Dissolved Oxygen Criteria Attainment Assessment

DRAFT PT JUNE 2013. Chapter 1

Assessing Dissolved Oxygen Criteria Attainment:

A Focus on Short-Duration Criteria Attainment Assessments

BACKGROUND

Early in the 1990s, experts first identified DO concentrations necessary to protect the Chesapeake Bay's aquatic living resources. With the publication of the *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for Chesapeake Bay and its tributaries* (U.S. EPA 2003a), a suite of dissolved oxygen criteria was codified supporting living resources growth and survival protections across life stages using multiple space and time scales (Table 1). There was recognition that the temporal scale of the Chesapeake Bay long term, fixed station tidal water quality monitoring program was designed to capture long-term trends, as well as seasonal and interannual variation in water quality conditions. The existing program would support 30-day mean assessments, however, it was considered insufficient on its own to assess short-duration dissolved oxygen criteria (U.S. EPA 2003a, CBP-STAC 2012 Umbrella Criteria Report), i.e., it was poorly suited for supporting Clean Water Act 303d listing assessments of the new water quality criteria that included 7-day mean, 1-day mean and instantaneous minimum DO criteria (p.177, U.S. EPA 2003a). "Short-duration" as defined here will refer to a criterion with a temporal scale of less than the 30-day mean dissolved oxygen criterion used to support assessments of water quality standards in the Chesapeake Bay. In U.S. EPA (2003a), it was suggested that assessment of short-duration criteria might be accomplished using statistical methods that estimate probable attainment (p.179). Further evaluation of water quality monitoring and assessment options to support measurements of Chesapeake Bay short-duration water quality standards attainment was needed.

The Chesapeake Bay tidal water quality monitoring network has undergone adaptive changes during almost three decades coincident with the ongoing evolution of the Chesapeake Bay Program partnership's information requests for supporting their management needs (CBP 1989a, 1989b, USEPA 2003a, MRAT 2009). Each component of the tidal monitoring network has been designed to support the four Bay jurisdictions' tidal water Bay Clean Water Act section 303d listing decisions for dissolved oxygen, water clarity/submerged aquatic vegetation, chlorophyll *a* and benthic infaunal community-based impairments (USEPA 2003a, 2004a, 2007a, 2007b, 2008a, 2010a, TMDL 2010). The Chesapeake Bay tidal water quality monitoring network has included the following:

Table 1. Chesapeake Bay Water Quality Criteria (modified presentation from USEPA 2003a).
Gray box text indicates the criteria being assessed by the existing CBP monitoring program.
Yellow box text indicates criteria assumed to be covered in the less critical seasons outside of the summer season.

Designated Use	**Criteria Concentration/Duration**	**Protection Provided**	**Temporal Application**
Migratory fish spawning and nursery use	7-day mean ≥ 6 mg/L (tidal habitats with 0-0.5 salinity)	Survival/growth of larval/juvenile tidal-fresh resident fish; protective of threatened/endangered species	February 1-May 31
	Instantaneous minimum ≥ 5 mg/L	Survival and growth of larval/juvenile migratory fish; protective of threatened/endangered species	
	Open-water fish and shellfish designated use criteria apply		June 1 – January 31
Shallow-water bay grass use	Open–water fish and shellfish designated criteria apply		Year-round
Open-water fish and shellfish use[1]	30-day mean ≥ 5.5 mg/L (tidal habitats with ≤0.5 salinity)	Growth of tidal-fresh juvenile and adult fish; protective of threatened/endangered species	Year-round
	30-day mean ≥ 5 mg/L (tidal habitats with >0.5 salinity)	Growth of larval, juvenile and adult fish and shellfish; protective of threatened/endangered species	
	7-day mean ≥ 4 mg/L	Survival of open-water fish larvae.	
	Instantaneous minimum ≥ 3.2 mg/L	Survival of threatened/endangered sturgeon species[1]	
Deep-water seasonal fish and shellfish use	30-day mean ≥ 3 mg/L	Survival and recruitment of bay anchovy eggs and larvae.	June 1 – September 30
	1-day mean ≥ 2.3 mg/L	Survival of open-water juvenile and adult fish	
	Instantaneous minimum ≥ 1.7 mg/L	Survival of bay anchovy eggs and larvae	
	Open-water fish and shellfish designated-use criteria apply		October 1 – May 31
Deep-channel seasonal refuge use	Instantaneous minimum ≥ 1 mg/L	Survival of bottom-dwelling worms and clams	June 1 – September 30
	Open-water fish and shellfish designated use criteria apply		October 1 – May 31

- Tidal water quality monitoring sampling for 26 parameters at over 150 long-term, fixed-site stations distributed over 92 Chesapeake Bay tidal management segments across Delaware, the District of Columbia, Maryland and Virginia
- Shallow-water monitoring assessing water quality annually, within subsets of the 92 Bay segments, on a rotational basis

- Living resource assessments (e.g. underwater Bay grasses, benthic infaunal communities, phytoplankton communities, zooplankton communities, fisheries independent population monitoring programs)

The Chesapeake Bay Program long-term tidal water quality monitoring program is designed to represent the complexities of the estuary. Every 2-4 weeks, a three dimensional view of the Bay water quality is obtained by sampling various depths from the surface to the bottom of the water column at each station. Each of the Bay's 92 management segments has one or more monitoring stations. Sites are scheduled for sampling at least once a month. Standardized sampling and analytical methods are used and were approved by EPA in 1986 and are still used today (TMDL 2010).

AN EXPLORATION OF SHORT-DURATION DISSOLVED OXYGEN CRITERIA ASSESSMENT APPROACHES TO INFORM CHESAPEAKE BAY WATER QUALITY MONITORING AND ASSESSMENT STRATEGY OPTIONS

Spectral analysis and logistic regression were two statistical methods previously highlighted for consideration in supporting Chesapeake Bay short-duration dissolved oxygen criteria attainment assessments (U.S. EPA 2004, 2007a). An alternative decision framework was developed called the "Umbrella Criterion" approach. The "Umbrella" approach suggested that measuring and attaining a single criterion could be used to ensure protection for one or more other scales of the associated criteria. In practice then, the application of an Umbrella Criterion assessment for protecting short-duration dissolved oxygen criteria in Chesapeake Bay tidal waters could be based on one measured criterion such as the 30-day mean. The two statistical methods and the Umbrella Criterion approach have been explored by Chesapeake Bay community analysts evaluating 1) the technical capacity to apply the techniques with available Bay monitoring data and 2) the appropriateness of the method to fulfill dissolved oxygen criteria assessment needs.

