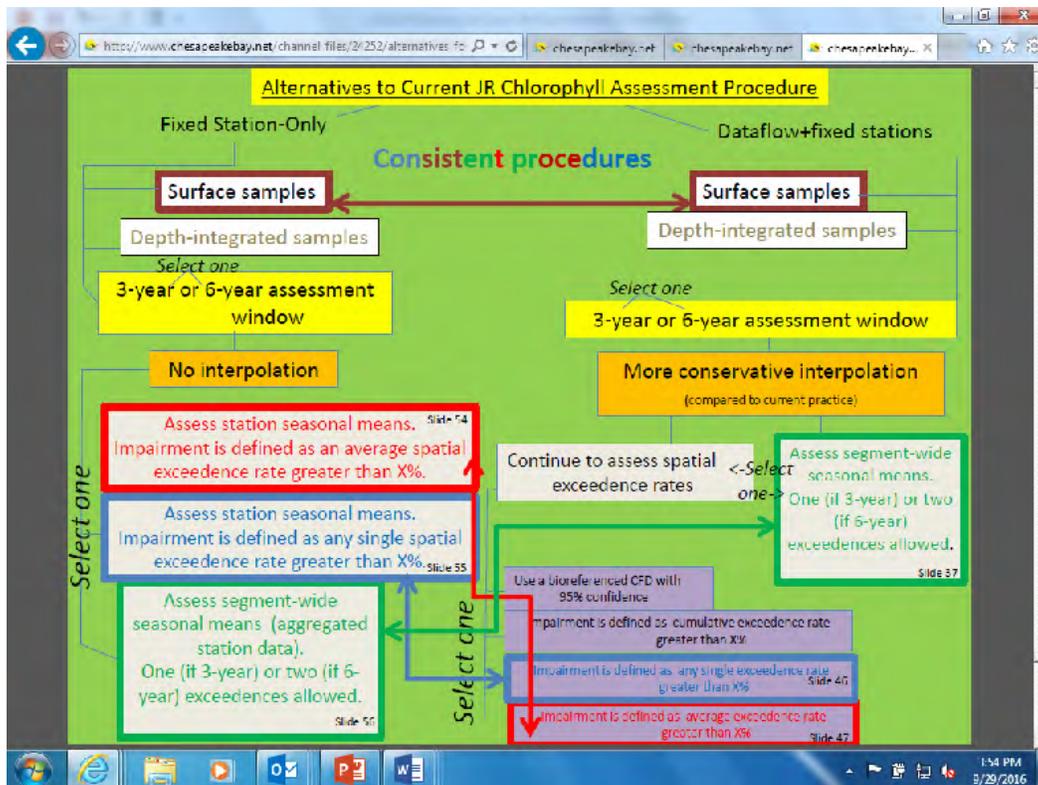


Figure 3. Many potential decision rules were considered by the CAP WG. Source: Robertson 2016 PPT.

- Recommendation:** *With the continued use of the seasonal mean criterion for chlorophyll a water quality standard attainment assessment, and under the longer assessment period of 5-6 years, the CAP WG has only settled on a range of suggested decisions. The range of recommended decision rules include 1 allowable exceedance year in 5-6 years up to a 50% allowable exceedance such that beyond 50% is considered an impairment. The CAP WG supports an annual level of decision (i.e., based on the season mean comparison to the criterion, is the segment in or out of attainment), however, the CAP WG needs feedback from the James River RAP on the proportion of years in a period that are deemed out of attainment that translates to the definition of impairment.*

Background

The following overview is derived as background from U.S.EPA (2003a) regarding the decision rule based on applications of the CFD assessment of the monitoring data such that any time the assessment curve is show to exceed the reference curve a segment would be designated out of attainment. If the curve is not exceeded, the segment would be in attainment.

Per U.S. EPA (2003a), the cumulative frequency distribution methodology for defining criteria attainment addresses the circumstances under which the criteria may be exceeded in a small percentage of instances, by integrating the five elements of criteria definition and attainment: magnitude, duration, return frequency, space and time. The methodology summarizes the frequency of instances in which the water quality threshold (e.g., dissolved oxygen concentration) is exceeded, as a function of the area or volume affected at a given place and over a defined period of time. Acceptable and protective combinations of the frequency and spatial extent of such instances are defined using a biologically based reference curve. Without a biologically based reference curve, a default 10% reference curve is used as the basis of comparing the monitoring CFD results to decide criteria attainment.

