Report of the NJDEP-Science Advisory Board

WATER QUALITY DATA EVALUATION AND INTERPRETATION

Prepared by the Science Advisory Board
Water Quality and Quantity
Standing Committee

Approved by the
NJDEP Science Advisory Board

Dr. Judith Weis, Ph.D. (chair)
Clinton J. Andrews, Ph.D., P.E.
John E. Dyksen, M.S., P.E.
Raymond A. Ferrara, Ph.D.
John T. Gannon, Ph.D.
Robert J. Laumbach, M.D., MPH
Peter B. Lederman, Ph.D., P.E.
Robert J. Lippencott, Ph.D.
Nancy C. Rothman, Ph.D.
Tavit Najarian, Ph.D.
Michael Weinstein, Ph.D
Anthony J. Broccoli, Ph.D.
Mark G. Robson, Ph.D.
David A. Vaccari, Ph.D., P.E.
Lily Young, Ph.D.

January 26, 2016

FINAL

WATER QUALITY DATA EVALUATION AND INTERPRETATION

Report to the NJDEP Science Advisory Board

By the Water Quality and Quantity Committee

Dr. David A. Vaccari, Stevens Institute of Technology (Chair)
Mr. Thomas W. Amidon, Kleinfelder
Dr. Brian T. Buckley, Rutgers University EOHSI
Dr. Keith R. Cooper, Rutgers University
Dr. William J. Cromartie, Richard Stockton College of NJ
Dr. Josh T. Kohut, Rutgers, The State University of New Jersey
Dr. Xiaoguang Meng, Stevens Institute of Technology
Dr. Andrew Michalski, Michalski & Associates, Inc. (resigned)
Dr. Anthony S. Navoy, United States Geological Survey (resigned)
Dr. Orren D. Schneider, American Water
Dr. Pen C. Tao, North Jersey Water Supply Commission
Dr. Christopher G. Uchrin, Rutgers University

January 2016

Executive Summary

The Water Quality and Quantity Committee (WQQC) of the New Jersey Department of Environmental Protection (NJDEP) Science Advisory Board (SAB) has the following charge questions before it:

1. How should continuous dissolved oxygen (DO) data be interpreted for compliance with the existing state water quality criteria? Should the existing criteria be refined to address the availability of continuous DO data?
2. How can the DO discrete, grab sampling results from the ambient stream network best be used to identify waters that are most likely to exceed these criteria?
3. What specific statistical methods or models should be used to evaluate the coastal water quality data collected along transects by either Underwater Automatic Vehicles or continuous recording devices attached to vessels?
4. The methodology in the EPA (coastal and estuarine) criteria document is broad and may be modified to protect local marine/estuarine species.
    a. Are the thresholds recommended by EPA and/or those modified by NYDEC adequately protective of the animals found in New Jersey’s coastal waters and specifically, Barnegat Bay?
    b. Should the time period for a low DO event be revised to protect NJ species? Are the criteria adequately protective for early life stages?
    c. What type of data are needed to determine if criteria are sufficient to protect NJ biota?

Water quality standards (WQS) for DO were developed in an era when compliance was almost exclusively determined by the use of grab samples taken at most once in a day during normal working hours, and usually in the morning. For example, one of the criteria applied to streams

classified as FW2-NT is that the DO should not be less than 4.0 mg/L at any time (24 hour average not less than 5.0 mg/L, but not less than 4.0 at any time). Using grab sampling, violations were cited only if the measurement was less than 4.0 mg/L. No attempt was typically made to determine if the DO had been lower than the grab value at other times of the day.

Recently, semi-continuous monitoring of water quality including DO has been instituted at a number of sites. These provide data equivalent to grab-sample measurements taken every hour, half-hour, or even every fifteen minutes [USGS, Maryland study]. These data show that the minimum DO occurs shortly before dawn on most days, and can be much less than grab values obtained as late as 11am or 12pm.

This raises two issues leading to questions 1 and 2: First, if the minimum of the daily continuous monitoring data were substituted for grab samples, the criterion would effectively become much stricter than has been the case historically. Second, as many sites still are sampled only with single grab samples, it becomes obvious that the effect of time of day should be somehow removed from the determination of whether a water quality violation has occurred. The following discussion details the findings and the bases for the findings that the WQQC recommends to the SAB for questions 1 and 2.

A similar situation arises in the marine environment due to the new technology of Underwater Autonomous Vehicles (UAVs). These add the spatial dimension to the problem just described, producing semicontinuous data as a function of both time and position. This compounds the problem of data interpretation and application of standards, leading to charge question 3.

Finally, there is the biological question of what levels of DO are protective of organisms in New Jersey's marine environment, coupled with time-and-concentration effects. If there is to be concern about the effect of low DO on short (~ 1 hour) time scales, what do the data say about such short-term effects? Also, are there species of special concern in New Jersey that may have distinct responses that should be taken into account? These concerns are suggested in charge question #4.

To summarize the findings:

> *Finding #1:* The available data do not support a determination of the biological response to short-term DO excursions. We recommend NJDEP utilize a statistical approach in which the Confidence of Compliance is computed for individual samples and for averages of several samples.
>
> *Finding #2:* Preliminary analysis shows a predicted value for the minimum or average DO could be estimated using two measurements nominally one hour apart taken between 9am and noon. The NJDEP should develop and validate a correlation or interpolation method to make this estimation. The estimated values could be used to screen for sites requiring more thorough sampling. We recommend that diurnal DO measurement be used more frequently to confirm possible violations of the Water Quality Standards.

Finding #3: NJDEP should adopt the use of a type of analysis called "spatial decorrelation models" to separate the temporal and spatial components of UAV data. In addition, fixed-location sensors should be deployed to provide data that will help with that analysis.

Finding #4a: The NYDEC model does not account for all species within New Jersey's estuaries but could be adopted once consideration of how constant monitoring data should be interpreted.

Finding #4b: NJDEP should develop its own frequency of exceedences for DO. However, NJDEP should not adopt criteria for bays until site-specific DO data become available.

Finding #4c: Measurements should be made of impacts of DO on sentinel species on an estuary-by-estuary basis, or by groups of estuaries, and in coastal areas. In addition the role of other measurements such as contaminant levels should also be considered.

Detailed Findings and Bases

Question 1:

How should continuous DO data be interpreted for compliance with the existing state water quality criteria? Should the existing criteria be refined to address the availability of continuous DO data?

In further discussions with the NJDEP, this question was clarified as follows:

> NJDEP is seeking a determination, whether the currently used one hour duration (NJDEP Integrated Water Quality Monitoring and Assessment Methods, June 2012) has documented ecological response considerations, statistical, or other mathematical relevance, for example, in terms of providing an acceptable level of confidence, to determine that the criterion has been violated. The premise here is that there might be an error related to equipment precision or accuracy, deployment or other procedural element of obtaining the data that would support discarding ephemeral readings that suggest that there has been a small dip of a relatively short duration below the criterion.
>
> In addition, is there anything in the foundational basis for the criterion, or assessment methods developed and employed by other states, that would inform the duration that should be considered for assessment using continuous data to ensure the intent of the criterion (support of aquatic life use) is met. Would a different duration be more appropriate, and if so, what would it be and what would be the basis for this alternative duration?