The logistic regression approach utilizes a well established statistical procedure and has been applied to Chesapeake Bay data (U.S. EPA 2007). Results were used to analyze relationships of mutual protections for water quality targets across multiple time scales (Jordan et al. 1992). Jordan et al. (1992) concluded that knowing the seasonal mean DO concentration for a given region of the Bay permitted "a good estimate of what proportion of actual DO observations are likely to meet, or fail to meet," target dissolved oxygen concentrations such as the instantaneous minimum. This approach conceptually reflected an umbrella criterion-style assessment method where attaining a criterion for one time scale provided an additional, implicit protection against violating a criterion assessed at one or more other time scales. U.S. EPA (2007a Appendix E) detailed advances and limitations in applying the logistic regression approach for Chesapeake Bay water quality standards assessments. Additional support was expressed for the continued development and eventual application of statistical approaches to assess short-duration dissolved oxygen criteria. The work of the Umbrella Criteria Assessment Team focused on advancing the use of spectral analysis to address the gaps in short-duration criteria attainment assessments and no further advances were made by the team in applying logistic regression solutions to this issue.

Spectral analysis is a time series approach and was considered an alternative to logistic regression for potentially filling gaps in short-term DO criteria attainment assessments supporting Chesapeake Bay 303d water quality standards listing and delisting decisions (U.S. EPA 2007). The foundation for using spectral analysis to estimate DO behavior at short time scales was published by Neerchal et al. (1992). The approach provided an analytical method for developing an estimate of high frequency DO behavior at a location where only low frequency measurements were available. Spectral analysis, therefore, provided a potential option to effectively combine two types of water quality monitoring data: high frequency (temporally dense, e.g. sub-hourly to sub-daily) and low frequency (temporally sparse, e.g. biweekly to monthly) DO measurements. The resulting time series was considered to leverage the power of the two types of DO behavior information available and enhance the power of the monitoring network for supporting DO criteria attainment assessments. An early caveat to thoroughly testing the method for Bay assessment applications, however, was a lack of season long, temporally dense water quality monitoring data sets. With the advent of the Chesapeake Bay Program Shallow-water monitoring network in 2004, dozens of shallow-water sensors have now been used throughout Chesapeake Bay and its tidal tributaries to generate many season-long to multi-season length data records. Well over a million observations are now available tracking water quality behavior from mostly nearshore and some offshore monitoring sites located across tidal fresh to polyhaline habitats. Bay analysts demonstrated the technical feasibility of conducting spectral analysis to create temporally dense ‘synthetic’ DO time series using the Chesapeake Bay Program DO data (U.S. EPA 2003a, 2007a, MRAT 2009). The method was modified for use by Bay analysts and termed “Spectral casting” (CBP-STAC 2012). With the new, robust sets of data available, a more complete evaluation of the spectral casting application to dissolved oxygen criteria assessments was pursued by Chesapeake Bay Program partner analysts during the Umbrella Criteria Assessment (CBP-STAC 2012).

The **Umbrella Criteria Concept** was explored as an alternative to, or a complement of, adopting the statistical approaches to overcome gaps in assessing short-duration DO criteria in Chesapeake Bay. The idea of an umbrella criterion was borrowed from conservation biology’s use of the term “umbrella species”, first used by Wilcox (1984) and with additional applications over recent decades (Launer and Murphy 1994, Roberge and Per Angelstam 2004). Some scientists have found that the umbrella effect provides a simpler way to manage ecological communities (e.g., Dunk et al. 2006). Specific to Chesapeake Bay water quality criteria assessments for Clean Water Act water quality standards evaluations then, the single most protective DO criterion being measured was termed an “Umbrella Criterion”. The condition of mutual criteria protection for multiple criteria by a single measured criterion meeting its standards threshold then was termed the “Umbrella Criterion Assumption”. The Umbrella Criterion Assumption surmises that attainment of one dissolved oxygen criterion can serve as an "umbrella" assessment protective of the remaining dissolved oxygen criteria in a designated use. Demonstrating support for the application of the Umbrella Criterion Assumption using Chesapeake Bay water quality data could simplify assessment of multi-tiered dissolved oxygen water quality standards in Chesapeake Bay. Supporting evidence was needed to show that applying an umbrella approach can be used to effectively and simultaneously assess multiple criteria protections with a single DO assessment result.

An initial assessment of multi-scale protection for Chesapeake Bay DO criterion attainment was conducted using multiple lines of evidence and showing a strong but not exclusive relationship between achieving or not achieving criteria for the 30-day mean and 7-day mean (U.S. EPA 2004). U.S. EPA (2004) findings further recommended that for a majority of Chesapeake Bay Program management segments, dissolved oxygen concentration data collected through the monthly to biweekly sampling across the Chesapeake Bay long-term water quality monitoring program fixed-station network could be used to assess attainment of all higher frequency (i.e. short-duration) dissolved oxygen criteria. At that time, however, the Chesapeake Bay Program tidal Bay partners did not adopt this umbrella criterion approach into their water quality standards for addressing assessment of the short-duration dissolved oxygen criteria. Further demonstration and quantification of the umbrella approach to fully support the Clean Water Act Chesapeake Bay dissolved oxygen criteria assessments would be needed.

In the course of developing the Chesapeake Bay Total Maximum Daily Loads (TMDL), analysts at the USEPA's Chesapeake Bay Program Office (CBPO) conducted an assessment of how well DO criteria that are already measured with the current Chesapeake Bay long term water quality monitoring program mutually protected the unmeasured, short-duration criteria (Shenk and Batiuk 2010). Using hourly output from a calibration run of the Chesapeake Bay Water Quality Sediment Transport Model (WQSTM), the CBPO analysts produced a summer season evaluation of the Umbrella Criterion Assumption. Note that for the purposes of developing the Chesapeake Bay TMDL, the summer season (June – September) is assumed to be the limiting season in all designated uses being assessed for DO impairment (i.e. Open Water, Deep Water and Deep Channel). CBPO analysts determined that evaluation of the 30-day mean DO criteria was sufficient to determine attainment of the open-water and deep-water designated uses of the Bay (Shenk and Batiuk 2010). Furthermore, in segments containing a Summer Deep Channel designated use (8 of the 92 segments in Chesapeake Bay), non-attainment rates of the summer instantaneous minimum DO criterion for the Deep Channel were higher than for any other criterion in the Open Water and Deep Water designated uses of the same segment. *Thus, the criteria currently being assessed by the Chesapeake Bay long term water quality monitoring program appear to be "**umbrella criteria**" – the most restrictive of all available criteria protective of the full range of criteria by designated use*.

