



DATA INTERPRETATION AND SYNTHESIS METHODS MANUAL

CMC
Chesapeake Monitoring
Cooperative





Produced by the Chesapeake Monitoring Cooperative

Working together to understand the health of our waters

January 30, 2018

The development of this manual was supported through a cooperative agreement with the EPA Chesapeake Bay Program (#CB-963349).

Contributors

This document was created by the Integration and Application Network at the University of Maryland Center for Environmental Science. The authors want to thank Peter Tango for his expert review of these methods.

Integration and Application Network

UMCES

429 Fourth Street

Annapolis, MD 21403

chesapeakemonitoringcoop.org/

Table of contents

Explaining what your data means	7
Kinds of data	9
Accuracy and precision	11
Cleaning your data	12
Descriptive statistics	15
Data in tables	19
Data in graphs	21
Data in figures	29
Data interpretation sets the stage for synthesis	33
Visualization and storytelling help synthesize data into information	35
Data visualization	36
Storytelling	44
Report cards	51
Tidal Indicators	52
Nontidal indicators	75
Cleaning your data to analyze in a report card	97
Communicating science effectively	111
Resources and references	112

Introduction

We rely heavily on the Chesapeake Bay and all its tributaries for our drinking water, food sources, recreation, and navigation. Since the initiation of the Chesapeake Bay Program in 1983 the communities in the watershed have been working towards improving the health of these waters. A significant portion of that work is tracking our progress through water quality and benthic macroinvertebrate monitoring. There are many sources of water quality and benthic macroinvertebrate data, including data collected by volunteers, local governments, conservation districts, and nongovernmental groups such as academia and watershed organizations that are not currently being used by the Chesapeake Bay Program to track Bay health.

The Alliance for the Chesapeake Bay, Izaak Walton League of America, Dickinson College's Alliance for Aquatic Resource Monitoring, and the University of Maryland Center for Environmental Science's Integration and Application Network, have partnered to create the Chesapeake Monitoring Cooperative (CMC). The CMC will provide technical, logistical, and outreach support for the integration of citizen-based and non-traditional water quality and macroinvertebrate monitoring data into the Chesapeake Bay Program (CBP) partnership.



This is the first effort to integrate citizen science water quality and benthic macroinvertebrate data into a federal program in order to inform policy management and water quality assessments. The contributions of data by volunteer and non-traditional monitoring groups to the CMC and CBP monitoring network will provide valuable information that supports shared decision-making, adaptive management, and measuring progress towards the 2014 Chesapeake Bay Watershed Agreement. Additionally, providing technical support to volunteer and non-traditional monitoring groups is a cornerstone of the CMC.

One way in which the CMC is providing support is through training opportunities in data interpretation and science communication. Data interpretation—adding context and meaning to data—is a critical step after data collection and is oftentimes overlooked by organizations when determining their staff skill needs. Science communication—synthesizing results into readable, easy-to-understand products, such as report cards—requires more training and analysis than data interpretation.

Purpose of this manual

This manual hopes to provide an overview and examples of data interpretation and science communication for nontraditional monitoring groups. This manual is one tool that will help take raw water quality and benthic macroinvertebrate data and transform it into meaningful and useful

information for all audiences. Other tools in the toolbox include hands-on training in data interpretation techniques, graphic design, and supported software as well as reference materials and extended training opportunities.

This manual was developed to support the wide variety of monitoring groups throughout the Chesapeake Bay watershed and the communication needs that they have. The manual is intended to support the work of the Chesapeake Monitoring Cooperative and supplies methods for monitoring groups to use after they have collected data and are moving on to the analysis and synthesis stage of communication. The manual is not meant to be a comprehensive and detailed instruction manual on how to *analyze* and *present* data collected by monitoring programs. It is meant to introduce useful ways monitoring groups can integrate data they have collected, and then help them to synthesize that data into clear products such as report cards and other visual reports, for personal understanding and public consumption.



Participants in UMCES-IAN Data Interpretation workshops will use this manual to help interpret and analyze water quality and benthic macroinvertebrate data. Participants will work both together as team members and as individuals on their own data.



UMCES-IAN workshops include hands-on activities that are used to describe and understand data, provide context to that data, and build communication products.

Supporting materials

The following books and websites are critical supporting material for what is covered in this manual. Using these reference materials while interpreting and synthesizing your data will ensure high quality analyses that convey meaningful information about your data.

Books

Longstaff, BJ, Carruthers TJB, Dennison WC, Lookingbill TR, Hawkey JM, Thomas JE, Wicks EC, Woerner J. eds. 2010. *Integrating and applying science: A practical handbook for effective coastal ecosystem assessment*. IAN Press, Cambridge, Maryland, USA. 244pp. http://ian.umces.edu/press/books/publication/259/integrating_and_applying_science_a_handbook_for_effective_coastal_ecosystem_assessment_2010-05-10/

Thomas, JE, Saxby TA, Jones AB, Carruthers TJB, Abal EG, Dennison WC. 2006. *Communicating science effectively: A practical handbook for integrating visual elements*. IWA Publishing, London, England. 132pp. Free PDF download: <https://explorer.ian.umces.edu/public/7604ed>. Buy online: http://ian.umces.edu/press/books/publication/65/communicating_science_effectively_a_practical_handbook_for_integrating_visual_elements_2006-07-01/

Townend, J. 2002. *Practical statistics for environmental and biological scientists*. John Wiley & Sons, Ltd. West Sussex, England. 276pp.

Tufte, E. 1997-2006. A variety of books on visualizing science. <https://www.edwardtufte.com/tufte/>

Valiela, I. 2001. *Doing Science: Design, Analysis, and Communication of Scientific Research*. Oxford University Press, New York, New York. 294pp.

Websites

Science Visualization semester-long course <https://sciencevisualizationclass.wordpress.com/>

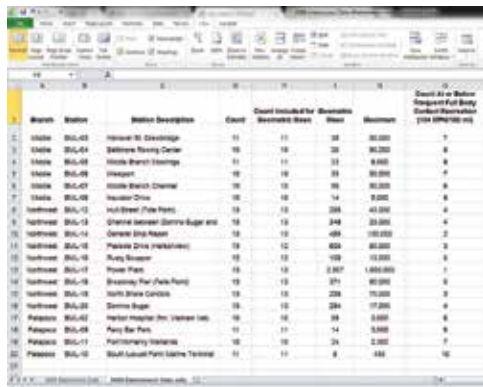
INTERPRETATION

EXPLAINING WHAT YOUR DATA MEANS

For our purposes, data interpretation is defined as the process by which you evaluate and analyze your data so that it can be communicated in a meaningful way to your selected audience. The types of analyses you will do to interpret your data is based on the sampling methods used during field collection. A critical assumption before using this manual is that your sampling methods, whether for water quality or benthic macroinvertebrate parameters, are appropriate to your overall monitoring objectives and are supported by a study design process.

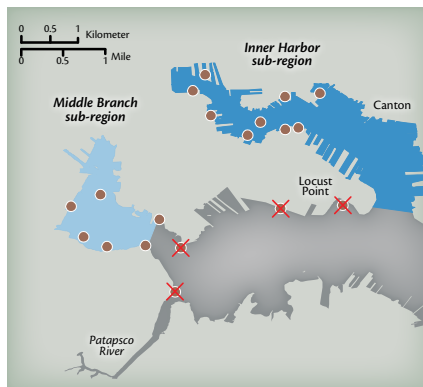
Data interpretation is a broad subject that can be applied in different ways. We will focus on basic interpretation of water quality and benthic macroinvertebrate data that were collected in the field. Sampling water quality or benthic macroinvertebrate parameters, such as dissolved oxygen or percent clingers, yields data in the form of qualitative or quantitative measurements. For example, a dissolved oxygen measurement might read 5.2 milligrams per liter (mg/l). Interpreting that measurement includes giving it meaning—is that particular measurement at a healthy dissolved oxygen level? Does it mean that the entire stream being measured is healthy or just that portion of that stream where the dissolved oxygen was measured? How does that measurement fit into the context of the entire stream you're monitoring? Does it seem unusual to get a reading of 5.2 mg/l and if so, what might be causing such an irregular number?