Finding #1:

The available data do not support a determination of the biological response to short-term DO excursions. We recommend NJDEP utilize a statistical approach in which the Confidence of Compliance is computed for individual samples and for averages of several samples.

Basis for Finding #1:

Use of the minimum DO would best align with the current wording of the DO water quality standard (WQS), but would effectively make the standard much stricter than previously applied using mid-morning grab sampling.

The WQQC considered that the standard makes no note of the duration of a low-DO event. Toxicology testing used in the development of standards typically uses durations of at least 24 hours, and sometimes 48 or even 96 hours (USEPA, 2000). However, data are lacking for the biological impact of shorter duration periods such as are of concern for this question.

Lacking such data, the WQQC recommends a statistical approach to determine the probability that a given measurement or set of measurements suggest that the minimum DO water quality standard (e.g. dissolved oxygen not less than 4.0 mg/L at any time in FW2-NT waters) has been violated. The WQQC conducted a preliminary study of the statistical behavior of diurnal DO profiles using data collected by the USGS for the Pequannock River at Oak Ridge (01382210). These data were selected because they had a near absence of diurnal variation, and thus provide a conservative indication of variance in measurement error. (The actual analytical variance is likely to be lower than the computed value.)

This approach will also apply to determining whether the mean of *n* samples is below the water quality standard for average *DO* over a 24-hour period.

The Confidence of Compliance (*CoC*) is the probability that the water quality standard (WQS) has been met, given the available data [Jagupilla *et al*, 2008, 2013]. Appendix 1 shows how to compute the *CoC* and gives several example calculations.

For example, if the average *DO* of four samples taken over a one-hour period is 3.95 mg/L, there is a 4.9% probability that the water quality standard of a minimum *DO* of 4.0 mg/L has been met.
References for Question 1:

NEW JERSEY ADMINISTRATIVE CODE 7:9B-1.14(d) General Surface Water Quality Criteria for FW2, SE and SC Waters

U.S. Environmental Protection Agency, “Ambient Aquatic Life Water Quality Criteria for Dissolved Oxygen (Saltwater): Cape Cod to Cape Hatteras”, November 2000

Jagupilla, S.C.K., D. Vaccari, R. Miskewitz, R Hires, T Su. 2013. TMDL Report – Passaic River Pathogen TMDL. *New Jersey Department of Environmental Protection*. Trenton, NJ.

Jagupilla, S.C.K., D. Vaccari, R. Hires. 2008. “Computing the Bacteriological TMDL Rollbacks by the Statistical Rollback Method”. *Protection and Restoration of the Environment Conference*. Kefalonia, Greece.

USGS National Water Information System, Water Data for the Nation Automated Retrievals, http://nwis.waterdata.usgs.gov/nwis/?automated_retrieval_info, Retrieved: 2012-06-12: USGS 01382210 Pequannock River at Oak Ridge NJ

Question #2:

How can the DO discrete, grab sampling results from the ambient stream network best be used to identify waters that are most likely to exceed these criteria?

Finding #2:

Preliminary analysis shows a predicted value for the minimum or average DO could be estimated using two measurements nominally one hour apart taken between 9am and noon. The NJDEP should develop and validate a correlation or interpolation method to make this estimation. The estimated values could be used to screen for sites requiring more thorough sampling. We recommend that diurnal DO measurement be used more frequently to confirm possible violations of the Water Quality Standards.

Basis for Finding #2:

The committee decided to propose the use of predicted C_{min} instead of a fixed time of day (such as dawn or 9am). The preliminary analysis of diurnal DO data shown in Appendix 2 demonstrates that DO swing during the growing season exhibits strong regularities that can be captured in linear models. This demonstrates the feasibility of the proposed approach.

In the analysis shown in Appendix 2, multilinear models were developed to accurately estimate the minimum DO (C_{min}) and the DO swing (ΔDO), or the difference between the maximum and minimum DO on a day. A similar modeling effort should be able to develop a validated model to predict the minimum DO value from any two measurements taken between 9am and noon on the same day, separated by at least one hour.

The development of such a model would require a more thorough study than shown in the Appendix 2. The analysis in the appendix was conducted with data from a single year. Furthermore, it is effectively a curve-fit, and has not been validated with independent data. A complete development should include multiple years of data, and should validate its applicability to varying conditions, such as hydrologic parameters (e.g. flow and stream geometry), temperature, and water quality parameters (e.g. measures of eutrophication or of sediment oxygen demand), and should include sites that exhibit either little or inverse diurnal swing (such as the lower Millstone River, which is dominated by sediment oxygen demand (SOD).

References for Question #2:

See Appendix 2.

Question #3:

What specific statistical methods or models should be used to evaluate the coastal water quality data collected along transects by either Underwater Automatic Vehicles or continuous recording devices attached to vessels?

Finding #3:

Preliminary analysis of autonomous vehicle observations of dissolved oxygen off the New Jersey coast show that while there are persistent patterns in the dissolved oxygen fields associated with water depth and stratification off our coasts, rapid changes can occur with varied responses across the region. Therefore the committee recommends that autonomous vehicle data be analyzed with spatial decorrelation models in both space and time. The committee also notes that data analysis would benefit from concurrent fixed sensors to help differentiate the contribution of temporal and spatial variability.

Basis for Finding #3:

The coastal ocean is a highly variable system with processes that have significant implications on the hydrographic and oxygen characteristics of the water column. The spatial and temporal variability of these fields can cause dramatic changes to water quality and in turn the health of the ecosystem. While low Dissolved Oxygen (DO) concentrations are not uncommon in the coastal ocean, what is less understood is how the location and size of these low DO regions vary and what impact that variability has on ecosystem health. Therefore alternative sampling strategies are needed to continuously map these low DO areas in a way that quantifies this variability. These strategies include continuous underway measurements from either autonomous underwater vehicles or surface vessels. Interpretation and analysis of these continuous data sampled by these platforms must consider the variability in time and space. In response to question 3 posed by NJDEP related to the analysis techniques and interpretation of these types of data, we cite recent results from a coastal monitoring project focused on the dynamics of DO off the New Jersey coast sampled through a series of six glider deployments [Kohut et al., 2014].

The strongest gradients in DO off the New Jersey coast are typically across the thermocline with surface waters usually much more oxygenated than the bottom waters. These gradients tend to be weaker closer to the coast and can be significantly weakened following several strong wind events like nor'easters and tropical storms. Recent results from autonomous platform missions along the New Jersey coast indicate that spatial variability explained most of the variability with more mixed conditions in the shallow waters near the coast and more stratified conditions in the deeper water offshore. It was in the deeper waters offshore that most of the lower DO concentrations can be found below the thermocline. The scale of this variability observed over two seasons is on the order of 60-80 km in space and 3-4 days in time. From these data, we conclude that these decorrelation scales are representative of the distance over which the water depth varied. The time scale is more an indicator of the time it takes the glider to cover this distance rather than a change across all space in time.

