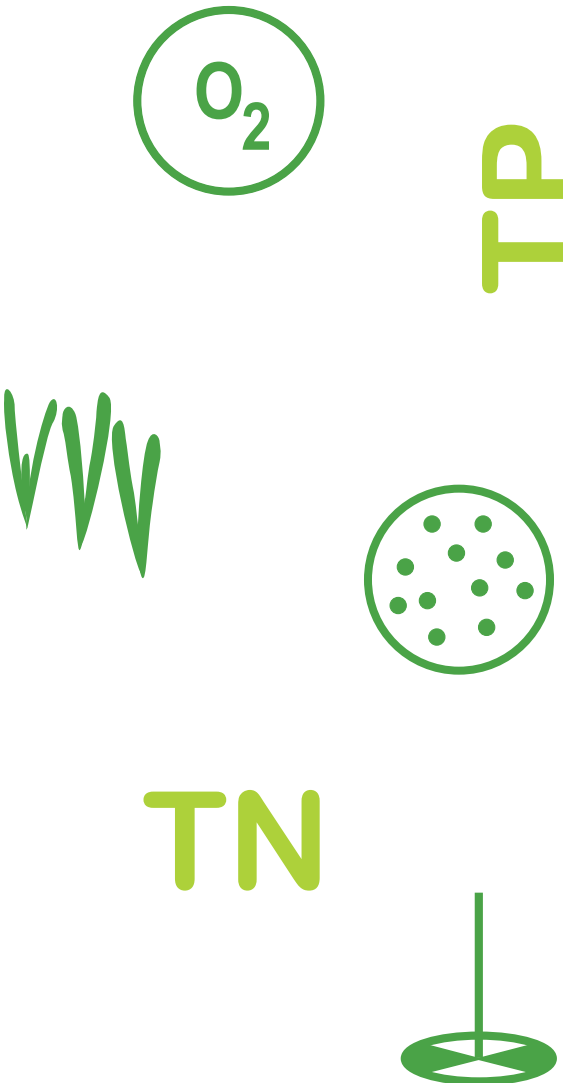
Sampling and data analysis protocols for Mid-Atlantic tidal tributary indicators



MTAC
Mid-Atlantic
Tributary
Assessment
Coalition

Preferred citation: EcoCheck. (2011). *Sampling and data analysis protocols for Mid-Atlantic tidal tributary indicators.* Wicks EC, Andreychek ML, Kelsey RH, Powell SL (eds). IAN Press, Cambridge, Maryland, USA.

Contributors

The protocols discussed in this document represent the collective efforts of many individuals and organizations working within the Mid-Atlantic Tributary Assessment Coalition (MTAC). The following individuals contributed significantly to the development of these protocols: Peter Bergstrom, Jamie Brunkow, Carol Cain, Jana Davis, Katie Foreman, Jane Hawkey, Tom Leigh, Carol McCollough, Ron Melcer, Andrew Muller, Diana Muller, Rupert Rossetti, Chris Trumbauer, Megan Ward, Cathy Wazniak.

PO Box 775
Cambridge, MD 21613
U.S.A.
www.ian.umces.edu
ianpress@umces.edu

First published in 2011
Set in Minion Pro

IAN PRESS

University of Maryland
CENTER FOR ENVIRONMENTAL SCIENCE

IAN Press is committed to producing practical, user-centered communications that foster a better understanding of science and enable readers to pursue new opportunities in research, education, and environmental problem-solving. IAN Press is the publication division of the Integration and Application Network at the University of Maryland Center for Environmental Science (UMCES). Visit *www.ian.umces.edu* for information on our publications and access to downloadable PDFs of our reports, newsletters, posters, and presentations. Contact IAN Press at *ianpress@umces.edu*.

The Integration and Application Network (IAN) is a collection of scientists interested in solving, not just studying environmental problems. IAN seeks to inspire, manage, and produce timely syntheses and assessments on key environmental issues, with a special emphasis on Chesapeake Bay and its watershed. IAN is an initiative of the University of Maryland Center for Environmental Science, but links with other academic institutions, resource management agencies, and non-governmental organizations.

Beginning as a small college laboratory and a state research and education agency, UMCES has developed into a multi-campus institution of Maryland's university system. UMCES continues its rich tradition of discovery, integration, application, and teaching at its three laboratories: Chesapeake Biological Laboratory (1925), Appalachian Laboratory (1962), and Horn Point Laboratory (1973), as well as Maryland Sea Grant in College Park and the Annapolis Synthesis Center in downtown Annapolis. The Integration and Application Network was established in 2002 to allow UMCES to apply the scientific knowledge of its faculty and staff to the environmental challenges we face today.

MTAC
Mid-Atlantic
Tributary
Assessment
Coalition

CHESAPEAKE EcoCheck
Assessing and forecasting ecosystem status
NOAA
University of Maryland
CENTER FOR ENVIRONMENTAL SCIENCE

CHESAPEAKE BAY TRUST

Table of contents

Chapter 1: Coordinating tributary monitoring efforts

Environmental health report cards (Figure 1.1) are detailed ecosystem health assessments that have proven to be important outreach tools for generating community interest and increasing citizen understanding of ecosystem health, water quality, and watershed issues. Report cards provide useful and timely information on environmental issues to local decision-makers and can highlight actions that residents can take to become involved in the improvement and conservation of their communities.

Although report cards are proven tools, their effectiveness can be enhanced by increasing the consistency of water quality monitoring, data analysis, and communication efforts among report card-producing organizations. This protocol document, developed by EcoCheck through consensus of Mid-Atlantic Tributary Assessment Coalition (MTAC) members, will substantially improve the utility of report cards across water systems (e.g., tributaries and estuaries).

The overall objective of this protocol document is to encourage and enable comparisons of monitoring results from report card-producing organizations and to increase the scientific validity of report cards as outreach tools. This document is intended for use in tidal areas only, as the ecosystem health indicators (see blue text box below) and thresholds discussed are pertinent only to tidal ecosystems. A companion document, which will present non-tidal protocols, will be started in summer 2011: visit the EcoCheck website (www.eco-check.org) for updates and more information.

Report card indicators

Report cards provide scores for individual ecosystem health indicators, such as dissolved oxygen and water clarity, that are averaged into an overall report card grade. An indicator is a measure, an index of measures, or a model that characterizes an ecosystem or one of its critical components (Longstaff et al. 2010). Indicators relay complex messages, potentially from numerous sources, in a simplified and useful manner.

The primary uses of indicators are to characterize current condition status, to track or predict significant condition changes, and to identify condition trends.

Local-scale monitoring provides a detailed picture of health

Report cards have been very successful as local outreach tools for individual water systems. However, comparisons of monitoring data and results among various water systems, or from a tributary to the estuary into which it drains (such as the Chester River to Chesapeake Bay), are neither valid nor informative because the report cards are likely based on different indicators and methods for monitoring and analysis.

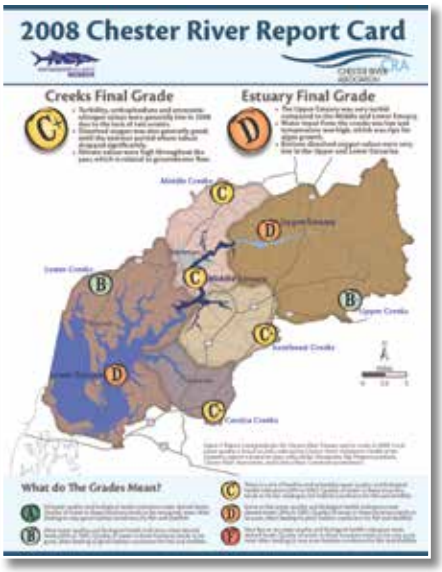


Figure 1.1. An example of a report card. This one is produced by the Chester River Association.

Historically, state and federal government agencies have monitored the health of water systems for management and regulatory purposes. For example, Maryland's Department of Natural Resources and Virginia's Department of Environmental Quality perform most monitoring activities that support the management and regulation of the Chesapeake Bay. Additionally, the Chesapeake Bay Program—a regional partnership mandated with management and regulation of the Chesapeake Bay—works closely with state and federal agencies, such as the U.S. Geological Survey, to evaluate environmental impacts on the Bay.

Unfortunately, it is not economically or logistically feasible to place monitoring stations in all focus areas of an estuary's waters, most often because the area is large and crosses multi-jurisdictional boundaries. In the Chesapeake Bay, the Chesapeake Bay Program has carefully chosen sampling site locations to maximize coverage, so as to adequately assess Bay-wide health conditions across the entire Bay (Figure 1.2a). However, this may mean that there are only one or two monitoring stations within each tributary. Despite more than two decades of intense monitoring and assessment at a Bay-wide scale, more information is needed at finer scales (i.e., targeted regions within the Bay-wide reporting waters, such as the Upper Eastern Shore region) to evaluate the effectiveness of management actions taken at localized levels (Figure 1.2b). This is particularly important in light of the new

Chesapeake Bay Total Maximum Daily Load (TMDL) and state Watershed Implementation Plans (WIPs).

Data collection at the scale needed for these types of assessments is currently being carried out by many watershed associations and citizen monitoring programs. These data are very useful for providing detailed assessments of local environments (Figure 1.2c).

However, these watershed associations and citizen monitoring programs may choose to monitor different indicators based on unique local issues. Varied indicators and methods for monitoring and analysis make it difficult to compare data and results across water systems. This diminishes the overall power of report cards.

This protocol document addresses the need for a common framework of monitoring, analysis, and communication efforts among watershed associations and citizen monitoring programs. It will add substantial value to the data collected and reported by individual groups by allowing direct comparisons of results from one tributary to another. In doing so, these protocols will also greatly enhance the value of information synthesized from existing and planned report card projects. Monitoring data may then also be integrated into additional, Chesapeake Bay-wide assessments, such as the Chesapeake Bay report card.

Standardization of sampling and data analysis methods

MTAC was formed to better organize and coordinate mid-Atlantic citizen monitoring programs that are interested in producing tributary or regional report cards. MTAC conducts monthly group meetings to share information and work toward reaching consensus on ecosystem health indicators for tidal and non-tidal areas, sampling methodology for measuring these parameters, and data analysis procedures for calculating report card scores. Current participating groups represent the Chester, Magothy, South, West/Rhode, Patuxent, Nanticoke, Sassafras, Gunpowder, and Anacostia Rivers, Maryland's Coastal Bays, and Baltimore's Inner Harbor. Other agencies and organizations involved in this effort include the Chesapeake Bay Program, National Oceanic and Atmospheric Administration (NOAA), University of Maryland Center for Environmental Science (UMCES), and Chesapeake Bay Trust.

This protocol document was developed by MTAC (Figure 1.3), with technical guidance provided by scientists from UMCES, which has extensive experience identifying indicators, analyzing data, and developing integrated assessments of ecosystem health. UMCES regularly partners with watershed organizations and citizen monitoring programs to assist in the production of tributary report cards.

This document provides guidelines for the successful production of tidal ecosystem health report cards. Specifically, this document develops clear and consistent protocols for the identification, collection, and analysis of indicators to be used by report card-producing organizations in the mid-Atlantic region. These water systems have similar physical, chemical, and biological characteristics that allow for the application of standardized protocols. Methods for different types of

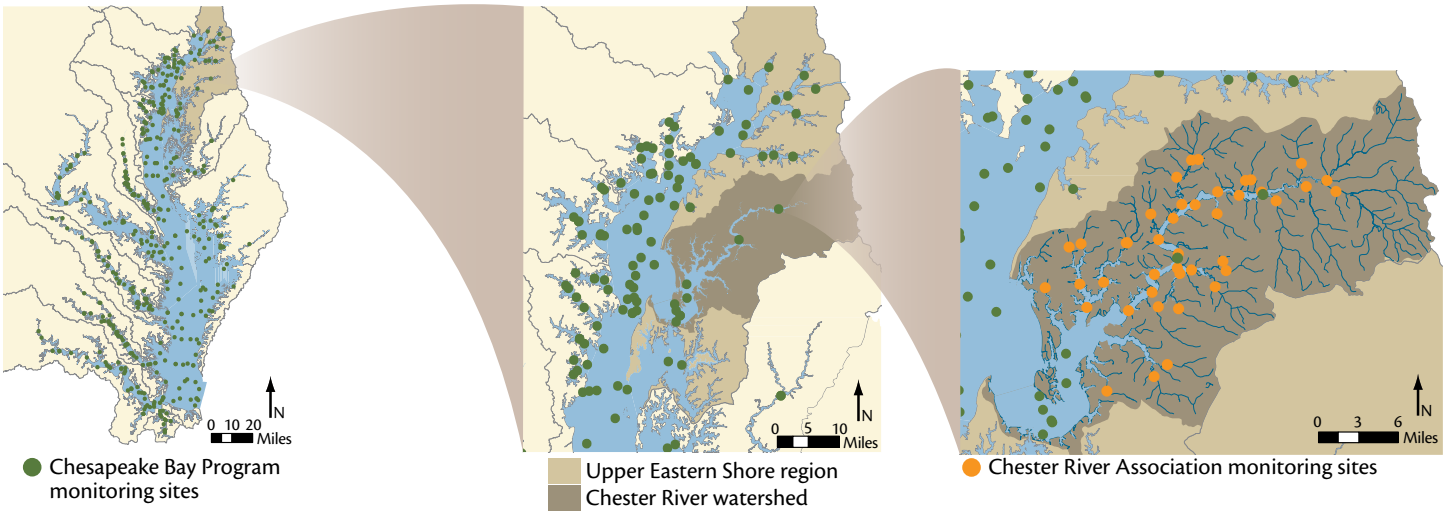


Figure 1.2a: The Chesapeake Bay Program sampling site locations are located throughout the tidal Bay area. Information at this scale is used for Bay-wide health assessments (e.g., annual Chesapeake Bay report card).

Figure 1.2b: The Upper Eastern Shore region, used in Bay-wide assessments, groups several smaller watersheds together because few Chesapeake Bay Program sites are located in this area of the Bay.

Figure 1.2c: The Chester River Association focuses on monitoring just the Chester River and its tributaries, and therefore has a higher data density in that watershed than is provided by Bay-wide monitoring efforts.

Figure adapted from EcoCheck. (2010). A guide to the mid-Atlantic tributary report cards [newsletter].

Ken Barton

Figure 1.3. Members of the Mid-Atlantic Tributary Assessment Coalition (MTAC) meet once a month to discuss sampling methodology, data analysis, and report cards.

systems (e.g., coastal lagoons) are not directly addressed in this document. Coastal lagoons have different physical, chemical, and biological characteristics that necessitate their own set of indicators, some of which overlap with this document. Additionally, a different set of thresholds against which the indicator measurements are compared is also necessary. See the References and further reading section for more information. Protocols for non-tidal indicators will be presented in a companion document.

The methods in this document are recommended steps for watershed organizations and by no means preclude groups that are already monitoring and analyzing water quality data. While the hope is that participating groups can adjust their methods so that standardization across water systems can occur, it is not mandatory. These protocols are designed to be sustainable by establishing consensus among the many groups currently producing report cards, and by transferring its protocols to new organizations that are interested in producing report cards. Sampling and data analysis for all indicators were discussed and agreed upon at the MTAC monthly meetings.

General conclusions and recommendations for indicator sampling and analysis include:

- Six core indicators: dissolved oxygen, chlorophyll *a*, water clarity, total nitrogen, total phosphorus, and aquatic grasses. These indicators are relatively easy to measure, have reasonable lab costs, and are pertinent to most tidal water systems. See following chapters for details on each indicator.
- Elective indicators, which may be chosen by each reporting organization, include those that may be difficult to measure, costly, or of particular importance to regional groups.
- A sampling regime that allows sub-regions based on strata (e.g., mainstem vs. creeks) and salinity regime.
- Minimum twice monthly sampling for a minimum total of 14 samples per year during the relevant sampling period.
- Use of relevant seasons for each indicator's data analysis calculations.

Organization of document

This document begins with a brief introduction to successful monitoring programs, followed by a discussion of QA/QC procedures. It then details each of the six core indicators, including field sampling methods and techniques, laboratory analyses, and data analyses. Finally, synthesis of data into a report card and communication of results is discussed.

Chapter 2: Organizing a successful monitoring program

This chapter addresses monitoring programs that seek to assess the ecosystem health of a local, tidal water system. Monitoring the health of these ecosystems is important because it informs management, local decision-makers, and residents and also provides direction to research. Integrating monitoring results via a report card or other communication product builds community knowledge, which is usually a cornerstone objective of watershed associations and citizen monitoring groups.

Additional information on the importance of standardizing sampling, analysis, and communication efforts is discussed in Chapter 1 of this protocol document.

Establishing goals & objectives

The most critical step in the planning process for a monitoring group is to establish the goals and objectives of the program because every decision and action that follows will stem from this initial framework. For the purposes of this document, the overall goal of an effective monitoring program is to accurately assess the ecosystem health of a tidal water system. In developing this kind of program, the following considerations must first be taken into account:

- Capacity (e.g., number of volunteers, availability of volunteers, accessibility of sampling site locations and sampling equipment, financial support) and
- Specific program objectives (e.g., to produce a tributary report card, to contribute to larger assessments or mandated regulatory programs, to establish a baseline condition assessment).

Understanding an organization's capacity is critical to the program's success. The total number of employees in the program may be small or large, but the number of people that help with the actual monitoring, both paid and volunteer, is key (volunteers are discussed specifically later in this chapter). First, it must be determined if there are enough people to perform the monitoring. Next, staff and volunteers need to be properly trained, provided with appropriate equipment, and a determination of where to sample must be made (Figure 2.1). A major part this process includes assessing the financial support needed for monitoring. To keep costs reasonably low, many of the indicators discussed in this document require only basic equipment and can be easily measured by volunteers from docks and piers rather than from boats.

Objectives of a monitoring program may include providing a general picture of the ecosystem, establishing a baseline assessment against which to evaluate the impact of future changes, providing an early warning system (forecast) for threats or future changes, and/or evaluating if management actions (e.g., restoration, nutrient controls) result in a measurable difference (Longstaff et al. 2010).

Properly trained staff and volunteers, appropriate oversight and management, reliable and well-maintained instrumentation, and a valued and usable end product are all features of a successful monitoring program that matches its capacity to its objectives (Longstaff et al. 2010).

Recruiting, training, & retaining volunteers

In order to successfully recruit and retain volunteers, it is important to understand what motivates people to volunteer in the first place. Some people volunteer because they believe in the cause (in this case, ecosystem health, water quality, and watershed issues) and think it's important to be involved. Others volunteer because they enjoy the social interaction with like-minded people, and others still because they enjoy learning new skills and knowledge that might help them grow in their career and/or personal lives.

Few aspects of a monitoring program are more important than the training of volunteers because proper training provides the background needed for a scientifically-sound and well-designed data collection effort.

There are three broad types of volunteer training:

- introductory,
- quality assurance and quality control (QA/QC), and
- motivational sessions.

Nanticoke Watershed Alliance

Figure 2.1. Volunteers monitor water quality in a creek that drains into the Nanticoke River. Volunteers are a critical part of a monitoring program.

Introductory training should describe the monitoring program and teach standard methods for collecting and analyzing samples. Training on how to collect field samples should take place in the field to prepare volunteers for conditions that may be less predictable than those to which they are accustomed.

QA/QC training will help ensure consistency and reliability of data collected by volunteers. Such sessions should focus on proper techniques and ideally be offered two times per year, depending on the length of the sampling season.

Volunteer monitoring program "Do"s and "Don't"s

DO:

- *Have a team leader (and give him or her time).* It takes time to build the proper framework for a successful program. But remember: don't re-invent the wheel! Call upon experts and draw from programs already in place.
- *Understand your volunteers' motivations.* Volunteers want what they're doing to be meaningful. Collecting and managing consistent, reliable data will ensure that.
- *Explain what you are asking your volunteers to do.* Prepare them by explaining that monitoring programs are highly involved and scientifically rigorous. They need not be scientists, but willing to learn and follow the protocols in place.
- *Put it in writing.* Studies have shown that volunteers often want to be treated like paid staff. Provide them with formal descriptions and clear expectations.
- *Provide regular communications and support.* It is the program's responsibility to train, and provide ongoing support to, volunteers. Express appreciation for relevant contributions, address any widespread issues, and generally applaud what the volunteers are doing.
- *Get—and use—feedback from your volunteers.* Volunteers are a program's on-the-ground eyes and ears. They will know what's working and what's not.
- *Let volunteers grow or diversify.* Use their broad range of skills!
- *Keep in touch with "retired" volunteers.* Keep former volunteers on your mailing lists. Allow them the opportunity to return.