Data

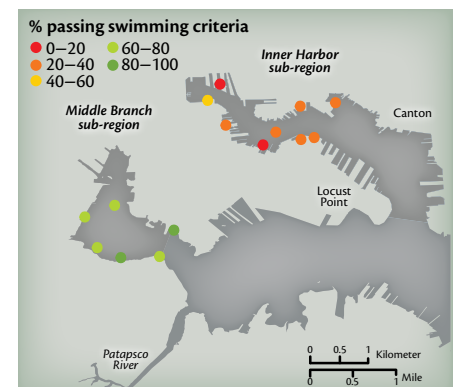


Station	Station Description	Event	Count Included for Benthos	Count Excluded for Benthos	Percent	Count At or Below Background for Benthos	Count Excluded for Benthos	Percent
1	Station 01	Station 01	10	10	100	0	0	0
2	Station 02	Station 02	10	10	100	0	0	0
3	Station 03	Station 03	10	10	100	0	0	0
4	Station 04	Station 04	10	10	100	0	0	0
5	Station 05	Station 05	10	10	100	0	0	0
6	Station 06	Station 06	10	10	100	0	0	0
7	Station 07	Station 07	10	10	100	0	0	0
8	Station 08	Station 08	10	10	100	0	0	0
9	Station 09	Station 09	10	10	100	0	0	0
10	Station 10	Station 10	10	10	100	0	0	0
11	Station 11	Station 11	10	10	100	0	0	0
12	Station 12	Station 12	10	10	100	0	0	0
13	Station 13	Station 13	10	10	100	0	0	0
14	Station 14	Station 14	10	10	100	0	0	0
15	Station 15	Station 15	10	10	100	0	0	0
16	Station 16	Station 16	10	10	100	0	0	0
17	Station 17	Station 17	10	10	100	0	0	0
18	Station 18	Station 18	10	10	100	0	0	0
19	Station 19	Station 19	10	10	100	0	0	0
20	Station 20	Station 20	10	10	100	0	0	0
21	Station 21	Station 21	10	10	100	0	0	0
22	Station 22	Station 22	10	10	100	0	0	0
23	Station 23	Station 23	10	10	100	0	0	0
24	Station 24	Station 24	10	10	100	0	0	0
25	Station 25	Station 25	10	10	100	0	0	0
26	Station 26	Station 26	10	10	100	0	0	0
27	Station 27	Station 27	10	10	100	0	0	0
28	Station 28	Station 28	10	10	100	0	0	0
29	Station 29	Station 29	10	10	100	0	0	0
30	Station 30	Station 30	10	10	100	0	0	0

Interpretation



Synthesis



This section of the manual focuses on Interpretation—giving your data meaning.

The first step for any data analyst is to look at the data values and check for errors. The next steps are to calculate basic statistics of the data and graph or map the data values. These types of interpretation must occur before advanced data visualization and assessment (e.g., report cards) and is a critical missing link between measuring the parameters in the field and using the data to assess ecosystem health. This manual is intended to provide a basic understanding of the different types of data, give you confidence in that data, and interpret it in the right way for your intended audience. Introductory

communication products, such as a factsheet or PowerPoint presentation, can be compiled based on these data interpretation techniques, but in-depth assessment (e.g., report cards) and advanced products, such as printed and digital data visualizations, require more advanced and in-depth training.

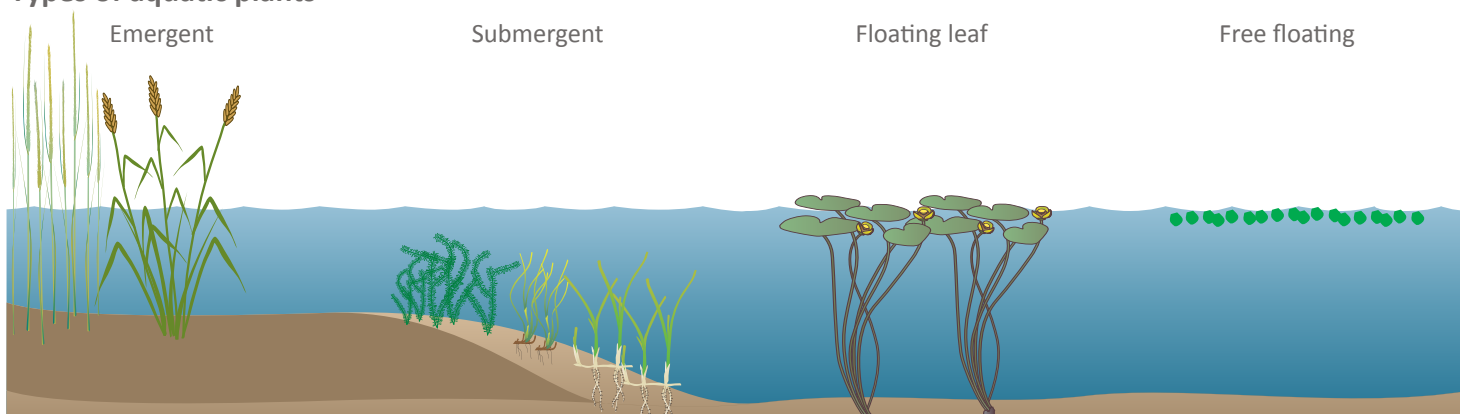
KINDS OF DATA

There are four forms of data that this methods document addresses: nominal (non-numerical), ordinal (non-numerical), interval (numerical), and ratio (numerical). Each type of data is defined below and an example relevant to interpreting and understanding water quality or benthic macroinvertebrate results is given. More in-depth discussion on types of data can be found in *Doing Science* by Ivan Valiela, an excellent reference for data interpretation.

Nominal data

Nominal data are non-numerical and are sometimes known as qualitative data. Nominal data can be separated into categories, such as types of aquatic vegetation or kinds of benthic macroinvertebrates in a sample. Thinking about your data, which would you label as nominal data?

Types of aquatic plants

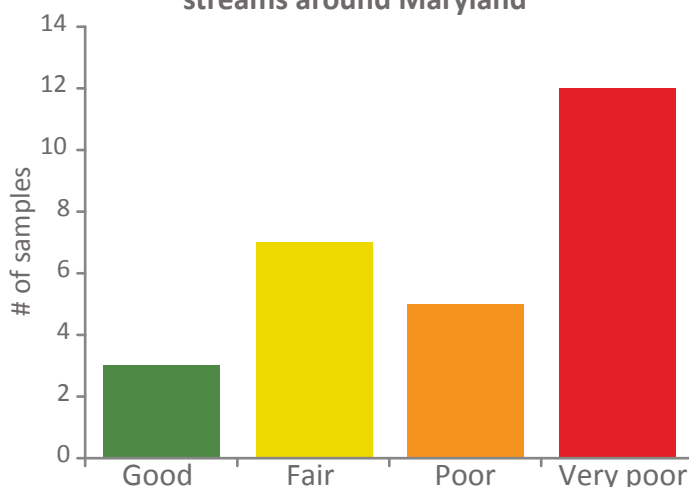


Nominal data displaying types of aquatic vegetation. In this example, data is displayed in a drawing, whereas data can be displayed as photographs as well as listed in a table.

Ordinal data

Ordinal data are non-numerical, but can be ranked based on values, such as whether benthic macroinvertebrate samples are considered good, fair, poor, or very poor.

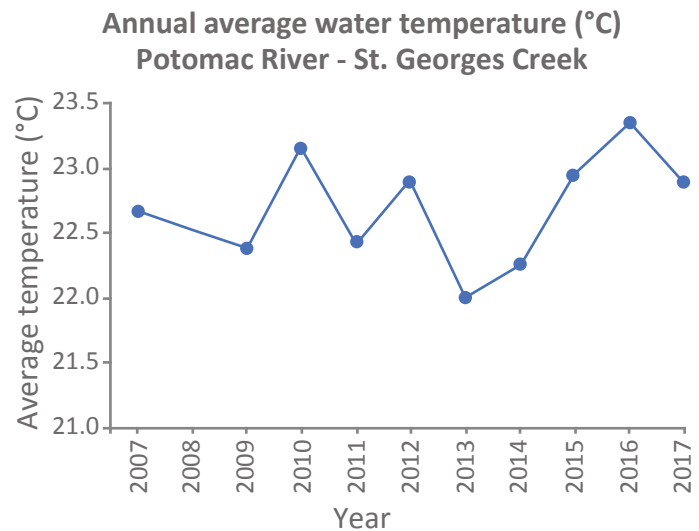
Health of benthic macroinvertebrates collected in streams around Maryland



Ordinal data displaying health of macroinvertebrates collected by high school students in streams around MD. Adapted from the 2017 Maryland Watershed Report Card by the Howard County Conservancy.

Interval data

Interval data are the most common type of data and are based on known intervals, such as numbers. The distance between two numbers is important and says something about the value of the data. Some examples include basic water quality data, such as water temperature and pH measurements. Thinking of your data, which data would you label as interval data?



Interval data displaying water temperature at one station in the Potomac River, Maryland. Data from eyesonthebay.net.

Ratio data

Ratio data are numerical data similar to interval data and are characterized by an absolute zero that means there is nothing of that property. One common example is dissolved oxygen measurements. The amount of dissolved oxygen is relative to when there is none. Other examples are fish weight or length or size of macroinvertebrates.