These findings can have significant implications for monitoring and assessment of the full suite of dissolved oxygen water quality standards applicable to the Chesapeake Bay's tidal waters. Chesapeake Bay Program partners further requested additional testing of the concept using

Chesapeake Bay Program water quality monitoring data to validate the model-based results. The request led to the 2010-2012 “Umbrella Criteria Assessment” process. The Umbrella Criteria Assessment Team, a group of Chesapeake Bay community analysts, was charged with providing further tests of the umbrella concept as well as any of the previously mentioned statistical approaches, and providing options and recommendations toward supporting assessments of short-duration dissolved oxygen standards attainment (CBP-STAC 2012).

CHESAPEAKE BAY WATER QUALITY DATA SUPPORTING DEVELOPMENT AND TESTING OF SHORT DURATION DISSOLVED OXYGEN CRITERIA ASSESSMENTS

Quality assured, quality controlled water quality data sets were targeted by the Umbrella Criteria Assessment Team to conduct their method evaluations (Table 2). The nearly three decades-long Chesapeake Bay Program long-term water quality monitoring network data set formed the foundation of the low frequency monitoring data needs. During the U.S. EPA (2004) analyses of criteria protections, the temporally dense, high frequency monitoring data sets were largely limited to U.S. EPA EMAP short-term buoy deployments (Table 2). At that time, season-long continuous dissolved oxygen monitoring data sets from tidal waters of Chesapeake Bay were not widely available. The focus on high frequency dissolved oxygen data collection was on the threshold of being incorporated into the new, shallow-water focused network of stations in an expanded Bay monitoring framework for the Chesapeake Bay Program. In 2004, the Chesapeake Bay Program formalized this monitoring network expansion in the tidal waters of the Bay and invested in what is now known as the Shallow-water Monitoring Program. During the 2000s, Federal, State and local agencies along with academic institutions further made investments into nearshore and offshore water quality monitoring technologies. The application of the new technologies have provided near-realtime data collection capabilities and produced water quality time series with temporally dense dissolved oxygen measurements at fixed depth or in vertical profile. Local level pilot studies using towed sensors behind a boat that can swim up and down through the water column have provided an early demonstration of alternative monitoring opportunities for collecting high temporal resolution water quality observations across space and for multiple depths (Table 2).

Table 2. Data sources serving the Umbrella Criterion Assumption analyses.

Program Description	Data Collection and Availability	Sampling Locations and Habitats
CBP long-term water quality monitoring program: Low temporal frequency and spatial spatial resolution, good vertical profile resolution of the data collection.	1985-present. Biweekly to monthly sampling. Water column profiles taken with grab samples and sensors. Web accessible data: *CBP CIMS* accessible.	Fixed site, mid-channel, approximately 178 stations. Covers tidal fresh to polyhaline habitat conditions.
USEPA EMAP: Historical short-term buoy deployments with high temporal frequency at a station. Single depth sensor evaluations.	Mix of short term (days to weeks) time series with high temporal frequencies by sensor. See USEPA (2004).	Fixed site, off shore locations, varied depths. Tidal fresh to polyhaline habitat conditions.
CBP Shallow Water Monitoring Program, **Con**tinuous **Mon**itoring (CONMON): High temporal frequency at moored locations.	Approximately 2000-present. Mostly seasonally, near continuous (15 min interval) time series April-October. Fixed depth sensor, usually 1m off bottom. Web accessible data: *Eyes on the Bay* in MD, *VECOS* in Virginia.	Fixed site, shallow water, nearshore locations, approximately 70 sites Baywide with 1-9 yrs of data. Tidal fresh to mesohaline conditions.
VIMS, MD DNR Vertical Profilers: High temporal frequency in 2 dimensions. **VIMS**: Bottom sonde .	Approximately 2006-present. Limited seasons. Sensors provide water column profiles at sub-daily scales. Bottom sonde. Web accessible data: MD DNR and VADEQ.	Fixed sites (n<5), offshore locations in MD (Potomac River) and VA (York and Rappahannock Rivers). Dominantly mesohaline lower tidal tributary data.
CBP Shallow Water Monitoring Program, surface water quality mapping with DATAFLOW: High Spatial resolution along temporally dense collection track.	Approximately 2000-present. Biweekly to monthly mapping assessments within April-October season. Multi-year assessments (3 yr sets). Sensor 0.5m below surface Web accessible data: *Eyes on the Bay* in MD, *VECOS* in Virginia.	Chesapeake Bay Program management segments. Approximately 40 of 92 segments assessed to date. Tidal fresh to polyhaline habitats.
VIMS Volumetric Assessment with ACROBAT (towed sensor underwater at variable depths). High spatial resolution -	Approximately 2003-present Limited seasons. 3-dimensional sensor assessment of water column water quality. *VIMS data*, Brush et al.	York and Rappahannock Rivers (VA) study sites, deep water reaches. Dominantly mesohaline habitat.

METHODS

ANALYSES SUPPORTING THE EVALUATION OF THE UMBRELLA CRITERION CONCEPT FOR APPLICATION IN CHESAPEAKE BAY SHORT-DURACTION CRITERIA ASSESSMENT

The issues to address gaps in short-duration criteria assessment under the existing Chesapeake Bay water quality criteria assessment framework included 1) testing methods for using high frequency data to supply new information on dissolved oxygen behavior at long term water quality monitoring locations, 2) testing the assumption of simultaneous criteria protections using Bay data, and 3) testing the effect of sampling variability on the Umbrella Effect. The following approaches have occasionally been piloted or suggested from earlier work (U.S. EPA 2003a, 2004, 2007a) while other approaches were new investigations conducted by the Umbrella Criteria Assessment Team.

SPECTRAL CASTING AS A METHOD FOR GENERATING HIGH FREQUENCY DATA AT MONITORING STATIONS WITH LOW FREQUENCY WATER QUALITY MONITORING.

The synthesis of a temporally dense, high frequency dissolved oxygen data set (e.g., every 15 minutes to 1 hour time step) for a low frequency, fixed station monitoring location (e.g., tidal water, mid-channel Chesapeake Bay Program monitoring network sites, generally biweekly sampling in the summer season) is one potentially new step in the Chesapeake Bay dissolved oxygen criteria assessment process to support short-duration criteria assessments. Season-long, high frequency vertical profiles of water quality in offshore habitats of Chesapeake Bay are rare (e.g. VIMS York and Rappahannock River locations, MD DNR Potomac River). In lieu of not having high frequency vertical profile measurements of water quality conditions for most regions of the Bay and its tributaries, an estimated time series could potentially help fill the gap in short time-scale dissolved oxygen assessment needs. Mid-channel locations without high frequency monitoring profiling technology could now be linked in the spatial assessments of water quality using temporally dense measurements of nearshore water quality conditions, a complement to these coincidently available continuous monitoring data records. The resulting data sets could be used to interpolate the dissolved oxygen patterns from shoreline to shoreline or shoreline to management segment boundary. Enhanced temporal resolution would provide decreased uncertainty in time and improve spatial resolution for estimating patterns in dissolved oxygen concentrations giving the Chesapeake Bay partnership the capacity to better assesses short-duration criteria. The approach was recommended for further evaluation (U.S. EPA 2003a, 2007a).