CAP WG Discussion

The CAP WG recognized multiple impairment decision options. The discussions present defensibility of the varied recommendations. The CAP WG favored a longer assessment period, however, how many violations of the season-mean criterion has been considered a subjective. Final decisions of this nature typically depend on the opinion of regulatory agencies with input from stakeholders informed by the science.

Discussions of the CAP WG with regard to the basis of a change in the decision rule recognized that the existing approach uses a 3 year period where allowable violations must meet a particular violation structure. The CFD application of attainment assessment has a 10% allowable exceedance that defines the decision-rule. This 10% spatio-temporal allowable exceedance structure, however, translates to the following: 1 year can be no more than 11% out of attainment, while a second year in the 3 year assessment period can be no more than 4% out of attainment and the third year can be no more than 1.4% out of attainment. Thus far the CAP WG community of analysts has not been able to develop seasonal mean-based bioreference curve that supports that an ecosystem is impaired or unimpaired based on such a violation rate structure. This presentation of the results is further a significant challenge to explain to stakeholders.

The present CAP WG recommended approach continues to assume the use of seasonal means for the chlorophyll *a* water quality criteria attainment assessments. The recommended approach is further based on an assumption of sampling within declared homogeneous sub-segments (based on chlorophyll *a* behavior *per the* VADEQ 2016 “Proposed criteria assessment” approach and Jassby et al. 1997 clustering assessments to support network designs). Under these proposed approach, seasonal means of subsegments represent the area of the subsegment and there is no other form of interpolation. This approach continues to pool the conditions assessment over space and time. However, violation rates are now computed as a proportion of the assessment period with “year” as the basis for decision rather than a cruise by cruise violation rate grid assessment. The change here simplifies the steps in the assessment and the explainability of the decision on attainment characterization and status of attainment.

5 to 6 year assessment period for attainment decisions

USEPA (2003b) states “EPA guidance recommends use of a 1 in 3 year maximum allowable excursion recurrence frequency – number of times conditions in a water are worse than those specified by the concentration and duration components of a freshwater life criterion for a toxic chemical”. A key basis for this recommendation was a 1989 literature survey of over 150 studies looking at recovery rates of freshwater ecosystems from various kinds of natural disturbances and anthropogenic stressors. The vast majority of macroinvertebrate and fish metric endpoints recovered in 2 years or less. However, some recoveries were incomplete even after decades. Nevertheless, EPA ORD recommended adoption of a 1 in 3 year maximum recurrence level based on the scientific literature. EPA further states that the frequency decision rule should be based on scientific rationales and other relevant information.

A more recent review on recovery rates specific to estuarine and coastal ecosystems was published by Borja et al. 2010 in *Estuaries and Coasts*. There were 51 examples used to evaluate recovery patterns as a function of various stressors. Similar to the 1989 EPA review, some studies showed nearterm (months to a few years) recoveries of certain taxonomic groups. The lower boundaries for the summary (see Borja et al. Table 2) is frequently 2-3 years with many recoveries taking minimally over 6 years. Therefore, while the conclusions based on the 1989 review pointed to a 1 in 3 allowable frequency, the Borja et al. review would suggest longer recoveries are common. Specifically with respect to eutrophication, Borja et al. classify the recovery rates at >3 to >6 years. With 1) an estuarine and coastal habitat specific set of 51 impact-recovery studies, 2) broad spectrum assessment of stressor response showing medium to long range recoveries, and 3) a recovery of typically 4 or more years specified from the effects of eutrophication, a decision rule based on an assessment period of no less than 5 years and having no more than one exceedance of the seasonal mean would appear appropriately supported compared to a 1 in 3 rule.