There are also observed changes in time, predominately caused by strong (Hurricane Irene) and moderate wind events that mixed the more oxygenated surface water with the deeper less oxygenated water. During Hurricane Irene there was rapid mixing of the more oxygenated surface waters across the thermocline and into the bottom waters. In addition, events like Irene and the coastal bloom in 2011 highlighted the capability to adapt pre-determined missions to respond to these events. This ensures that observations are taken relative to the bloom throughout the storm.

The results show that while there are persistent patterns in the dissolved oxygen fields associated with water depth and stratification off our coasts, rapid changes can occur with varied responses across the region. These results highlight the need to coordinate the high-resolution data sampled along the gliders path with strategic point measurements in time. Based on these missions, a line of at least two moored bottom DO time series stations oriented across the shelf would help to distinguish the variability observed by the glider in space and time. These point observations combined with the coast wide coverage of the glider missions would be able to identify regions of low DO and characterize how they evolve through time.

Since the glider is a non-stationary platform it is important to state that it is simultaneously sampling temporal and spatial change. It is difficult to differentiate a measured change in DO concentration as a change in time or a change in space when looking at the glider data in isolation. Using autocorrelation we calculated the decorrelation time and length scales for each deployment. The decorrelation scale is defined as the scale, in time or space, in which the autocovariance coefficient falls below 0. These scales describe the time and space over which the DO variability becomes uncorrelated. For example, a decorrelation length scale of 50 km indicates that the DO observations at any point are correlated with DO observations within 50 km. Similarly, a decorrelation time scale of 5 hours indicates that the DO observations at a particular time are correlated with DO observations at that point for 5 hours before and after the measurement. These scales can be used to guide the sampling required in time and space to capture the variability of DO along the coast.

Appendix C provides the method to be used in the software package 'R' to calculate decorrelation scales in time and space for a data series sampled by the glider in the upper (3-4m below the surface) and lower water column (3-4 m above the bottom). This code requires the base 'R' library packages. No additional libraries need to be installed. In this example, the resolution for the determination of the spatial and temporal decorrelation scales is set to 0.1 km and 5 minutes, respectively.

References for Question 3:

Kohut, J., C. Haldeman, and J. Kerfoot. Monitoring Dissolved Oxygen in New Jersey Coastal Waters Using Autonomous Gliders. U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-13/180, 2014.

Question #4:

The methodology in the EPA (coastal and estuarine) criteria document is broad and may be modified to protect local marine/estuarine species.

a. *Are the thresholds recommended by EPA and/or those modified by NYDEC adequately protective of the animals found in New Jersey's coastal waters and specifically, Barnegat Bay?*
b. *Should the time period for a low DO event be revised to protect NJ species? Are the criteria adequately protective for early life stages?*
c. *What type of data is needed to determine if criteria are sufficient to protect NJ biota?*

Finding #4a:

The NYDEC model does not account for all species within New Jersey's estuaries but could be adopted once consideration of how constant monitoring data should be interpreted.

Finding #4b:

NJDEP should develop its own frequency of exceedences for DO. However, NJDEP should not adopt criteria for bays until site-specific DO data become available.

Finding #4c:

Measurements should be made of impacts of DO on sentinel species on an estuary-by-estuary basis, or by groups of estuaries, and in coastal areas. In addition the role of other measurements such as contaminant levels should also be considered.

Basis for #4:

The USEPA has proposed an approach for the Virginian Province (Cape Cod to Cape Hatteras) to protect organisms that live in saltwater. New York State has modified this approach and will apply it to the NY-NJ Harbor. Some considerations for answering question 4 include:

- Should NJ adopt the approach developed by NY?
- Can the approach developed for the NY-NJ Harbor be used for the NJ coast, and or the Delaware Bay?, or
- Are there specific species that NJ should be protecting in the coastal areas, and Delaware Bay?

Basis for #4a:

There are very few species unique to NJ, not present in the NYDEC model. One example is the horseshoe crab. Horseshoe crabs are found primarily on the ocean side of the barrier islands and lay their eggs in specific areas on the beaches. Although they are a protected species it is unlikely that they will be heavily impacted because of their ability to move out of these low DO events. A species ability to move outside of a low DO event will govern much of the decision process in how to use constant monitoring data.

Barnegat species most affected by low DO are going to be the benthic communities because of the sedimentation characteristics of the Bay. Species diversity, chemical pollutants and physical chemical characteristics of the sediments and pore waters are all critical data in evaluating the potential risk to species in the Bay. *Byproducts* of a low DO environment (e.g. evolution of ammonia) may actually create a greater risk to biota than DO itself which would otherwise not be viewed as detrimental. The time of year is also a critical component both from the temperature but also for the specific species that are undergoing reproduction and have sensitive life-stages. For some species the change in the makeup of plant species (i.e. loss of eel grass) present can also drastically impact species survival which in many cases is related to nutrient loads. Although we deal with these parameters as individual values there are important interactions that can drastically impact aquatic toxicity.

One of the problems with barrier island bays is adequate water circulation to maintain removal of nutrients that directly affect algal blooms and DO levels in the water column. Other anthropogenic sources (land applied fertilizers and recreational uses) will also contribute to the low DO events in these waters. A more protective measure other than a direct DO reading could include regulation of watershed use or a unified (statewide) approach to regulation of allowed contaminant contribution into the bay rather than the current community based standards. Alternatively, increased flow patterns into the Bay (including man made) might correct some of the restrictive hydrologic flow pattern into and out of the Bays.

Basis for #4b:

The important question is – How should the DO criterion be modified in light of continuous monitoring data? The committee felt that three factors should be included: seasonality, time, and duration of exposure. An additional consideration is upwelling. It was noted that upwelling tends to occur in the same locations every year.

DO risk assessments have largely been carried out for periods of time of 24 hours and due to tidal cycles this should be protective, exception is for a large event for multiple days caused meteorological (high temps with an algal bloom) and hydrological events along the Jersey Coast previously reported. These extraordinary events cannot be predicted and will not be amenable to standards unless there is an anthropogenic cause to the increased occurrence of these low DO events.

Motile pelagic larval and juvenile stages should not be heavily impacted by low DO events since they are dependent on tidal movement and water column DO. The benthic communities both adult and recently set species such as bivalves surf-clams, quahogs and other benthic crustaceans and annelid worms are likely the most sensitive to prolonged DO events, because of their limited mobility. It is for this reason that a more careful consideration of the DO studies must be examined. The discussion below describes some of the current thinking on DO in estuarine environments. It should be pointed out that single measurements are of limited use so efforts should be made to have more frequent or continuous monitoring stations.