Between 2010 and 2012, the Umbrella Criteria Assessment Team conducted analyses validating a modified use of spectral analysis, i.e. spectral casting, for developing high frequency dissolved oxygen time series at monitoring sites where only low frequency (biweekly to monthly) monitoring data exists in Chesapeake Bay and its tidal tributaries (Appendix 1). The approach provides a method to statistically transfer information about the variation in dissolved oxygen behavior at short time scales from a location with high frequency measurements (e.g. nearshore, continuous monitoring stations) to fill in or estimate dissolved oxygen behavior at a different location where measurements are more temporally sparse.

Three elements of the spectral casting method were evaluated and validated: 1) statistical methods to pass information about water quality behavior between monitoring sites, 2) assessing the ability of the new, estimated dissolved oxygen data to match actual, high frequency measures of dissolved oxygen at a near-surface depth in an offshore location, and 3) assessing the accuracy of spectral casting estimates of dissolved oxygen behavior in matching high frequency, vertical water column profiles of water quality conditions at offshore locations. Addressing the first element involved evaluating multiple statistical approaches for passing information about dissolved oxygen behavior from a temporally dense, high frequency data location (e.g. a continuous monitoring sensor) over to a temporally sparse data location (e.g., Chesapeake Bay tidal monitoring network mid-channel monitoring stations). Applying the process creates the new, estimated high frequency dissolved oxygen records at the temporally sparse data locations (U.S. EPA 2003a, 2007, Appendix 1). The temporal interpolation methods evaluated for this potentially new step in the criterion attainment assessment process were Fast Fourier Transformation (FFT), cubic spline and linear approaches. Benefits of FFT interpolation is that it is computationally fast, allows cycle trimming, deals with cyclical prediction and preserves autocorrelation structure in the data. Limitations to FFT include meeting assumptions of cyclical behavior, a need for equally spaced inputs in time and equally spaced outputs. By comparison the cubic spline and linear interpolation approaches had fewer implementation constraints.

The second and third elements of spectral casting evaluations by the Umbrella Criteria Assessment Team used the new, estimated (i.e. 'synthesized') high frequency dissolved oxygen concentrations time series to compare with the actual time series of the offshore Chesapeake Bay monitoring locations. Offshore, high frequency time series measurements of water quality in vertical profile are rare. However, a small number of highly valuable, high frequency vertical profiler stations collecting continuous water quality measurements have provided the support needed to inform the assessment results (e.g York River, Rappahannock River, Potomac River). Comparisons of dissolved oxygen violation rates were made between synthesized and measured time series, and sources of uncertainty in estimating the offshore time series for dissolved oxygen concentrations were assessed.

STATISTICAL ASSESSMENT OF HIGH FREQUENCY SHALLOW-WATER CONTINUOUS MONITORING DATA TO INFORM CRITERIA PROTECTION ACROSS TIME SCALES

Water quality standards nonattainment rates were computed directly from nearshore shallow water monitoring continuous monitoring data records or high frequency offshore vertical water quality profiler data using the range of Chesapeake Bay dissolved oxygen criteria attainment thresholds (Table 3). Additional details of methods are in Appendix 2. Results were provided in table and graph formats (Appendix 2).

Table 3. Example of a nonattainment assessment across temporal scales for one Chesapeake Bay continuous water quality monitoring station, St. George's Island, 2006-2008.

St. George's Island (XBF7904)

Year	Method	Available Annual Dataset Mean	June through August Mean	July Mean	Available Annual Dataset % Non-Attainment	June through August % Non-Attainment	July % Non-Attainment
2006	Instantaneous	6.69	5.78	5.68	4	8	10
	Daily Mean				1	2	3
	7 Day Moving Average (15 min. increment)				0	0	0
	1 Average per 7 Days				0	0	0
	30 Day Moving Average (15 min. increment)				0	0	0
	1 Average per 30 Days				0	0	0
2007	Instantaneous	7.05	5.73	5.35	5	9	17
	Daily Mean				2	4	13
	7 Day Moving Average (15 min. increment)				0	0	0
	1 Average per 7 Days				0	0	0
	30 Day Moving Average (15 min. increment)				0	0	0
	1 Average per 30 Days				0	0	0
2008	Instantaneous	7.11	5.33	5.07	10	21	27
	Daily Mean				4	9	17
	7 Day Moving Average (15 min. increment)				1	1	4
	1 Average per 7 Days				4	8	25
	30 Day Moving Average (15 min. increment)				12	25	40
	1 Average per 30 Days				0	0	0

BIPLOT ASSESSMENT OF COMPARATIVE PROTECTION FOR DIFFERENT TIME SCALES OF THE DISSOLVED OXYGEN CRITERIA

Plots were created that compared 30-day mean measurements on the X-axis with violation rates of a shorter duration criteria such as the 7-day mean on the Y-axis. The resulting graphics illustrated the coincident behavior of DO conditions at two temporal scales and allowed for an assessment of violation rate thresholds between those scales (Figure 1,2). Sequential 30-day means and coincident short duration (e.g. 7-day) criteria means were computed from Chesapeake Bay water quality monitoring data for the summer season period (June-September). Sequential means are used under the current protocols of criteria attainment assessment. The use of rolling means within the context of the current Chesapeake Bay CFD criteria assessment procedures was examined by the Umbrella Criteria Assessment Team. However, the Team evaluation of both approaches supported continued use of sequential means and found rolling means to be inappropriate for use with the present CFD criteria assessment methodology (Appendix 3).

Data used to generate the nonattainment rates were derived from high frequency near-continuous (i.e. every 15 minutes) data records of water quality sensors located in offshore (U.S. EPA 2004) and shallow water (CBP-STAC 2012) habitats. Open water designated use criterion thresholds were applied to the data to calculate the percent non-attainment. Means were computed based on the full record data set for a criterion period (e.g. 30-day, 7-day, 1-day). Unless otherwise noted, it is important to note that data were not sub-sampled to mimic the low frequency, biweekly to monthly, Chesapeake Bay long term water quality monitoring program sampling scheme.

