appendix I

# Analytical Approaches for Assessing Short-Duration Dissolved Oxygen Criteria

The Chesapeake Bay dissolved oxygen criteria have several different durations: 30-day mean, 7-day mean, daily mean and instantaneous minimum. Users' ability to assess these criteria and to have certainty in the results depends on the time scale of available data and on the ability of models to estimate conditions at those time scales. At present, long-term, fixed-station, midchannel water quality monitoring in the Chesapeake Bay and its tidal tributaries provides dissolved oxygen measurements twice monthly at most or approximately every 15 days between April and August. Proposed enhancements to the tidal water quality monitoring program include shallow-water monitoring, as well as high-resolution spatial and temporal monitoring in selected locations. However, these new components are only in the planning and early implementation stages at this point, and because of financial constraints or limitations to current technology, direct monitoring at the scales of the criteria may not be possible in the foreseeable future. Therefore, the assessment of attainment for some geographic regions and for some short-term criteria elements must be waived for the time being or must be based on statistical methods that estimate probable attainment. Several approaches to addressing the duration issue are described below in more detail.

## LOGISTIC REGRESSION MODELS USING ROUTINE FIXED-STATION MONITORING DATA

This method is a modification and significant update of a method developed originally to measure attainment of the 1992 Chesapeake Bay dissolved oxygen restoration goal (Jordan et al. 1992). The early work demonstrated predictable relationships, on a segment by segment basis, between seasonal mean dissolved oxygen concentrations and the percent of observations above a target concentration in areas where dissolved oxygen concentrations ranged above and below goal target concentrations (figures I-1 and I-2).

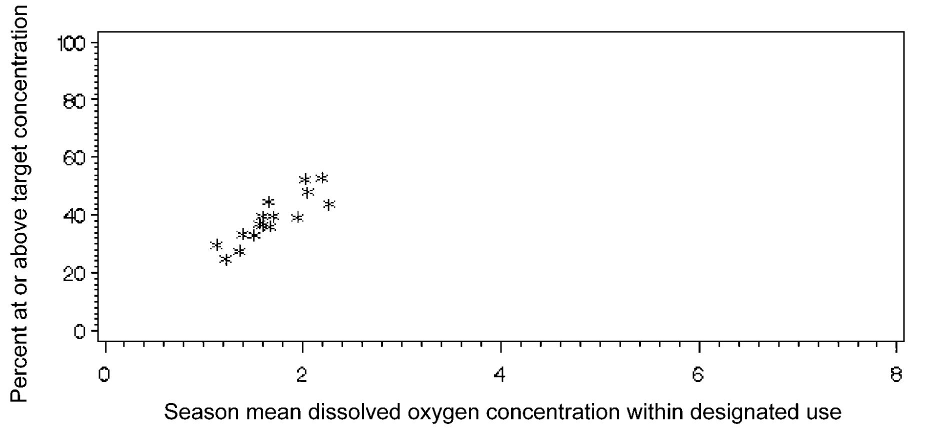


**Figure I-1:** Percent of Summer dissolved oxygen concentrations above 1.7 mg liter$^{-1}$ in segment CB3MH deep-water designated use habitat.

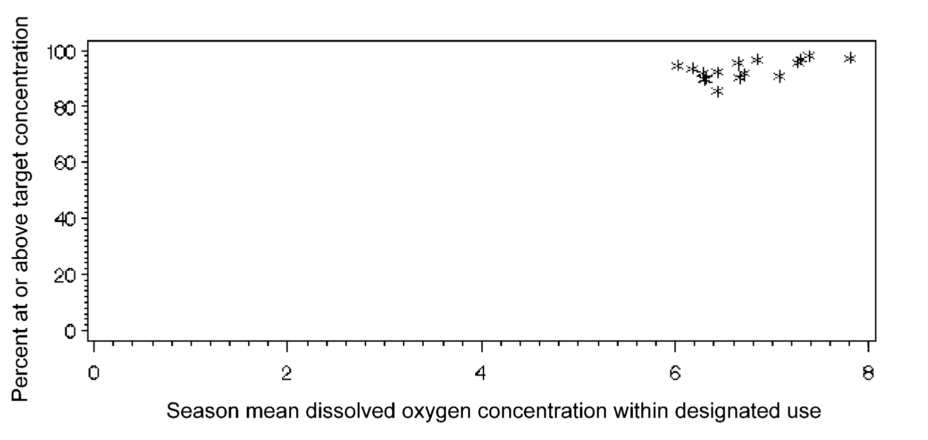


**Figure I-2.** Percent of Summer dissolved oxygen concentrations above 3 mg liter$^{-1}$ in segment CB3MH open-water designated use habitat.