“The juvenile/adult survival and the growth criteria provide boundaries within which to judge the DO status of a given site. If DO conditions are above the chronic growth criterion (4.8 mg/L), then this site would meet objectives for protection. If DO conditions are below the juvenile/adult survival criterion (2.3 mg/L), then this site would not meet objectives for protection. When the DO conditions are between these two values, then the site would require evaluation using the larval recruitment model that integrates duration and intensity of hypoxia to determine suitability of habitat for the larval recruitment objective.” (Federal Register: November 30, 2000)

Basis for #4c:

The DO criteria that are discussed here are based entirely on laboratory findings (Figure 1). Field observations on the impact of low DO levels support the findings of laboratory studies. Field acute effects occurred in juvenile and adult animals at 2.0 mg/L, which would be predicted based on the 2.3 mg/L juvenile/adult criterion. In the field, behavioral effects generally occurred within the range where many of the laboratory sub-lethal effects occurred. These behavioral responses can be protective for the species (i.e. shell closure and suspended filtering) for short periods of time. Species that are mobile may avoid low DO areas, however species confined to these areas will demonstrate behavioral stress responses (flared gills, surface gulping or erratic swimming).

In addition, as with all criteria, these criteria do not account for changes in sensitivity to low DO that accompany other stresses, such as high temperature, extremes of salinity, or toxicants (Diaz and Rosenberg 2008). Chief among these concerns would be high temperature because high temperature and low DO often appear together. Generally, low DO would be more lethal at water temperatures approaching the upper thermal limit for each species. The USEPA reports that the DO limits provided in their document are sufficiently protective under most conditions where aquatic organisms are not otherwise unduly stressed. We would tend to agree with this assertion, however, site specific conditions may warrant actions being taken due to impacts that will reduce survival due to a shift in the dose response curves shown in Figure 1. For example, using 3.5 as a cut off for not meeting objectives for protection would be prudent since this value would protect ~ 90% of the *Homarus* and >50% of the *Dyspanopeus*.

The recommended criteria described here apply to both continuous (persistent) and cyclic (diel, tidal, or episodic) hypoxia. If DO exceeds the chronic protective value for growth (4.8 mg/L), the site meets objectives for protection. If DO is below the limit for juvenile and adult survival (2.3 mg/L), the site does not meet objectives for protection. When the DO is between these values, the site requires evaluation of duration and intensity of hypoxia to determine suitability of habitat for the larval recruitment objective. These criteria would appear to be sufficiently protective for estuarine waters along the New Jersey coast with adequate hydrodynamic turnover. In certain growing areas where intense aquaculture is ongoing these criteria may need to be modified based on the individual species.

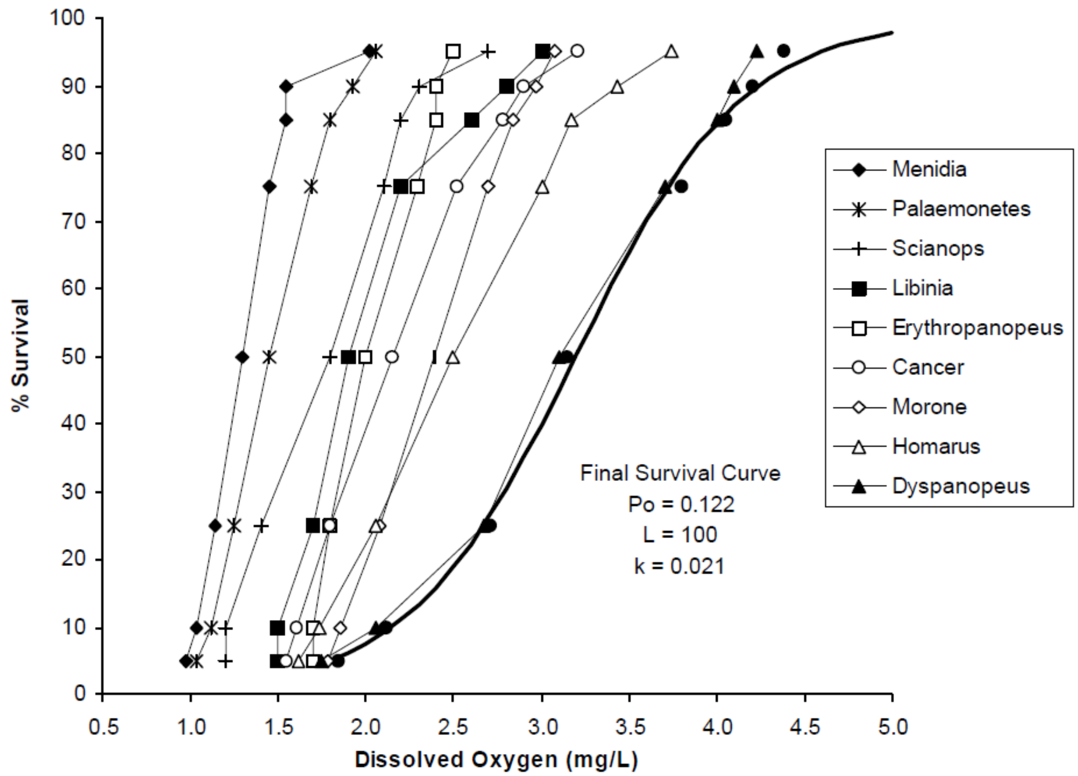


Figure 1. Twenty-four hour dose-response curves for nine genera used in the larval recruitment model. Dark solid line is the regression line of best fit for the FSC. See text for explanation of FSC and of P0, L, and k. The Solver routine in Microsoft® Excel 97 was used to determine P0 and k. *Source: [Federal Register: November 30, 2000 (Volume 65, Number 231)] [Notices] [Page 71317-71321] From the Federal Register Online via GPO Access [wais.access.gpo.gov][DOCID:fr30no00-51]. The equation for the Final Survival Curve (FSC) in Figure 1 (bold line) was derived by an iterative process of fitting the best fit sigmoidal line through the points generated by the survival from the most sensitive species as a variable of dissolved oxygen.

Within the Figures 2 and 3 are shown the temperature and dissolved oxygen levels in Barnegat Bay, which demonstrate that both of these parameters vary over quite large ranges and that temperature is an important stressor for organisms and would change the dose response curves for estuarine aquatic organisms. Generally high temperatures in this shallow system would result in additional stress to both benthic invertebrates and fish species. The occurrence of diurnal algal blooms in these shallow water systems would also cause large fluctuations in the DO that would result in poor survival rates depending on the duration of the event.