(PLACE HOLDER PIC FOR A LARGER, CLEANER GRAPHIC FROM USEPA 2004)

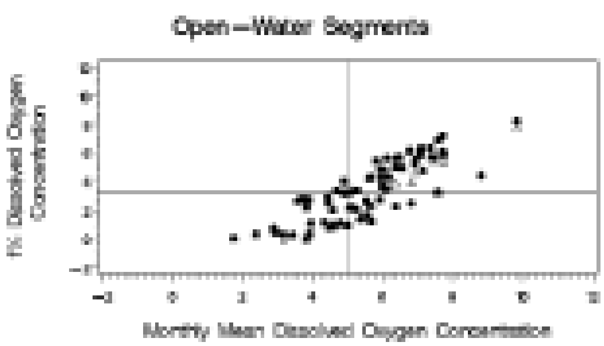

Figure 1. Example biplot (Placeholder) examining mutual criteria protections between monthly mean measures and the 1% dissolved oxygen concentration as an estimate of instantaneous minimum. Graphic from U.S. EPA (2004), p53.

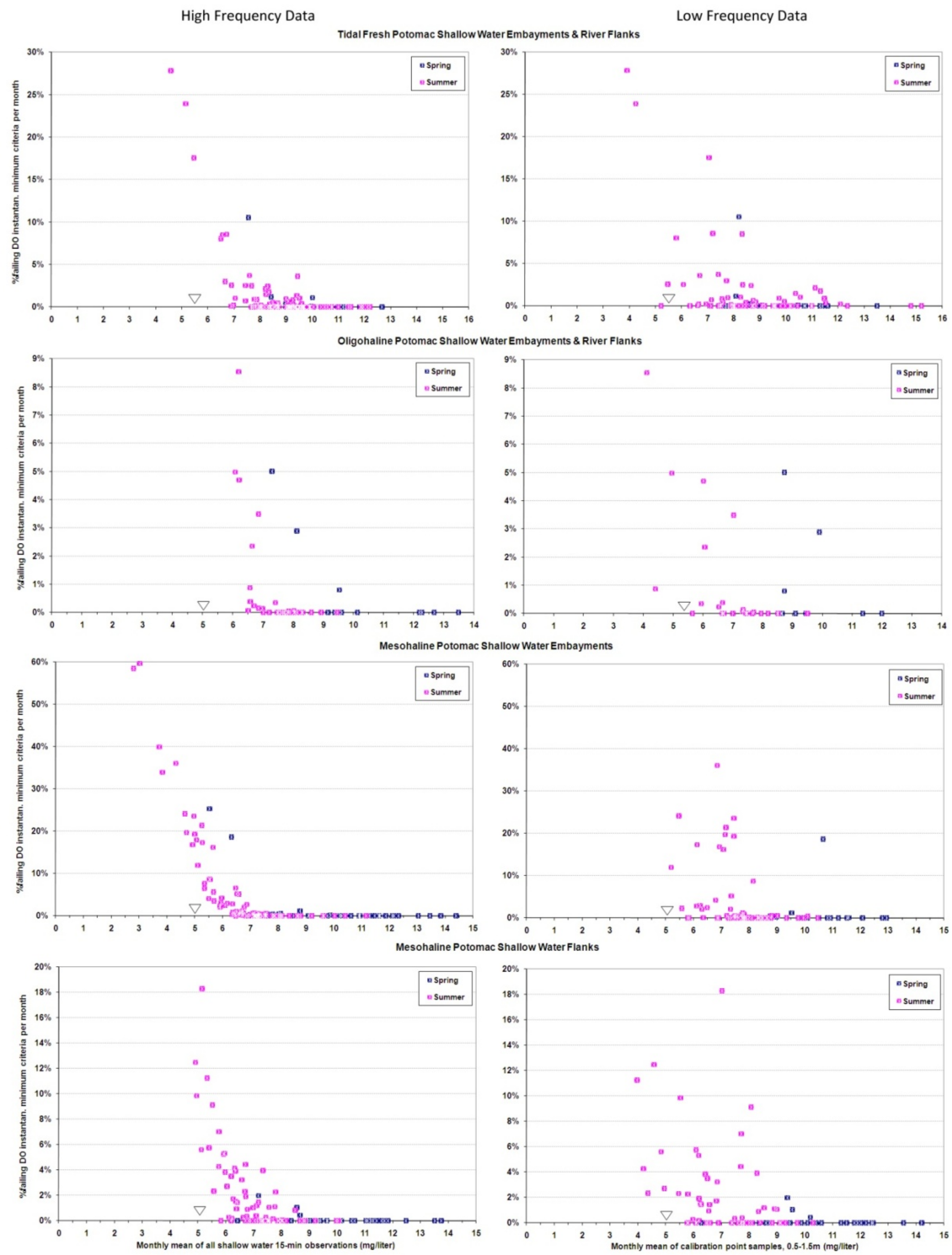

Figure 2. Example of biplots graphics. Using high frequency shallow water time series, criteria failure frequency was computed by month for summer season. Failure calculations for the 7-day mean DO criteria used rolling 7-day period with a 1-day steps. The results are plotted against the corresponding 30-day mean DO. Overall, 175 of the 415 months (42.2%) represented in the 20 tidal Potomac shallow water stations between 2004 and 2008 had failures of the instantaneous minimum DO criteria. Most instantaneous minimum criteria failures occurred in months where the 30-day mean criteria are met. (Appendix 3)

CONDITIONAL PROBABILITY ANALYSIS: PROTECTION OF THE 30-DAY MEAN FOR THE 7-DAY MEAN.

The method employed is based on the basic approach that if the variability of the 7-day mean dissolved oxygen concentration about the 30-day mean has a standard deviation less that 0.7805, then we can expect that the 7-day criterion will be violated less than ten percent of the time if the 30-day criterion is met. To use this approach, an estimate of the standard deviation of the 7-day mean for dissolved oxygen about the 30-day mean is needed (Figure 3). To estimate this quantity, Potomac River Shallow water Continuous Monitoring data was used. Further details of the methods are provided in Appendix 4.