The relationships were then expressed as regression equations, which could be used to predict the percentage of observations above or below target for any seasonal mean. By extension, the ‘percentage of observations’ applied both to space and time within a segment. Most of this pilot work was done for mainstem Chesapeake Bay segments that were relatively densely populated with fixed monitoring stations laterally and longitudinally. Because of the spatial density of stations, the range of potential dissolved oxygen exposure that any particular point might experience over a tidal cycle was captured in the models. Contemporaneous, semicontinuous dissolved oxygen measurements made with in situ sensors deployed on buoys were used to validate the model estimates. The models did not predict the extreme minima recorded in the continuous record, but were able accurately to predict the frequency of observations below the mean, $5^{th}$ and $95^{th}$ percentile concentrations over a month's time (Jordan et al. 1992).

With benefit of the long (more than 16 yrs) record of the Chesapeake Bay tidal-water quality monitoring program and the density of measurements (the vertical dissolved oxygen profile is characterized at 1- to 2-meter intervals), the simple 1992 regression models have been improved. These enhanced models use logistic regression, which is better suited to percent distributions (i.e., distributions between 0 and 100). The models are now month- and depth-specific in many segments. These models can be adapted and applied to estimate attainment of the instantaneous minimum dissolved oxygen criterion, if the user considers that the minimum criterion is not met if the dissolved oxygen concentration is below the criterion value *at any time and anywhere* in the segment-designated use.

The first step is to reconstruct the models using the cruise-by-cruise three-dimensional interpolations of dissolved oxygen monitoring data. That is, collect the percent volume passing/failing the criterion at each depth in a segment month by month from, for example, 1985 through 1998, and model the relationship of percent volume failing/passing the criterion as a function of the monthly mean of that segment/depth as represented by all the cells in the grid. Using the interpolated data should improve the spatial representation within the segment.

To assess current attainment, for 1999-2001 for example, the user would interpolate the monthly average dissolved oxygen concentrations across all tidal waters, as before, for each month of the season to be evaluated in the assessment period, e.g., the summer period including June through September. Then the month/segment/depth-specific model appropriate for the designated use and cell location would be applied to estimate the percent of time each cell was likely to be below the instantaneous minimum, based on the cell's interpolated monthly average. If the model predicts that the cell is above the minimum dissolved oxygen level less than 99-100 percent of the time, then the cell is not in attainment. Each cell is assessed in this way and its volume added to the 'failing' or 'passing' category. Ultimately, the percent of total volume failing or passing the criterion within the segment and designated use is calculated for each month. The monthly percentages are tallied over all months in the season in the assessment period and the cumulative frequency distribution is calculated. Except for the use of the logistic regression model, each of the steps is consistent with assessment methods of the other criteria (see Chapter VI for details).

The following is a sample attainment model for Chesapeake Bay Program segment CB3MH for the open-water instantaneous minimum 3.2 mg liter[1] criterion. This model was based on fixed station data, not on interpolated data as proposed above.

> If the time frame is September through March, attainment is likely to be 100 percent, regardless of depth. For other months, the estimated percent attainment is estimated by

$$\begin{aligned} LGT = {} & 1.0757 \times (\text{mean monthly dissolved oxygen concentration}) \\ & -0.0724 \times (\text{depth in meters}) \\ & -1.8576 \text{ for April} \\ & -2.9219 \text{ for May} \\ & -2.7982 \text{ for June} \\ & -2.8341 \text{ for July} \\ & -2.5443 \text{ for August.} \end{aligned}$$

$$\text{Percent attainment} = 100 \exp(lgt) / (1 + \exp(lgt)).$$

The figures below illustrate attainment curves for segment CB3MH for summer deep-water designated use, where the instantaneous minimum criterion is 1.7 mg liter$^{-1}$ (Figure I-3) and open-water designated use, where the instantaneous minimum of 3.2 mg liter$^{-1}$ criterion applies (Fig. I-4).