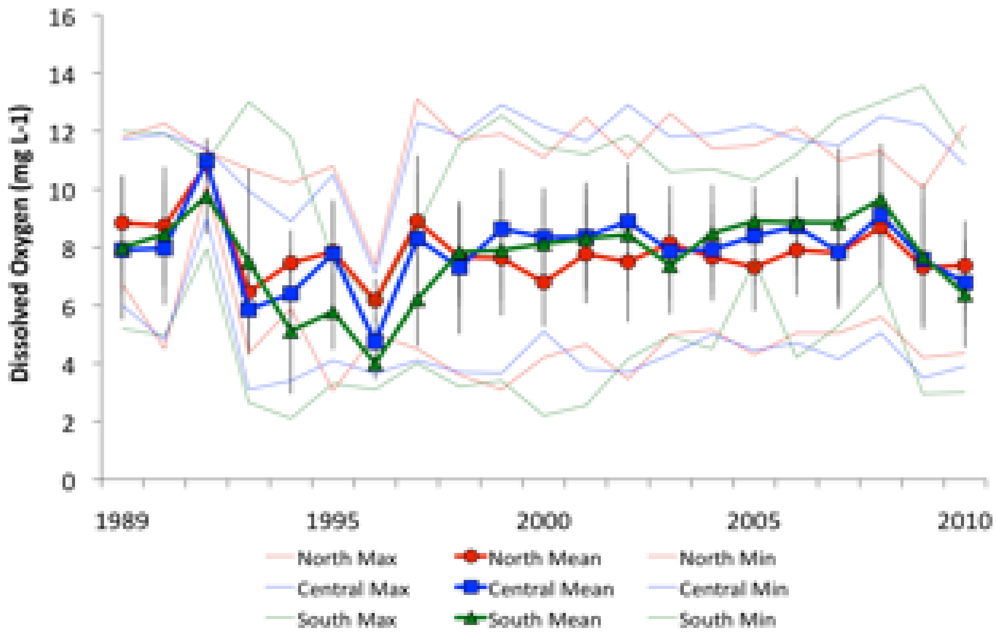


Figure 1. Minimum, mean, and maximum dissolved oxygen values recorded in the BB-LEH Estuary from 1989 to 2010. Data from the New Jersey Department of Environmental Protection.

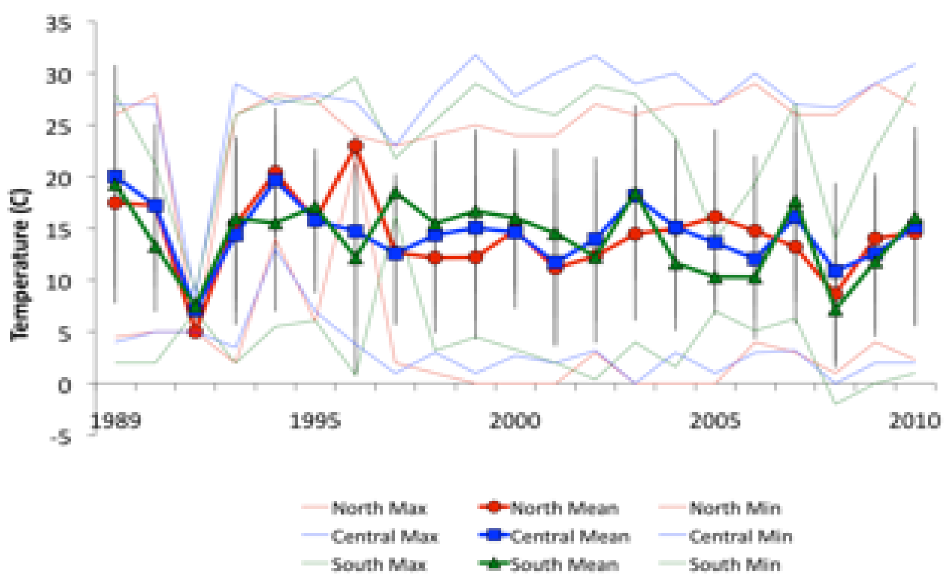


Figure 3. Minimum, mean, and maximum temperatures recorded in the BB-LEH Estuary from 1989-2010. Data from the New Jersey Department of Environmental Protection.

DO is due in part to the overlying waters but also the sediment composition. Sediments containing high organic matter generally have a higher potential for anoxic conditions due to organic matter breakdown. Bays such as Barnegat have diverse bottom make-up going from predominantly sand to heavy organic carbon content which will result in lower DO concentrations at the sediment surface. The physical characteristics of the benthic sediments will often dictate the makeup of the benthic communities.

References for Question 4:

1. Federal Register: November 30, 2000 (Volume 65, Number 231) and EPA document EPA-822-R-00-012 **Ambient Aquatic Life Water Quality Criteria for Dissolved Oxygen (Saltwater): Cape Cod to Cape Hatteras,** November 2000

2. Michael Kennish and Benjamin Fertig, Assessment of Nutrient Loading and Eutrophication in Barnegat Bay-Little Egg Harbor, New Jersey in Support of Nutrient Management Planning Prepared for: New England Interstate Water Pollution Control Commission, Fig 2-1 and 2-2 pg 169. http://www.neiwpcc.org/nynj-assessment.asp

APPENDIX 1

Confidence of Compliance of DO Measurements

David A. Vaccari and Sarath Chandra K. Jagupilla
Stevens Institute of Technology

November, 2015

The WQQC conducted a preliminary study of the statistical behavior of diurnal DO profiles using data collected by the USGS for the Pequannock River at Oak Ridge (01382210) at 15-minute intervals from 0:00 on 7/29/11 to 23:45 on 8/8/11. These data were selected because they had a near absence of diurnal variation, and thus provide an indication of variance in measurement error. The 1056 data points ranged from 3.6 to 4.0 mg/L, and had a standard deviation of 0.0671 mg/L. We assume that this is the population standard deviation, which is justified given the large number of data points. It may be assumed that 90% of this variability is due to analytical error, conservatively yielding a standard deviation (σ) of 0.0604 mg/L due to this effect.

The Confidence of Compliance (*CoC*) is the probability that the water quality standard (WQS) has been met, given the available data [Jagupilla *et al*, 2008, 2013].

We computed the difference between the measured dissolved oxygen (*DO*) and the WQS (Δ_m):

$$\Delta_m = DO - WQS \qquad \text{(A1)}$$

Note that *DO* may be the average of *n* individual measurements. In that case, the standard deviation of the mean is computed as:

$$\sigma_{\acute{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{(A2)}$$

To compute the value of Δ_m corresponding to any given *CoC*, we need to know the number of standard deviations from the mean for that level (the *z* statistic). For example, for *CoC* = 10%, *z* = -1.282. Put another way, if the *DO* is 1.282 standard deviations below the *WQS,* then there is a 90% probability that the standard has been violated. In mathematical terms:

$$\Delta_m = \frac{\sigma}{\sqrt{n}} \cdot z(CoC) \qquad \text{(A3)}$$

The value of *z* can be computed from the *CoC* using the Microsoft Excel function:

=NORM.S.INV(*CoC*) (A4)

Table A-1 gives the values for *z* for several given values of *CoC*, and the associated Δ_m for various values of *n* and for the value of σ given above (0.0604 mg/L).

Table A-1

	N =	1	4	24	96
CoC	**z**	**Δ_m**	**Δ_m**	**Δ_m**	**Δ_m**
1%	-2.326	-0.1405	-0.0703	-0.0287	-0.0143
5%	-1.645	-0.0993	-0.0497	-0.0203	-0.0101
10%	-1.282	-0.0774	-0.0387	-0.0158	-0.0079
20%	-0.842	-0.0508	-0.0254	-0.0104	-0.0052
50%	0.000	0.0000	0.0000	0.0000	0.0000
90%	1.282	0.0774	0.0387	0.0158	0.0079
95%	1.645	0.0993	0.0497	0.0203	0.0101
99%	2.326	0.0774	0.0703	0.0287	0.0143

Note that Δ_m can be computed for values of *n* from the value given in column 3 of the Table simply by dividing by $\sqrt{n}$.