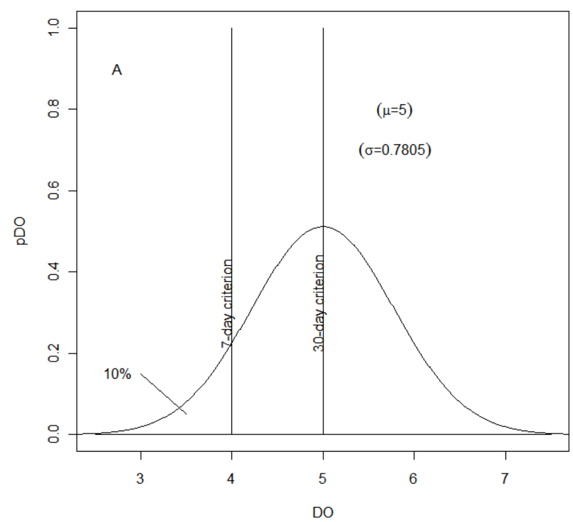

Figure 3. Illustration of the level of variability of the 7-day mean about the 30-day mean that results in up to 10 % violations of the 7-day mean criterion when the 30-day mean criterion is met.

A PARAMETRIC SIMILATION APPROACH TO ASSESSING THE UMBRELLA CONCEPT WHEN ASSESSING THE 30-DAY MEAN AND ITS ABILITY TO PROTECT THE OPEN WATER 7-DAY MEAN AND INSTANTANEOUS MINIMUM CRITERIA.

The basic approach of the simulation is to generate time series that have properties similar to observed DO concentration time series. Autoregressive (AR) modeling is a statistical technique that has been used to describe certain time-varying processes in nature. Perry (CBP STAC 2012) used a specific case of autoregressive models, an AR(2) model, for simulating Chesapeake Bay dissolved oxygen dynamics The data used for this exercise are the open water buoy data compiled by Olson (per Chapter 5 in U.S. EPA 2004b). To illustrate the level of protectiveness for short-duration criteria provided by a

30-day mean based on Bay specific data, the results were presented to show a gradient of monthly mean dissolved oxygen concentrations and the estimated level of criteria non-attainment for related 7-day mean and instantaneous minimums dissolved oxygen levels associated with a 30-day mean dissolved oxygen concentration. For details of the autoregressive modeling approach, refer to Appendix 5 in this document.

ASSESSMENT OF SAMPLING VARIABILTY ON THE ABILITY OF THE 30-DAY MEAN TO SERVE AS AN UMBRELLA FOR THE 7-DAY MEAN CRITERION

The Umbrella Criteria Assessment Team examined the additional uncertainty that is created by the use of small sample size to estimate DO 30-day means and further evaluated the consequences of this uncertainty for the application of the umbrella criterion concept. In many parts of the bay, the monthly mean is estimated from as few as one or two observations per month by the Chesapeake Bay Partnership's long term water quality monitoring program. Because the uncertainty of a monthly mean of two observations is much greater than the uncertainty of a monthly mean from near continuous data, it is reasonable to expect that effectiveness of the umbrella effect of the 30-day criterion for protecting other criteria will diminish when the low sample size mean is employed. Appendix 6 provides details of the analysis approach.

RESULTS

APPLYING ANALYSIS OUTCOMES AND SUPPORTING MONITORING AND ASSESSMENT OPTIONS TO ASSESSING CHESAPEAKE BAY DISSOLVED OXYGEN CRITERIA.

SPECTRAL CASTING VALIDATION

Results showed that *application of the spectral casting technique to address short-duration criteria attainment assessment needs for the Chesapeake Bay DO criteria is limited by the large uncertainty of the low frequency sampling used by the Chesapeake Bay long-term water quality monitoring program network*. Validation tests for introducing the estimated and temporally dense, high frequency DO concentration time series at mid-channel water quality monitoring sites as a potentially new step in DO criteria attainment assessments is not recommended. The Umbrella Criteria Assessment Team tests showed application of variations available within the spectral casting method process were technically feasible to implement by Bay analysts (CBP-STAC 2012). Linear interpolation provided a better fit than FFT or cubic spline methods in the comparisons. The technical capacity to apply the technique to Chesapeake Bay water quality criteria assessments was not a limitation.

Comparisons were made between estimated and measured dissolved oxygen conditions for an offshore monitoring site in the York River. The Umbrella Criteria Assessment Team found that dissolved oxygen varied similarly on a weekly scale. However, 24-hour periodicity was found to explain more dissolved oxygen variability in the time series and this time scale of comparison for nearshore and offshore dissolved oxygen patterning was not supported using the spectral casting approach.

MUTUAL PROTECTIVENESS OF DISSOLVED OXYGEN CRITERIA: 30-DAY MEAN CRITERION PROTECTION OF SHORT-DURATION (7-day mean, 1-day mean and instantaneous minimum) CRITERIA

An Umbrella Effect exists between the 30-day mean criterion and shorter-duration criteria in Chesapeake Bay. To apply the Umbrella Criterion Approach:

- *the sampling effort used in estimating the Umbrella Criterion (30-day mean) value must be accounted for,*
- *A level or levels of protection supporting mutual criteria protection from the single measured criteria must be stated, and,*
- *The associated dissolved oxygen threshold(s) supporting the stated level(s) of protection must be met to protect an unmeasured criterion.*

For example, if the Chesapeake Bay Open Water 30-day mean dissolved oxygen criterion is the measured criterion and is satisfied by meeting the water quality standards criterion threshold of 5.0 mg O_2/L, the Umbrella Criteria Assessment Team showed that there can be a less than a 10% chance that the 7-day dissolved oxygen criterion will be violated. However, it is necessary to understand that this particular result was determined from near continuous, high temporal frequency time series of dissolved oxygen concentrations in Chesapeake Bay (i.e Continuous Monitoring sensor data sets of Shallow-water Monitoring Program, temporally dense vertical water quality profiler time series) that allow for very accurate estimates of both a monthly mean and its associated weekly means.

The impact of small sample size estimation of the primary measure being estimated (e.g. the 30-day mean) when applying an Umbrella Criterion approach is to diminish the Umbrella Effect compared with the greater certainty of protection levels that are provided by using high temporal density time series assessments. In the practice of conducting Chesapeake Bay dissolved oxygen criteria attainment assessments under the existing Chesapeake Bay long-term, fixed station water quality monitoring program, the 30-day mean dissolved oxygen concentration is estimated from as few as one or two observations per month. The uncertainty of estimating the 30-day mean dissolved oxygen concentration using so few observations is much greater than the uncertainty of a monthly mean from near continuous sensor data.