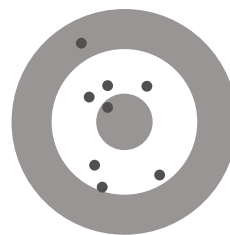
Sampling location	Depth (meters)	Dissolved oxygen (mg/l)
Choptank River 1	surface (0.3)	3.4
Choptank River 1	1.0	2.5
Choptank River 1	2.0	1.5
Choptank River 1	3.0	0.3
Choptank River 1	surface (0.3)	6.2
Choptank River 2	1.0	6.1
Choptank River 2	2.0	5.7
Choptank River 2	3.0	3.1
Choptank River 2	4.0	2.2
Choptank River 2	5.0	1.1
Choptank River 2	6.0	0.4
Choptank River 2	7.0	0.0

Example data to show that ratio data usually are characterized by zero amount of that parameter means there is none.

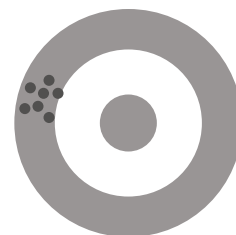
ACCURACY AND PRECISION

Accuracy and precision are both important to understanding your data. Determining what accuracy you want from your equipment is also important to determine before buying equipment. First, let's define each term.

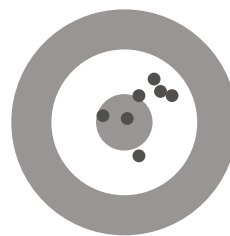
- Accuracy is when a measurement is as close to the real value as possible. This is ensured by calibrating your equipment before going into the field so that the piece of equipment measures the real value in the water column.
- Precision is when repeated measurements closely match each other. This would be if you measured Secchi depth multiple times in the same location and the measurements closely matched each other. This is less practical in water quality monitoring when the properties of the water are constantly changing due to the movement of water.



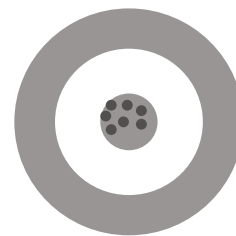
Low accuracy
Low precision



Low accuracy
High precision



High accuracy
Low precision



High accuracy
High precision

Remember that measurements may be precise but not accurate. Measurements may be accurate but not very precise.

This type of information for each piece of equipment or measurement should be included in a Quality Assurance Project Plan or Standard Operating Procedure and easily referenced. Something to keep in mind with accuracy is that it is dependent on the ability of the piece of equipment you are using to measure to that level of accuracy. The accuracy of the equipment should be listed in the manufacturer's manual.

CLEANING YOUR DATA

Now that you know what kind of data you are working with and are reasonably confident in the accuracy and precision of the data, you need to organize ("clean") your data in a spreadsheet or database. It is important to clean your data before doing any interpretation because errors will affect the basic statistical calculations that you will do next.

What does it mean to clean your data exactly? This encompasses everything from formatting your spreadsheet so that it's consistent and usable to flagging unusual or duplicate data as potential errors that must be corrected. Below are steps you can take to clean the data before interpretation. Specific examples of what to look for are given in the following figures.

- Open the spreadsheet and just start looking at the data.
 - Are the correct headers with correct units associated with the data values?
 - Do the rows and columns have the correct widths and heights to view all the data or is that even needed?
 - Are there any missing data? Are there data in another spreadsheet or on another tab that need to be incorporated here?

Should this be Conductivity?

Why are there two blank lines?

The screenshot shows an Excel spreadsheet with the following columns: Site, Date, Depth, Salinity, Temperature, O₂, O₂ Saturation, Conductivity, Light Trans, Sigma-T, Fluorometer, pH400-700nm light, PAR, Oxidation Reduction Poter, and LabID. The data is organized by date, with rows for each survey date. Annotations include:

- A red box highlights two blank lines at the top of the data section.
- A red circle highlights the 'Conductivity' column header, with a note asking 'Should this be Conductivity?'.
- A red circle highlights the 'Harbor Survey' tab name at the bottom.
- A red circle highlights a cell containing '60.1 NS' in the 'Conductivity' column, with a note asking 'How are you going to address data values that are not numbers?'.

Should you add more tabs so that the data is organized by indicator? Or by surface vs bottom measurements? Or something else?

How are you going to address data values that are not numbers?

- How are the data organized? By sampling date or sampling station? Which way is the best way to look at and interpret the data?
- Are there any duplicate entries? Why? Do you delete them altogether or save them “just in case”? How do you organize and structure your files to do this?
- Are there any unusual data? You can sort the data from high to low and determine if any values are outside the expected range for a particular variable. This could be due to typing errors, instrument error, or they could be genuine outliers.
- Are there any cells that need to be changed from numbers to text or vice versa so that Excel can read them correctly?
- How are your latitude and longitude written? Is it in a format that works for you or for someone who will be doing GIS mapping?

Setting up an error checklist to use for each spreadsheet can be helpful.

The column headers are bolded and each label makes sense and has appropriate units