DON'T:

- *Let a recruitment opportunity pass you by.* Concerned citizens that call your program with questions are potential volunteers. Let them know what you're doing and ask if they'd like to help.
- *Take your volunteers for granted.* Provide enrichment activities or other gatherings on a regular basis to show your appreciation.

Motivational sessions may be held as needed to encourage the exchange of information between volunteers, identify any problems, and, of course, to motivate! Supplemental continuing education and re-training sessions are often also helpful.

If sampling is conducted on a seasonal basis, training sessions for new volunteers and re-training for returning volunteers can be held before the sampling period begins, with a QA/QC session scheduled for the middle of the season, and motivational sessions as needed.

Retaining volunteers is also important to the success of a monitoring program. Finding and training volunteers takes time and effort, so losing volunteers can be a drain on the program's resources. In order for volunteers to feel compelled to continue with a monitoring program, volunteers must feel that their efforts are recognized, respected, and appreciated, and that their work is producing tangible, useful results. Producing a report card is a great way to use volunteer-collected data to achieve all of these things. Additionally, a retention plan can include incentives for longer-term volunteers, volunteer appreciation days, and other related activities.

Types of sampling

Once the goals and objectives of a monitoring program are established and volunteers have signed up to help monitor, a program needs to determine what indicators to measure and where to monitor. Water quality monitoring can be used to assess a wide variety of indicators, depending on the goals of the program and the type of water system. Regardless of the type of water system, a monitoring program should sample a set of core indicators that can be used to assess the health of that system.

Types of sampling include:

- basic water quality monitoring,
- nutrient monitoring,
- biological monitoring,
- sediment monitoring, and
- bacteria monitoring.

Basic water quality monitoring refers to indicators—such as temperature, salinity, dissolved oxygen, water clarity, and pH—that can typically be measured instantaneously with a multi-parameter instrument. Nutrient monitoring—chiefly nitrogen and phosphorus—usually involves collecting water samples to be processed later by a laboratory. Biological monitoring involves sampling living resources, such as fish, shellfish, or benthic macroinvertebrates. Sediment monitoring requires taking samples from the bottom to analyze the content and make-up of the sediment. Bacteria monitoring usually requires water samples to be collected and analyzed at a laboratory.

Sampling considerations

Determining what to monitor and where to collect data is often decided upon in meetings and workshops, and by summarizing existing literature or monitoring programs. Drawing on experts not only in the field of environmental science but also in the community will help monitoring programs to be well-rounded. Many sampling considerations focus on what pressures are occurring in the ecosystem and on what management actions are being taken to correct them. Measuring the health of an ecosystem is important for tracking changes resulting from these pressures and actions.

Spatial sampling scheme

Another sampling consideration is the sub-regions of the water system being monitored (Figure 2.2). When locating sampling sites, try to identify sites that represent any and all distinct zones within the water system. For instance, be sure to locate sites in both shallow and deep water. If there is a gradient of salinity levels, be sure to locate sites within all different salinity zones. The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt).

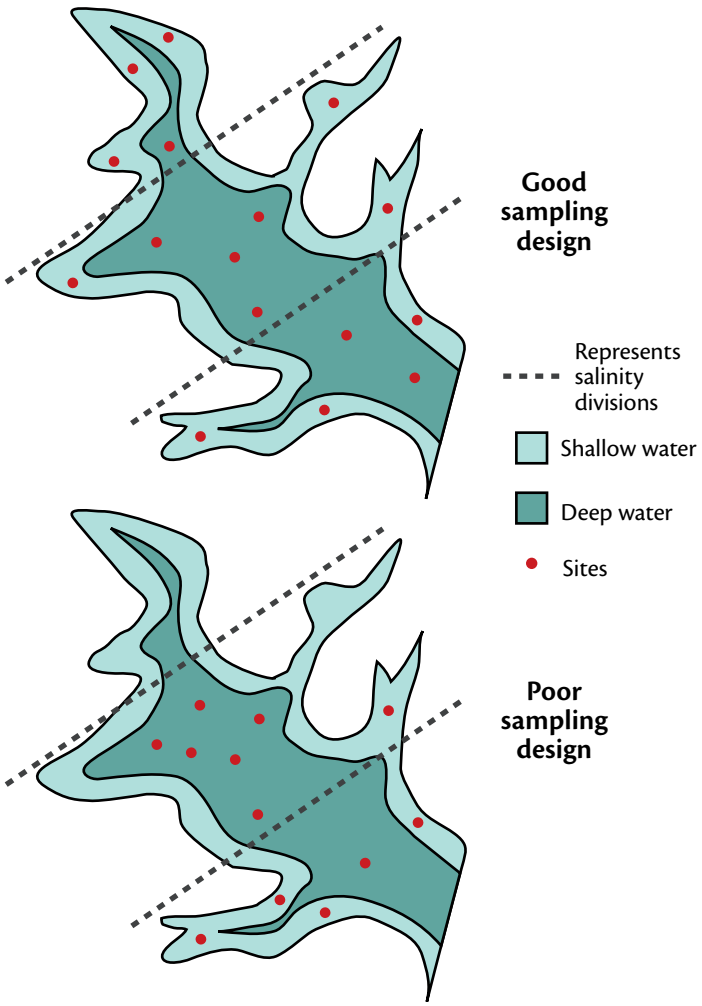


Figure 2.2. Site locations can be based on salinity regime, depth, and convenient access for volunteers, but should be evenly spaced throughout sub-regions.

When sampling a water system with tributaries (i.e., a river with many creeks feeding into it), consider locating at least one site in each of the tributaries in order to allow those tributaries to be evaluated. It may be tempting to focus only on problem areas; however, a targeted sampling regime will not accurately depict the true health of the ecosystem. The health of a sub-region may be affected by its physical characteristics; for example, the shape of a basin may affect the length of retention times, which in turn affects dissolved oxygen levels.

Sub-regions are important for calculation of indicator scores and for the report card product. Determining the sub-regions at the outset of the program is best, but not always possible. Some programs have changed their sampling scheme after establishing their monitoring program, usually due to a higher capacity (e.g., more volunteers, more funding) to sample more areas. This is fine to do, as long as the new sampling scheme is communicated to the public and is well documented.

In general, a greater number of sites will result in a more accurate assessment. Ideally, it is best to have three or more sites per sub-region, but due to funding, staffing, and geographic constraints, this is not always possible. The goal is to maximize available resources, so do not go overboard with sampling. If volunteers commit to two hours a week, do not design a sampling run that will take four hours.

Many volunteer organizations do not have access to a boat, so all data must be collected from piers, docks, bridges, or the shoreline. This often leaves out any deep locations within the mainstem channel, which describes part of the story of overall system health. Low dissolved oxygen very frequently occurs in deeper waters where there is less mixing with more oxygenated surface waters, for example. This is generally not seen at shallow water sampling sites.

If this type of limited sampling (e.g., sampling in shallow areas only) occurs, it is important to communicate this strategy as well as its drawbacks in the resulting report card and/or supporting, technical documentation.

Temporal sampling regime

To determine an appropriate sampling frequency, consider the variability of the indicators being assessed. Some indicators, such as dissolved oxygen and temperature, have daily cycles, while others, such as salinity, may change with the tide. Other indicators (e.g., water clarity, bacteria) tend to be episodic—they follow a pattern in that they are generally affected by precipitation events (e.g., precipitation often flushes increased sediment and bacteria loads into water systems).

Modern technology allows near-continuous monitoring of many indicators (e.g., multi-parameter data sondes can be deployed unattended and, with proper maintenance, can provide very accurate data). However, data are most valuable when they are properly captured, analyzed and interpreted, so be sure to reference your program's capacity for data collection and analysis.

For this protocol document, the recommended sampling frequency is weekly from March to November for chlorophyll *a*, water clarity, and nutrients, and April to October for dissolved oxygen. However, recognizing funding and staffing constraints, the minimum sampling recommendation is twice monthly during the same time periods. Additionally, each indicator has a minimum sampling season. See Chapter 4 for more information and chapters 5–9 for specific indicator recommendations.

It is best to evenly distribute sampling within the specified time period. For example, it is not desirable to have four samples in one month and zero in the following month. Furthermore, monitoring should occur on the same day of the week at the same time each week. This increases the likelihood that data are consistent and reliable, and are not biased due to weather events or other influences on the measurement. Varied spatial and temporal scales among programs are resolved when comparisons are made at the region or sub-region level for individual indicators and health index scores.

Once the monitoring program, including goals and objectives, volunteer recruiting, and the sampling regime, are planned, monitoring can begin! Sampling, however, needs to follow strict quality management plans, which are discussed in the next chapter.

Chapter 3: Ensuring quality management

> "A Quality Management Plan is a management tool that documents an organization's quality system for planning, implementing, documenting, and assessing the effectiveness of activities supporting environmental data operations and other environmental programs." —U.S. EPA

Collecting data according to a scientifically-credible method is necessary for the success of any water quality monitoring program. Good sampling practices ensure data validity and that data collected can be used to meet the program's goals and objectives.

If followed consistently, these methods can help guarantee uniform data quality and enhance the value of data collected and information synthesized by monitoring programs.

A quality assurance project plan is a key element of monitoring programs

A Quality Assurance Project Plan (QAPP) is an essential component of monitoring programs' sampling and reporting efforts (Figure 3.1). Every monitoring program should prepare a QAPP and revise it periodically to ensure that procedural changes are documented and that quality assurance is considered when these changes are made.

QAPPs should provide details for the following elements:

- project management,
- data generation and acquisition (i.e., sampling and sample analysis),
- assessment and oversight, and
- data validation.

All the guidelines presented in this chapter are produced by experts and practitioners that use quality assurance and quality control (QA/QC) procedures on a regular basis. QA/QC procedures, which are common for monitoring programs, are intended to ensure that data are not lost or corrupted during transcription or analysis.


Quality Assurance Project Plan
For
Nanticoke Watershed Alliance's
Water Quality Monitoring Program
Chemical and Physical Properties

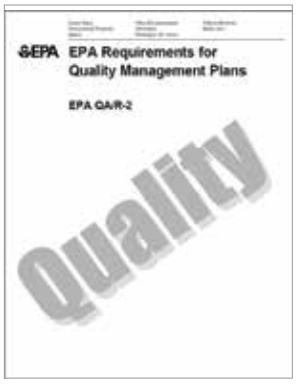


Figure 3.1. Left: Example of a QAPP from the Nanticoke Watershed Alliance. Right: EPA protocols help monitoring programs ensure that their data can be incorporated into the EPA's regulatory process.

Many of the recommendations presented in this chapter are intended to help create QAPPs for acceptance for EPA-funded projects—a QAPP may also be necessary if data are to be included into the Chesapeake Bay Program's databases or used in Chesapeake Bay-wide ecosystem health assessments.

Please visit the Chesapeake Bay Program website (www.chesapeakebay.net) for more information on requirements for data acceptance into the program database. This database is used for Bay-wide criteria assessment—if data from other organizations are to be included in criteria assessments, they must first be accepted into the database. Additional links to helpful documents are included at the end of this chapter.

High quality data are necessary to achieve objectives

A main objective of ecosystem health assessments is to inform citizens, local decision-makers, and other resource managers of scientific discovery in light of management objectives. If ecosystem health assessments are to influence public policy or citizen behavior, they must first be grounded in reliable, high quality data.

Many monitoring programs that provide data to constituents and resource managers may also make it available to wider scientific audiences. Only data with clear and rigorous quality management procedures will be acceptable if it is to be useful in these larger contexts. Samples should be collected using consistent, accepted methodologies and analyzed using scientifically accepted methods. It is essential to write a QAPP before implementing a monitoring program for these reasons. Strategies for standardizing sampling schemes and methodologies are discussed in subsequent chapters of this document.

Pre- and post-sampling calibration of monitoring instruments is also necessary for consistent, reliable data. Data cannot be considered valid or acceptable without a rigorous instrument calibration protocol, and these calibration results should be recorded. Calibration procedures can normally be found within the instrument documentation.

Good data management is important for quality assurance

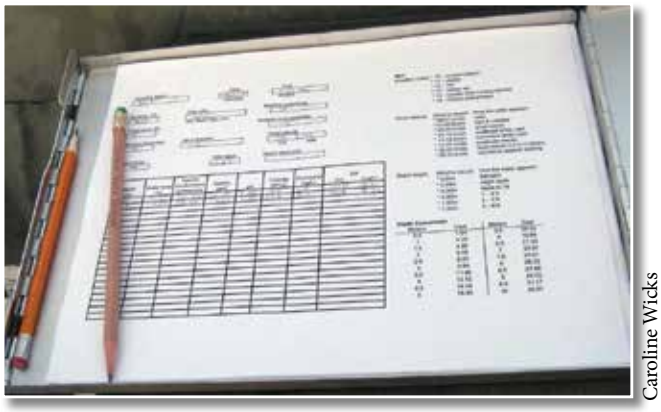
After data are collected, they must be recorded and analyzed appropriately. This may involve:

- transferring data from data sheets or data loggers to spreadsheets or databases,

- grouping data for analysis to extract information, and
- integrating data to calculate scores and synthesize information.

Each of these steps provides an opportunity for mechanical, human, or computational errors and requires attention to quality assurance measures (Figures 3.2 and 3.3). To maintain reliable data, monitoring programs should:

- manually review data sheets and transferred data,
- keep unaltered, original data sheets in a secure location,
- flag unusual, blank, or out-of-range data values, and
- document analytical and integration objectives and methods.



Caroline Wicks

Figure 3.2. Data sheets are filled out in the field and need to be entered into a computer spreadsheet or database program (e.g., Microsoft Excel or Access). It is critical to avoid numerical errors when transferring data.

Avoiding errors is critical when transferring data

Transferring data from data sheets to spreadsheets or databases can be a tedious process—and one of the most common sources of errors. Poor handwriting, smudges and smears from field conditions, and tired samplers can all contribute to errors in data transcription. Care should be taken to ensure that the data are transcribed accurately. Make notes of handwriting questions or obscured values and ensure that they are addressed as soon as possible, while information is fresh in the field crews' minds. It is important to note that original data sheets are considered records and must not be altered in any way (Table 3.1). Original data should be stored on a CD or external hard drive and write-protected for security.

Transferring data from data loggers to spreadsheets or databases (Table 3.2) also can be problematic, and errors are often invisible to the field or office personnel performing the transfer. Always confirm that the data are recorded in the spreadsheet or database correctly. One quick way to do this is to double-check that the correct number of records is present. Twelve sampling stations should be accompanied by 12 data records—if these numbers do not match, there should be an explanation why. It is also a good idea to have another person with "fresh eyes" check the data for inconsistencies and/or incompleteness.

Nanticoke Watershed Alliance

Figure 3.3. Data are recorded on a field data sheet before being taken back to the office and put into a spreadsheet.

Most monitoring programs store and work with their data in a basic spreadsheet application (e.g., Microsoft Excel). Although spreadsheets are relatively easy to use, errors can nevertheless be quickly created and compounded. If unnoticed, even small errors lead to lost time and, in the worst cases, incorrect or misleading interpretations of data. To prevent misinterpretations or permanent data loss from spreadsheet errors, always save the original spreadsheet in multiple locations before working with the data, and "lock" the original spreadsheet so that it is protected and cannot be changed. Copy the

Table 3.1. It is important to note the difference between documents and records. Original data sheets are considered records.

Document	Record
A document is a living thing.	A record, on the other hand, is history.
The information contained within a document is subject to change; it can be revised.	The information contained within a record cannot be changed, because it simply states what's already happened.

Table 3.2. Spreadsheets and databases are computer software programs that help manage and process data.

Spreadsheet	Database
A spreadsheet is a computer software program that simulates a piece of paper with rows and columns, with each cell containing either alphanumeric text or numeric values.	A database is a computer software program that stores, retrieves, and manipulates a collection of organized information in a regular structure.
e.g., Microsoft Excel	e.g., Microsoft Access

spreadsheet to a new location for analysis and calculation of report card scores, and periodically refer back to the original, secured data to ensure that errors are detected and corrected if necessary.

Alternatively, databases may also be used to store original data. Databases are generally more stable than spreadsheets (i.e., they are less likely to be affected by small errors in aligning data, or wholesale changes to columns or rows), and data can be extracted from databases to work within spreadsheet applications.

Once data are organized and ready for analysis, a good practice is to "flag" data values that are suspicious (e.g., extremely high or low values, or values completely out of the range of possibility). Expressions, or mathematical functions that use equations to determine certain outcomes, are useful tools to help identify unusually high or low values. In a spreadsheet, it is relatively easy to write an expression that searches a column for values exceeding a user-specified range. These values can be checked against data sheets or data loggers and investigated for accuracy.

Decisions to include or remove data are made on an individual basis, but in general, data should be excluded only if values are clearly outside the range of possible values, or if there are clear reasons to suspect that the data are incorrect (e.g., inconsistent or abnormal calibration information from data sheets). Because of data quality issues, only the most senior data analysts or program staff should decide if individual observations should be included in analyses. Decisions to include or exclude data should be clearly documented.

Chapter 4: Measuring core indicators

The previous chapters have provided a general overview of monitoring programs, spatial and temporal sampling considerations, and quality assurance/quality control procedures. These are critical steps that support the production of a report card. This and the following chapters discuss in detail how to sample and analyze the core indicators that should be incorporated into a report card.

The six core indicators in this protocol document (see Figure 4.1 and Table 4.1) were chosen by the Mid-Atlantic Tributary Assessment Coalition (MTAC) to be used by report card-producing organizations in the mid-Atlantic region for tidal water system assessments. The indicators and the methods for evaluation are specifically targeted at tidal rivers and estuaries, not coastal lagoons. Coastal lagoons require a different set of indicators and thresholds (See References and further reading).

The indicators for this protocol were chosen due to their ease of collection and communication, low costs, and, most importantly, the amount of information they convey about the ecosystem. They answer the question: "How is the system doing—is it healthy or unhealthy?" The core indicators are:

- dissolved oxygen,
- chlorophyll *a*,
- water clarity,
- total nitrogen,
- total phosphorus, and
- aquatic grasses.

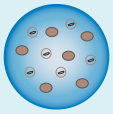
The core indicators should be measured and analyzed by all monitoring programs that wish to compare the health of their water system with adjacent systems, and who wish to incorporate their data into the state and federal regulatory system.

Sampling and data analysis

An overview and methods for sampling and data analysis are provided for each indicator in the following chapters. A summary table of preferred and minimum recommendations is provided here (Table 4.1). The recommended sampling period is sometimes longer than the sampling period needed to perform the data analysis for scoring. This is because it is important to measure

Elective indicators

The six core indicators discussed in this document provide a consistent base for data comparisons among water systems. However, elective indicators, such as phytoplankton community, benthic community, impervious surface, bacteria, and hard clams, may also be measured if organizations have a particular interest in them. For example, bacteria is a commonly measured human health indicator, though sampling procedures and data analysis evaluations have not been scientifically validated for incorporation into overall report card scores. Nevertheless, it is an important indicator to measure, especially in areas with high human impacts.