As previously stated, many of the studies cited by Borja et al. (2010) pertained to systems that had experienced severe, long-term degradation from a variety of stressors, as opposed to episodic impacts associated with individual seasons of elevated chlorophyll *a*. Hence, it cannot be presumed that individual years of elevated chlorophyll *a* would lead to system collapse-and recovery as represented in those studies. Actual recovery times would be related to both the nature of the impacts and the affected taxa.

A decision rule based on 1 in 5 years allowable exceedance technically translates to an allowable exceedance of 20% and not declaring an impairment unless the violation rate is 40% (2 or more years in a 5 year period). A 1 in 3 rule by comparison means a violation rate is 67% before an impairment is declared, however, the opportunity to act is nearer term than with a 5 year or longer period. A decision rule for 1 in 6 years provides for an allowable exceedance rate of 16.7% and impairment decision of 33.4%. The allowable exceedance rate with 6 years is not significantly different than the cumulative allowable spatial exceedance of the CFD (equals 16.4%, see VADEQ 2016). By comparison, the State of Georgia uses a decision rule with a 5 year assessment period for lake assessments of chlorophyll *a* where 1 in 5 years moves the lake to Category 3 and additional information is then used to evaluate attainment or impairment). For

chlorophyll *a* assessment in Beaver Lake, Arkansas, Scott and Haggard (2015) suggested one alternative to assessing changes in average condition in 5 year windows may be to use a window as large as 10 years. 1 year in 10 would be a 10% allowable exceedance.

Other decision rules were provided by the CAP WG as options with additional and valid support. For example, a “2 of 6 rule” is basically equivalent to the suggestion by U.S. EPA (2003b) for a 1 in 3 ratio of allowable impairment. Virginia currently uses a version of a 1 of 3 ratio for assessing chlorophyll *a* in lakes and reservoirs. Similarly, Florida’s extensive work on nutrient criteria resulted in adoption of a 1 in 3 ratio of allowable attainment for assessment (Florida DEP, 2012). Further, VADEQ (2016b) suggests another alternative for a 50% rule where allowable impairment is less than or equal to 50% failure of meeting the criterion as acceptable and more than 50% is impaired. CAP WG members did however express concern that more violations of the season mean that are allowed will increase the stress of algal blooms since the 90th percentile of chlorophyll *a* distributions tends to increase with increasing means. However, Scott and Haggard (2015) similarly support the 50% rule with chlorophyll *a* assessments but do consider a wide range of options as potential alternatives pending regulatory agency and stakeholder support, informed by the science.

In summary, with the continued use of the seasonal mean criterion for chlorophyll *a* water quality standard attainment assessment, and under the longer assessment period of 5-6 years, the CAP WG has only settled on a range of suggested decisions. The range of recommended decision rules include 1 allowable exceedance year in 5-6 years up to an option where 50% allowable exceedance such that beyond 50% is considered an impairment. The CAP WG supports an annual level of decision (i.e., based on the season mean comparison to the criterion, is the segment in or out of attainment?), however, the CAP WG needs feedback from the James River Chlorophyll *a* Regulatory Advisory Panel (RAP) on the proportion of years in a period that are deemed out of attainment that translates to the definition of impairment.

CRITERIA ASSESSMENT DECISIONS APPLICATION TO THE ASSESSMENT PROTOCOL

The CAP WG focused on five key issues in the water quality standards attainment procedure for chlorophyll *a*. VADEQ (2016b) developed a support document on a new proposed assessment. The CAP WG has not vetted the full details of the overall procedure. The following section highlights where CAP WG decisions dovetail with the methods while noting where final decisions may be beyond the scope of the CAP WG review.

1. Applying the assessment: Chesapeake Bay Segmentation.

- To date, the James River and DC tidal waters assessments for chlorophyll *a* have occurred at the segment level.
- Under the proposed revised protocol, CAP WG supported a subsegmenting of existing segments for assessment purposes. VADEQ (2016b) *Proposed Assessment Methodology for James River Chlorophyll Criteria* highlights U.S. EPA (2005) that supports States subdividing complex waters into discrete assessment units to maintain homogeneity

within the boundary conditions of the unit. Chesapeake Bay assessment of bay grasses has applied subsegmentation to their assessments demonstrating precedence for subsegmenting assessments supported by U.S. EPA. The best available explanation of the proposed subsegmenting process is what has been presented to the CAP WG by T. Robertson and is summarized in VADEQ (2016b).