There are combined impacts to the Umbrella Criterion approach that deal with water quality sampling intensity and sample variability resulting from sampling frequency that affect the level of mutual criteria protections (Table 4). First, the Umbrella Criteria Assessment Team showed that assessing one criterion using high temporal resolution measurements compared with a two-sample estimate of the same criterion provides greater certainty about criteria attainment for an accompanying unmeasured criterion. This effect is not linear. As the 30-day mean increased above the passing value of 5.0 mg O_2/L, the risk of nonattainment for the 7-day mean declined more rapidly than when the 30-day mean is based on two samples. Second, the sample variability associated with estimating a 30-day mean dissolved oxygen concentration with two samples had an effect of 3-4% on the risk of nonattainment for the unmeasured criterion. Small sample risk was evaluated from a range of standard deviations extracted from Chesapeake Bay data to illustrate the sample variability effects. Third, applying the findings of sample intensity and sample variability to an Umbrella criterion Assessment, in order to reduce the risk of nonattainment for an unmeasured criterion, you either need to measure at a greater frequency than two samples a month in the water column, or the measurement of the mean must meet an additional higher threshold dissolved oxygen value for a desired level of mutual criteria protection.

Table 4. Estimates of risk of violating the 7-day criterion given the monthly mean estimate (column 1), assessment of near true risk based on continuous monitoring data (Column 2) and and four levels of sampling variation (column 3). Adapted from Appendix 4.

	Risk of violating 7-day criterion	
Monthly Mean DO	Near True Risk based on high frequency data	Small Sample Risk on Nonattainment SD=1.7358[2] SD=1.6054[3] SD=1.9287[4]
5.0	16%	27%-30%
5.1	14%	25-28%
5.2	12%	23-27%
5.3	10%	21-25%
5.4	8%	19-24%
5.5	7%	18-22%
5.6	6%	16-20%
5.7	5%	14-19%
5.8	4%	13-18%
5.9	3%	12-16%
6.0	2%	11-15%
6.1	2%	10-14%
6.2	1%	9-13%
6.3	1%	8-12%
6.4	<1%	7-11%
6.5	<1%	6-10%

1 standard deviation of true weekly mean from true monthly mean
2 standard deviation base on pooling 20 resampling estimates
3 standard deviation based on minimum of 20 resampling estimates
4 standard deviation based on maximum of 20 resampling estimates

Without offshore, high frequency water quality profile data to compensate for the uncertainty of low sample size in estimating the 30-day mean, dissolved oxygen thresholds were developed supporting protection of the 7-day mean in Open Water (Table 4). The dissolved oxygen thresholds are higher than the 30-day mean criterion. As the measured mean from two samples gets higher above the 30-day criterion, the ability to insure protection of short-duration criteria improves. Using two samples a month, a 30-day mean of 5.0 mg/L is in attainment of the open water summer 30-day mean criterion. However, to insure that water quality conditions also have a less than 10% risk of not attaining the 7-day mean criterion, the two-sample mean must further be at least 6.5 mg/L. Just passing the 30-day mean criterion leaves a 27-30% risk of nonattainment for protecting the 7-day mean dissolved oxygen criterion; this is between a 1-in-3 and a 1-in-4 chance of degraded dissolved oxygen conditions. Continuous monitoring data records provide greater certainty that meeting the 30-day mean at values closer to the criterion value also insures protection of the 7-day mean criterion.

The 30-day mean can also protect the instantaneous minimum, however, it is easier to protect the 7-day mean for a comparable level of nonattainment risk (Table 5). The Umbrella Criteria Assessment Team (Appendix 5) used Open Water Chesapeake Bay data to explore the simultaneous protection of the three criteria. In order to also protect the instantaneous minimum at the same certainty as protecting the 7-day mean based only on the 30-day mean assessment, the 30-day mean will have to meet even higher dissolved oxygen thresholds that it does for protecting the 7-day mean alone. To have a low risk of nonattainment for both the 7-day mean and the instantaneous minimum may require an unrealistically and potentially unattainable high 30-day mean dissolved oxygen value. The example derived in this particular Umbrella Criteria Assessment study using high frequency observations showed that passing the 30-day mean had a 17.7% nonattainment rate for the 7-day mean but a 44.9% rate of nonattainment beyond a presumed 10% allowable exceedance. In order to approach a comparable level of nonattainment for 17% nonattainment the 30-day mean needs to be over 7 mg/L. This 30-day mean would need to be higher yet based on two sample estimation of the 30-day mean. Alternative options in establishing mutual protectiveness will be to endorse different levels of certainty for passing all three criteria simultaneously or measuring more frequently to reduce uncertainty and have a more accurate estimate of the 30-day mean.

Table 5. An evaluation of comparable protectiveness within the Open Water designated use when meeting the 30-day mean and the risk of criteria nonattainment for the 7-day mean and instantaneous minimum criteria. See Appendix 5 for more details.

30-day Mean DO[1]	7 day criterion failure rate	rate of instantaneous criterion > 10%
5.0058	17.7%	44.9%
5.6732	5.5%	34.3 %
6.3407	2.1%	25.4 %
7.0082	2%	16.6%

1. The Dissolved oxygen means were derived from sensor data over multiple depths of 6,5,4 and 3m respectively in this column as a surrogate for exploring data across many sites to develop the same gradient of monthly mean values

SUMMARY

MONITORING AND EVALUATION OPTIONS FOR ASSESSING SHORT-DURATION DISSOLVED OXYGEN CRITERIA ATTAINMENT

The Umbrella Criteria Assessment Team findings point to two primary monitoring options for assessing Chesapeake Bay short-duration water quality criteria:

- *Enhanced Monitoring Approach*: Measure water quality profiles at high temporal frequency. Increase spatial resolution when possible. Evaluate the high resolution data against the suite of water quality criteria using the present Chesapeake Bay water quality criteria attainment assessment methods.
  - *This approach provides support to improve the estimate of the 30-day mean and address gaps in short-duration criteria assessment of 7-day mean, 1-day mean, instantaneous minimums for Open Water and Deep Water designated uses.*

- *Umbrella Criterion Approach- A Risk-based Assessment*:
  - *Use the Existing Chesapeake Bay Partnership long-term water quality monitoring program sampling strategy:* Use twice-a-month sampling as a minimum assessment of the 30-day mean. Define and apply an acceptable risk associated with meeting multiple criteria in a designated use when the basis of the Umbrella assessment is a single 30-day mean criterion assessment
    - *This approach provides support to address gaps in short-duration criteria assessment of 7-day mean, instantaneous minimums for Open Water.*
      - *This approach has had very limited evaluation for the Deep Water designated use which has a different 30-day mean criterion than the Open Water designated use. Further data and assessment would be needed to develop the relationships of mutual protection for criteria in the Deep Water designated use.*
  - *Enhanced Monitoring Strategy*: Increasing monthly sampling intensity to greater

than two samples per month lowers the uncertainty for risk of nonattainment in unmeasured short duration criteria. Employ nearshore and offshore sampling strategies that collect high temporal frequency data and capture DO behavior throughout the water column for computing the 30-day mean assessment. Meeting the 30-day mean based on high frequency data (e.g. 15 minutes to 1 hour time steps) was shown to effectively lower the risk of violating the short-duration criteria protections assumption for 7-day mean when the 30-day mean meets the water quality standards criterion. However, when meeting the 7-day mean there remained a greater risk of violating the instantaneous minimum. Therefore, a threshold for an acceptable level of risk for meeting the instantaneous minimum can be defined, or higher temporal and spatial density monitoring can be conducted.