Extra decimal point

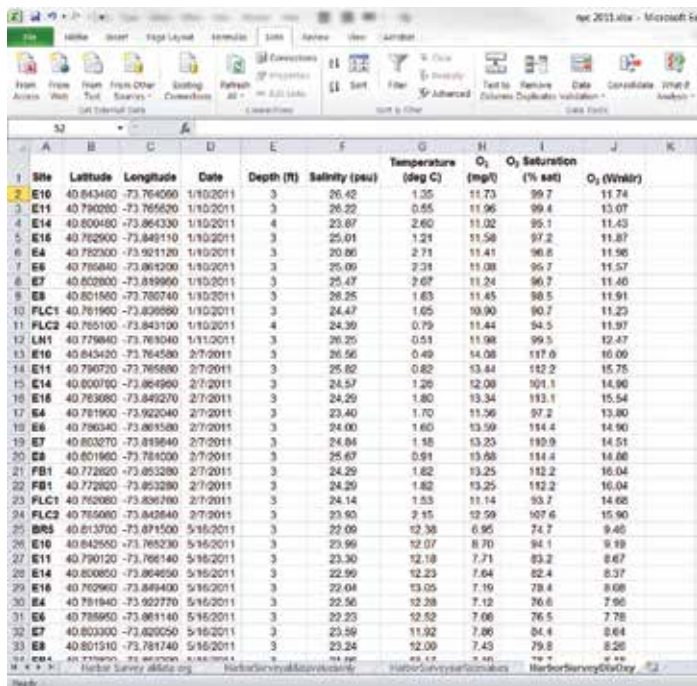
	Site	Latitude	Longitude	Date	Depth (ft)	Salinity (psu)	Temperature (deg C)	O ₂ (mg/L)	Saturation (% sat)	Conductivity (uS/cm)	Light Trans (%)	Sigma-T (kg/m ³)	Fluorometer (mg/L)	pH	PAR Ref (400-700nm light, uE/m ²)	PAR (uE/m ²)	Oxidation Potential (mV)	Hr:Min	Act:Dep (ft)	Current speed (knots)	Current direction
2	E15	40.762900	-73.849110	1/10/2011	3	25.01	1.21	1.58	97.2	2.22	54.9	20.02	8.12	8.99	1638	12	9.614E+01	13:21	16.00	0.3	S
3	FLC1	40.761980	-73.836660	1/10/2011	3	24.47	1.05	10.90	90.7	2.16	50.2	19.59	10.29	8.88	1492	79	8.283E+01	13:05	16.00	0.4	S
4	FLC2	40.765100	-73.843100	1/10/2011	4	24.39	0.79	11.44	94.5	2.14	47.9	19.53	9.44	8.95	1571	48	8.678E+01	13:14	18.00	0.7	E
5	FB1	40.772620	-73.853060	1/10/2011	3	25.03	2.33.195	11.26	97.5	2.29	70.3	19.96	7.83	8.98	1528	11	9.914E+01	13:30	21.00	0.6	S
6	E6	40.765840	-73.861200	1/10/2011	3	25.09	2.31	11.08	95.7	2.29	73.1	20.03	8.80	8.98	1652	52	9.218E+01	13:40	22.00	0.6	S
7	E14	40.800480	-73.864330	1/10/2011	4	23.67	2.60	11.02	95.1	2.21	68.4	19.04	8.49	8.96	1603	36	9.602E+01	13:50	24.00	0.5	E
8	E7	40.802800	-73.819980	1/10/2011	3	25.47	2.07	11.24	96.7	2.31	73.2	20.34	8.82	8.98	1771	40	9.378E+01	12:30	50.00	1.8	E
9	E4	40.762300	-73.821120	1/10/2011	3	20.60	2.71	11.41	96.8	1.96	58.3	16.63	9.09	9.01	1315	6	9.338E+01	14:13	72.00	0.4	W
10	E8	40.801560	-73.780740	1/10/2011	3	26.25	1.83	11.45	98.5	2.36	75.1	20.98	9.63	9.03	1843	11	1.043E+02	12:13	91.00	0.7	NE
11	E10	40.843400	-73.764080	1/10/2011	3	26.42	1.35	11.73	99.7	2.54	84.1	21.14	10.53	9.01	1707	60	1.150E+02	11:21	91.00	0.8	E
12	E11	40.790200	-73.765620	1/10/2011	3	26.22	0.55	11.90	99.4	2.27	10.3	21.01	14.16	8.99	1885	36	8.036E+01	11:49	8.00	0.5	S
13	LN1	40.779840	-73.761040	1/11/2011	3	26.25	0.51	11.96	99.5	2.27	5.6	21.04	14.03	8.99	1857	3	9.094E+01	11:57	7.00	0.5	S
14	E15	40.763080	-73.849270	2/7/2011	3	24.29	1.80	13.34	113.1	2.20	54.9	19.41	63.02	9.56	1071	43	1.207E+02	13:09	17.00	WEAK	NS
15	FLC1	40.762080	-73.836790	2/7/2011	3	24.14	1.53	11.14	93.7	2.17	51.8	19.30	68.01	9.46	1118	47	2.282E+02	12:56	17.00	0.3	N
16	FLC2	40.765080	-73.842840	2/7/2011	3	23.93	2.15	12.59	107.6	2.19	54.0	19.11	61.33	9.55	970	36	8.449E+01	13:03	20.00	0.2	W
17	E14	40.800720	-73.864960	2/7/2011	3	24.57	1.26	12.06	101.1	2.18	50.3	19.06	58.99	9.66	1516	54	7.278E+01	13:36	23.00	0.8	W
18	FB1	40.772620	-73.853280	2/7/2011	3	24.29	1.82	13.25	112.2	2.20	55.3	19.41	61.06	9.55	1155	48	1.557E+02	13:17	25.00	0.5	N
19	FB1	40.772620	-73.853280	2/7/2011	3	24.29	1.82	13.25	112.2	2.20	55.3	19.41	61.06	9.55	1155	48	1.557E+02	13:17	25.00	0.5	N
20	E6	40.766340	-73.861550	2/7/2011	3	24.00	1.60	13.59	114.4	2.16	58.2	19.18	49.14	9.42	1282	45	1.623E+02	13:26	25.00	0.5	N
21	E7	40.800270	-73.819980	2/7/2011	3	24.84	1.18	13.23	110.9	2.20	56.4	19.86	58.91	9.50	1376	55	1.703E+02	12:19	53.00	0.7	E
22	E4	40.761900	-73.822940	2/7/2011	3	23.40	1.70	11.56	97.2	2.12	57.7	19.70	46.39	9.32	1007	39	1.326E+02	13:54	72.00	3.2	W
23	E8	40.801980	-73.781000	2/7/2011	3	25.67	0.91	13.68	114.4	2.25	58.8	20.55	60.47	9.57	1825	86	1.326E+02	12:03	95.00	0.8	N
24	E10	40.843420	-73.764580	2/7/2011	3	26.56	0.49	14.06	117.0	2.29	61.9	21.26	53.58	9.64	1135	48	1.130E+02	11:20	98.00	0.6	S
25	E11	40.790720	-73.765880	2/7/2011	3	25.82	0.82	13.44	112.2	2.26	58.2	20.68	61.09	9.64	2337	29	3.798E+01	11:44	11.00	0.5	N
26	LN1			2/7/2011														NS	NS	NS	NS
27	BR5	40.813700	-73.871500	5/16/2011	3	22.09	12.38	6.95	74.7	2.66	71.4	18.52	22.35	7.65	964	NS	1.792E+02	14:02	13.00	0.6	S
28	E16	40.762960	-73.849400	5/16/2011	3	22.04	13.05	7.19	78.4	2.70	70.9	18.37	25.97	7.67	571	NS	2.124E+02	13:18	16.00	WEAK	S
29	FLC1	40.762100	-73.836700	5/16/2011	3	21.65	13.49	6.20	68.2	2.69	68.9	18.90	26.38	7.57	946	NS	1.705E+02	13:33	16.00	0.2	N
30	E14	40.800850	-73.864650	5/16/2011	3	22.99	12.23	7.64	82.4	2.75	71.2	17.24	20.10	7.79	613	NS	1.980E+02	14:13	22.00	1.3	W
31	FLC2	40.765340	-73.842860	5/16/2011	3	22.14	12.61	6.32	68.5	2.69	71.7	18.50	19.68	7.59	723	NS	2.041E+02	13:26	22.00	0.5	W
32	FB1	40.772620	-73.853200	5/16/2011	3	21.94	11.17	7.10	78.7	2.70	70.3	18.38	24.10	7.64	801	NS	2.219E+02	13:10	24.00	WEAK	S

Is this the correct sampling date? Check original field sheets or verify with field staff/volunteer monitor.

This row is duplicated. Do you delete it or keep it and flag it?

Once you have looked through the spreadsheet and have determined that the data are correct and organized, you can start interpreting the data with basic statistics and graphs.

This spreadsheet has been cleaned and simplified. Only columns pertaining to dissolved oxygen were kept. The tab is labeled appropriately. The columns have been sorted first by date, then by station.



	A	B	C	D	E	F	G	H	I	J	K
	Site	Latitude	Longitude	Date	Depth (ft)	Salinity (psu)	Temperature (deg C)	O ₂ (mg/l)	O ₂ Saturation (% sat)	O ₂ (Winkler)	
1	E10	40.843400	-73.764000	1/10/2011	3	26.42	1.35	11.73	99.7	11.74	
2	E11	40.790200	-73.765620	1/10/2011	3	26.22	0.55	11.96	99.4	13.07	
3	E14	40.800480	-73.864330	1/10/2011	4	23.87	2.60	11.02	95.1	11.43	
4	E16	40.702900	-73.848110	1/10/2011	3	25.01	1.21	11.58	97.2	11.87	
5	E4	40.752300	-73.921120	1/10/2011	3	20.86	2.71	11.41	96.8	11.96	
6	E6	40.796840	-73.861200	1/10/2011	3	25.09	2.31	11.08	95.7	11.57	
7	E7	40.802800	-73.819960	1/10/2011	3	25.47	2.67	11.24	96.7	11.40	
8	E8	40.801960	-73.780740	1/10/2011	3	26.25	1.63	11.45	98.5	11.91	
9	FLC1	40.781960	-73.828690	1/10/2011	3	24.47	1.05	10.90	90.7	11.23	
10	FLC2	40.755100	-73.843100	1/10/2011	4	24.39	0.79	11.44	94.5	11.37	
11	LN1	40.779840	-73.761040	1/11/2011	3	26.25	0.51	11.98	99.5	12.47	
12	E10	40.843400	-73.764580	2/7/2011	3	26.56	0.49	14.08	117.0	10.00	
13	E11	40.790720	-73.765880	2/7/2011	3	25.82	0.82	13.44	112.2	15.75	
14	E14	40.800700	-73.864990	2/7/2011	3	24.57	1.26	12.00	101.1	14.90	
15	E16	40.763080	-73.849270	2/7/2011	3	24.29	1.80	13.34	113.1	15.54	
16	E4	40.781900	-73.922040	2/7/2011	3	23.40	1.70	11.56	97.2	13.80	
17	E6	40.796340	-73.861580	2/7/2011	3	24.00	1.60	13.59	114.4	14.90	
18	E7	40.803270	-73.819640	2/7/2011	3	24.84	1.58	13.23	110.9	14.51	
19	E8	40.801960	-73.781000	2/7/2011	3	25.67	0.91	13.68	114.4	14.88	
20	FB1	40.772620	-73.853280	2/7/2011	3	24.29	1.62	13.25	112.2	16.04	
21	FB1	40.772620	-73.853280	2/7/2011	3	24.29	1.82	13.25	112.2	16.04	
22	FLC1	40.752050	-73.826190	2/7/2011	3	24.14	1.53	11.14	93.7	14.68	
23	FLC2	40.755080	-73.842840	2/7/2011	3	23.93	2.15	12.59	107.6	15.90	
24	BR5	40.813700	-73.871500	5/16/2011	3	22.09	12.38	6.95	74.7	9.48	
25	E10	40.842650	-73.765230	5/16/2011	3	23.99	12.07	8.70	94.1	9.19	
26	E11	40.790120	-73.766140	5/16/2011	3	23.30	12.18	7.71	83.2	8.67	
27	E14	40.800850	-73.864050	5/16/2011	3	22.90	12.23	7.64	82.4	8.37	
28	E16	40.762960	-73.849400	5/16/2011	3	22.04	13.05	7.19	78.4	8.68	
29	E4	40.781840	-73.922770	5/16/2011	3	22.56	12.38	7.12	76.6	7.96	
30	E6	40.788950	-73.861140	5/16/2011	3	22.23	12.52	7.66	76.5	7.78	
31	E7	40.803000	-73.820250	5/16/2011	3	23.59	11.92	7.86	84.4	8.64	
32	E8	40.801510	-73.781740	5/16/2011	3	23.24	12.09	7.43	79.8	8.26	
33	BR5	40.813700	-73.871500	5/16/2011	3	22.09	12.38	6.95	74.7	9.48	
34	BR5	40.813700	-73.871500	5/16/2011	3	22.09	12.38	6.95	74.7	9.48	

DESCRIPTIVE STATISTICS

Descriptive statistics provide basic summarized information about your data and allow you to dip your toe into the data interpretation pool. There are many good statistics websites and books that can walk you through the calculations for all descriptive statistics. Below are just the basic definitions of calculations you need to perform in order to evaluate your data for its use in interpretation and synthesis.