Example A-1:

What average *DO* corresponds to a 1% Confidence of Compliance for four samples where the *WQS* is that the *DO* be no less than 4.0 mg/L? The null hypothesis (H_0) is $\mu \geq$ 4.0 (compliance); the alternative hypothesis (H_1) is $\mu <$ 4.0 (violation of the *WQS* for minimum *DO*).

The *z* value from the Table above at 1% *CoC* is -2.326. Using this value with σ = 0.0604 and *n* = 4 in equation (A3), we obtain:

$$\Delta_m = \frac{0.0604}{\sqrt{4}} \cdot (-2.326) = -0.070\, mg/L$$

Thus, the average *DO* can be found by rearranging equation (A1):

DO = *WQS* + Δ_m = 4.00 – 0.070 = <u>3.93 mg/L</u>

The interpretation of this result is that an average of 4 *DO* values below 3.93 mg/L would have less than a 1% chance of occurring if the water quality standard were not being violated.

Consider the inverse problem: What is the *CoC* for a given *DO* (the average of *n* individual measurements)? The above relations can be inverted to the following:

$$CoC = cpd(z) \qquad \text{Where: } z = \frac{DO - WQS}{\sigma/\sqrt{n}} \qquad \text{(A5)}$$

The function *cpd*(*z*) is the cumulative normal probability distribution, and can be computed using the Microsoft Excel function:

=NORM.S.DIST(z,1) (A6)

Example A-2:

Suppose four samples (e.g. 4.0, 4.0, 3.9, 3.9) are averaged to produce a mean *DO* = 3.95 mg/L. First compute *z* using equation (A5):

$$z = \frac{3.95 - 4.00}{0.0604/\sqrt{4}} = -1.656$$

Then, use equation (A6) in Microsoft Excel to compute the *CoC*:

CoC = NORM.S.DIST(-1.656,1) = <u>4.9%</u>

The interpretation is that the probability that the water quality standard has not been violated given these data is 4.9%. I.e, the probability of a violation is 95.1%.

If a *DO* value of 3.95 mg/L were measured with a single sample, the *CoC* would be 20.4%.

In a third example, if the four *DO* measurements were 3.9, 4.0, 4.0 and 4.0 mg/L, for an average of 3.975 mg/L, the *CoC* would also be 20.4%.

For the case of the *average* water quality standard (rather than the minimum), then a full day's worth of data are available. If the data were collected at 15-minute intervals, then *n* = 96 and the *CoC* is computed using equation A5 in the same way as in example A-2. As a further example for this situation, consider 96 values are collected over a 24-hour period, and a *WQS* = 5.0 mg/L. If the mean of the 96 measurements is 4.99 mg/L, the *CoC* will be 5.2%. If the mean of the 96 measurements is 4.98 mg/L, the *CoC* will be 0.059%.

<u>*References:*</u>

U.S. Environmental Protection Agency, "Ambient Aquatic Life Water Quality Criteria for Dissolved Oxygen (Saltwater): Cape Cod to Cape Hatteras", November 2000

Jagupilla, S.C.K., D. Vaccari, R. Miskewitz, R Hires, T Su. 2013. TMDL Report – Passaic River Pathogen TMDL. *New Jersey Department of Environmental Protection*. Trenton, NJ.

Jagupilla, S.C.K., D. Vaccari, R. Hires. 2008. "Computing the Bacteriological TMDL Rollbacks by the Statistical Rollback Method". *Protection and Restoration of the Environment Conference*. Kefalonia, Greece.

USGS National Water Information System, Water Data for the Nation Automated Retrievals, http://nwis.waterdata.usgs.gov/nwis/?automated_retrieval_info, Retrieved: 2012-06-12:
USGS 01382210 Pequannock River at Oak Ridge NJ

APPENDIX 2

Analysis of DO swing patterns

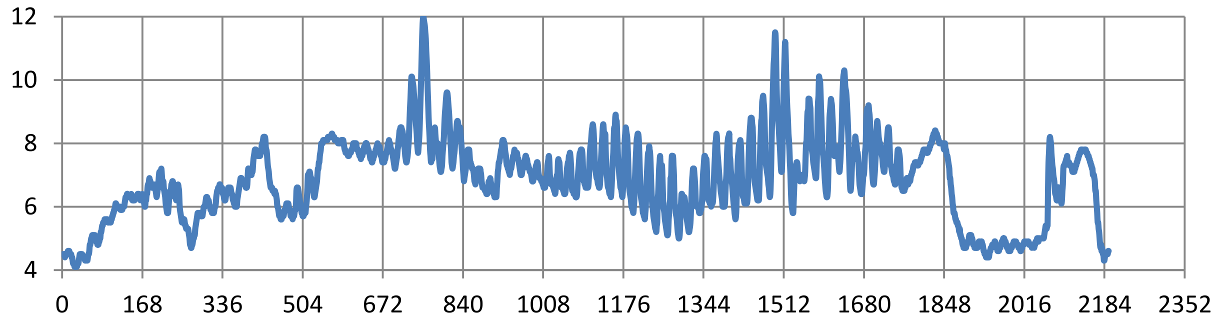
David A. Vaccari

Stevens Institute of Technology

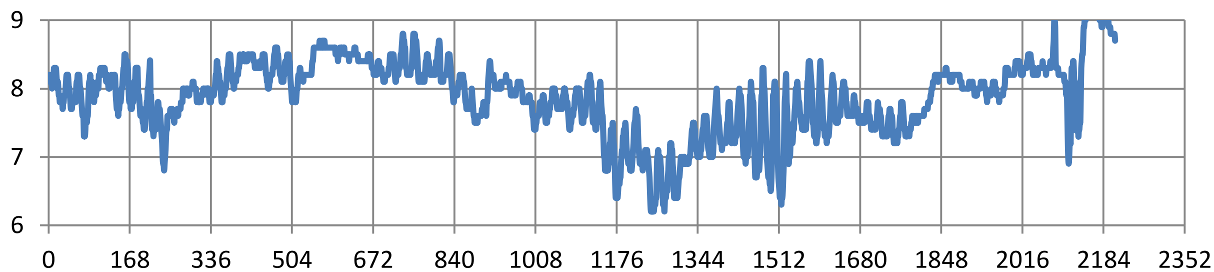
August 15, 2012

Continuous DO data were obtained for the following sample locations (USGS 2012). The corresponding time periods were selected to represent growing season trends.