At this time the method has not been adequately validated in areas other than the mainstem Chesapeake Bay. It is, therefore, premature to recommend its implementation for formally assessing criteria attainment. The issue is not so much the method itself, but how well the midchannel stations represent the flanks and surrounding waters where station density is low. The models are only as good as the information on which they are developed. The day-time sampling schedule cannot detect nocturnal lows in shallow areas or other areas where dissolved oxygen is quickly regenerated. In the mainstem Chesapeake Bay, however, the dissolved oxygen 'memory' in the deep waters represents, to some extent, the night-time dissolved oxygen sags and the station density captures exposure diversity. In other areas and designated uses, hypoxia is more ephemeral temporally and the density of fixed-monitoring stations is reduced spatially so the models are likely to be weaker.

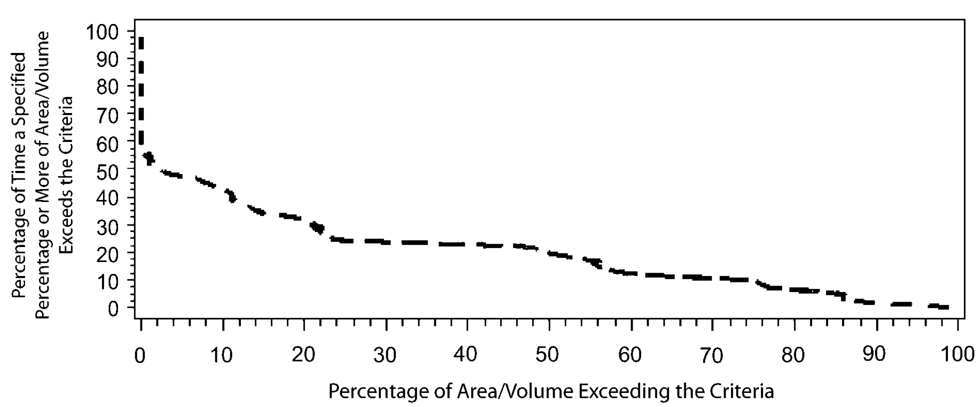


**Figure I-3.** Summer instantaneous minimum deep-water dissolved oxygen 1.7 mg liter$^{-1}$ criterion attainment curve for segment CB3MH based on application of the logistic regression model.

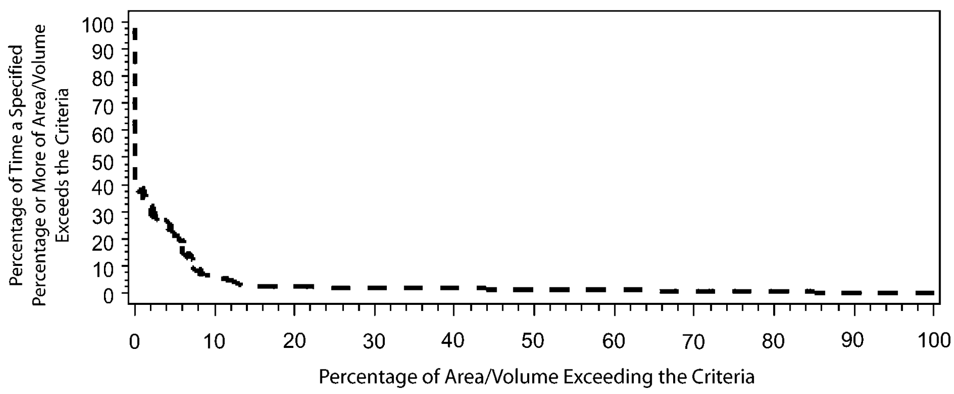


**Figure I-4.** Summer instantaneous minimum open-water dissolved oxygen 3.2 mg liter$^{-1}$ criterion attainment curve for segment CB3MH based on application of the logistic regression model.