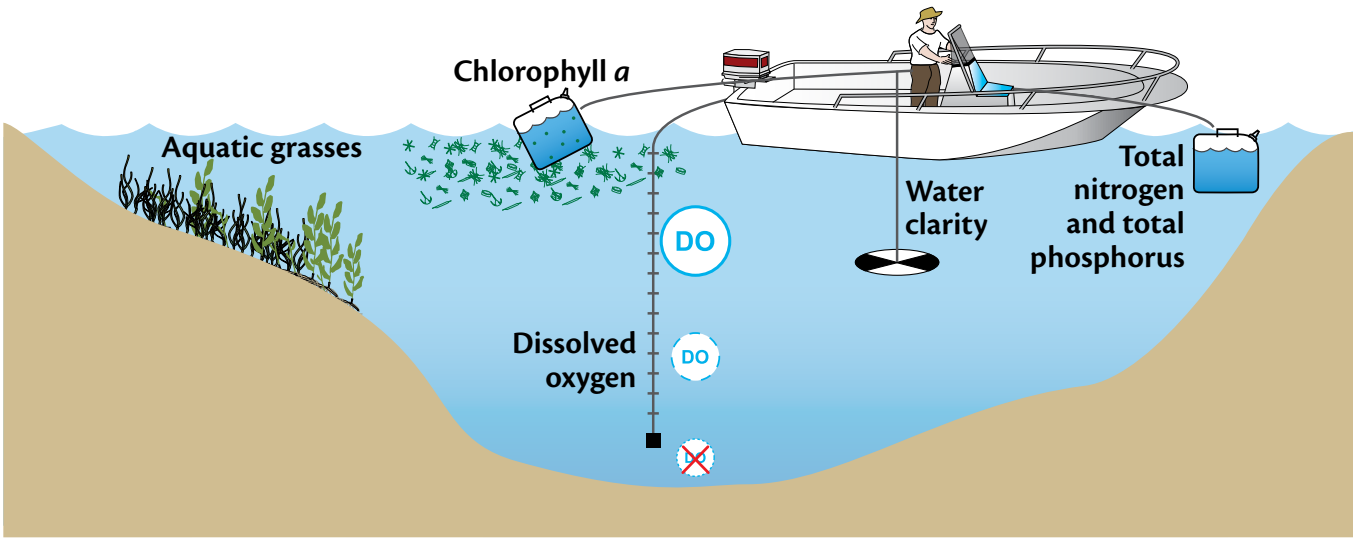


Figure 4.1. This conceptual diagram illustrates the six core indicators discussed in this document. Most indicators are measured by monitoring programs, although aquatic grasses in the Chesapeake Bay are also measured and provided to groups, by the Virginia Institute of Marine Sciences. Water samples are collected at sites so that chlorophyll *a*, total nitrogen, and total phosphorus can be analyzed in the lab.

Table 4.1. Summary of preferred and minimum sampling recommendations for five of the six core indicators. Aquatic grasses are not included as they are not measured in the field by watershed organizations.

Indicator	Preferred sampling period	Preferred sampling resolution	Minimum sampling period (needed for data analysis)	Minimum sampling resolution	Salinity regime (needed for data analysis)
Dissolved oxygen	April–October	Weekly	June–September	Twice monthly	No
Chlorophyll *a*	March–October	Weekly	March–May; July–September	Twice monthly	Yes
Water clarity	March–November	Weekly	April–October; March–November for polyhaline	Twice monthly	Yes
Total nitrogen	March–October	Weekly	April–October	Twice monthly	Yes
Total phosphorus	March–October	Weekly	April–October	Twice monthly	Yes

these indicators for the entire season of interest, be it year-round or only for certain months, so that inter-annual variability can be determined and long-term trends can be analyzed. However, due to funding and time constraints on watershed organizations, it is understood that a group may only have enough capacity to sample during just the most ecologically relevant months. Therefore, this protocol also provides a minimum sampling effort that is required to adequately assess and score the indicators.

A minimum of 14 samples during the relevant season is recommended. This is approximately twice monthly sampling from April to October. However, if a group follows this standard, there will only be eight dissolved oxygen samples measured during the June to September period, which is the relevant sampling period for dissolved oxygen. Each watershed organization must decide if eight sampling points (multiplied by total number of stations within a sub-region) is enough to characterize dissolved oxygen in their system. That is why weekly sampling is preferred. Also, as mentioned in Chapter 2, samples should not be clumped within part of a sampling period (e.g., four samples measured in June and zero in July) because this does not adequately represent the conditions throughout the season.

Thresholds

Assessment thresholds are determined using information from previous studies

The reporting framework used in this protocol is similar to other assessments done by the University of Maryland Center for Environmental Science, and requires that data values be assessed in relation to specific ecological thresholds of significance (Table 4.2). The thresholds are significant because they represent the point where prolonged exposure to unhealthy conditions leads to a negative response (Longstaff et al. 2010). Thresholds were derived from peer-reviewed scientific articles and years-long development of health indicators of Chesapeake Bay via the Chesapeake Bay Program (US EPA 2003, Williams et al. 2009). Additionally, the multiple thresholds described in this document for chlorophyll *a*, water clarity, total nitrogen, and total phosphorus were developed during monthly MTAC meetings.

These recommendations provide one way of measuring the indicators and analyzing data so that each system's results are comparable. Exceptions and other unforeseen reasons that an indicator could be measured or analyzed in a way different than recommended are explained in breakout boxes throughout the rest of the document, or in an addendum, as necessary.

Scoring of data

In addition to data threshold values, appropriate temporal periods over which to assess the data must also be established. It is not informative to assess data from periods when values consistently fall below threshold values, for example. Including such data may skew results toward unrealistically high scores; it is more informative to evaluate data during periods when the exceedances would have significant ecological consequences. To determine the appropriate temporal periods for data assessment, evaluation of time series data in relation to specific thresholds can be useful (Figure 4.2).

Table 4.2. The core indicators used in this protocol and examples of threshold values used to compare observed data to the reference community.

Health indicator →	Example threshold value	→ Comparison of data to threshold
Chlorophyll *a*	≥20.9 $\mu g \cdot l^{-1}$	
Dissolved oxygen	≥5.0 $mg \cdot l^{-1}$	+ = Proportion of data that meets threshold values for each indicator
Water clarity	≥1.8 meters	
TN Total nitrogen	≤0.48 $mg \cdot l^{-1}$	
TP Total phosphorus	≤0.02 $mg \cdot l^{-1}$	
Aquatic grasses	Area (hectares)	= Area compared to goal

Once thresholds and relevant assessment time periods have been identified, data are scored using either a pass/fail or multiple threshold method. Ideally, multiple thresholds are used to provide some gradation of results from poor to excellent, rather than just pass or fail, but this may not be appropriate for all indicators.

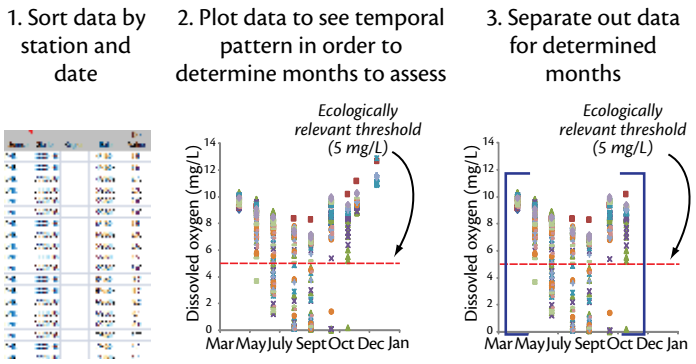


Figure 4.2. Examining data over time in relation to relevant thresholds helps determine the appropriate temporal period for evaluation.

Pass/Fail scoring method

A pass/fail scoring method is a simple method used to calculate indicator scores based on whether or not an ecologically relevant threshold was met. The process is outlined in Figure 4.3, using dissolved oxygen as an example, and results in a score on a scale of 0 to 100%, where the higher percentage values represent more healthy conditions (Williams et al. 2008).

One disadvantage of using a pass/fail method is that there is no way to know how close a failing value is to passing. In other words, if a dissolved oxygen measurement is 4.9 $mg \cdot l^{-1}$, it fails because the threshold is 5.0 $mg \cdot l^{-1}$. However, it is much closer to passing than a value of 1.0 $mg \cdot l^{-1}$. Therefore, using a pass/fail method does not allow for any knowledge of how close or far values are from the threshold criteria.

Multiple thresholds

Multiple thresholds are used to score indicators based on a gradient of healthy to unhealthy conditions (Table 4.3). For example, total phosphorus is an indicator of the amount of phosphorus in the water system. However, the amount of phosphorus, from acceptable levels, to just a little bit too much, to a truly excessive amount, can have different effects on the ecosystem. Therefore, when the measured value of total phosphorus is compared to multiple thresholds, it can score low, medium, or high. This is similar to a grading scale, in which an A is excellent, a B is good, and a C is average. In this way, indicators

Table 4.3. With the exception of aquatic grasses and dissolved oxygen, core indicators discussed in this document are measured against multiple thresholds, then scored from zero to five. The score is then converted into a grade scale.

	Multiple Thresholds		Grade	% Score
Measured indicator value	5	Pristine condition	A	80–100
	4		B	60–<80
	3		C	40–<60
	2		D	20–<40
	1		F	<20
	0	Impaired condition		

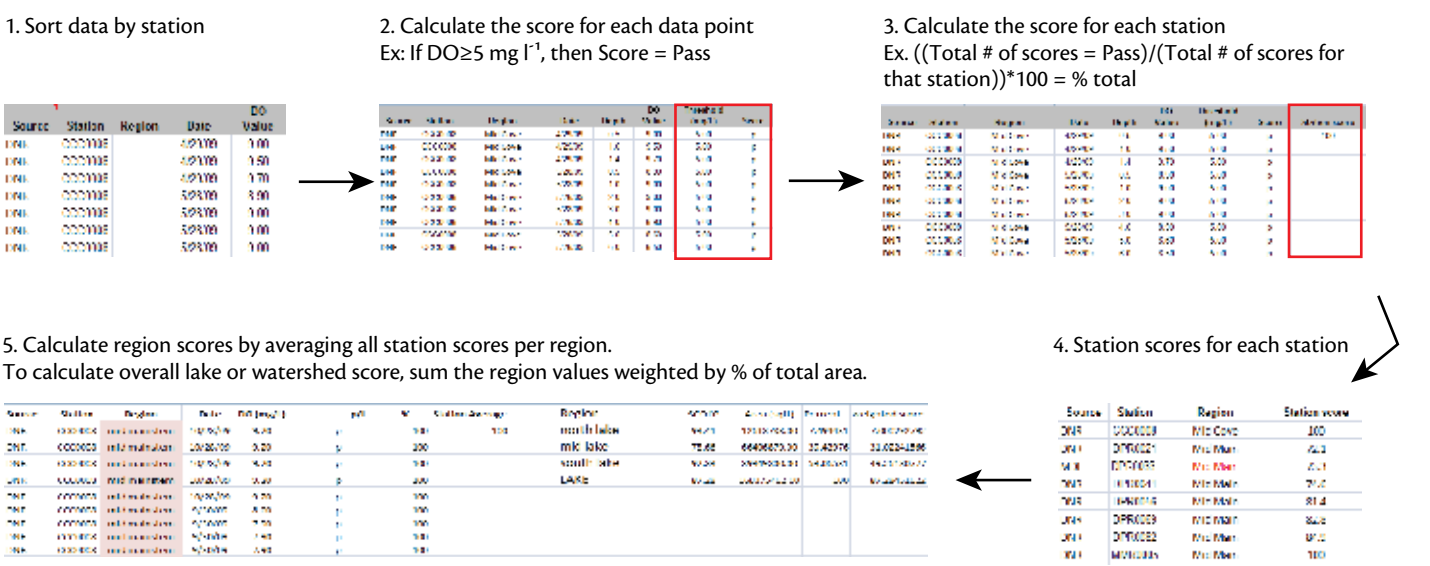


Figure 4.3. A pass/fail scoring method is a simple way to score some indicators.

can be assessed with greater precision than using a pass/fail method.

Applications of multiple thresholds work well if divided into several categories, corresponding to specific percentiles in the frequency distribution of the data (Figure 4.4, top). This creates a scoring scheme based on intervals within the frequency distribution such that the lowest and highest 5% of measurements represent the very worst and best scores.

Scores between the highest and lowest 5% are divided into regular intervals. If a particular value is identified as a standard or ecologically significant criterion, this value can be used to "anchor" the distribution of scores (Figure 4.4, bottom). Previous applications of these types of thresholds have used the preferred or goal value as the next-to-highest score so that this value scores highly, but values that are within the top 5% of the distribution receive the best score.

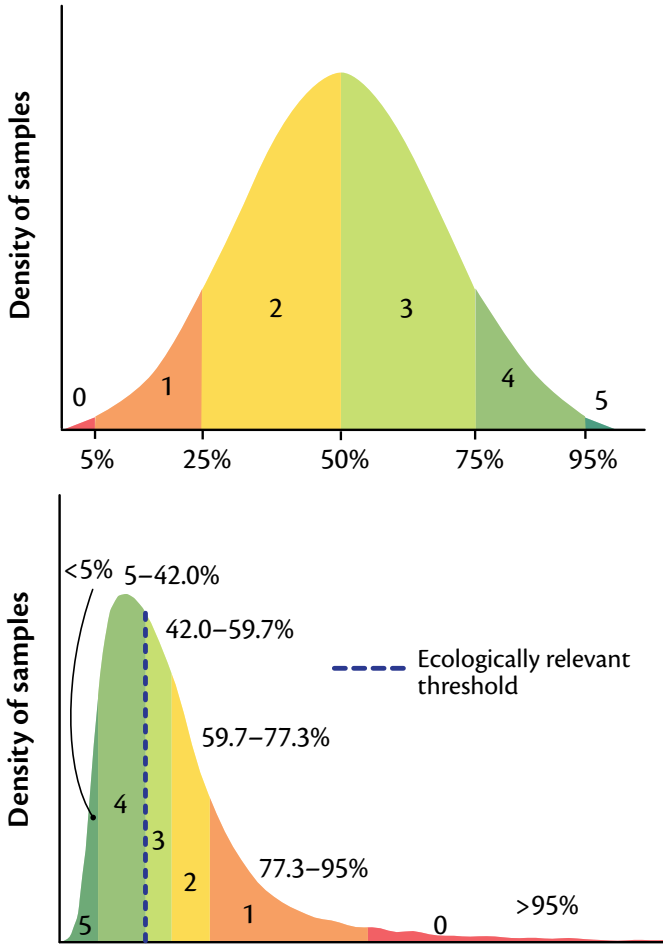


Figure 4.4. *Top,* Example frequency distribution—scores are divided equally among percentiles. *Bottom,* Example frequency distribution—scores are anchored by an ecologically relevant threshold, then divided equally among percentiles.

Scores are standardized to 0–100% scale

In order to integrate individual indicator scores into a more encompassing index (e.g., aquatic habitat or swimming quality), scores are standardized to a 0–100% scale. This allows indicators with different score classes to be easily combined. For instance, one indicator may have three appropriate thresholds that are useful, while others may have five. By converting each to 0–100%, the results can be combined into an overall index.

A score for a reporting region is calculated by averaging all station scores within the region. An overall (i.e., water system-wide) score can be calculated as the area weighted average of regional scores.

Grading scale

Once each indicator is compared against the multiple threshold table, assigned a score, then averaged into the sub-region score (see individual indicator chapters), a grade can be assigned. For this protocol, the grading scale follows the Chesapeake Baywide report card scale of 0–100%, with equal interval breaks (Table 4.4). This was determined through consensus meetings of the Chesapeake Bay Program. The reason the grades are equally divided is to provide a clearer picture of health. Conversely, following the typical grading scale of <60% = F, many of the indicators and sub-regions would fail. This does not tell us as much information as an equally divided scale. A narrative description of the major categories are provided, which relate the grade to ecological health (Figure 4.5).

Summary

This overview of the core indicators, sampling specifications, and thresholds should provide a general understanding of this protocol. The following chapters provide much more detail and step-by-step instructions for collecting, analyzing, and assessing each indicator.

Table 4.4. A grade and description are assigned based on the score that the indicator or sub-region achieves.

Score (%)	Grade	Description
≥0 to <20	F	Very poor
≥20 to <25	D–	Poor
≥25 to <35	D	Poor
≥35 to <40	D+	Poor
≥40 to <45	C–	Moderately Poor
≥45 to <55	C	Moderate
≥55 to <60	C+	Moderate
≥60 to <65	B–	Moderately Good
≥65 to <75	B	Moderately Good
≥75 to <80	B+	Moderately Good
≥80 to <85	A–	Good
≥85 to <95	A	Good
≥95 to <100	A+	Good
=100	A+	Very Good

All water quality and biological health indicators meet desired levels. Water quality in these locations tends to be very good, most often leading to very good habitat conditions for fish and shellfish.

Most water quality and biological health indicators meet desired levels. Water quality in these locations tends to be good, often leading to good habitat conditions for fish and shellfish.

There is a mix of good and poor levels of water quality and biological health indicators. Water quality in these locations tends to be fair, often leading to fair habitat conditions for fish and shellfish.

Some or few water quality and biological health indicators meet desired levels. Water quality in these locations tends to be poor, often leading to poor habitat conditions for fish and shellfish.

Very few or no water quality and biological health indicators meet desired levels. Water quality in these locations tends to be very poor, most often leading to very poor habitat conditions for fish and shellfish.

Figure 4.5. Descriptions of ecological health that correspond with each grade.

Chapter 5: Measuring dissolved oxygen

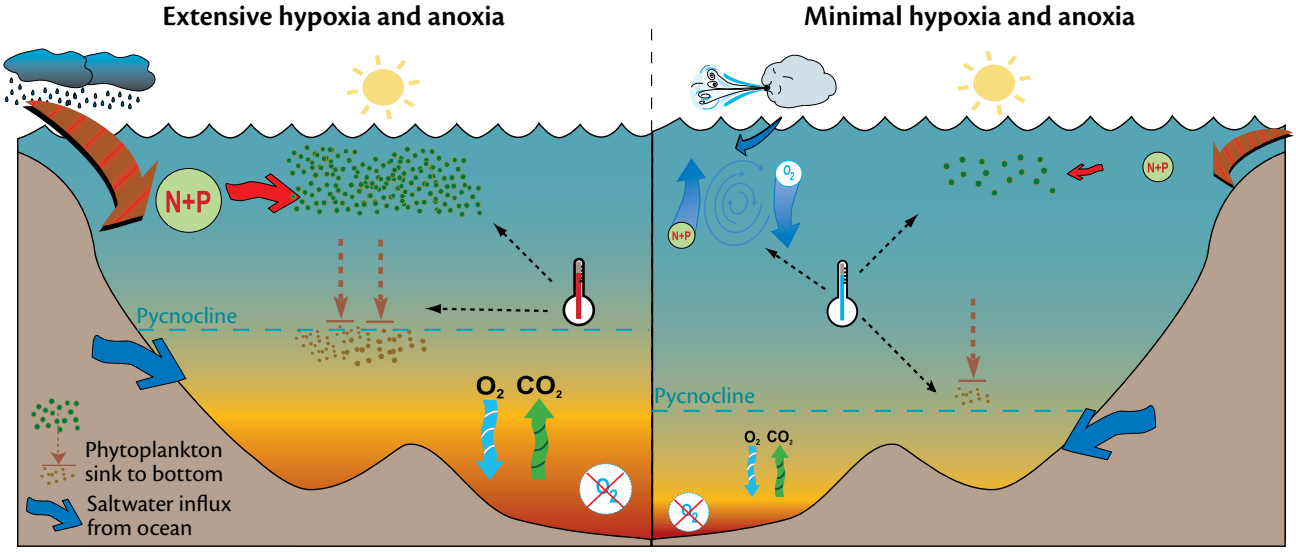
Dissolved oxygen (DO) is a key indicator of ecosystem health, especially during the summer. Nearly all aquatic animals need adequate DO in the water to survive (Figure 5.1)—even aquatic plants can be harmed if the water around their roots is low in DO. Low dissolved oxygen levels (hypoxia is DO <2.0 $mg \cdot l^{-1}$; anoxia is DO <0.2 $mg \cdot l^{-1}$) can also cause changes in water chemistry that may trigger the release of nutrients from sediments into the water column.

Low DO is often a result of eutrophication—excess nutrients in the water fuel algal blooms, and when the algae die and decompose, the decomposition process uses up DO. Most problems with low DO occur during the summer due to increased temperatures (warm water holds less oxygen) and higher biological activity in the water column. Thus, summer is the key time to measure DO. Additionally, low DO is more common in deeper waters for three reasons (Figure 5.2):

1) Deep water is less easily aerated by diffusion from the air, wind-driven mixing, and oxygen from photosynthesis,
2) Dead organic matter tends to fall to the bottom, and its decomposition is one of the main causes of low dissolved oxygen, and
3) In stratified estuaries (usually ones with more river flow), lighter, fresher river water tends to stay on top of heavier, saltier, and often cooler ocean water, further separating surface from bottom waters. This separation creates a barrier to mixing with the more oxygenated surface waters.