- *This approach provides support to improve the estimate of the 30-day mean and address gaps in short-duration criteria assessment of 7-day mean, and instantaneous minimums for Open Water.*
    - *This approach would provide data to support development of an Umbrella Criterion relationship to address gaps in short-duration criteria assessment of 1-day mean and instantaneous minimums for* the *Deep Water designated uses.*

The Umbrella Criteria Assessment Team results (CBP-STAC 2012) support the U.S. EPA (2004) recommendation that site-specific buoy deployments may be necessary to either better quantify statistical relationships or assess dissolved oxygen criteria attainment. U.S. EPA (2007a) also suggested collection of continuous measures of dissolved oxygen to resolve gaps in assessing short-duration dissolved oxygen criteria with statistical options that included logistic regression and time series analysis (i.e. spectral analysis) . As demonstrated through evaluation of the spectral analysis approach, statistically filling in the gaps in the need for high frequency data at long term monitoring sites is limited by the uncertainty introduced from estimating the 30-day mean dissolved oxygen conditions from one or two samples. Logistic regression remains a viable option for further exploration acknowledging the strengths of the approach expressed in U.S. EPA (2007a) but with its limitations that are also linked to the temporal frequency of data available at the long term fixed station monitoring locations. An additional refinement of the recommendation for site specific buoy deployments is to target management segments near attainment with limited monitoring resources available. Activate monitoring approaches that provide measurements in high temporal frequency of water quality profiles to support short-duration criteria assessments.

CHESAPEAKE BAY CRITERIA ATTAINMENT ASSESSMENT USING HIGH FREQUENCY MONITORING DATA: The Umbrella Criteria Assessment Team demonstrated the technical capacity available for using the present Chesapeake Bay Program CFD criteria attainment method for assessing attainment when mid-channel, high frequency data are available (i.e. Lane and Robertson *in* CBP-STAC 2012). Given the Umbrella Criteria Assessment Team findings that did not support use of synthetic time series at mid-channel locations through spectral casting, the limitation of conducting the assessment over the entire

Bay remain data driven. Offshore, season-long, high frequency vertical profiles of water quality conditions are rare for Chesapeake Bay. There have been data collected in offshore locations by vertical profilers in single locations and therefore a limited distribution of segments (e.g. the lower York River, Rappahannock River and Potomac River) where shallow water data can be leveraged for greater spatial resolution. Given the large proportion of shallow water habitat in Chesapeake Bay, coupling nearshore water quality measurements with offshore sensors collecting data throughout the water column, it is feasible to produce measurements of short-duration water quality criteria suitable for supporting Chesapeake Bay assessments. The greater technical challenge has been placement and maintenance of vertical profiler technology to the support such water quality assessments.

Alternative sampling strategies to obtaining high frequency measurements could be considered depending on the limiting conditions of a particular segment. Individuals could manually collect water column profile data in mid-channel locations multiple times per day throughout a season or the year. Autonomous Underwater Vehicles could be flown in multiple missions per day to add vertical and horizontal resolution to an assessment. Vertical profiler instruments are being used around Chesapeake Bay in relatively sheltered locations. Vertical profiler and related technology applied to monitoring water quality in the diversity of Bay habitats should continue to evolve to overcome present hydrodynamic challenges for use in many offshore regions. The eventual collection of high frequency water quality profile data will be applicable and amenable to use with the Chesapeake Bay water quality criteria attainment assessment strategy.

LITERATURE CITED (IN PROGRESS)

CBP 1989a,

CBP 1989b

CBP-STAC 2012. Umbrella Criteria Assessment.

Dunk, Jeffrey R., William J. Zielinkski and Hartwell H. Walsh, Jr. 2006. “Evaluating reserves for species richness and representation in northern California.” *Diversity and Distributions*, Vol. 12, 434-442.

Jordan, J., C. Stenger, M. Olson, R. Batiuk, and K. Mountford. 1992. Chesapeake Bay dissolved oxygen goal for restoration of living resource habitats. CBP/TRS 88/93. Chesapeake Bay Program, Annapolis, MD.

Launer, Alan E. and Dennis D. Murphy. 1994. "Umbrella Species and the Conservation of Habitat Fragments: A Case of a Threatened Butterfly and a Vanishing Grassland Ecosystem." *Biological Conservation,* Vol. 69, No. 2, 145-153

MRAT 2009. The Monitoring Realignment.

Neerchal, N.K., G. Papush, and R. Shafer. 1992. Statistical method of measuring DO restoration goals by combining monitoring station and buoy data. Chesapeake Bay Program, Annapolis, MD.

Roberge, Jean-Michael and Per Angelstam. 2004. "Usefulness of the Umbrella Species Concept as a Conservation Tool." *Conservation Biology*, Vol. 18, No. 1, 76-85

Shenk and Batiuk 2010

TMDL 2010

U.S. EPA. 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. 2004a. *Technical Support Document for Identification of Chesapeake Bay Designated Uses and Attainability – 2004 Addendum.* October 2004. Region III. Chesapeake Bay Program Office. EPA 903-R-04-006. Annapolis, MD.

U.S. EPA. 2004b. *Chesapeake Bay Program Analytical Segmentation Scheme: Revisions, Decisions and Rationales 1983-2003. October 2004*. Region III Chesapeake Bay Program Office, Annapolis, MD. EPA 903-R-04-008.

U.S. EPA 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 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-07-00*. Region III Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA. 2008a. *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. CBP/TRS 290-08. U.S. Environmental Protection Agency, Region 3, Chesapeake Bay Program Office, Annapolis, MD.

U.S. EPA. 2010a. *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. CBP/TRS 301-10. U.S. Environmental Protection Agency, Region 3 Chesapeake Bay Program Office, Annapolis, MD.

USEPA (U.S. Environmental Protection Agency). 2010b. Chesapeake Bay Total Maximum Daily Loads for Nitrogen, Phosphorus, and Sediment…

Wilcox, Bruce A. 1984. "In situ conservation of genetic resources: Determinants of minimum area requirements." *In National Parks, Conservation and Development, Proceedings of the World Congress on National Parks*. J.A. McNeely and K.R. Miller, Smithsonian Institution Press, pp. 18-30.