## SYNTHETIC, CLOSE-INTERVAL DATA SETS CREATED BY COMBINING SHORT- AND LONG-TERM PATTERNS OF VARIATION IDENTIFIED THROUGH SPECTRAL ANALYSIS

It is critical to obtain real time information about the ephemeral and episodic events of low dissolved oxygen. However, it is not possible for the states or other partners to collect such information over vast regions of the Chesapeake Bay and over the length of time that would be required to address the short duration (i.e., the 7-day mean, 1-day mean, instantaneous minimum) criteria directly. The following methodology is proposed to address the short interval criteria by integrating information from the long-term, low-frequency monitoring program with short-term, high-frequency monitoring that can be accomplished using in-situ semi-continuous data recorders.

The method, still in development, is an adaptation of work by Neerchal (1992). The method combines temporal variability information from the long-term, low-frequency monitoring data with that from short-term, high-frequency data such as collected with in situ continuous recording devices. Since these devices are often deployed using moored buoys, the associated data are referred to as 'buoy data'.

The method uses spectral analysis to extract the cyclical components of the long- and short-term dissolved oxygen time series records and combines them to create a synthesized time-series data set with data synthesized at user-specified time steps. The synthetic data have the annual and seasonal cyclic and trend characteristics of the long-term record as well as the tidal, diurnal and other periodic characteristics of the short-term, high-frequency record. At present, the synthetic data are hourly, with cyclic components limited to two cycles per day. The synthetic data are then analyzed like any other data set relative to the specific elements of the criteria.

The spectral equation for the long-term data (such as the Chesapeake Bay water quality monitoring program data) is

a) $$x_t^{lt} = \bar{x}^{lt} + \sum_{k=1}^{ltm} a_k^{lt} \cos(2\pi f_k^{lt} t) + b_k^{lt} \sin(2\pi f_k^{lt} t) \qquad \text{(Equation I-1)}$$

where ltm n/2, $f^{lt}{}_k$ = fourier frequencies, t = time in months, a,b = spectral coefficients, and $x^{lt}{}_t$ = data.

The spectral equation for the *short-term* data (for example, from an *in situ* dissolved oxygen data recorder) is

b) $$x_t^{st} = \bar{x}^{st} + \sum_{k=1}^{stm} a_k^{st} \cos(2\pi f_k^{st} t) + b_k^{st} \sin(2\pi f_k^{st} t) \qquad \text{(Equation I-2)}$$

where stm n/2, $f^{st}{}_k$ = fourier frequencies, t = time in hours, a,b = spectral coefficients, and $x^{st}{}_t$ = data.

The equation for the spectral forecast (for the synthetic data set) is

c) $$x_t^{lt} = \bar{x}^{lt} + \Sigma_{k=1}^{m1} a_k^{lt} \cos(2\pi f_k^{lt} t) + b_k^{lt} \sin(2\pi f_k^{lt} t) + \sum_{k=1}^{m2} a_k^{st} \cos(2\pi f_k^{st} t) + b_k^{st} \sin(2\pi f_k^{st} t) \qquad \text{(Equation I-3)}$$

where t = time scaled to suit f, $m^1 < ltm$, is chosen to exclude high frequencies that would be duplicated in the short-term equation (>½ cycle per month), and $m^2 < stm$ is chosen to exclude frequencies too high to be important (>2 cycles per day).

## A SAMPLE APPLICATION OF THE METHOD

The example application below uses long-term data from station CB4.2C, a monitoring station in the mid-region of the Chesapeake Bay, and a two-month series of continuous dissolved oxygen measurements at a buoy deployment in the vicinity of that station at approximately 9 meters below the surface. Figure I-5 shows the observed monthly dissolved oxygen concentrations (asterisks) at station CB4.2C (8-10-meter depth) and the long-term forecast (line) from the spectral equation.

The synthetic data record is obtained by combining the long- and short-term equations (Figure I- 6). A sample two-month period, August-September, 1987, indicated by the two vertical parallel reference lines in Figure I-5, is expanded in Figure I-6.