2. Sampling:

- Surface sampling remains the CAP WG supported method. The CAP WG supports research into applying water column measures for criteria development and assessment.
- The present VADEQ funding supports biweekly monitoring in the summer at fixed stations, however, additional sources of information may support greater temporal or spatial resolutions (e.g. HRSD DATAFLOW, Citizen-based monitoring).
 - Greater fixed station density has been proposed for some segments to improve representativeness of spatial behavior of chlorophyll dynamics in the James River segments.
 - DATAFLOW data will continue to be used as available.

3. Computing the season mean for comparison with the respective season criterion.

- Supported by U.S. EPA (2010) and outlined in VADEQ (2016b), the measure of comparison for the water quality standard remains a geometric season mean.
- The CAP WG did not discuss how the season mean should be derived from the data.
- The best treatment of this computation is now in VADEQ (2016b). VADEQ (2016b) highlights how the U.S. EPA (2010) assessment of chlorophyll a distributions were assessment of temporal data; the statistical form of spatial data distributions, and in particular those of the James River, were not evaluated in U.S. EPA (2010) or any previous U.S. EPA documentation. The VADEQ (2016b) document reports on James River spatial chlorophyll a distributions following a normal distribution more than a log normal distribution. Therefore, the cruise specific means for a segment are suggested to reflect normal statistics.
- VADEQ recommends computing cruise specific measures of central tendency based on medians. If there are only 2 fixed stations in the segment, the cruise specific mean equals the median.
- VADEQ (2016b) continues to target a final step consistent with the existing criteria attainment assessment that generates a seasonal geometric mean for a segment, in this case, computed on the medians for the season assessment period.

4. Decision rule

- The CAP WG has not agreed on a specific rule for declaring a water body impaired.
- The CAP WG supports an annual-base for decisions (i.e., based on the season mean comparison to the criterion, is the segment in or out of attainment when some % of non-attaining years is met).
- With the continued use of the seasonal mean criterion for chlorophyll a water quality standard attainment assessment, and under the longer assessment period of 5-6 years, the CAP WG has only settled on a range of suggested decisions. The range of recommended decision rule options include:

- 1 allowable exceedance-year in 5-6 years
- 2 allowable exceedance-years where 3 or more years of non-attainment translate to a definition of impairment
- 50% or less of years impaired is still in attainment while greater than 50% of years impaired is out of attainment.
- The CAP WG needs feedback from EPA on the proportion of years in the longer assessment period that are deemed out of attainment as a translation to support a definition of chlorophyll *a* impairment.

LITERATURE CITED

Alden, R. W. III and E. S. Perry 1997. *Presenting Measurements of Status: Report to the Chesapeake Bay Program Monitoring Subcommittee's Data Analysis Workgroup*. Chesapeake Bay Program, Annapolis, Maryland.

Arianna Johns and Tish Robertson. 2016. VADEQ PPT from September 21 CAP WG meeting. Available upon request from the author.

Bahner, L. 2001. The Chesapeake Bay Interpolator. User guide for the Chesapeake Bay and tidal tributary interpolator. Software version 4.6, August 2006.

Borja, A., D. M. Dauer, M. Elliott and C.A. Simenstad. 2010. Medium and long-term recovery of estuarine and coastal ecosystems: patterns, rates and restoration effectiveness. 33:1249-1260.

Florida Department of Environmental Protection. 2010. Technical Support Document: Development of Numeric Nutrient Criteria for Florida Lakes, Spring Vents and Streams. 227 pp.

Haan, C.T. 1977. *Statistical Methods in Hydrology*. Iowa State University Press. Ames, Iowa. 378 pp.

Helsel, D. R. and R. M. Hirsch. 1992. *Statistical Methods in Water Resources*. Elsevier Science Publishing Company, Inc. New York. 522 pp.