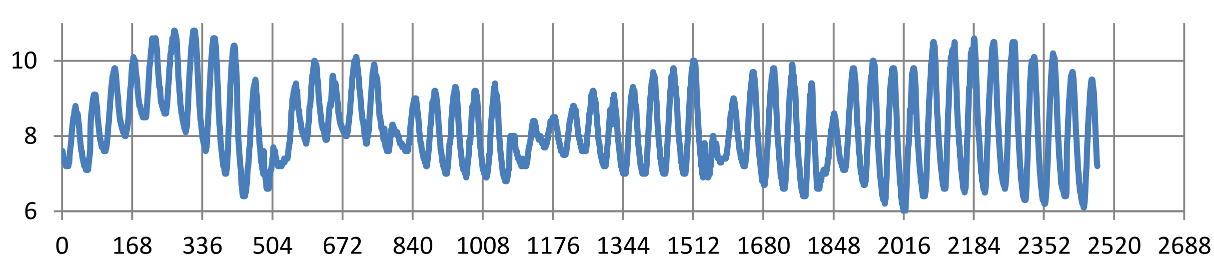
- USGS 01389005 Passaic River below Pompton Riv at Two Bridges NJ
  - 7/18/11 to 8/6/11 (1120 to 1596 hours in the graph below)



- USGS 01388000 Ramapo River at Pompton Lakes NJ
  - 7/27/11 to 8/13/11 (1345 to 1796 hours)



- USGS 01400500 Raritan River at Manville NJ
  - 7/4/11 to 7/22/11 (1577 to 2484 hours)



The data were used to determine three questions:

1. What is the duration of night-time low DO events?
2. Can multiple measurements of DO be used to predict the minimum night-time DO?
3. Can multiple measurements of DO be used to predict the DO swing?

Duration of night-time low-DO events

The duration was defined as follows: The DO at 9am on each day was obtained by the data. The low-DO duration was defined as the number of hours since the the DO was last equal to or exceeded the 9am DO. There were 53 durations that resulted. The statistics of these values are as follows:

	Minimum	**1.0**
25%	**Percentile**	**10.1**
	Median	**12.0**
75%	**Percentile**	**14.0**
	Maximum	**18.0**
	Avg	**11.9**

Thus, half the low-DO duration values were between 10.1 and 14.0 hours. The median value of 12.0 would seem to be both representative and convenient for use as an estimate for the low-DO duration.

Prediction of minimum DO

The minimum DO was found for the period from 12pm to 12am for a combined 77 days for the three sample locations. A strong multilinear correlation was found to predict minimum DO (C_{min}) from the DO at 9am (C_{9am}) and the rate of increase of DO (r, mg/L/hr) from t_1 to t_2 where t_1 was 9am and t_2 was either 10am or 11am.

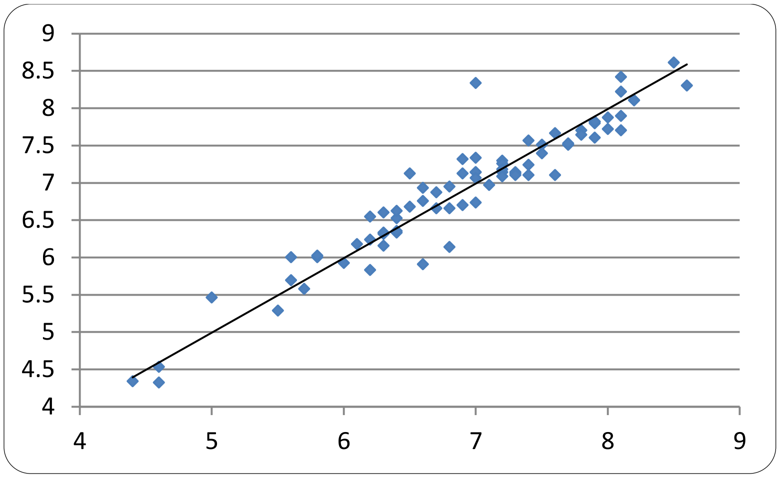
$$r = \frac{C(t_2) - C(t_1)}{t_2 - t_1}$$

In the cases of all the models described here, the intercept was not significant.

For the case of t_2 = 10am the resulting model was:

$$C_{min} = -1.156 \cdot C_{9am} + 0.965 \cdot r \qquad R^2 = 0.998$$

The plot of predicted versus observed follows:



For the case of t_2 = 11am the resulting model was:

$$C_{min} = -1.452 \cdot C_{9am} + 0.972 \cdot r \qquad R^2 = 0.999$$

Thus, minimum DO could be well predicted using the 9am DO and the rate of DO change over the next one or two hours.

Prediction of DO Swing

The DO Swing (ΔDO) was computed as:

$$\Delta DO = C_{max} - C_{min}$$

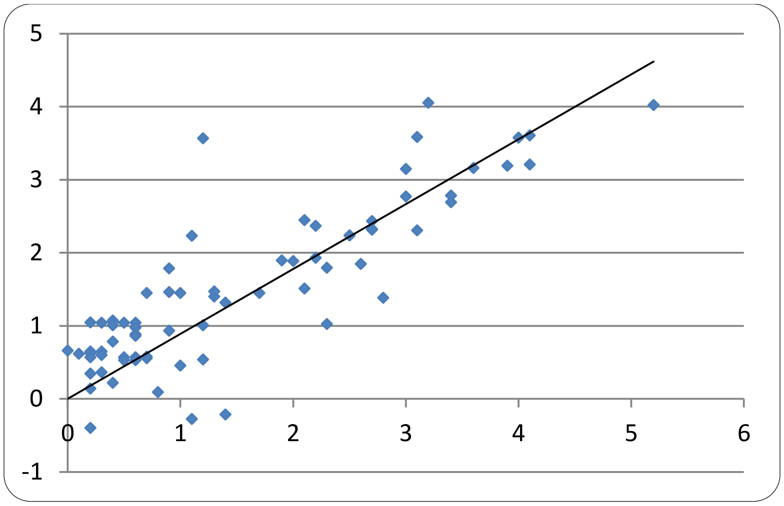
where C_{max} was the maximum DO between the hours of 9am and 9pm.

As before, a strong multilinear correlation was found to predict minimum ΔDO from the DO at 9am (C_{9am}) and the rate of increase of DO (r, mg/L/hr) from t_1 to t_2 where t_1 was 9am and t_2 was either 10am or 11am.

For the case of t_2 = 10am the resulting model was:

$$\Delta DO = 4.305 \cdot C_{9am} + 0.0772 \cdot r \qquad R^2 = 0.888$$

The plot of predicted versus observed follows:



For the case of t_2 = 11am the resulting model was:

$$\Delta DO = 5.256 \cdot C_{9am} + 0.0580 \cdot r \qquad R^2 = 0.923$$

Thus, DO swing could be well predicted using the 9am DO and the rate of DO change over the next one or two hours.

Generalized model for AM data

A more general model could be developed to estimate C_{min} and DO swing from DO values measured at any two times in the morning.

The predicted values in this case would be:

$$C_{min} = a_1 \cdot C_{t2} + r \cdot [a_1 \cdot (9.0 - t_2) + a_2]$$

where a_1 = -1.282 and a_2 = 0.968

$$\Delta DO = a_3 \cdot C_{t2} + r \cdot [a_3 \cdot (9.0 - t_2) + a_4]$$

where a_3 = 4.710 and a_4 = 0.0692, and as before:

$$r = \frac{C(t_2) - C(t_1)}{t_2 - t_1}$$

The two sampling times, t_1 and t_2, should be between the hours of 9am and 11am, and should be at least one hour apart. It's possible that the time constraint for this type of model could be reduced. For example, it might be possible to extend the range to from, say, 8am to noon.