Mean, median, mode

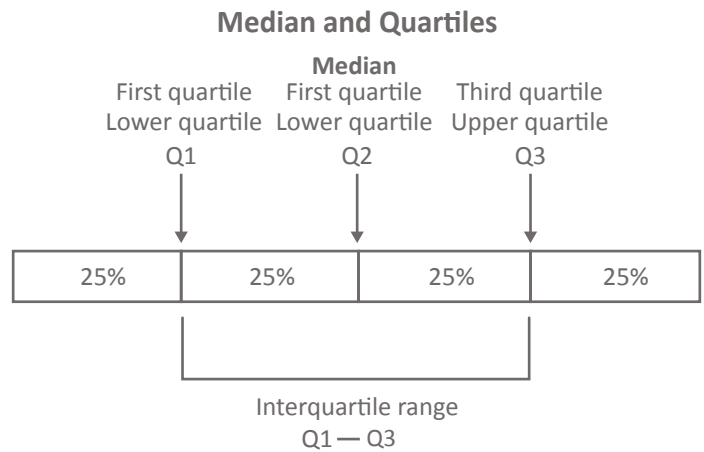
- The mean is the calculated average of all values in a dataset.
- The geometric mean is a calculated average of all values in a dataset that takes into account highly variable data. Instead of summing all the numbers and then dividing, the geometric mean multiplies all the numbers and takes the square root. Bacteria data often use geometric mean.
- The median is the center value of the data once it's ranked in order from smallest to largest. The median is in the middle of the dataset.
- The mode is the value that is measured the most. It is not used often and may not apply when interpreting your data.

In the following table, an example of a table of raw data as well as the mean, median, and mode for each type of benthic macroinvertebrate is given. These types of tables are useful when you are looking at your own data and want basic summary statistics to get an idea of the data in your stream. Codes such as N/A (not applicable) can be used, but need to be included in the metadata or legend of any figure.

Sample site	Caddisfly (count)	Mayfly (count)	Stonefly (count)	Clams (count)	Aquatic worms (count)	Crayfish (count)	% Sensitive to Pollution
A1	2	3	1	15	2	4	22
A2	1	1	1	5	7	11	12
GP1	0	0	0	13	16	9	0
GP2	4	7	4	8	12	3	40
GP3	4	5	2	12	10	9	33
D1	6	6	9	11	7	4	49
B11	0	1	2	8	9	12	1
mean	2.4	3.3	2.7	10.3	9.0	7.4	22.4%
median	2	3	2	11	9	9	22%
mode	0	1	1	8	7	4	N/A

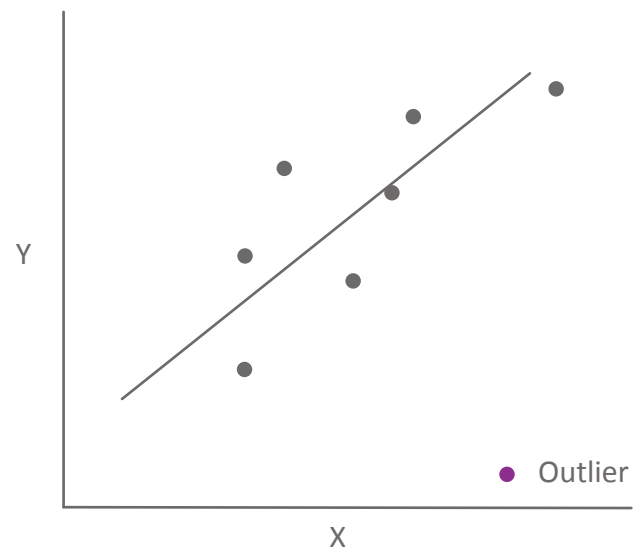
Range

- The range is simply the total spread of all values in the dataset.
- A quartile is one of four segments of your data. Quartiles divide your data into four segments based on the minimum value, the maximum value, and the median.
- The interquartile range is the range of data between Quartile 1 and Quartile 3 and spans 50% of the data, including the mean. You will want to know the interquartile range to determine outliers.



Outliers

- Outliers are data values that fall outside the general distribution of the data. Outliers often represent a measurement or user error that needs to be taken out of the dataset. Outliers will skew a normally distributed dataset (explained on the next page), so knowing whether an outlier is true or not is very valuable to interpreting the data.
 - For example, a volunteer monitor measures Secchi depth at the same site once a week for the entire calendar year. The range of the data falls between 1.2 m and 1.5 m most of the time except one measurement which is 0.2 m. Is this a true measurement (perhaps during an algal bloom) or an error? You can calculate whether it is an outlier using the interquartile range to find out.



Standard deviation

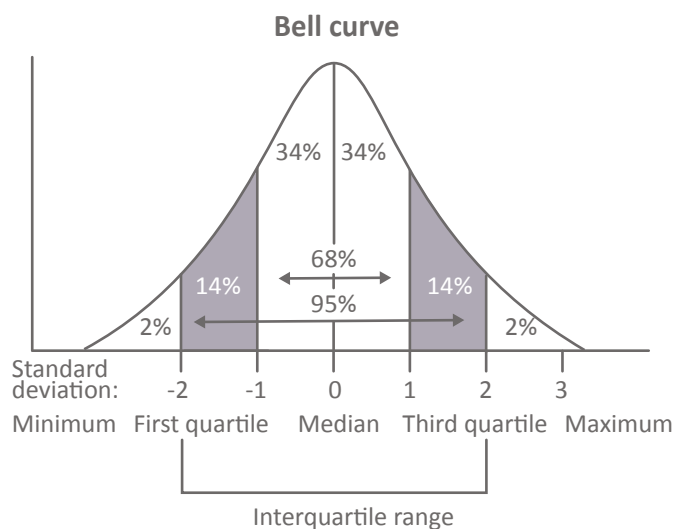
- The standard deviation is a calculated value that shows the amount of variation in the population. This is used more often than the range because it incorporates both large variation and small deviations from the mean. It performs a similar function as the range in that it tells you how spread out the data is from the mean.

Standard error

- The standard error is very similar to the standard deviation. The standard error is a calculation of how far from the true population mean the sample mean is. You will often calculate standard error for your samples to determine how close to the sample mean all the samples are.

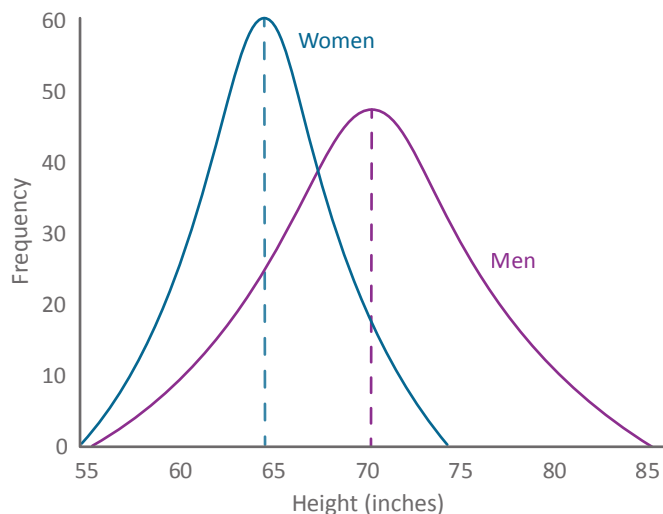
Distributions

Once you have calculated the median, range, and standard deviation of the data and eliminated outliers, you should be able to evaluate what type of distribution the data have. There are many types of data distributions, including unimodal, bimodal, symmetric, asymmetric, and skewed, to name a few. A complete discussion of different distributions and how to use them is outside the scope of this manual. However, it is worth noting the basic difference between normal and non-normal data.



Normal

Normally distributed data follows a bell-shaped curve when a frequency distribution is plotted. Examples of normally distributed data include human height measurements and test scores of students, but these are example cases that rarely occur in the real world. Furthermore, determining whether a dataset is normal or non-normal is most relevant to what types of statistical tests you will perform to infer information about the data.