Maryland Department of the Environment

Figure 5.1. A fish kill due to near-zero DO levels in Baltimore Harbor, MD during 2008.



	Loads	Phytoplankton	Decomposition	Temperature	Wind event
Extensive hypoxia and anoxia	Large nitrogen and phosphorus loads	Elevated nutrients cause large phytoplankton blooms	High oxygen consumption by decaying phytoplankton	Warm water a) Stimulates decomposition b) Stratifies water column c) Stimulates phytoplankton	No wind event: water column remains stratified
Minimal hypoxia and anoxia	Small nitrogen and phosphorus loads	Less nutrients lead to small or no phytoplankton blooms	Low oxygen consumption by decaying phytoplankton	Cool water: a) Slow decomposition b) Mixed water column c) Slow phytoplankton growth	Wind events destratify water column: a) Bottom water aerated b) Nutrients move to surface

Figure 5.2. This conceptual diagram illustrates some of the ecosystem conditions that can lead to anoxic or hypoxic conditions in estuarine water systems.

Field sampling procedures

Multi-parameter meters, such as YSI sondes, are typically used to measure DO (Figure 5.3). Further information regarding instrumentation can be found in Addendum 1. The general procedure for measuring DO in the field using a meter is as follows:

- The DO probe must be calibrated prior to use. Typically, the calibration/storage cup over the probe is loosened until just barely engaged with the probe body. Make sure there is water in the cup but that it does not cover the probe. Turn on the meter and toggle to calibration mode. Wait for the DO reading to stabilize and then use the buttons to accept calibration as 100% oxygen concentration. Refer to the manufacturer's instructions for the proper calibration procedure.
- Upon reaching the sampling location, remove the protective calibration cup and replace it with the probe guard.
- Place the probe in the water to desired depth.
- Wait for the reading to stabilize.
- Record the reading on the field datasheet and/or in the YSI computer, making sure to name the station and date correctly.
- Proceed to the next depth.
- Replace the protective calibration cup to prevent damage to the probes during transition, and proceed to the next sampling location.

If, at any point, the probe touches the bottom, raise the probe to the desired depth above the bottom and wait several minutes for the disturbed sediment to settle or to flow away from probe. If the probe is equipped with a turbidity probe, wait until the turbidity reading returns to appropriate range before recording DO. This is an indication that any disturbance caused by the probe hitting the bottom has passed.

Caroline Wicks

Figure 5.3. A scientist measures dissolved oxygen in the West River, Maryland, using a YSI probe.

Troubleshooting

- If the recorded DO value is impossible (e.g., less than zero) or highly improbable (e.g., thousands of milligrams per liter), or the reading takes a very long time to stabilize, the probe likely needs to be re-calibrated or the DO membrane needs to be replaced.
- If the meter is not equipped with a pressure gauge for depth estimation and the current is strong enough to pull the meter so that the cable is at an angle noticeably different than vertical, estimation of depth will have to be corrected. Weighted probe guards may help prevent displacement by current.

Sampling scale

In addition to instrumentation, other important issues should also be considered when sampling DO, including temporal and spatial scales and depth profiles.

Temporal scale

Most monitoring programs sample DO starting in April or May, when most low DO events begin to occur. This protocol recommends measuring DO from April to October. While DO can be measured year round, low DO is usually only a problem during the summer months. This protocol also recommends sampling at least twice a month in the summer (see Table 4.1 in previous chapter), though weekly sampling is more likely to detect extreme conditions.

Time of day can affect surface DO levels because more DO is produced by photosynthesis as the day progresses, and more DO is used up overnight during respiration. Analysis of historical data has shown that time of day does not affect the overall scores for dissolved oxygen; however, as a general rule it is best to measure DO at the same time of day on each sampling trip to be consistent.

Spatial scale

The location of DO measurements within a water system is very important. DO tends to vary more over space than other common measures of water quality (such as temperature and salinity), so sampling locations may have more of an effect on the results. Compared to other parameters, sampling at a larger number of sites is necessary to effectively characterize DO status.

Some guidance regarding site characteristic effects on DO includes:

- Deeper sites tend to have more instances of low DO,
- Shallow sites on the edges of the mainstem will tend to have higher DO than sites of the same depths in small creeks, because the sites on the mainstem tend to be better mixed (from winds and waves traveling over the larger expanse of water), and

- Upstream sites in creeks and tidal creeks may have lower DO, because they tend to be narrower and thus have less wind and wave generated mixing. Additionally, these sites are generally closer to the sources of nutrients that promote algae blooms, which cause low DO when they die and decompose.

Depth profiles

It is important to measure DO at multiple depths: surface and bottom at minimum, and profiles (equally spaced intervals) at deeper sites (Figure 5.4). Recommended profile measurements are:

- Surface and bottom measurements at sites ≤3 m deep,
- One meter profiles at sites >3 m deep,
- Measure 0.3 m above bottom, then 1 meter intervals to 0.3 m below surface,
  - Example: At 3.4 m deep site, measure at 3.1, 3.0, 2.0, 1.0, and 0.3 m;
  - Example: At 3.0 m deep site, measure at 2.7 and 0.3 m (would not capture measurements at 2 and 1 m).
- Record mg·l^{-1} value only; % saturation can be calculated from it if needed, and
- Record temperature and salinity from the same depths.

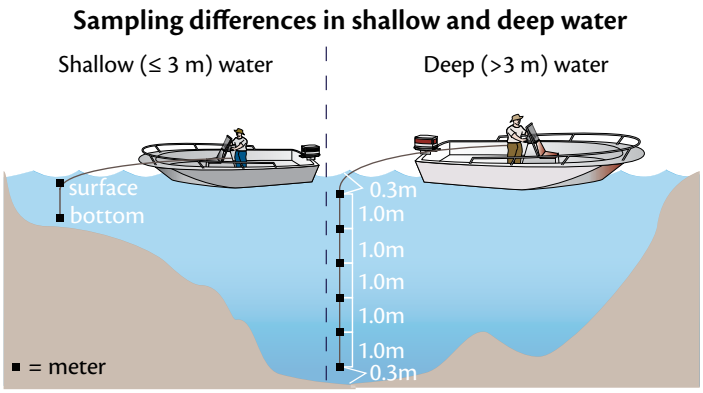


Figure 5.4. When sampling in deep water compared to shallow water, an equal interval from top to bottom is needed.

Data analysis

DO data are compared against ecologically relevant criteria and assigned as passing or failing.

Several issues relate to the analysis of DO data, including designated use determination, stratification of the water column, and assessment of appropriate thresholds for each measurement. For tidal Chesapeake Bay tributaries, the Chesapeake Bay Program has predetermined designated use areas by analyzing historical DO data and water depth (Figure 5.5a, b, and Table 5.1). For areas outside of the tidal Chesapeake Bay, an assessment of expected stratification must be made on a case-by-case basis using historical DO data and bathymetry. If stratification and designated uses are not determined, an open water criteria of 5 mg·l^{-1} should be used.

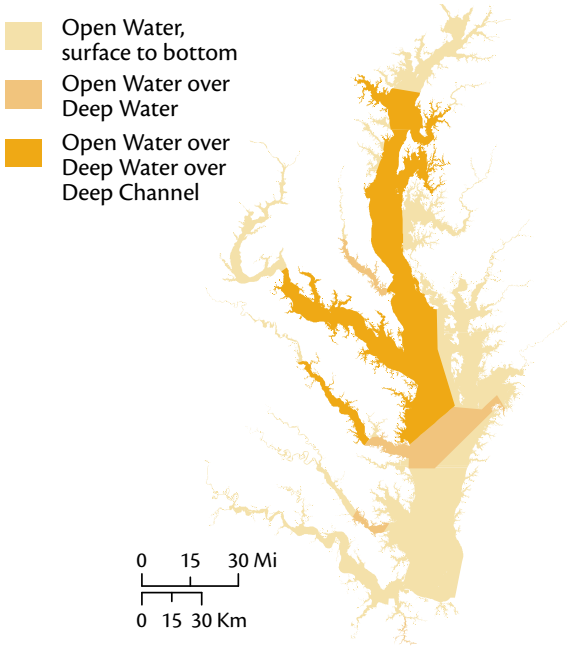


Figure 5.5a. A map showing a plan view of designated use areas for Chesapeake Bay and its tributaries.

Stratification

Some areas within a water system, such as a deep, mainstem channel of a river, are expected to have frequent water column stratification during the summer. In estuaries, stratification occurs based on water density and is a naturally-occurring phenomenon that can be exacerbated by eutrophication effects (Figure 5.2, page 16). Temperature and salinity are used to calculate density, which in turn is used to calculate pycnocline (i.e., change in density) boundaries. For each measurement of temperature and salinity, the existence of upper and lower pycnocline boundaries is determined by looking for the shallowest robust vertical change in density of 0.1 kg·m^{-3}·m^{-1} for the upper boundary and deepest change of 0.2 kg·m^{-3}·m^{-1} for the lower boundary. To be considered robust, the density gradient must not reverse direction at the next measurement and must be accompanied by a change in salinity, not just temperature. See detailed calculation methods for determining stratification in Addendum II.

In Chesapeake Bay tributaries that have deep water and deep channel designated use zones, pycnocline depths must be calculated to determine which DO criteria apply where (see Addendum II). At the time this document went to press, the Chesapeake Bay Program was updating its pycnocline calculations and therefore its designated use areas for Chesapeake Bay tributaries. This document will be updated with any new information as it becomes available. Please contact the Chesapeake Bay Program for the latest pycnocline and designated use areas.

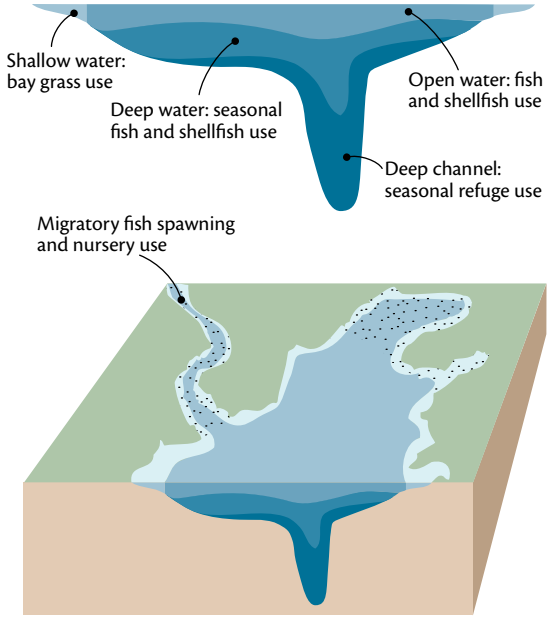


Figure 5.5b. Depth view of designated use areas for Chesapeake Bay and its tributaries.

change—they can be one designated use in one month and another designated use the next month. Therefore, *criteria applied to DO data are determined by designated use and stratification at each site on each sampling instance.* Each individual data point is then compared to the appropriate criterion and scored as pass or fail.

Each individual measurement is assigned a 100 (pass) or a zero (fail) and a station score is calculated by averaging all measurements taken at that station during the relevant time period. Then, station scores are averaged into a sub-region score. An overall score is calculated as an area-weighted average of the sub-region scores. A summary of the data analysis steps are listed below:

1) Calculate upper and lower boundaries of pycnocline.
2) Assign threshold values to appropriate designated use layers (5 $mg \cdot l^{-1}$ for open water, 3 $mg \cdot l^{-1}$ for deep water, or 1 $mg \cdot l^{-1}$ for deep channel).
3) Compare measured DO value at each depth to the appropriate threshold and assign it pass/fail. This can be done using an If/Then statement (Figure 5.7).

Comparison to criteria

Once stratification (or its absence) has been determined, the appropriate criteria for the different layers (i.e., the designated uses) can be applied to the data (Figure 5.6). Sites where no stratification is expected are open water designated use areas, and all measurements in those areas use a 5.0 $mg \cdot l^{-1}$ criterion.

Portions of the water column that are deeper than stratification boundaries (pycnocline) are expected to have lower DO. Criteria for measurements below stratified layers therefore are lower than 5.0 $mg \cdot l^{-1}$.

For example, where a single stratification layer is evident (deep water designated use areas), the 5.0 $mg \cdot l^{-1}$ criterion will apply to samples above the pycnocline, and a 3.0 $mg \cdot l^{-1}$ criteria will apply to measurements below the pycnocline.

Likewise, where two stratification layers are evident (deep water and deep channel designated use areas), a criteria of 1.0 $mg \cdot l^{-1}$ is applied to measurements below the lower pycnocline boundary. If a measurement is above the criterion it has passed, and if it is below it has failed.

It is important to remember that data points located near the pycnocline can

Table 5.1. A sample of river systems with designated uses assigned by the Chesapeake Bay Program. Most rivers only have one designated use (Open Water) and therefore can use just 5.0 $mg \cdot l^{-1}$ as the criterion for all dissolved oxygen measurements. Some rivers, however, have multiple designated uses; therefore, the pycnocline will need to be calculated and the correct criteria assigned for the assigned designated use at the corresponding depth.

River	Designated Uses	Dissolved oxygen criteria ($mg \cdot l^{-1}$)
Patapsco	Open water fish and shellfish habitat	5
	Deep-water seasonal fish and shellfish habitat	3
	Deep-channel seasonal refuge habitat	1
Magothy	Open water fish and shellfish habitat	5
Severn	Open water fish and shellfish habitat	5
South	Open water fish and shellfish habitat	5
West and Rhode	Open water fish and shellfish habitat	5
Patuxent	Open water fish and shellfish habitat	5
	Deep-water seasonal fish and shellfish habitat	3
Sassafras	Open water fish and shellfish habitat	5
Chester	Open water fish and shellfish habitat	5
	Deep-water seasonal fish and shellfish habitat	3
	Deep-channel seasonal refuge habitat	1
Choptank	Open water fish and shellfish habitat	5
Nanticoke	Open water fish and shellfish habitat	5
Wicomico	Open water fish and shellfish habitat	5

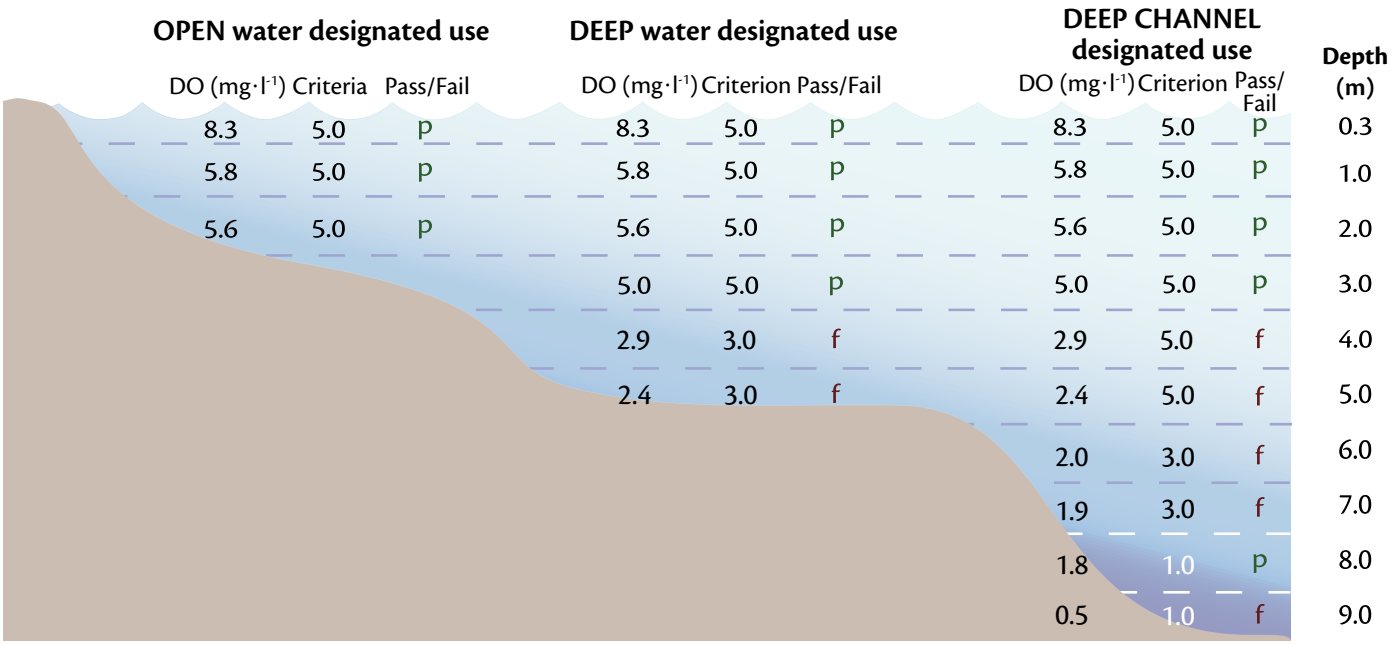


Figure 5.6. Measuring DO using profiles is especially important where there is a pycnocline. This figure shows an example of how to compare measured values against the appropriate threshold to determine pass/fail values.

4) For each pass value, assign it a 100 (one hundred), and for a fail, a 0 (zero).
5) Average the 100s and 0s (zeroes) for each station. This is the average % passing, and therefore the score, for each station.
6) Average the station scores into an average sub-region score.
7) Based on the average % score, assign a grade for each sub-region (see Chapter 4 for grade scale).

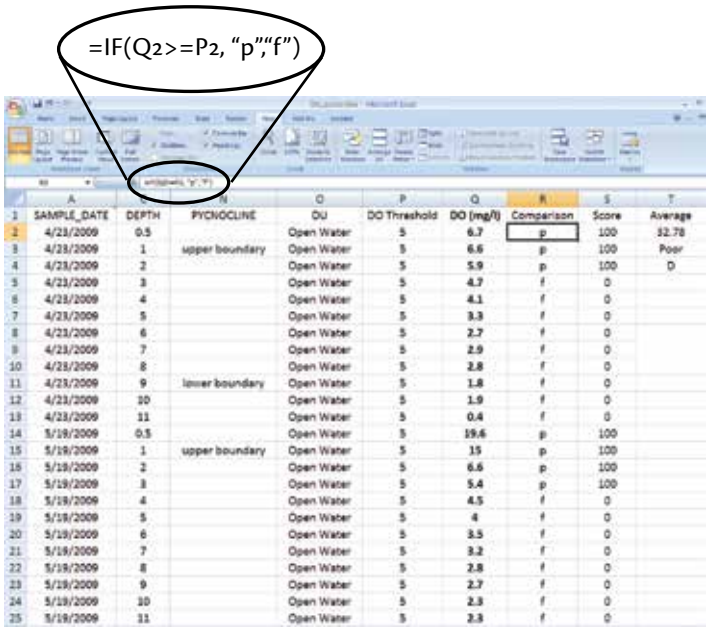


Figure 5.7. An example of using an If/Then statement in a spreadsheet program (e.g., Excel).

Next, determine the average % score and grade for the overall water system.

1) Calculate the area of each sub-region and area-weight the sub-region average before calculating the average DO score for the entire ecosystem. (Example: DO=45% for sub-region 1, sub-region 1 area = 5 km^2 out of a total 20 km^2= 0.25, 45% x 0.25 = 11.25%.
2) Sum the resulting sub-region area-weighted DO scores into an overall score.
3) Assign a grade to the total % score for the entire waterbody (see Chapter 4 for grade scale).

For health assessments, it is recommended that DO measurements for each station are scored and the % passing for each station is calculated. This method is followed so that a station that has many more measurements than others is not weighted more heavily than others. For example, if one site has 20 measurements and another site has 10, the site with 20 measurements would have more influence on the final average DO score than the site with 10 measurements if the values were averaged over the whole region. However, if the percent passing is calculated for each station, the % passing scores are equally weighted.