It should be emphasized that the predicted values of C_{min} (or similar predictions that could be made for the average *DO*) should not be used for enforcement purposes. Rather they should be treated as screening values to indicate the possibility of a violation of the Water Quality Standard. If the predicted values are found to be in violation of the standard, then a more thorough sampling method should be employed, such as installation of a sonde for semicontinous diurnal *DO* measurement.

Reference:

USGS National Water Information System, Water Data for the Nation Automated Retrievals, http://nwis.waterdata.usgs.gov/nwis/?automated_retrieval_info, Retrieved: 2012-06-12:
USGS 01382210 Pequannock River at Oak Ridge NJ
USGS 01389005 Passaic River below Pompton Riv at Two Bridges NJ
USGS 01388000 Ramapo River at Pompton Lakes NJ
USGS 01400500 Raritan River at Manville NJ

Appendix C:

```
# clear workspace
rm(list==ls())

# define path to csv files and list all
fpath<-"/Users/palamara/Documents/gliders/csv/ru28-422/"
fnames<-list.files(path=fpath,pattern=".csv")

# initialize data frame
all.data<-
data.frame(time=numeric(0),dist=numeric(0),depth=numeric(0),surfDO=numeric(0),botDO=numeric(0))
all.data.names<-names(all.data)

# read in data
for (n in 1:length(fnames)){
	# check if file is empty
	goodfile<-TRUE
	f1test<-read.csv(paste(fpath,fnames[n],sep=""),header=FALSE,nrows=14,row.names=NULL)
	for (test_good in 1:dim(f1test)[1]){
		if (grepl("%num_records::0",f1test[test_good,])){
			goodfile<-FALSE
		}
	}
	if (goodfile){
	# header with variable names
	f1head<-
read.csv(paste(fpath,fnames[n],sep=""),header=FALSE,skip=14,nrows=1,row.names=NULL)
	# data
	f1<-read.csv(paste(fpath,fnames[n],sep=""),header=FALSE,skip=16,row.names=NULL)
	t1<-NA
	dist1<-NA
	depth1<-NA
	sDO1<-NA
	bDO1<-NA
	DO1<-NA
	maxdepth1<-NA
	for (n1 in 1:length(f1head)){
		if (grepl("time",f1head[,n1])){
			t1<-mean(f1[,n1],na.rm=TRUE)
		}
		if (grepl("pressure",f1head[,n1])){
			depth1<-f1[,n1]
		}
		if ((grepl("oxy",f1head[,n1]))&(grepl("mass_cor",f1head[,n1]))){
			DO1<-f1[,n1]
		}
		if (grepl("water_depth",f1head[,n1])){
			maxdepth1<-mean(f1[,n1],na.rm=TRUE)
		}
		if (grepl("distance",f1head[,n1])){
			dist1<-mean(f1[,n1],na.rm=TRUE)
		}
	}
	# find surface and bottom indices, calculate DO
	surfind<-which(depth1>3&depth1<4)
```

```
		if (!is.na(maxdepth1)){
			alt<-maxdepth1-depth1
			botind<-which(alt>3&alt<4)
		} else{
			maxdepth1<-max(depth1,na.rm=T)
			alt<-maxdepth1-depth1
			botind<-which(alt>0&alt<1)
		}
		sDO1<-mean(DO1[surfind],na.rm=TRUE)
		bDO1<-mean(DO1[botind],na.rm=TRUE)
		# add row to all.data
		all.data<-rbind(all.data,c(t1,dist1,maxdepth1,sDO1,bDO1))
		}
}
# correct the column names in all.data
names(all.data)<-all.data.names


# define resolutions for autocovariance calculations, and calculate number of steps at that resolution
spatial_resolution<-.1
spatial_N<-ceiling((max(all.data$dist,na.rm=T)-min(all.data$dist,na.rm=T))/spatial_resolution)
temporal_resolution<-5*60
temporal_N<-ceiling((max(all.data$time,na.rm=T)-min(all.data$time,na.rm=T))/temporal_resolution)

# bin temporal data
bin.temp<-
as.data.frame(seq(from=min(all.data$time,na.rm=T)+temporal_resolution/2,to=min(all.data$time,na.rm=T)
+temporal_resolution*temporal_N-temporal_resolution/2,by=temporal_resolution))
bin.temp<-cbind(bin.temp,NA)
bin.temp<-cbind(bin.temp,NA)
names(bin.temp)<-c("time","surfDO","botDO")
for (n in 1:length(bin.temp$time)){
		ind<-
which((all.data$time<bin.temp$time[n]+temporal_resolution/2)&(all.data$time>=bin.temp$time[n]-
temporal_resolution/2))
		bin.temp$surfDO[n]<-mean(all.data$surfDO[ind],na.rm=T)
		bin.temp$botDO[n]<-mean(all.data$botDO[ind],na.rm=T)
}

# bin spatial data
bin.spatial$dist<-
as.data.frame(seq(from=min(all.data$dist,na.rm=T)+spatial_resolution/2,to=min(all.data$dist,na.rm=T)+sp
atial_resolution*spatial_N-spatial_resolution/2,by=spatial_resolution))
bin.spatial<-cbind(bin.temp,NA)
bin.spatial<-cbind(bin.temp,NA)
names(bin.temp)<-c("dist","surfDO","botDO")
for (n in 1:length(bin.spatial$dist)){
		ind<-
which((all.data$dist<bin.spatial$dist[n]+spatial_resolution/2)&(all.data$dist>=bin.spatial$dist[n]-
spatial_resolution/2))
		bin.spatial$surfDO[n]<-mean(all.data$surfDO[ind],na.rm=T)
		bin.spatial$botDO[n]<-mean(all.data$botDO[ind],na.rm=T)
}

# run autocovariance function
surfDO.temporal.cov<-acf(bin.temp$surfDO,type="covariance",na.action=na.pass,lag.max=temporal_N-1)
```

```
botDO.temporal.cov<-acf(bin.temp$botDO,type="covariance",na.action=na.pass,lag.max=temporal_N-1)
surfDO.spatial.cov<-acf(bin.spatial$surfDO,type="covariance",na.action=na.pass,lag.max=spatial_N-1)
botDO.spatial.cov<-acf(bin.spatial$botDO,type="covariance",na.action=na.pass,lag.max=spatial_N-1)

# find point of zero-crossing for temporal autocovariance of surface DO
ind<-which(surfDO.temporal.cov$acf<0)
ind<-min(ind)
surfDO.temporal.lag<-temporal_resolution*surfDO.temporal.cov$lag[ind]

# find point of zero-crossing for temporal autocovariance of bottom DO
ind<-which(botDO.temporal.cov$acf<0)
ind<-min(ind)
botDO.temporal.lag<-temporal_resolution*botDO.temporal.cov$lag[ind]

# find point of zero-crossing for spatial autocovariance of surface DO
ind<-which(surfDO.spatial.cov$acf<0)
ind<-min(ind)
surfDO.spatial.lag<-spatial_resolution*surfDO.spatial.cov$lag[ind]

# find point of zero-crossing for spatial autocovariance of bottom DO
ind<-which(botDO.spatial.cov$acf<0)
ind<-min(ind)
botDO.spatial.lag<-spatial_resolution*botDO.spatial.cov$lag[ind]
```