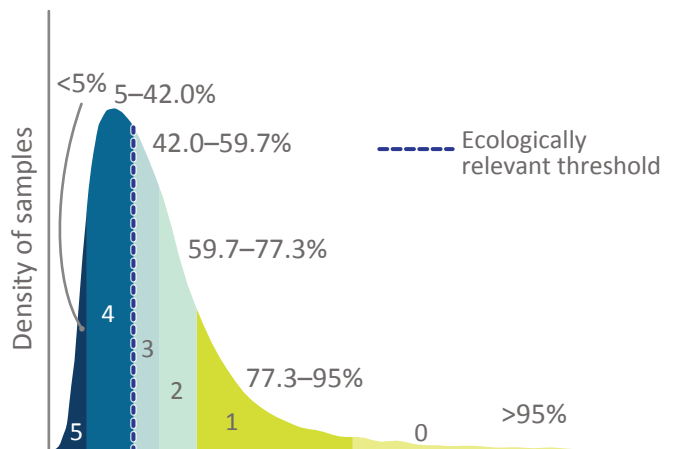
These data are normally distributed curves showing the average height of women and men throughout the United States.

Non-normal

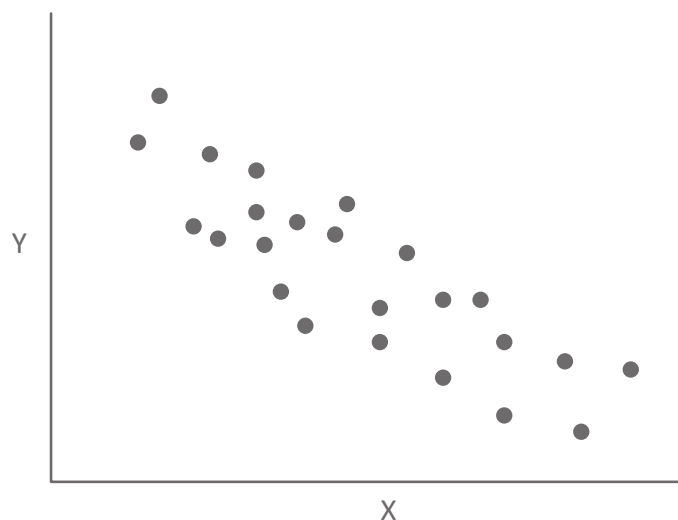
Despite the name, non-normal data is very common—most data does not follow a bell shaped curve. In order to achieve a normal distribution and therefore perform certain statistical tests, you may need to transform your non-normal data using logarithmic transformation or other common transformations. More in-depth discussion on distributions can be found in *Doing Science* by Ivan Valiela, an excellent reference for data interpretation.

Correlation

Statistical tests are beyond the scope of this manual, but one basic test that should be completed on your data is correlation analysis. Correlations tell you whether two parameters, such as water temperature and number of clingers are related. This does not mean that one parameter is affecting the other, but rather that they are changing together, either in a positive correlation (both increasing) or in a negative correlation (one increasing while other decreasing).



These data are chlorophyll a samples from the mesohaline (salinity of 5-18) portion of Chesapeake Bay. There are many more samples of low chlorophyll amounts than high amounts. This is typical of chlorophyll a as an indicator.



DATA IN TABLES

Why use tables?

Using tables helps to compare and contrast basic data values and information in a linear format. By putting data in a table format, it is easier to read and decipher than if it were in a list or paragraph of text. A table of data is a summary that provides a sense of variability, outliers, and whether the data is normal or non-normal. However, tables are most useful when you are looking at your own data and becoming familiar with the detail of each measurement. Tables are not used as often in presentations or communication products to an outside audience.

Parts of a table

A table consists of *cells* made into *rows* and *columns*. Most tables have a header row that lists out each column's description (with units). A table can also have a header column that lists out each row's description. In this way, tables have a hierarchy and can structure sub-categories of data.

In the table above, the orange space is a column, the yellow space is a row, and the green space is a cell.

Examples of tables and their many uses

In these example tables, the data are typical of a volunteer monitor collecting benthic macroinvertebrate samples in freshwater streams. Each sample site is listed as well as the counts of the number of each type of macroinvertebrate. The last column is a calculation of the number of pollution sensitive organisms compared to the total amount.







EXAMPLE 1: BASIC TABLE

Sample site	Caddisfly	Mayfly	Stonefly	Clams	Aquatic worms	Crayfish	% Sensitive to Pollution
A1	2	3	1	15	2	4	22
A2	1	1	1	5	7	11	12
GP1	0	0	0	13	16	9	0
GP2	4	7	4	8	12	3	40
GP3	4	5	2	12	10	9	33
D1	6	6	9	11	7	4	49
B11	0	1	2	8	9	12	1

EXAMPLE 2: BASIC TABLE WITH SOME FORMATTING

Sample site	Caddisfly	Mayfly	Stonefly	Clams	Aquatic worms	Crayfish	% Sensitive to Pollution
A1	2	3	1	15	2	4	22
A2	1	1	1	5	7	11	12
GP1	0	0	0	13	16	9	0
GP2	4	7	4	8	12	3	40
GP3	4	5	2	12	10	9	33
D1	6	6	9	11	7	4	49
B11	0	1	2	8	9	12	1

EXAMPLE 3: BASIC TABLE WITH SCIENCE COMMUNICATION PRINCIPLES EMBEDDED

Sample site	Caddisfly 	Mayfly 	Stonefly 	Clams 	Aquatic worms 	Crayfish 	% Sensitive to Pollution
Anacostia 1	2	3	1	15	2	4	22
Anacostia 2	1	1	1	5	7	11	12
Gunpowder 1	0	0	0	13	16	9	0
Gunpowder 2	4	7	4	8	12	3	40
Gunpowder 3	4	5	2	12	10	9	33
Davis Branch 1	6	6	9	11	7	4	49
Big Elk 11	0	1	2	8	9	12	1

DATA IN GRAPHS

Why use figures and graphs?

Graphing data is the easiest way to visualize your data. It allows you to see potential relationships between different measurements and different data as well as helps communicate that information to your audience. Graphs can also help determine if there are any outliers in your data or which measurement could be potential errors that need to be corrected. Finally, graphs can help visualize trends in the data.

Figures, such as maps and diagrams, are used to help readers visually connect information. While graphs usually show a lot of numbers, figures can help connect those numbers to general patterns and trends or show information on a geographic scale. Graphs and figures complement each other and you will often use both in a presentation or document.

Parts of a graph

- **Data**—these data usually come in the form of a spreadsheet. You will make graphs directly in the spreadsheet that has the data.
- **Axes**—The x-axis is the horizontal axis, usually at the bottom of the graph. The y-axis is the vertical axis and can be on one or both sides of the graph.
- **Labels**—Everything on a graph must be labeled. The data will be labeled in the legend. The axes can be labeled directly adjacent to them and include the units of measurement of the data.
- **Legend**—the legend is used to describe the different types of data that you may be displaying. If there are multiple sets of data, different colors, lines, and shapes can be used to distinguish among them.
- **Caption**—The caption is usually an expanded title or legend that describes the data and some inferences about the data. Captions are usually used in a document, not in a PowerPoint presentation.
- **Title**—The title of the graph should be a quick way to orient the reader to what they are looking at.

Design of a graph

Designing a graph starts in the spreadsheet. The first step is to choose the type of graph you want to use to best illustrate your data.