Chapter 6: Measuring chlorophyll *a*

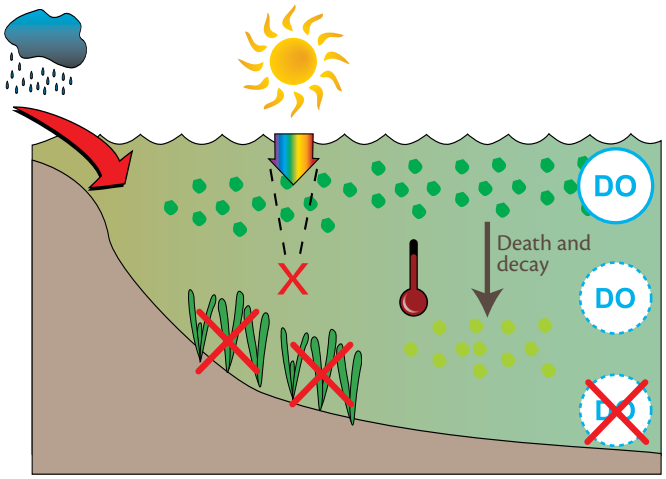
Chlorophyll is essential to the health and diversity of estuaries. It is the green pigment that allows plants to convert sunlight into organic compounds during photosynthesis. Of the several kinds of chlorophyll, chlorophyll *a* is the predominant type found in microalgae in fresh and saltwater ecosystems. Therefore, chlorophyll *a* is used as a measure of microalgae biomass, which is controlled by factors such as water temperature and light and nutrient availability. Too much algae leads to large algal blooms that can reduce water clarity, thereby threatening aquatic grasses, an important habitat for fish, invertebrates, and other organisms. Additionally, once an algal bloom dies, the algae cells sink to deeper water, where they decay and deplete waters of oxygen (Figure 6.1). Lower algae levels promote cleaner, clearer water, more available habitat, and fewer harmful algal bloom effects. Having the right level of chlorophyll *a* generally means there are enough algae to fuel the food web.

Hilary Stevens

Figure 6.2. Sampling chlorophyll *a* in the field can be done in several different ways. In this photo, a handheld vacuum pump is used to pump water through a filter. The filter is labeled and placed on ice before sending to an analytical lab.

Field sampling procedures

Measuring chlorophyll concentrations in water (Figure 6.2) is a surrogate for an actual measurement of algae biomass, which is far more expensive and time consuming. While there are several techniques for measuring chlorophyll *a*, the recommended procedure for monitoring programs is to use grab samples. Grab sampling is just what it sounds like: all of the test material is collected at one time. As such, a grab sample reflects ecosystem condition only at the point in time that the sample was collected. Water samples can be preserved and sent to a lab for analysis, simplifying the collection process and, therefore, any errors in collection.



Nutrients from land run off into water, and along with warm water temperatures in the spring and summer, fuel algal blooms. The algae blocks sunlight, which aquatic grasses need to grow. The algal bloom eventually dies and decays, a process that uses up dissolved oxygen.

Figure 6.1. This conceptual diagram illustrates the role of chlorophyll *a* in an ecosystem.

For this protocol, a minimum of twice monthly sampling from March to September is needed. However, preferably chlorophyll *a* should be measured weekly from March to October. The health of the chlorophyll indicator is assessed using just spring (March–May) and summer (July–September) samples. June is left out of the data analysis because chlorophyll *a* is highly variable during this month and it is uncertain if the spring or summer threshold should apply to the data. Therefore, if a group is very tight on funding and personnel, they can decide to skip the two June sampling dates for chlorophyll *a*.

The salinity regime of a system must also be determined before sampling begins. In an estuarine water system, freshwater mixes with salty waters, leading to a salinity gradient that affects chlorophyll *a*. Consequently, different thresholds are applied to different salinity regimes during different seasons. The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt). See Data Analysis section for more details.

Example field supplies

- Sampling pole (an ideal sampling pole is a 12-foot extendable pole with clamp)
- 500-mL polypropylene (PP) sample bottles
- 50-mL syringes
- filter bodies with fourteen (14) filter caps
- 25-mm 0.7-µm porosity GF/F filter membranes
- Handheld vacuum pump

- Opaque towels
- Aluminum foil
- Filter forceps

Before going out to collect samples, prepare equipment and supplies according to the recommended sampling procedure of the laboratory where the samples will be analyzed. This can include syringe filtering or a handheld vacuum pump and filters.

Sample collection

1) Using the sampling pole, rinse the 500-mL labeled site-specific bottle and syringe three times.
2) Fill the 500-mL bottle with water just beneath the surface (Figure 6.3).
3) Follow the recommended filtering procedure by the analytical laboratory where the samples will be analyzed. Color on the filter generally indicates a sufficient sample for analysis (Figure 6.4).
4) Record the volume of water pushed through the filter on the data collection sheet.
5) Store samples in cooler. Samples must be kept cool and out of sunlight for the duration of field sampling.
6) Cap 500-mL bottle retaining sampled water and store in dark location to bring back to lab. This sample will serve as a back-up sample should there be a filter problem.

Peter Bergstrom

Figure 6.3. If sampling from a pier, use an extendable grab pole to hold the bottle under the surface. Make sure to rinse the bottle three time before collecting the sample.

Laboratory preparation

Following the procedures laid out by the analytical lab that will process the samples is important. Here are a few general steps:

1) Prepare pieces of aluminum foil.
2) Fold in half again, then unfold, creating a crease.
3) Create labels using labeling tape noting site number, date, and volume pressed through filter.
4) Place filter in aluminum foil with the center of the filter centered on the crease, with side containing the intercept chlorophyll up (should have slight color to it). Folding foil and gently assisting with forceps if necessary by pressing on filter fold the filter in half.
5) Double over edges of fold, displacing air and create a little pocket in which the folded filter is located.
6) Repeat for all samples.
7) Label foil packets.
8) Place foil packets in locking plastic bag and then double bag with another locking plastic bag.
9) Place in freezer to await shipment to the analytical laboratory.
10) Rinse all filter holders and 500-mL bottles with tap water and allow to air dry.

It is critical that the chlorophyll water samples and foil packets remain dry. The samples in foil should be double bagged and packed with ice in portable Styrofoam transport coolers with surrounding cardboard box. Samples should be mailed overnight to arrive at the analytical laboratory as soon as possible. If properly packaged and frozen (sampled filters should be stored frozen, at at least -20˚C, in the dark), chlorophyll *a* samples can be stored for up to three and a half weeks. The package should also be marked to indicate "chlorophyll samples" as contents.

Lab analysis

Chlorophyll *a* is measured using a spectrophotometer. Spectrophotometry is conducted by the analytical lab that the samples are taken to. The chlorophyll *a* detection limit is 0.62 $\mu g \cdot l^{-1}$. For detailed methods on the spectrophotometry procedure, visit the analytical laboratory's website or request the written standard operating procedure. There are several labs within the Chesapeake region that can be used, for example, Chesapeake Biological Laboratory in Solomons Island, Maryland, Horn Point Laboratory in Cambridge, Maryland, and the Maryland Department of Health and Mental Hygiene in Baltimore, Maryland.

Laura Fabien

Figure 6.4. Filters with chlorophyll residue on them. The different colors represent the different amounts of chlorophyll in the water sample.

Data analysis

Once samples have been analyzed in the lab, a spreadsheet of data will be provided. These data should be compared to ecologically relevant thresholds that, for chlorophyll *a*, are based on levels of dissolved inorganic nitrogen and orthophosphate that are low enough to limit the formation of algal blooms and on light penetration (Secchi depth) that is deep enough to promote healthy plant growth and favor a positive energy balance between photosynthesis and respiration (Table 6.1, Buchanan et al 2005).

For chlorophyll *a*, each data point is separated into season and salinity regime and compared to a corresponding threshold. The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt). For example, a data point collected in March and in a tidal fresh area would be compared to a threshold in Table 6.1a.

Each data point is compared to the thresholds in the appropriate table and scored from 0 to 5. Each measurement score (0–5) is averaged into a station score for the entire season. Then, station scores are averaged into a sub-region score. Once the score for the sub-region is calculated, calculate a total overall score by area-weighting each sub-region score and averaging them for an overall score. A summary of steps for calculating the chlorophyll *a* scores is:

1) Make sure the data used for data analysis is from the relevant months. For chlorophyll *a*, this is March to May and July to September.
2) Filter data by salinity regime and season.
3) Compare individual measurements to relevant thresholds for that salinity regime and season.
4) Score all measurements from 0 to 5 (see multiple thresholds tables).

Table 6.1a. Ecologically relevant multiple thresholds for chlorophyll *a* for spring and summer for the TIDAL FRESH salinity regime.

Score	Spring (Mar–May) thresholds ($\mu g \cdot l^{-1}$)	Summer (Jul–Sept) thresholds ($\mu g \cdot l^{-1}$)
5	≤1.0	≤1.8
4	>1.0–≤14.0	>1.8–≤12.0
3	>14.0–≤18.7	>12.0–≤22.4
2	>18.7–≤24.8	>22.4–≤37.1
1	>24.8–≤35.6	>37.1–≤65.4
0	>35.6	>65.4

Table 6.1b. Ecologically relevant multiple thresholds for chlorophyll *a* for spring and summer for the OLIGOHALINE salinity regime.

Score	Spring (Mar–May) thresholds ($\mu g \cdot l^{-1}$)	Summer (Jul–Sept) thresholds ($\mu g \cdot l^{-1}$)
5	≤1.5	≤3.0
4	>1.5–≤20.9	>3.0–≤9.5
3	>20.9–≤27.7	>9.5–≤16.4
2	>27.7–≤39.4	>16.4–≤29.9
1	>39.4–≤62.3	>29.9–≤76.8
0	>62.3	>76.8

Table 6.1c. Ecologically relevant multiple thresholds for chlorophyll *a* for spring and summer for the MESOHALINE salinity regime.

Score	Spring (Mar–May) thresholds ($\mu g \cdot l^{-1}$)	Summer (Jul–Sept) thresholds ($\mu g \cdot l^{-1}$)
5	≤2.09	≤1.7
4	>2.09–≤6.2	>1.7–≤7.7
3	>6.2–≤11.1	>7.7–≤11.0
2	>11.1–≤19.1	>11.0–≤15.8
1	>19.1–≤49.8	>15.8–≤35.8
0	>49.8	>35.8

Table 6.1d. Ecologically relevant multiple thresholds for chlorophyll *a* for spring and summer for the POLYHALINE salinity regime.

Score	Spring (Mar–May) thresholds ($\mu g \cdot l^{-1}$)	Summer (Jul–Sept) thresholds ($\mu g \cdot l^{-1}$)
5	≤2.5	≤2.9
4	>2.5–≤2.8	>2.9–≤4.5
3	>2.8–≤6.9	>4.5–≤7.7
2	>6.9–≤12.6	>7.7–≤11.2
1	>12.6–≤31.7	>11.2–≤25.0
0	>31.7	>25.0

5) Calculate the percent score for each station by averaging all the scored (0 to 5) measurements at each station, and then divide the average score by 5 and multiply by 100 (e.g., station 1 average chlorophyll *a* score = 3.8/5.0 = 0.76 x 100 = 76%).
6) Calculate sub-region scores by averaging the scores of the stations in each sub-region.
7) Assign a grade to each sub-region score (see Chapter 4 for grade scale).

Now you have a score and grade for each sub-region. Next, you want to determine the average % score and grade for the overall water system.

1) Calculate the area of each sub-region and area-weight the sub-region average before calculating the average chlorophyll *a* score for the entire waterbody (e.g., chlorophyll *a* = 76% for sub-region 1, sub-region area = 5 km^2 out of a total 20 km^2= 0.25, 76% x 0.25 = 19%).
2) Sum the resulting sub-region scores into an overall score.
3) Based on the overall score, assign a grade for the entire waterbody.

For health assessments, it is recommended that chlorophyll *a* measurements for each station are scored, the % passing for each station is calculated, and region scores are calculated as the average of the station average scores. This method is followed so that a station that has more measurements than others is not weighted more heavily than others. For chlorophyll *a*, this happens if one station is not sampled during a routine field day, perhaps due to time constraints, missing filters, or because the sampling site is very shallow and sampling occurs during extreme low tides.

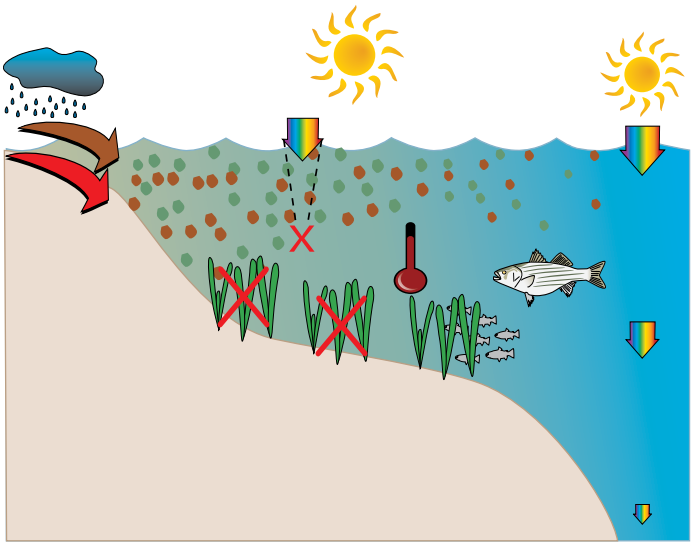
Chapter 7: Measuring water clarity

Water clarity is a measure of how much light penetrates though the water column. It is dependent upon the amount of suspended particles (e.g., sediment and plankton) and colored organic matter present. Clear water is critical for the growth and survival of aquatic grasses, as well as fish, crabs, and other aquatic organisms (Figure 7.1).

However, clear water should not be confused with the color of the water. Black water systems, for example, have highly colored water, but that is a natural phenomenon and is not an indication of eutrophication. Accordingly, the type of system (brown, black, or blue water) will determine what thresholds to use for Secchi depth. For example, in black water systems, where the water color is naturally dark, Secchi values are significantly lower than in blue water systems, such as tropical waters. Additionally, other water clarity measurements, such as Total Suspended Solids or light attenuation, may be more appropriate than Secchi in blackwater systems (see Addendum III).

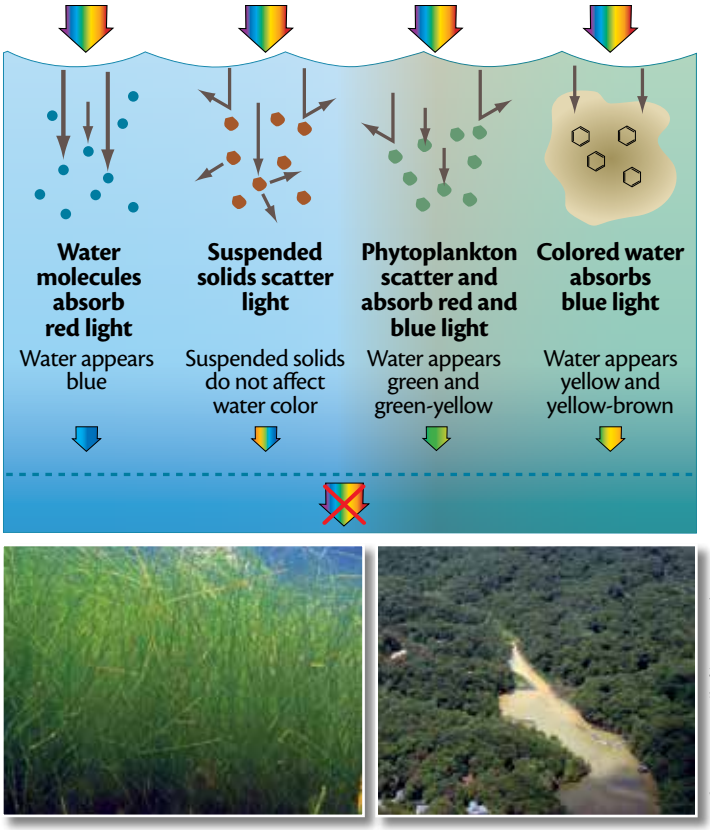
Poor water clarity is usually caused by a combination of excess suspended sediments from runoff from the land and the growth of phytoplankton, which is fueled by nutrients (see Figure 7.2 and Chapter 8).

For this protocol, a minimum of twice monthly sampling from April to October for tidal fresh, oligohaline, and mesohaline areas is needed. A minimum of twice monthly sampling from March to November for polyhaline is needed. However, preferably water clarity should be measured weekly from March to November for all systems. These time ranges have been established because they have shown to be periods when water clarity can be low enough to negatively affect biological processes in Chesapeake Bay, which is a brown water system.



Nutrients and sediments from the land run off into the water, and along with warm water temperatures in the spring and summer, fuel algal blooms. The algae blocks sunlight, which aquatic grasses need to grow. Water clarity also influences how well predators can see their prey.

Figure 7.1. This conceptual diagram illustrates the role of water clarity in an ecosystem.



Adrian Jones (left); Jane Thomas

Figure 7.2. When light enters water, there are several factors that affect how far into the water column the light penetrates (top). Aquatic grasses need light to grow, but sediment runoff from land can block light and smother the grasses (bottom).

The salinity regime of a system is also important. In an estuarine, brown water system, freshwater mixes with salty waters, leading to a salinity gradient that affects water clarity. Therefore, different thresholds are applied to different salinity regimes during different seasons. The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt). See Data Analysis section for details.

Field sampling procedures

Water clarity is measured by Secchi depth. This is determined in meters using a standard Secchi disk (Figure 7.3), which has alternating black and white quadrants. The size of the disk is dependent on the sampling area. The disk is lowered into the water column with a calibrated rope to the point where the disk is just visible. Specific steps are provided in the next section.

Adrian Jones

Figure 7.3. A Secchi disk being lowered into the water to measure water clarity.

Sample collection

Observations should be made on the shady side of a boat or dock, as close as possible to noon (ideally between 10 a.m. and 2 p.m.; Figure 7.4).

1) Lower the Secchi disk into the water until it just cannot be seen.
2) Raise the Secchi disk until it just becomes visible.
3) Raise and lower the Secchi disk several times to 'zero in' on the depth(s) where the disk becomes invisible and visible.
4) Record the midpoint of these two depths as the Secchi depth.
5) If the Secchi disk is still visible on the bottom, record the Secchi depth as ">[depth of disk on bottom]" and make a note for that observation that the Secchi disk was on the bottom.

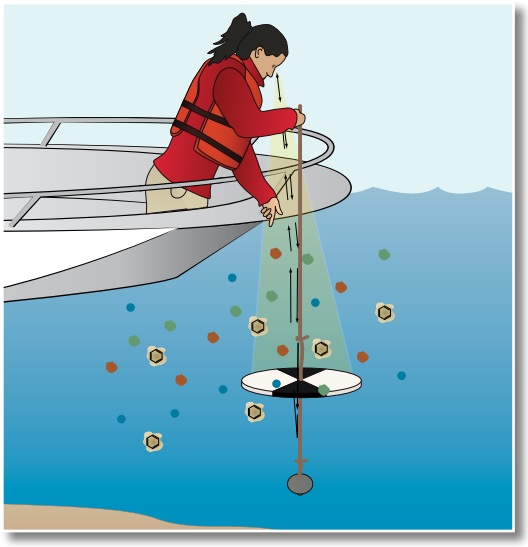


Figure 7.4. Water clarity is measured by lowering a Secchi disk over the edge of a boat or dock and observing when the white part of the disk can no longer be seen by the observer.

Data analysis

Field sampling measurements should be marked on a field data sheet, then entered in a spreadsheet or database.

For analysis, each data observation is filtered by salinity regime and compared to a corresponding threshold (Table 7.1). The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt). Each Secchi depth observation is measured against a multiple threshold criteria set and assigned a score from 0 to 5. Each measurement score (0–5) is averaged into a station score for the entire season. Then, station scores are averaged into a sub-region score. Once the score for the sub-region is calculated, calculate a total overall score by area-weighting each sub-region score and averaging them for an overall score.

If the Secchi measurement indicates that the depth was so shallow that the Secchi disk lay on the bottom, this protocol recommends scoring that measurement as a 4. Other options are: not including those measurements in the scoring process at all or using a Secchi tube to determine Secchi depth.