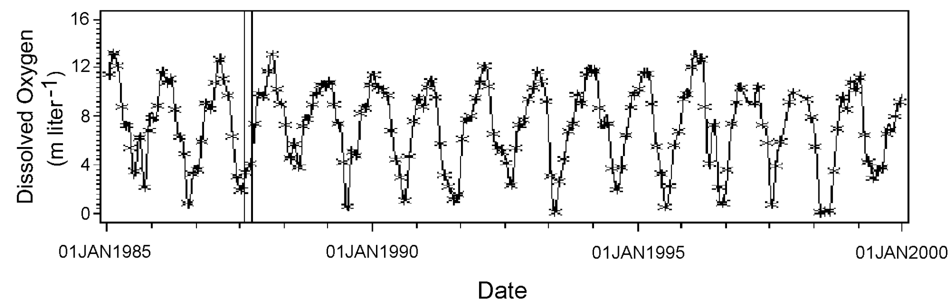


**Figure I-5.** Observed monthly dissolved oxygen concentrations (*) at Chesapeake Bay Monitoring Program Station CB4.2C (at the 8- to 10-meter depth) from January 1985 to January 2000 and the long-term 'forecast' (—) from application of the spectral equation.

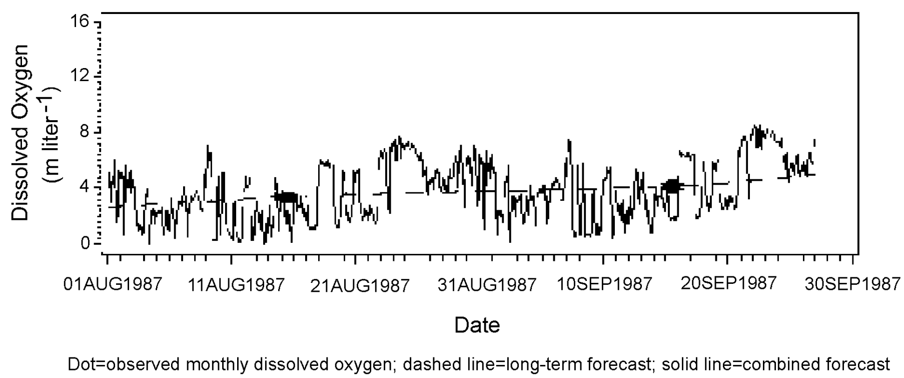


**Figure I-6.** Expanded view from Figure VI-23 of the two-month period August-September 1987 synthetic data record obtained by combining the long- and short-term spectral equations.

Some major difficulties beyond the cost and labor involved in deploying sensors, which are substantial to begin with, must be overcome to implement this methodology. The spectral forecasts are essentially temporal interpolations that can be sampled analytically. The forecasts do not lend themselves easily to an analysis of spatial extent of criteria attainment. For other criteria assessments, the direct measures of dissolved oxygen in the environment—temporal snap shots—are interpolated using the Chesapeake Bay Program interpolator, and the spatial extent of attainment is assessed. Then, the frequencies of the many spatial extent measurements are collected into a cumulative frequency distribution. The spectral analysis methodology developed thus far has yet to incorporate the assessment of spatial extent. Three-dimensional spatial interpolations at the 10-, 20- and 30-minute short-interval frequencies is a computational impracticality. Furthermore, more information is needed to determine the sphere of representativeness for the short-term, high-frequency patterns, both vertically and horizontally.

Some compromises will doubtless be required. New pilot projects, some of which are already underway, and additional analysis of current data already at hand will answer some of these questions and provide some basic underpinnings. Also, due to the rapidly developing technology in this area, changes and new, unforeseen opportunities are likely to present themselves. The importance of the short duration criteria to the protection of many of the Chesapeake Bay's target species and communities has been demonstrated. As Bay scientists and managers move forward with developing assessment tools, it would be prudent to seize new opportunities as technology evolves.

## LITERATURE CITED

Jordan, S. J., C. Stenger, M. Olson, R. Batiuk and K. Mountford. 1992. *Chesapeake Bay Dissolved Oxygen Goal for Restoration of Living Resource Habitats: A Synthesis of Living Resource Requirements with Guidelines for Their Use in Evaluating Model Results and Monitoring Information*. CBP/TRS 88/93. Chesapeake Bay Program, Annapolis, MD.

Neerchal, N. K., G. Papush and R. W. Shafer. 1992. *Statistical Method for Measuring DO Restoration Goals by Combining Monitoring Station and Buoy Data*. Chesapeake Bay Program, Annapolis, Maryland.