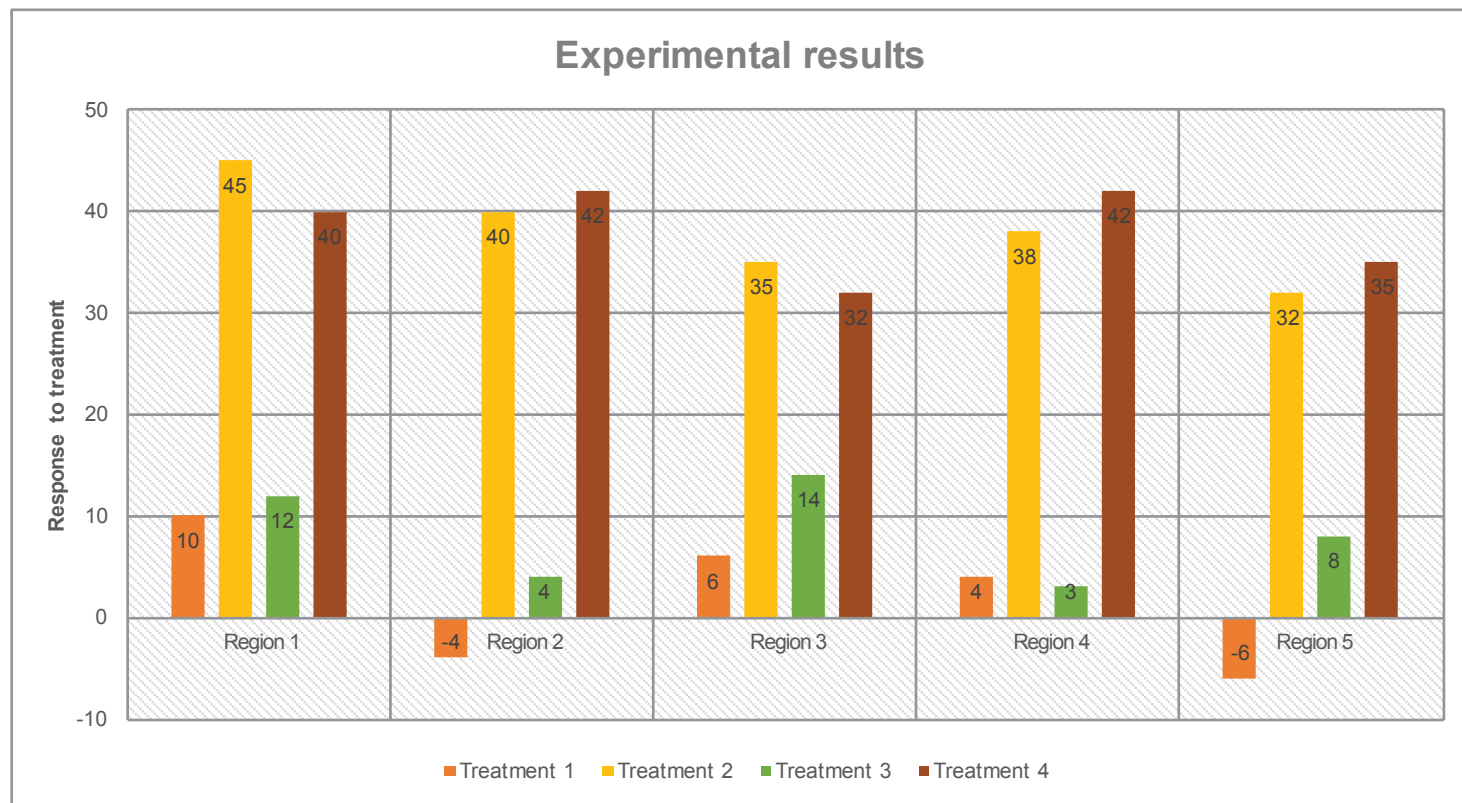
Types of graphs

- Bar graph—bar graphs compare different measurements to each other. They are most often used where the x-axis data are categories, such as sampling sites or different experimental treatments. The y-axis is usually the quantity of the thing you're measuring, like water depth or amount of nutrients.
- Line graph—line graphs are most often used when depicting data through time. The x-axis data is time and the y-axis is how the thing you're measuring is changing over that time period.
- Pie graph—pie graphs are used to compare data that make up a whole. These are percentage of totals. For example, the different types of land use in the Chesapeake Bay watershed.
- Comparison bar graph—More complex graphs can be used to explain several different aspects of one study or experiment. Multiple bar graphs can be used for each sampling site, for example. Each of those individual bar graphs represents a different aspect of the study for each sampling site.

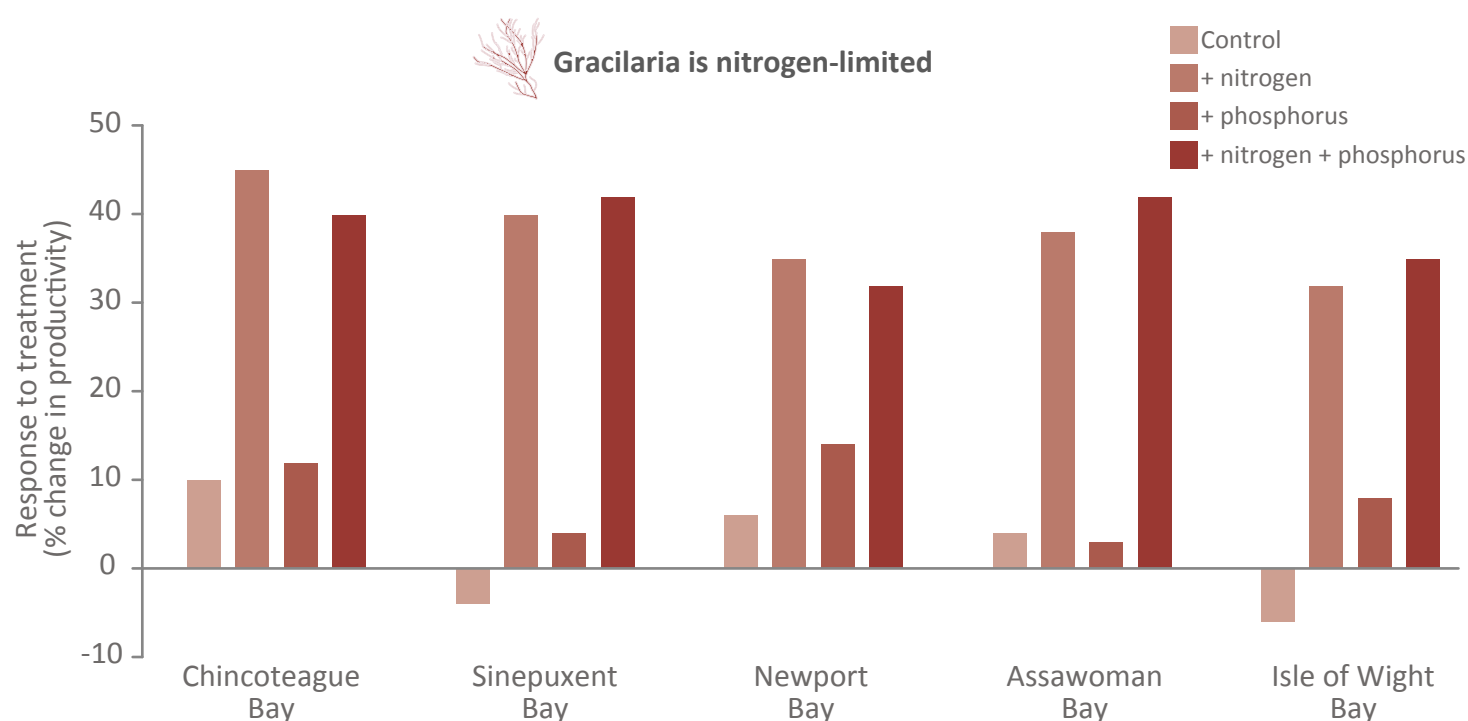
After choosing the type of graph you will use to illustrate your key points, you will need to spend time labeling the axes, the legend, and title. What colors make the most sense for each category of data? Or, is a black and white graph more useful? There are a lot of options in the spreadsheet that will make the graph clear to your audience. Besides the options of the graph itself, you can add arrows, text, symbols, and other descriptive objects directly in the spreadsheet. For example, to show the long-term average of dissolved oxygen, you can use a line across the graph that corresponds with the long-term dissolved oxygen value on the y-axis. Then, you can label just that line with a text box describing what it is. This information can also be included in the caption.

In these graphs the data is typical of an experimental design used by scientists to determine if certain treatments affect the parameter in question. Examples of these types of experiments include the effect of nutrients on plants, the effect of living shorelines on trapping sediment, or the effect of different water temperatures on the composition of benthic macroinvertebrates in streams.