A summary of steps for calculating water clarity scores is:

1) Make sure the data used for data analysis is from the relevant months. For water clarity, this is April to October for tidal fresh, oligohaline, and mesohaline and March to November for polyhaline.
2) Filter data by salinity regime.
3) Compare individual measurements to relevant threshold for each salinity regime.
4) Score all measurements from 0 to 5 (see multiple thresholds table).
5) Calculate the percent score for each station by averaging all the scored (0 to 5) measurements at each station, and then divide the average score by 5 and multiply by 100 (e.g., average water clarity at station 1 = 3.8/5.0 = 0.76 x 100 = 76%).
6) Calculate sub-region scores by averaging the scores of the stations in each sub-region.
7) Assign a grade to each sub-region score (see Chapter 4 for grade scale).

Table 7.1. Multiple thresholds based on salinity regime for water clarity calculations.

Score	Tidal Fresh	Oligohaline	Mesohaline	Polyhaline
5	≥1.3	≥0.9	≥1.8	≥2.1
4	≥0.9–<1.3	≥0.7–<0.9	≥1.6–<1.8	≥2.0–<2.1
3	≥0.6–<0.9	≥0.5–<0.7	≥1.0–<1.6	≥1.1–<2.0
2	≥0.4–<0.6	≥0.3–<0.5	≥0.6–<1.0	≥0.8–<1.1
1	≥0.2–<0.4	≥0.2–<0.3	≥0.3–<0.6	≥0.5–<0.8
0	<0.2	<0.2	<0.3	<0.5

Now you have a score and grade for each sub-region. Next, you want to determine the average % score and grade for the overall water system.

1) Calculate the area of each sub-region and area-weight the sub-region average before calculating the average water clarity score for the entire waterbody (e.g., water clarity = 76% for sub-region 1, sub-region area = 5 km^2 out of a total 20 km^2= 0.25, 76% x 0.25 = 19%).
2) Sum the resulting sub-region scores into an overall score.
3) Based on the overall score, assign a grade for the entire waterbody.

For health assessments, it is recommended that water clarity measurements for each station are scored and the % passing for each station is calculated. This method is followed so that a station that has more measurements than others is not weighted more heavily than others. For water clarity, this could happen if one station is not sampled during a routine field day, perhaps due to time constraints or because the sampling site is very shallow and sampling occurs during extreme low tides.

Additional considerations when measuring water clarity using a Secchi disk:

- Each group should define a standard procedure for wearing sunglasses during Secchi depth measurements. Consider recommending that sunglasses should not be worn since they can affect Secchi measurements.
- Use the standard size disk for the water clarity range expected in your region (e.g., in the Chesapeake Bay, it is recommended to use a 20-cm disc).
- If using a calibrated rope, measure it monthly to ensure that it has not stretched or shrunk, or use a surveyor's tape instead.
- Check to make sure the Secchi disk is clean and well maintained before each sampling trip to ensure consistent visibility.
- If a site is so shallow that the Secchi disk lays on the bottom, a Secchi tube could be used. Each group will have to research and decide if using a Secchi tube is an option for them.

Chapter 8: Measuring nutrients

Nutrients are essential to the health and diversity of estuaries. However, excessive nutrients in water systems can lead to algal blooms, which may then lead to human health issues. Many of the tidal tributaries in the Mid-Atlantic region, including all of those within the Chesapeake Bay watershed, are currently impaired for nutrients under the federal Clean Water Act. An impaired tributary is one that is too polluted or otherwise degraded to meet state water quality standards.

The primary nutrients of concern are nitrogen and phosphorus. Both are required for plants and animals to grow; however, when in excess, they can cause serious problems.

Total nitrogen

Nitrogen is an essential nutrient for all plants and animals and naturally occurs in the environment and water systems. However, due to human activities, nitrogen is entering water systems at unsustainably high rates. Nitrogen may enter water systems from sources such as power plants (through atmospheric deposition), agricultural practices, septic systems, sewer overflows, and wastewater treatment plants (Figure 8.1). As part of the nitrogen cycle, phytoplankton and macroalgae take up nitrogen and use it during photosynthesis for growth. Bacteria also use nitrogen for growth. When nitrogen is present in excess, algae overgrowth may occur, resulting in an algal bloom that eventually dies and decays—a process that uses up dissolved oxygen, which can lead to very low dissolved oxygen levels (Figure 8.2) and subsequent fish kills. Lower algae levels promote cleaner, clearer water, more available habitat, and fewer harmful algal bloom effects. Having the right level of nitrogen generally means there are enough algae to fuel the food web.

Total phosphorus

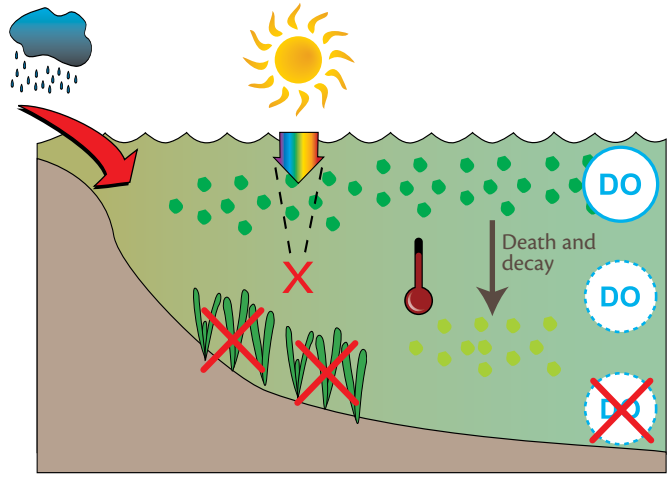
Phosphorus is also an essential nutrient for all plants and animals. However, poor land-use practices, untreated wastewater, leaking septic systems, sewer overflows, over-fertilizing, household cleaners, legacy sediments, and stormwater runoff can elevate concentrations of phosphorus in water systems. Algae, both macroalgae and phytoplankton, take up phosphorus and can grow in large, dense blooms, which can reduce water clarity, thereby threatening aquatic grasses, an important habitat for fish, invertebrates, and other organisms. Additionally, once an algal bloom dies, the algae cells sink to deeper water, where they decay and deplete waters of oxygen (Figure 8.2). Lower algae levels promote cleaner, clearer water, more available habitat, and fewer harmful bloom effects. Having the right level of phosphorus generally means there are enough algae to fuel the food web.

Jane Thomas

Figure 8.1. Fairfield Industrial Park, on the Patapsco River in Baltimore, Maryland, includes the Patapsco Wastewater Treatment Plant. Wastewater treatment plants contribute nutrients to local waterways.

Field sampling procedures

While there are many constituents of nitrogen and phosphorus in the water column, for this protocol, it is recommended that total nitrogen and total phosphorus be



Nitrogen and phosphorus from the land runs off into the water, and along with warm water temperatures in the spring and summer, fuel algal blooms . The algae blocks sunlight , which aquatic grasses need to grow. The algal bloom eventually dies and decays , a process that uses up dissolved oxygen .

Figure 8.2. This conceptual diagram illustrates the role of nutrients in some ecosystem processes.

measured and analyzed as indicators of water quality. This provides a picture of nutrients as a whole, and therefore the processes that they affect.

Nutrient indicators are assessed using April to October samples. A minimum of twice monthly sampling is needed. Preferably, however, nutrients should be measured weekly from March to October.

The salinity regime of a system must also be determined. In an estuarine, brown water system, freshwater mixes with salty waters, leading to a salinity gradient that affects nutrients and therefore the processes that occur in that area. Different thresholds are applied to different salinity regimes during different seasons. The four major salinity regimes are tidal fresh (0–0.5 ppt), oligohaline (>0.5–5 ppt), mesohaline (>5–18 ppt), and polyhaline (>18 ppt). See Data Analysis section for details.

Sampling equipment for boat

Each group should follow the procedures set by the analytical laboratory where their samples will be analyzed. However, here are some general guidelines:

- Kemmer or Van Dorn water sampler or automatic pump, with non-stretchable rope
- 500-ml Polyethylene bottles: chemically cleaned (with either dilute nitric or dilute hydrochloric acid) and dried. Safe handling and storage of chemicals should be followed at all times.
    - Dilute nitric and dilute hydrochloric acid are hazardous chemicals: please reference proper instructions for safe use and disposal
    - All sample bottles need to be labeled with sampling site identification: e.g., Station 1, surface or Station 1, bottom
- Cooler with ice

Sample collection from boat

1) Water samples should be taken one meter from the surface and one meter from the bottom of the water column at sites with depths greater than four meters. Because it is below the layer of mixing caused by wind, boating, and other activities, sampling one meter below the surface gives a better representation of the surface water column.
2) At sites with depths less than four meters, water samples should be taken one meter from the surface.
3) Facing upstream, extend the pole and bottle, rinse the bottle out three times, and take the sample the fourth time (Figure 8.3).
4) After samples are taken, immediately place the sample on ice up to the shoulders of the bottle. The lid should not be immersed under the ice, in case ice water leaks into the sample bottle, diluting the concentration of the sample. It is good procedure to put the samples in clean plastic bags that can be sealed while they are in the cooler, which prevents any contamination from the ice/water/slush in the cooler and/or other samples. It is important to note that when freezing samples the water contained in the container will expand. If too much water is placed in the sample container, the lid may pop off or the sample may squeeze out, and contamination may occur.
5) If possible, one set of duplicate samples should be taken for quality assurance/quality control purposes. An example of how this is labeled would be: Station 1, surface, duplicate.
6) On the field data sheet, record the time, date, and any other information about the water sampling event.

Sampling equipment for stream or river

- Sampling pole (an ideal sampling pole is a 12-foot extendable pole with clamp)
- 500-ml Polyethlylene bottles: chemically cleaned (with either dilute nitric or dilute hydrochloric acid) and dried. Safe handling and storage of chemicals should be followed at all times.
    - Dilute nitric and dilute hydrochloric acid are hazardous chemicals: please reference proper instructions for safe use and disposal
    - All sample bottles need to be labeled with sampling site identification: e.g., Station 1, surface or Station 1, bottom
- Cooler with ice

Caroline Wicks

Figure 8.3. After dipping the pole and bottle in the water, the bottle top is screwed on and the sample put on ice for transport to the laboratory for analysis.

Sample collection in a stream or river (side-sampling)

1) Attach the sample bottle to the sampling pole, making sure that the clamp is tight.
2) The sampling point in the stream or river should have a low to medium flow and not be in eddies or stagnant water. Entering the water by foot is not advisable due to the possibility of suspending the sediment in the river, which could potentially cause problems in the analysis.
3) Facing upstream, extend the pole and bottle, rinse the bottle out three times, and take the sample the fourth time.
4) Fill the bottle up to the shoulders and immediately cap and place on ice. The lid should not be immersed under the ice, in case ice water leaks into the sample bottle, diluting the concentration of the sample. It is good procedure to put the samples in clean plastic bags that can be sealed while they are in the cooler. This prevents any contamination from the ice/water/slush in the cooler and/or other samples. It is important to note that when freezing samples the water contained in the container will expand. If too much water is placed in the sample container, the lid may pop off or the sample may squeeze out and contamination may occur.
5) If possible, one set of duplicate samples should be taken for QA/QC purposes. An example of how this is labeled would be: Station 1, surface, duplicate.
6) On the field data sheet, record the time, date, and any other information about the water sampling event.

Laboratory analysis

Samples should be mailed overnight to arrive at the analytical laboratory as soon as possible. If properly packaged and frozen, nutrient samples can be stored for up to 28 days. The package should also be marked to indicate "nutrient samples" as contents.

Data analysis

Once samples have been analyzed in the lab, a spreadsheet of data will be provided. For nutrients, each data point is filtered by salinity regime and compared to a corresponding threshold (Table 8.1 for nitrogen and Table 8.2 for phosphorus). *The thresholds for total nitrogen and total phosphorus are different, so make sure the appropriate thresholds are being used.*

Each nutrient observation is measured against a multiple threshold criteria set under the appropriate salinity regime and assigned a score from 0–5. Each measurement score (0–5) is averaged into a station score for the entire season. Then, station scores are averaged into a sub-region score. Once the score for the sub-region is calculated, calculate a total overall score by area-weighting each sub-region score and averaging them for an overall score. A summary of steps for calculating nutrient scores is:

1) Make sure the data used for data analysis is from the relevant months. For nutrients, the minimum sampling period is April to October.
2) Filter data by salinity regime.
3) Compare individual measurements to relevant threshold for each salinity regime. (Make sure you are looking at the right table for nitrogen and phosphorus).
4) Score all measurements from 0 to 5 (see multiple thresholds tables).
5) Calculate the percent score for each station by averaging all the scored (0 to 5) measurements at each station, and then divide the average score by 5 and multiply by 100 (e.g., average total nitrogen score at station 1 = 3.8/5.0 = 0.76 x 100 = 76%).
6) Calculate sub-region scores by averaging the scores of the stations in each sub-region.
7) Assign a grade to each sub-region score (see Chapter 4 for grade scale).

Table 8.1. Total nitrogen ($mg \cdot l^{-1}$) multiple threshold table for determining scores. Thresholds are different for different salinity regimes.

Score	Tidal Fresh	Oligohaline	Mesohaline	Polyhaline
5	≤0.6	≤0.6	≤0.5	≤0.4
4	>0.6–≤0.9	>0.6–≤0.9	>0.5–≤0.6	>0.4–≤0.5
3	>0.9–≤1.3	>0.9–≤1.2	>0.6–≤0.8	>0.5–≤0.6
2	>1.3–≤1.8	>1.2–≤1.6	>0.8–≤1.0	>0.6–≤0.8
1	>1.8–≤2.8	>1.6–≤2.8	>1.0–≤1.5	>0.8–≤1.2
0	>2.8	>2.8	>1.5	>1.2

Table 8.2. Total phosphorus ($mg \cdot l^{-1}$) multiple threshold table for determining scores. Thresholds are different for different salinity regimes.

Score	Tidal Fresh	Oligohaline	Mesohaline	Polyhaline
5	≤0.04	≤0.04	≤0.02	≤0.03
4	>0.04–≤0.06	>0.04–≤0.07	>0.02–≤0.04	>0.03–≤0.05
3	>0.06–≤0.09	>0.07–≤0.10	>0.04–≤0.06	>0.05–≤0.07
2	>0.09–≤0.13	>0.10–≤0.15	>0.06–≤0.08	>0.07–≤0.09
1	>0.13–≤0.23	>0.15–≤0.28	>0.08–≤0.15	>0.09–≤0.13
0	>0.23	>0.28	>0.15	>0.13

Now you have a score and grade for each sub-region. Next, you want to determine the average % score and grade for the overall water system.

1) Calculate the area of each sub-region and area-weight the sub-region average before calculating the average nutrient score for the entire waterbody (e.g., total nitrogen = 76% for sub-region 1, sub-region area = 5 km^2 out of a total 20 km^2 = 0.25, 76% x 0.25 = 19%).
2) Sum the resulting sub-region scores into an overall score.
3) Based on the overall score, assign a grade for the entire waterbody.

For health assessments, it is recommended that nutrient measurements for each station are scored and the % passing for each station is calculated. This method is followed so that a station that has more measurements than others is not weighted more heavily than others. For nutrients, this happens if one station is not sampled during a routine field day, perhaps due to time constraints or because the sampling site is very shallow and sampling occurs during extreme low tides.

Chapter 9: Measuring aquatic grasses

Aquatic grasses are an important shallow water habitat for many aquatic organisms. They take up nutrients from the water column, which decreases the amount available for phytoplankton growth and thereby reduces the potential for harmful algal blooms. Aquatic grasses also increase water clarity and protect against shoreline erosion by dampening waves and currents (Figure 9.1 and 9.2). Their roots and rhizomes stabilize sediment, preventing it from being re-suspended. They provide critical habitat to organisms such as crabs and fish, which use aquatic grasses as nursery areas as well as protection against predators, and are a food source for waterfowl (Figure 9.3).

The amount of light reaching the bottom directly determines if aquatic grasses can survive. Therefore, aquatic grasses may be negatively affected by any increase in the amount of nutrients and sediments running off into the water that create algal blooms and smother grasses.

Field sampling procedures

Aquatic grasses are such a valued resource in Chesapeake Bay and the Maryland Coastal Bays that annual aerial surveys have been conducted since 1989. The total area of aquatic grass coverage is determined by the Virginia Institute of Marine Science's (VIMS) aerial surveys each year (Figure 9.4). Details of the sampling procedure and calculation of total hectares are available at www.vims.edu/bio/sav. All monitoring programs in the mid-Atlantic region should use the data provided by these annual aerial surveys if possible.

USGS 7.5 minute quadrangle maps are used to organize the mapping process. 258 quadrangles in the study area include all regions with potential for aquatic grasses growth. Data are aggregated to 116 tidal water segments

MDNR

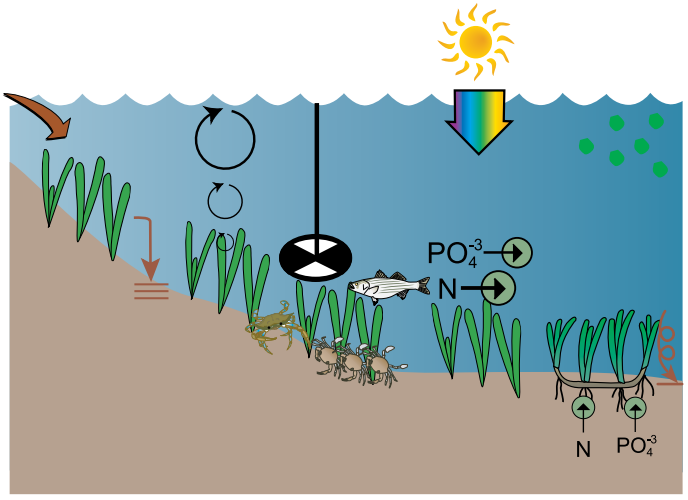
Figure 9.1. Water stargrass (*Heteranthera dubia*) in the Susquehanna Flats, in northern Chesapeake Bay. This aquatic grass bed is so large that it's visible from space.

MDNR

Figure 9.2. *Zostera marina* (eelgrass) is a common aquatic grass in higher salinity areas of the Chesapeake Bay.

for Chesapeake Bay. Different aggregations may be appropriate for different water systems and associated programs. For example, the Patapsco River is covered by three USGS quadrangle maps, but by only one Chesapeake Bay Program segment. Depending on the resolution that is needed for your water system, different spatial scales should be chosen. Also, it should be noted that the areas mapped include all regions with potential for aquatic grass growth.

Some areas were not assigned a restoration goal because historical assessments have never found grasses there or



Aquatic grasses take up nutrients PO_4^{-3} N from the water column, increase water clarity, and protect against shoreline erosion by dampening waves and currents. Their roots and rhizomes stabilize sediment. They provide critical habitat to key species, which use aquatic grasses as nursery areas as well as protection against predators.

Figure 9.3. This conceptual diagram illustrates the role of aquatic grasses in an ecosystem.

Figure 9.4, left to right. Small aircraft fly 173 flightlines around the Chesapeake Bay annually. A total of 2,033 photographs are taken at a scale of 1:24,000. Aquatic grass density is estimated from photographs and ground observations (diamonds). Aquatic grass maps are compiled by segment for the whole of Chesapeake Bay. *Figure adapted from Integration and Application Network. (2005). Bay grass restoration in Chesapeake Bay [newsletter].* Further information: http://www.vims.edu/bio/sav/.

because grasses have never been mapped there by VIMS's annual monitoring program. For example, the area from Calvert Beach north to Calvert Cliffs in Maryland is most likely either too deep or too exposed to waves and currents to support aquatic grasses. Therefore, aquatic grasses is not an appropriate indicator for that area.

There are several exceptions to these guidelines. One example is those water systems where aquatic grasses are considered a nuisance, perhaps due to invasion by exotic species, or are a detriment to boaters. Some measurements that could be used to assess conditions in such systems include biomass or acres removed and density. Each of these in turn require a different sampling procedure. This may be more of an issue in non-tidal water systems, such as lakes, and will be addressed in a forthcoming, companion non-tidal protocol document.