Hampton, S.E., L.R. Izmet'eva, M.V. Moore, S.L. Katz, B. Dennis, E.A. Silow. 2008. Sixty years of environmental change in the world's largest freshwater lake – Lake Baikal, Siberia. *Global Change Biology* 14:1947-1958. <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2486.2008.01616.x/full>

Harding, L.W. Jr., 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.J. Tango. 2013. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. *Estuaries and Coasts*. Published online 18 June 2013. DOI 10.1007/s12237-013-9656-6.

Katherine C. Filippino (VIMS). 2016. PPT for September 7, CAP WG meeting. Available upon request from the author.

Malone, T.C., L.H. Crocker, S.E. Pike, and B.W. Wendler. 1988. Influences of river flow on the dynamics of phytoplankton production in a partially stratified estuary. *Marine Ecology Progress Series* 48: 235–249.

Robertson, T. 2016. PPT. August 9 2016 CAP WG.

http://www.chesapeakebay.net/channel_files/24252/alternatives_for_chlprotocol_072816.pdf

Scott, J.T. and B.E. Haggard. 2015. Evaluating the assessment methodology for the chlorophyll-*a* and Secchi transparency criteria at Beaver Lake, Arkansas. Prepared for the Beaver Watershed Alliance. White paper. <http://www.beaverwatershedalliance.org/pdf/assessment-methodology-final-report.pdf>

STAC (Scientific Technical Advisory Committee). 2006. The Cumulative Frequency Diagram Method for Determining Water Quality Attainment: Report of the Chesapeake Bay Program STAC Panel to Review of Chesapeake Bay Program Analytical Tools. STAC Publication 06-003. 74 pp.

Tango, P. and W. Butler. 2008. Cyanotoxins in tidal waters of Chesapeake Bay. *Northeast Naturalist*. 15(3):403-416.

Testa, J.M. and W.M. Kemp. 2014. Spatial and temporal patterns of winter-spring oxygen depletion in Chesapeake Bay bottom water. *Estuaries and Coasts*. DOI 10.1007/s12237-014-9775-8.

U.S. EPA (U.S. Environmental Protection Agency). 2003a. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries (Regional Criteria Guidance)*. April 2003. EPA 903-R-03-002. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA (2003b). Guidance for 2004 assessment, listing and reporting requirements pursuant to sections 303d and 305b of the Clean Water Act. July 21, 2003. Watershed Branch, Assessment and watershed Protection Division, Office of Wetlands, Oceans and Watersheds, Office of Water, United States Environmental Protection Agency.

U.S. EPA (U.S. Environmental Protection Agency). 2004. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2004 Addendum*. October 2004. EPA 903-R-04-005. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA (U.S. Environmental Protection Agency). 2005. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act. U.S. Environmental Protection Agency, Office of Water, Office of Watershed, Oceans, and Wetlands, Assessment and Watershed Protection Division, Washington, DC.

U.S. EPA (U.S. Environmental Protection Agency). 2007a. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2007 Addendum*. July 2007. EPA 903-R-07-003. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA (U.S. Environmental Protection Agency). 2007b. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries –Chlorophyll a Addendum*. October 2007. EPA 903-R-07-005. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA (U.S. Environmental Protection Agency). 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*. September 2008. EPA 903-R-08-001. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA (U.S. Environmental Protection Agency). 2010. *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries – 2010 Technical Support for Criteria Assessment Protocols Addendum*. May 2010. EPA 903-R-10-002. Region III Chesapeake Bay Program Office, Annapolis, MD.

VADEQ. 2016a. VADEQ Critical Review of the Assessment Methodology for James River Chlorophyll Virginia Department of Environmental Quality. White paper. http://www.chesapeakebay.net/channel_files/24252/critical_review_of_jr_chlorophyll_assessment_methodology.pdf

VADEQ. 2016b. Proposed Assessment Methodology for James River Chlorophyll Criteria. White paper. http://www.chesapeakebay.net/channel_files/24252/2016proposed_jr_chlorophyll_procedure.pdf

VADEQ. Robertson, T. 2016. Power Point (PPT). August 9 2016 CAP WG. http://www.chesapeakebay.net/channel_files/24252/alternatives_for_chlprotocol_072816.pdf