For areas that do not have access to scientifically validated aerial surveys, there are a variety of methods that can be used to assess aquatic grass health. These depend on the resources (e.g., SCUBA, boat) and expertise available to the watershed organization. Please refer to References for field sampling alternatives.

Data analysis

For the Chesapeake Bay and Maryland Coastal Bays regions, the actual acres of aquatic grasses are compared to the goal for that region. These goals are set based on a Single Best Year (SBY) approach. For each of the 116 segments used by the Bay Program, aerial photographs from the long-term record were assessed for the single year with the most aquatic grasses coverage.

The SBY map was clipped to an application depth (i.e., how deep in that area aquatic grasses were expected to grow based on water clarity criteria). Finally, that result was clipped to a current shoreline GIS shapefile (due to shoreline change by development and erosion). This provides the total acres of aquatic grasses goal per segment.

Table 9.1. Examples of total aquatic grasses goal for reporting regions used in the Chesapeake Bay report card.

Reporting region	Aquatic grasses goal (acres)
Upper Western Shore	3,661
Lower Western Shore (MD)	1,811
Patuxent River	1,954
Potomac River	21,203
Rappahannock River	2,534
York River	3,304
James River	2,629
Upper Eastern Shore	12,866
Lower Eastern Shore	57,651
Choptank River	13,953

Examples of these goals for some river systems are provided in Table 9.1.

While aquatic grasses hectares are calculated by the Virginia Institute of Marine Science's group, the total number of acres for your waterbody need to calculated. Once the aquatic grasses current year numbers are obtained, the steps for calculating the scores for aquatic grasses are as follows:

1) The data from each polygon segment provided by VIMS is summed into the Bay Program segments. For example, there are 6 Patuxent Mesohaline polygons, each with their own number of total hectares mapped. Some small river systems may be located within one polygon segment and therefore you don't have to sum multiple segments together.
2) The total hectares for each Bay Program segment are converted into acres (1 hectare = 2.4710538147 acres).
3) An If/Then statement is used in Excel to determine if any of the current year's total acres for each segment is bigger than the total goal for that segment. For

example, if the PAXMH goal is 300 acres, but the current year has 325 acres, then use 300 acres as the number to compare to the goal. In this way, a segment cannot reach over 100% of the goal.

4) The total acres per Bay Program segment is summed into one number of total acres for the entire region for the current year. Again, small river systems may be located in one Bay Program segment and therefore you don't have to sum multiple Bay Program segments together.
5) The total goal acres per Bay Program segment is summed into one number for total goal acres for the region.
6) Using the acres obtained in step 3, calculate the region score by dividing the current year acres by the goal acres and multiplying by 100. For example, if the total Patuxent has 3000 acres, and the goal is 5000 acres, then (3000/5000)*100 = 60%.

Except for a few river systems (e.g., Chester River), there likely is only one total score for the entire waterbody. Therefore, the total acres in the current year compared to the goal is the overall % score for the entire waterbody. If the river system is large enough, there may be sub-region scores. These are averaged into an overall score for the entire waterbody, but NOT area-weighted like other indicators.

Current year aquatic grasses coverage is available as a GIS map from the VIMS group. A data map of where aquatic grass beds are located in your system is a good communication tool, especially in conjunction with water quality data maps.

Chapter 10: Synthesizing and communicating data

The previous chapters discussed in detail how to measure and analyze the core indicators that will be synthesized into a report card. To synthesize data is to combine and integrate large amounts of data into a single entity that generates meaningful information. Specifically, in the case of this protocol, it means to score a tributary and to give it a grade that is incorporated into a report card. Synthesizing data into one score for each indicator is an important step in answering the question, "How healthy is the tributary?" The audience does not need to see each measurement that goes into a year-long monitoring program's database. Rather, they need to know the ultimate outcome of those measurements, or "What do the data mean?".

One way to synthesize data is to roll up individual indicators into an overarching index. An index can combine similar types of indicators (e.g., chemical, physical, biological) into one index, or it can be an average of all measured indicators (Figure 10.1). Overarching indices give a much better integrated assessment (and therefore representative score) of an ecosystem's health than can be achieved using a single indicator. Additionally, comparing indices between different tributaries negates the need to resolve varying temporal and spatial sampling scales.

How to synthesize

Each of the six core indicators used in this protocol can be averaged together for a health score for a sub-region. Then, scores for each sub-region are area-weighted (i.e., the area of the sub-region divided by the total area of the tributary) and averaged for one tributary score. Each monitoring program will need to decide if it wants to provide sub-region scores, or if it wants to average all individual indicators into one health index for the tributary.

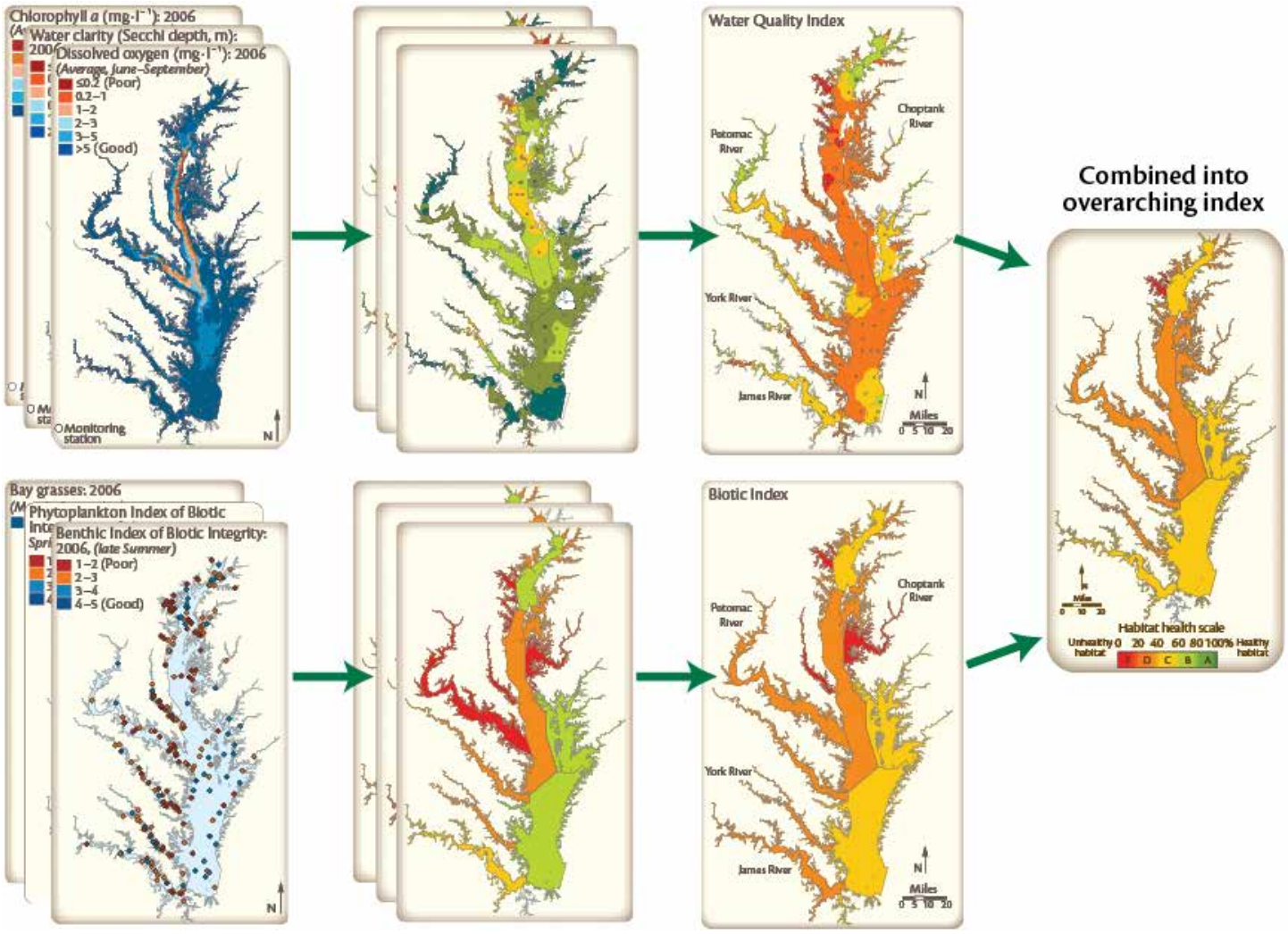


Figure 10.1. The Chesapeake Bay-wide report card: three water quality indicators and three biotic indicators are evaluated against threshold values. The water quality indicators are then averaged into a Water Quality Index, and the biotic indicators are averaged into a Biotic Index. These two sub-indices are then averaged into an overall Health Index and given a grade.

Once the core indicators have been analyzed against thresholds and averaged into an index, this index becomes the report card grade. Each reporting region has its own report card grade, and the entire tributary receives a report card grade, as well (see the grading scale in Chapter 4). For the Chesapeake Bay-wide report card, the main grade for the overall Bay and the individual sub-regions are communicated via a printed report card and a press release. Additionally, all supporting indicator and region scores and indicator and region maps are provided on a website.

To weight or not to weight

There are advantages and disadvantages of weighting indicators that depend upon whether you have chosen to use targets or relative ranking as your approach for measuring success or failure. For this protocol document, indicators will be weighted evenly. This means that each indicator is as important as all others when averaged into a health index score.

Selecting reporting regions

Sub-regions of your system may have already been determined to help clarify where to assign sampling sites (see Chapter 2). However, if they have not already been defined, it is one of the first tasks in developing a report card. There must be a sufficient number of sampling sites in a reporting region to provide a representative and accurate score for each indicator. Although there are no specific rules to follow when defining the boundaries, some considerations include the number of sub-regions (so the audience is not overwhelmed with too much detail) and alignment of regions with existing management and/or geophysical boundaries (e.g., counties, preservation areas or depth, salinity regimes).

Report cards

Report cards have become a popular and effective tool for promoting numerous issues, ranging from bacteria levels at swimming beaches to the ecosystem health of freshwater streams, because they have proven to be important outreach tools for generating community interest and increasing citizen understanding of ecosystem health, water quality, and watershed issues (Figure 10.2). Typically, after a report card has been released, awareness and responsiveness to a particular issue increases substantially, leading to a change in community and political knowledge and will. For example, the 2008 Chester River report card highlighted the impact of old septic systems on groundwater nutrients and promoted a free state program for system upgrades. This led to a increase in applications by citizens for new systems.

Communication strategy

A well-rounded communications strategy outlines key messages (i.e., what one wants to convey), identifies target audiences (i.e., with whom one wants to communicate), helps choose a spokesperson, and determines communication vehicles (i.e., the documents or techniques through which one communicates). At the same time communication products are being determined, the content of those products should also be decided.

The report card itself can be a printed product, such as a 4-page newsletter or double-sided trifold, or it can be produced as webpages on your organization's website. Often, the suite of communication products are determined at the beginning of a monitoring project during the proposal stage, so make sure that sufficient time and resources are allotted to complete the products to which the proposal will commit. Each communication product engages a different audience and requires different time commitments.

A website is now considered an essential science communication tool. It allows the widest possible audience to be reached in the most timely manner, without the normal delays of print media. The constant ability to edit and refine a website is one of the key features that makes them effective for science communication. However, this

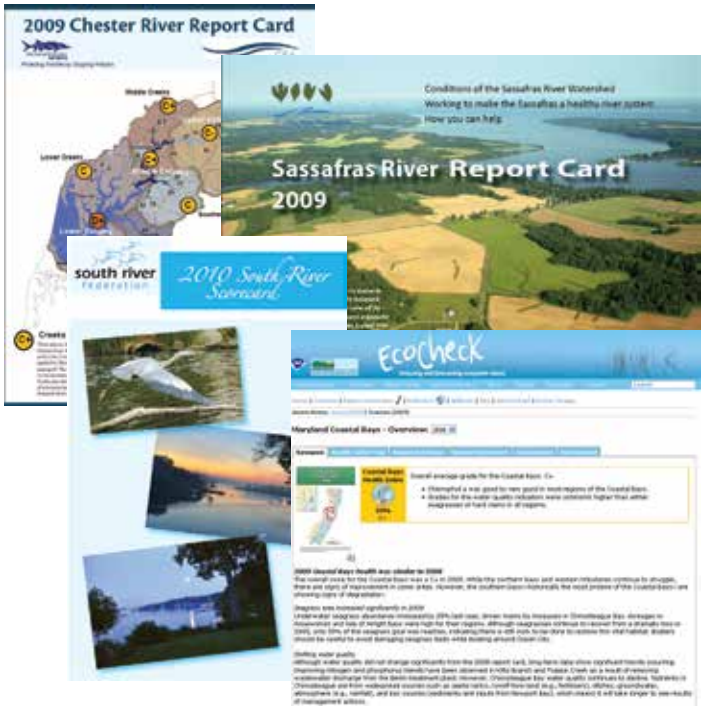


Figure 10.2. Examples of different report card products. Top to bottom: 2009 Chester River report card (4-page newsletter), 2009 Sassafras River report card (8-page brochure), 2010 South River scorecard (16-page pamphlet), and 2009 Maryland Coastal Bays website.

can also be a trap, because it is often too easy to publish something that is not well-designed, thinking that it can always be fixed later. The reality is often quite different, and as a result, the website can become a jumble of disjointed pages with a poorly designed structure and navigation system. Like other media, websites should follow the principles of effective science communication—they should be visually appealing and cleanly laid out with the right balance of meaningful graphics and informative text and also have a consistent look and feel. Some key features of an effective website are a clear and consistent navigation system and obvious hyperlinks. Above all, do not get too fancy—bells and whistles will not make up for poor content.

The high profile and sometimes controversial nature of report cards necessitates special attention to the communication strategy. A communication strategy needs to consider the main messages that the report card will deliver, how to best deliver the message, and how to reach a broad audience. In terms of messaging, a report card provides an opportunity to communicate the overall health of a region, how one region compares to another, and how health may have changed from one year to another. The report card also provides a vehicle to communicate other related messages such as restoration efforts being undertaken in the area or how the audience can become involved and help in restoration activities. Before releasing a report card, it is advisable to brief appropriate people and agencies about what the report card scores will be (with an embargo on their release until the chosen release date) so that they have the opportunity to prepare appropriate responses.

All of these products—a printed report card, website, and a general communication strategy—have varying amounts of time and effort associated with them. Discussion of these time constraints are beyond the scope of this protocol, but a thorough explanation of different communication products, time commitments, and audiences is provided in Longstaff et al. 2010.

Conclusions

Need for standardization

Ecosystem health report cards are proven outreach tools for engaging and educating citizens, stakeholders, managers, and elected officials about the health of their ecosystem. Many organizations in the Mid-Atlantic region have recognized the power of such report cards, and have begun to produce them on an annual basis. However, so far each organization has been producing a report card using indicators, data collection procedures, and analysis methods that are unique to their own monitoring program. This presents several potential issues and problems:

- Results from different report cards cannot be assessed in relation to each other. If each organization is using different indicators and methods, the results are not comparable. This limits the utility of report cards to present a larger picture of ecosystem health in a region.
- Data collected by individual organizations and/or citizen volunteers may not always adhere to rigorous scientific standards for quality control. Products derived from these data may be then viewed as less reliable and therefore not taken seriously.
- Data may also fail to be integrated into larger analyses or used in criteria assessments or management programs if the quality of the data are suspect.

Using this protocol to build scientific and public knowledge via report cards

The project that supported the development of this protocol was intended to alleviate some of the aforementioned concerns by developing standards for quality control, data collection, and analysis that would enhance the overall quality and utility of data collected and the resulting products produced (e.g., report cards). This document is intended to provide guidance for organizations as they develop monitoring programs and consider producing report cards. This protocol was designed to achieve two main objectives:

- Enhance the ability of organizations to produce effective ecosystem health report cards for tributaries in the Mid-Atlantic region, and
- Increase the utility of data collected by these organizations through standardization of data collection and analysis methods.

The availability of consistent, high quality data is the cornerstone of any assessment project, and one that cannot be achieved without rigorous and consistent guidelines for program design, quality control, and sampling and data analysis methodology. For report card style assessments, which are produced on an annual or semi-annual basis, it is especially important to have long-term consistency in the way data are collected and analyzed, and results presented.

This document represents the first attempt to develop consistent methods for these issues in the report card framework. By standardizing indicators and monitoring protocols, the scientific validity of the data collected will also be strengthened, thereby increasing the ability of groups to successfully reach and influence their audiences. Additionally, the overall utility of the data collected by individual groups will be enhanced by allowing direct comparison of results among regions.

Many organizations have contributed to this protocol document that already had sampling and data analysis procedures in place before this protocol was developed. Therefore, all current groups in MTAC do not necessarily follow every single guideline recommended here. The hope is that in time, all organizations will be able to adjust their monitoring and analysis procedures to be in keeping with the guidelines.

There are many critical elements to developing a report card as a successful communication tool. This protocol addresses many of these critical elements, but obviously cannot address all eventualities and scenarios. The editors and contributors sincerely hope that the guidelines presented in this document will help new organizations as they design their sampling programs and reporting frameworks.

References and further reading

Entire protocol

Alliance for the Chesapeake Bay. (2007). *RiverTrends Volunteer Water Quality Monitoring Program Manual.* Richmond, VA.

Burgan, B., Carpenter, D., Gould, D., Keeler, B., McGovern, C., & Miller, S. (2007). *Indicator Development for Estuaries.* (EPA Publication No. 842-B-07-004). Washington, D.C.: Government Printing Office.

Dunnette, D.A. (1979). *A geographically variable water quality index used in Oregon.* Journal of Water Pollution Control Federation, 51 (1), 543–61.

Integration and Application Network, Maryland Coastal Bays Program, Assateague National Seashore, & Maryland Department of Natural Resources. (2009). *Development of a spatially-explicit health index for the Coastal Bays of Maryland and Virginia.* Retrieved March 12, 2011, from www.eco-check.org/reportcard/mcb/2009/methods

Jackson, L.E., Kurtz, J.C., & Fisher, W.S. (Eds.). (2000). *Evaluation Guidelines for Ecological Indicators.* (EPA Publication No. 620-R-99-005). Research Triangle Park, NC: U.S. Government Printing Office. Retrieved February 18, 2011, from www.epa.gov/emap/html/pubs/docs/resdocs/ecol_ind.pdf

Jorgensen, S.E., Costanza, R., & Xu, F.L. (2006). Handbook of *Ecological Indicators for Assessment of Ecosystem Health.* Danvers, MA: CRC Press.

Longstaff, B.J., Carruthers, T.J.B., Dennison, W.C., Lookingbill, T.R., Hawkey, J.M., Thomas, J.E., et al. (Eds.). (2010). *Integrating and Applying Science: A Practical Handbook for Effective Coastal Ecosystem Assessment.* Cambridge, MD: IAN Press.

Maryland Department of Natural Resources. (2010). *Quality Assurance Project Plan for the Maryland Department of Natural Resources Chesapeake Bay Shallow Water Quality Monitoring Program. Annapolis, MD.* Retrieved February 18, 2011, from mddnr.chesapeakebay.net/eyesonthebay/documents/SWM_QAPP_2010_2011_FINALDraft1.pdf

Maryland Department of Natural Resources. (2010, May 27). *Quality Assurance Project Plan for the Maryland Department of Natural Resources Chesapeake Bay Water Quality Monitoring Program – Chemical and Physical Properties Component for the Period July 1, 2010–June 30, 2011. Annapolis, MD.* Retrived February 18, 2011, from mddnr.chesapeakebay.net/eyesonthebay/documents/QAPP_MainTrib_2010-2011_Draft1.pdf

Massachusetts Water Watch Partnership. (1994). *Volunteer Water Quality Monitoring Manual.* Amherst, MA.

McKenzie, D.H., Hyatt, D.E., & McDonald, V.J. (1992). *Ecological indicators: Volumes 1 and 2. Proceedings of the International Symposium on Ecological Indicators.* Essex, England: Elsevier Science Publishers.

Ott, W. R. (1978). *Environmental Indices: Theory and Practice.* Ann Arbor, MI: Ann Arbor Science Publishers.

Thomas, J.E., Saxby, T.A., Jones, A.B., Carruthers, T.J.B., Abal, E.G., & Dennison, W.C. (2006). *Communicating Science Effectively: A Practical Handbook for Integrating Visual Elements.* London, England: IWA Publishing.

U.S. Environmental Protection Agency. (1991). *Volunteer Lake Monitoring: A Methods Manual (2nd ed.).* (EPA Publication No. 440-4-91-002). Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency. (2003). *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and its Tributaries.* (EPA Publication No. 903-R-03-002). Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency. (2006). *Volunteer Estuary Monitoring: A Methods Manual (2nd ed.).* (EPA Publication No. 842-B-06-003). Washington, DC: U.S. Government Printing Office. Retrieved February 18, 2011, from water.epa.gov/type/oceb/nep/monitor_index.cfm

Wazniak, C.E., Hall, M.R., Carruthers, T.J.B., Sturgis, B., Dennison, W.C., & Orth, R.J. (2007). *Linking water quality to living resources in a mid-Atlantic lagoon system, USA.* Ecological Applications, 17 (5), S64–S78.

Williams, M., Longstaff, B., Buchanan, C., Llanso, R., Dennison, W. (2009). *Development and evaluation of a spatially-explicit index of Chesapeake Bay health*. Marine Pollution Bulletin, 59, 14–25.

Web resources

Chesapeake EcoCheck: www.eco-check.org
Chesapeake Bay Program: www.chesapeakebay.net
Chesapeake Bay Trust: www.cbtrust.org
NOAA: www.noaa.gov
UMCES: www.umces.edu

Chapter 1

Integration and Application Network, Maryland Coastal Bays Program, Assateague National Seashore, & Maryland Department of Natural Resources. (2009). *Development of a spatially-explicit health index for the Coastal Bays of Maryland and Virginia.* Retrieved March 12, 2011, from www.eco-check.org/reportcard/mcb/2009/methods

Wazniak, C.E., Hall, M.R., Carruthers, T.J.B., Sturgis, B., Dennison, W.C., & Orth, R.J. (2007). *Linking water quality to living resources in a mid-Atlantic lagoon system, USA.* Ecological Applications, 17 (5), S64–S78.

Chapter 2

Alliance for the Chesapeake Bay. (2002). *Volunteerism and Watershed Stewardship.* Annapolis, MD.

Chapter 3

Chesapeake Bay Program. (2011, January 5). *Quality Assurance Planning.* Retrieved February 15, 2011, from www.chesapeakebay.net/qualityassurance_planning.aspx

Nanticoke Watershed Alliance. (2009). *Quality Assurance Project Plan, Appendix I.* Vienna, MD.

U.S. Environmental Protection Agency (EPA). (2001). *EPA Requirements for Quality Management Plans.* (EPA Publication No. 240-B-01-002). Washington, DC: U.S. Government Printing Office. Retrieved March 12, 2001, from www.epa.gov/quality1/qs-docs/r2-final.pdf

U.S. Environmental Protection Agency (EPA). (2001). *EPA Requirements for Quality Assurance Project Plans.* (EPA Publication No. 240-B-01-003). Washington, DC: U.S. Government Printing Office. Retrieved March 12, 2011, from www.epa.gov/quality1/qs-docs/r5-final.pdf

Database options
Microsoft Access: office.microsoft.com/en-us/access/
Microsoft Excel: office.microsoft.com/en-us/excel/
SQLite: www.sqlite.org

Chapter 4

Wazniak, C.E., Hall, M.R., Carruthers, T.J.B., Sturgis, B., Dennison, W.C., & Orth, R.J. (2007). *Linking water quality to living resources in a mid-Atlantic lagoon system, USA*. Ecological Applications, 17 (5), S64–S78.

Chapter 5

U.S. Environmental Protection Agency. (2003). *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and its Tributaries.* (EPA Publication No. 903-R-03-002). Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency. (2004). *Technical Support Document for Identification of Chesapeake Bay Designated Uses and Attainability, 2004 Addendum.* Washington, D.C: U.S. Governmental Printing Office.

Chapter 6

Lane, L., Rhoades, S., Thomas, C., and Van Heukelem, L. (2000). *Analytical Services Laboratory standard operating procedures.* (Technical report No. TS-264-00). Cambridge, MD.

Maryland Department of Natural Resources. (2010). *Quality Assurance Project Plan for the Maryland Department of Natural Resources Chesapeake Bay Shallow Water Quality Monitoring Program. Annapolis, MD.* Retrieved February 18, 2011, from mddnr.chesapeakebay.net/eyesonthebay/documents/SWM_QAPP_2010_2011_FINALDraft1.pdf

Horn Point Laboratory. (2010). *Chlorophyll* a *Sampling Requirements for Submission to HPL Analytical Services for Analysis.* Cambridge, MD.

U.S. Environmental Protection Agency. (1997). *In Vitro Determination of Chlorophyll* a *and Pheophytin a in Marine and Freshwater Algae by Fluorescence.* Cincinnati, OH: National Exposure Research Laboratory, Office of Research and Development.

Chapter 7

Davies-Colley, R.J., Vant, W.N., & Smith, D.G. (1993). *Colour and Clarity of Natural Waters.* New York: Ellis Horwood.

Kent State University. (2011). *SECCHI Dipin.* Retrieved on April 6, 2011 from http://www.secchidipin.org/Wisconsin%20Method.htm.

U.S. Environmental Protection Agency. (2011). *Using a Secchi disk or Transparency Tube.* Retrieved on April 8, 2011 from http://water.epa.gov/type/rsl/monitoring/155.cfm

Chapter 8

Lane, L., Rhoades, S., Thomas, C., & Van Heukelem, L. (2000). *Analytical Services Laboratory standard operating procedures.* (Technical report No. TS-264-00). Cambridge, MD.

Sassafras River Association. (2010). *Sassafras Samplers Monitoring Program Standard Operating Procedures.* Georgetown, MD.

Sassafras River Association. (2010). *Tidal Monitoring Standard Operating Procedures.* Georgetown, MD.

U.S. Environmental Protection Agency. (2010). *Chesapeake Bay TMDL.* Retrieved March 3, 2011, from www.epa.gov/chesapeakebaytmdl/

Chapter 9

Moore, K.A., Wilcox, D.J., Anderson, B., Parham, T.A., & Naylor, M.D. (2004). *Historical Analysis of Submerged Aquatic Vegetation (SAV) in the Potomac River and Analysis of Bay-wide SAV Data to Establish a New Acreage Goal Final Report.* (Chesapeake Bay Program Publication No. 983627-01).

Short, F.T., McKenzie, L.J., Coles, R.G., Vidler, K.P., & Gaeckle, J.L. (2006). *SeagrassNet Manual for Scientific Monitoring of Seagrass Habitat, Worldwide edition.* University of New Hampshire Publication.

U.S. Environmental Protection Agency. (2003). *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and its Tributaries.* (EPA Publication No. 903-R-03-002). Washington, DC: U.S. Government Printing Office.

U.S. Environmental Protection Agency. (2007). *Ambient Water Quality Criteria for Dissolved Oxygen, Water Clarity and Chlorophyll a for the Chesapeake Bay and Its Tidal Tributaries, 2007 Addendum.* Retrieved March 12, 2011, from www.chesapeakebay.net/content/publications/cbp_27849.pdf

Virginia Institute of Marine Science. (2010). *Submerged aquatic Vegetation in Chesapeake Bay.* Retrieved March 12, 2011, from www.vims.edu/bio/sav

Addendum I: Dissolved oxygen instruments

There are a variety of considerations for instrumentation required for dissolved oxygen monitoring, Cost, ease of use, repeatability, stability, maintenance, and durability are all important considerations for choosing the most appropriate instrument for a monitoring program. Instrument recommendations are not intended to be endorsements, but rather suggestions based on previous users' experiences.

Recommended instruments

Some considerations when choosing a dissolved oxygen meter are:

- A meter used in estuaries for profiles must include temperature and salinity. (This excludes DO meters such as the YSI 55, 550A, DO200, and Pro20 that lack salinity readings.)
- Other parameters included (some lack pH).
- Probe size (some are quite large and hard to carry).
- Cord length (depends on deepest sampling site).
- Reliability.
- Stability/reproducibility.
- Ease of use, including calibration and DO membrane replacement.
- Cost.

The pros and cons of an entry level meter (YSI 85, about $1200) are:

- Pros: one of the cheapest dependable meters; easy to calibrate and replace DO membranes.
- Cons: No pH (can use separate pocket meter); calibrates for altitude only, not barometric pressure; does not measure depth, so marked line and weight need to be attached (see Figure 4.6); hard to view results (have to scroll through them in sequence); may have to move probe during sampling to get accurate DO readings.

The pros and cons of an intermediate level meter (YSI Professional Plus, about $2500):

- Pros: Interchangeable probes, space to add additional probes, two choices for DO probe (polarographic—the usual probe, or galvanic ; both use cap DO membranes which are easy to change), military grade connectors, offers 2-point DO calibration, has barometer, shows all results in one screen.
- Cons: DO probe still needs to be moved during sampling; no depth sensor so a marked line needs to be attached.

For the expected users of this protocol, no high level meter is recommended because the costs are outside the range typically allotted for equipment. Also, the basic information needed for the purposes of this protocol can be obtained using the entry and intermediate level recommendations.

Peter Bergstrom

A common YSI instrument. The computer with display is held in one hand, while the probe (black cord) is dipped into the water. The weights help the probe stay vertical in the water column, rather than being pulled by the current or tide.

Peter Bergstrom

YSI 85 probe with weights: the weights are attached below the probe to keep the probe from hitting the mud, but depth measurements start from the probe.

Calibration and maintenance

In general, calibration frequency is determined by the user and demands of a monitoring program. The calibration process should be performed as recommended by the manufacturer.

DO meter comparisons

Monitoring programs should hold regular meter comparison sessions, which could compare all meters used by the same program, or by different programs in the same area. All of the other parameters measured by the meters should be compared at the same time. Comparisons should be done at least once a year to see if any of the meters need service (e.g., manufacturer calibration or replacement of specific probes).

It is easiest to do this from a boat that holds all of the people involved, although this means the probes can not all be next to each other. Most piers are too small to do a large meter comparison from them. It is best to do the comparisons on a regular monitoring cruise at the regular sampling sites.

Methods

All of the readings should be taken at the same site at as close to the same time as possible. Readings should be compared from at least 6–8 different sites with a range of DO concentrations at two or more depths per site.

When comparing the results, look for any meters that had consistently higher or lower results than the others, and calculate the coefficient of variation (CV, the standard deviation divided by the mean) for all of the results from each site and sampling depth. The CV becomes unreliable when the means are very small, but this is usually not an issue for most parameters measured in estuarine systems.

There is no fixed problem level for CV, and CV will always be higher for some parameters than others, but it's best if it is 10% or less. For example, in one comparison of 3–4 meters at the same site over several days, the CV for salinity and temperature ranged from 1–5%, clearly within an acceptable range. The CV for DO and turbidity was higher (2–33%) and the CV for chlorophyll was even higher (14–137%). For DO, readings from one meter were consistently lower than the others, so more investigation, ideally using methods other than meters (e.g., Winkler titration) should be explored.

Peter Bergstrom

YSI Pro Plus meter, shown with probe cover for sampling in place and storage cover in upper right corner.

Maintenance procedures

When DO will not calibrate, or the readings are erratic or out of normal ranges, that usually means that the membrane needs to be replaced. If your meter uses a cap membrane, it is easy to replace, but if it uses a sheet membrane held in place with an O-ring, it can be challenging to replace it properly. You may need two people (one to stretch the membrane and one to roll down the O-ring) and a clamp to hold the probe while you do this. If the membrane does not seal properly, electrolyte will leak around the edges and the conductivity readings will jump around during calibration. If this occurs you need to repeat the membrane replacement, trying hard to avoid any leaks.

References

YSI DO Measurement FAQ: www.ysi.com/media/pdfs/FAQs-Dissolved-Oxygen-Measurement.pdf

Addendum II: Pycnocline calculations

Methods used by Chesapeake Bay Program as of March 2011

Some areas within a river system, such as a navigational channel, are expected to have frequent, summer water column stratification. In estuaries, stratification is based on water density and is a naturally-occurring phenomenon that may be exacerbated by eutrophication effects (see Chapter 5: Dissolved oxygen).

Temperature and salinity are used to calculate density, which in turn is used to calculate pycnocline (i.e., change in density) boundaries. For each column of temperature and salinity data, the existence of the upper and lower pycnocline boundaries is determined by looking for the shallowest robust vertical change in density of 0.1 $kg \cdot m^{-3} \cdot m^{-1}$ for the upper boundary and deepest change of 0.2 $kg \cdot m^{-3} \cdot m^{-1}$ for the lower boundary. *To be considered robust, the density gradient must not reverse direction at the next measurement and must be accompanied by a change in salinity, not just temperature.* The following steps are used to calculate the pycnocline boundaries for rivers within the Chesapeake Bay using measurements of water temperature and salinity.

Pycnocline calculations require a vertical profile of salinity and water temperature measurements at multiple depths

1) Sort the vertical profile of data from the surface downwards.
2) For each depth at which there are measurements, calculate a water density value as σT, or "sigma T", using water temperature and salinity measurements for that depth. Use the following method and equations:
   - $\sigma T = a(T) + b(T)*S$, where
   - T = temperature (°C); S = salinity; a & b are polynomial functions of T:
   - $a(T) = -9.22 \times 10^{-3} + 5.59 \times 10^{-2}*T - 7.88 \times 10^{-3}*T^2 + 4.18 \times 10^{-5}*T^3$
   - $b(T) = 8.04 \times 10^{-1} - 2.92 \times 10^{-3}*T + 3.12 \times 10^{-5}*T^2$
3) Look down through the profile. Wherever the difference between sequential depth measurements is <0.19 meters, average the two depth measurements and their corresponding salinity and density measurements.
4) Look down through the profile again. If there are still any depths (depth, salinity, temperature, and density measurements) <0.19 meters apart, then average them again. Continue until there are no depths <0.19 meters apart.

Calculation of change in salinity and density

1) Starting at the surface and continuing until the deepest measurement in the profile, calculate the change in salinity and density between each sampling depth. For example, for two density values at one meter depth (y_1) and two meters depth (y_2) respectively, change in density, or $\Delta\sigma T = y_2 - y_1$. Likewise, for salinity measurements $\Delta S = y_2 - y_1$.
2) Assign a depth measurement to each pair of Δ values (ΔS, ΔσT) equal to the average of two depths used to calculate the Δ values. Thus, for the two measurements y_2 and y_1, calculate accompanying depth as $(x_1 + x_2)/2$. You should now have a vertical profile of ΔS and ΔσT values with an accompanying depth.
3) To find the upper boundary of the pycnocline, look at the vertical profile of ΔσT, beginning with the 2nd value (from the surface), and excluding the two deepest values:
   IF ΔσT >0.1,
   AND IF ΔσT for the next depth is greater than zero,
   AND IF ΔS > 0.1,
   Then this depth represents the upper boundary of the pycnocline.

Determination of lower mixed layer

1) Identify whether there is a lower mixed layer: use the same vertical profile but examine it from the 2nd deepest value upward (exclude deepest value):
   IF change in density (ΔσT) at the 2nd deepest depth <0.2,
   OR IF ΔσT at the next depth (moving upwards, i.e., shallower) <0.2,
   THEN a lower mixed layer (i.e., a layer at depth where the density is not changing) below the pycnocline exists.
2) If a lower mixed layer exists, then look for the lower boundary of the pycnocline. Beginning at the 2nd deepest value, and moving up one row at the time to the depth immediately below the upper pycnocline boundary, for ΔS and ΔσT values at each depth:
   IF ΔσT > 0.2,
   AND IF ΔS > 0.1,
   Then this depth is the lower pycnocline boundary.
3) If a pycnocline exists, then the upper and lower (if present) boundaries of the pycnocline have now been identified.

Now that the upper and lower boundaries of the pycnocline have been identified, the relevant threshold (5, 3, or 1 $mg \cdot l^{-1}$) can be assigned to each depth measurement. Then, each measured value is compared to the associated threshold, and scored as passing or failing. See Chapter 5: Dissolved oxygen for the rest of the calculations.

Addendum III: Alternate thresholds for water clarity

Although extensive effort by the members of MTAC was made to reach consensus on all indicators and analysis methods, consensus on the use of one consistent set of water clarity (Secchi depth) data analysis thresholds was not achieved. The issue arises because there are two scientifically valid approaches to determining the thresholds, and these approaches result in different threshold values, especially in Chesapeake Bay mesohaline areas.

The method recommended in Chapter 7 uses thresholds that are established by evaluation of Secchi depth in relation to phytoplankton communities (Phytoplankton Index of Biological Integrity, or PIBI) in good condition (i.e., reference communities). Some groups use a set of thresholds that are derived from submerged aquatic vegetation (SAV) light requirements. Although there are arguments for and against both methods, only one method could be recommended in this protocol so that consistency within new report card projects could be maintained.

Arguments for using PIBI-based thresholds include (not comprehensive and in no particular order of importance):

- Thresholds should be ecologically derived and independent of regulatory goals. Regulatory requirements periodically change, creating the potential need to change thresholds based on them. Additionally, achieving regulatory-based thresholds may still not result in desired ecological condition. The PIBI-based thresholds are based on ecologically relevant reference communities, not regulatory goals.
- PIBI reference community condition is more strongly correlated with Secchi depth than SAV area is.
- Water clarity is an important water quality metric independent of other metrics.
- PIBI-based thresholds are less achievable given current conditions, but many feel that this better reflects the reality of water clarity problems in Chesapeake Bay.
- To evaluate the applicability of PIBI-based thresholds in tributary areas, extensive data evaluation was performed to examine score distributions from tributary data, including data from watershed organizations participating in MTAC. Data density distributions were similar for open water sites.
- SAV is measured directly by remote sensing, and using clarity thresholds based on SAV habitat is essentially over-weighting the SAV metric.

Arguments for SAV-based thresholds include (also not comprehensive and in no particular order of importance):

- Light penetration in the water column is an essential component of SAV habitat, and light penetration is directly related to water clarity as measured by Secchi disk.
- Some regulatory requirements are based on water clarity measures that are derived from light requirements for SAV.
- Several established groups are currently using the SAV-based thresholds; changes to methods would result in scoring schemes that are inconsistent from previous years, thereby creating confusion in communication products.
- Secchi depths derived from SAV-based thresholds are more easily achieved and result in scores that are not always failing, which is an important consideration for keeping the public engaged.
- PIBI-based Secchi depth thresholds are not as applicable in tributaries, where water clarity problems are more pronounced. PIBI-based thresholds were derived using data from deep, open water sampling locations and so are perhaps less useful in tributary settings.

Ultimately, the PIBI-based thresholds were chosen as the recommended thresholds because the scientific merits of the issues under consideration had been reviewed previously during the establishment of the Chesapeake Bay Report Card methodology. The resulting recommendation at that time was to use the PIBI-based water clarity thresholds. In the editors' judgment, there was not enough new information to justify a change in water clarity threshold recommendations for tributary report cards.

Moreover, a key MTAC goal is to enhance the utility of data collected by member watershed organizations. MTAC members and volunteers ultimately desire that their data be widely used, including in scientific applications and criteria assessments. To enhance data utility, it is important that the scientific community feel that the data and analysis methods are rigorous. As the issue of water clarity thresholds has been previously reviewed by the scientific community, and the resulting consensus was to suggest use of the PIBI-based thresholds, the editors felt that using the PIBI-based methods would result in greater confidence about the data and analysis methods.

SAV-based multiple threshold values are presented in the following table. There is one set for tidal fresh and oligohaline salinity regimes, and one set for mesohaline and polyhaline, for those groups that choose to use them. The score calculations are the same as discussed in Chapter 7.

Alternative threshold values using SAV-based measurements.

Score	Tidal Fresh and Oligohaline	Mesohaline and Polyhaline
5	≥1.2	≥1.8
4	≥0.7–<1.2	≥1.0–<1.8
3	≥0.4–<0.7	≥0.7–<1.0
2	≥0.3–<0.4	≥0.5–<0.7
1	≥0.2–<0.3	≥0.3–<0.5
0	<0.2	<0.3