EXAMPLE 1: BEFORE FORMATTING, BAR GRAPH

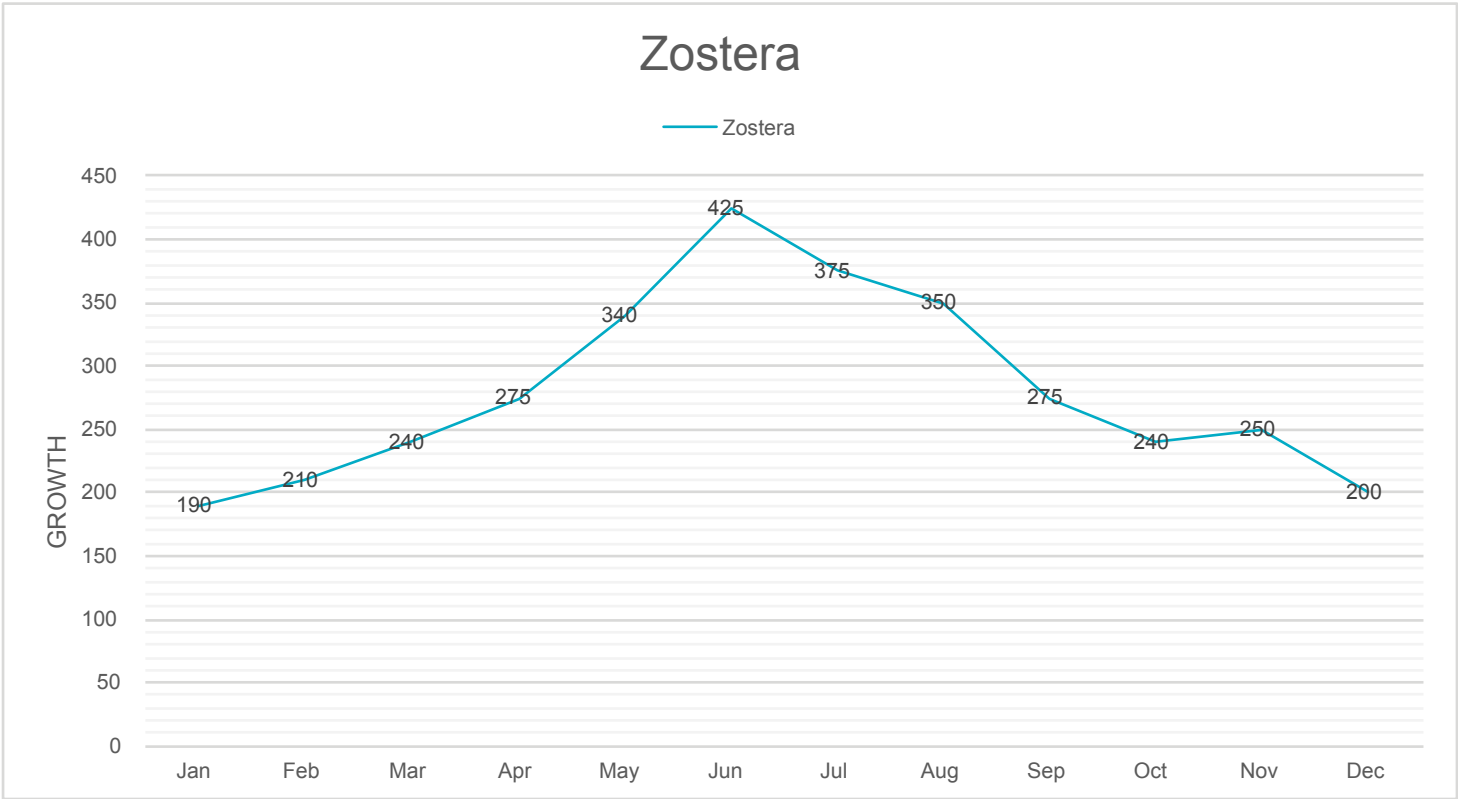


EXAMPLE 2: AFTER FORMATTING, BAR GRAPH

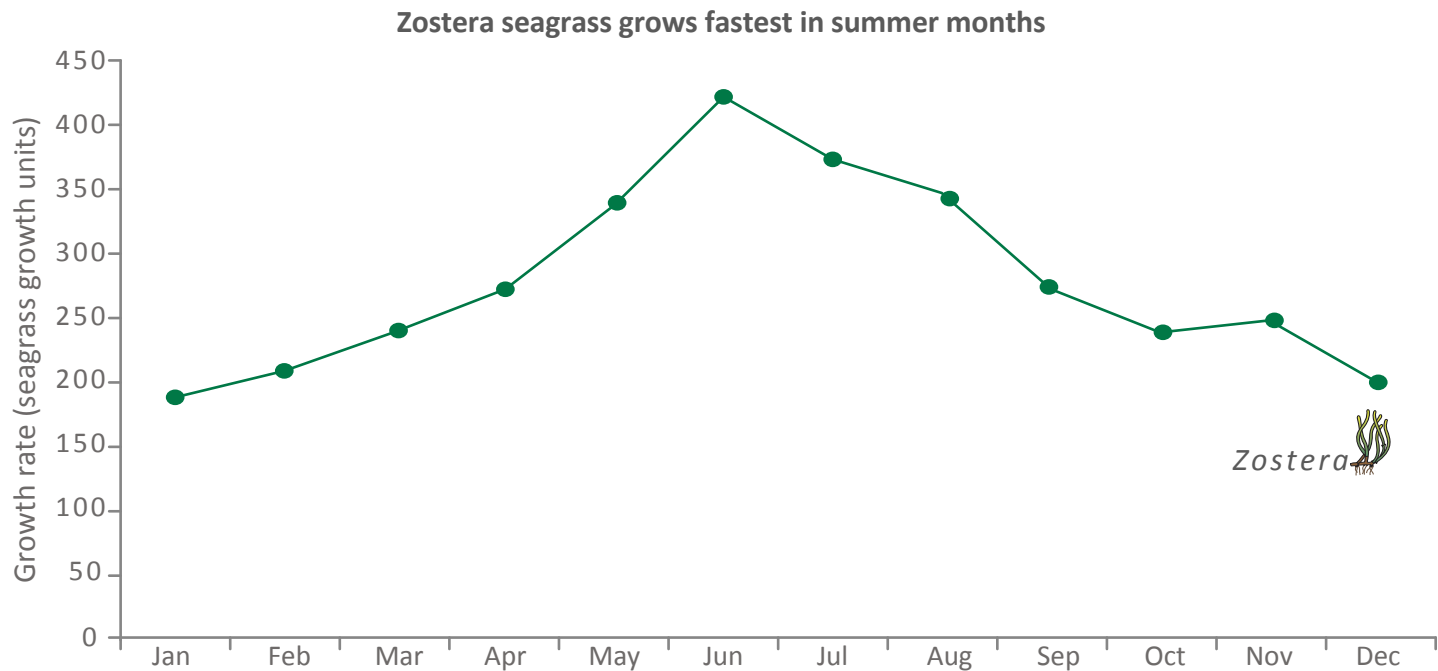


In these graphs the data is typical of measurements over time. These measurements can be discrete points, continuous monitoring data, or averages of multiple measurements. Examples of these types of data include seagrass growth over time, water temperature or dissolved oxygen over time, or number of caddisfly at one site over multiple sampling seasons.

EXAMPLE 3: BEFORE FORMATTING, LINE GRAPH

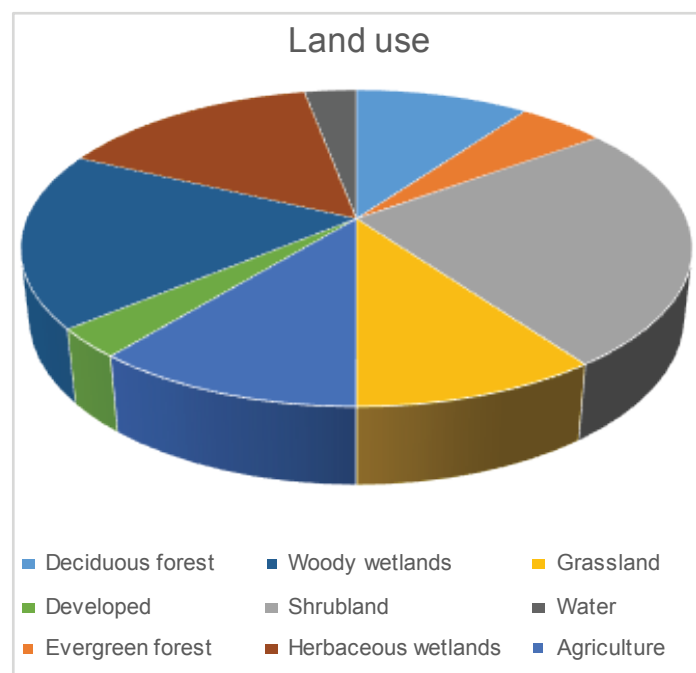


EXAMPLE 4: AFTER FORMATTING, LINE GRAPH



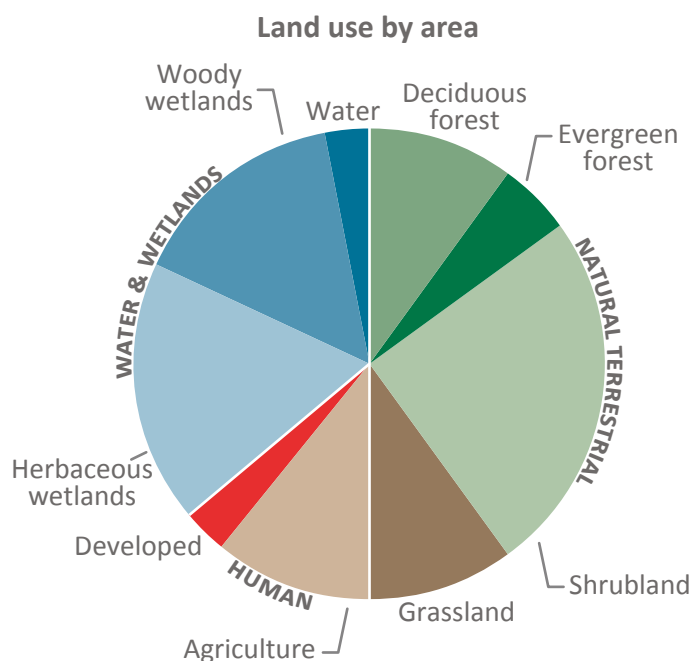
In these pie charts, the data are typically a fraction of a total. These measurements can be counts or areas and all the data can be presented as percentage of total. Examples of these types of data include land use area in a region, types of sediment sizes, and percent of samples that are types of macroinvertebrates.

EXAMPLE 5: BEFORE FORMATTING, PIE GRAPH



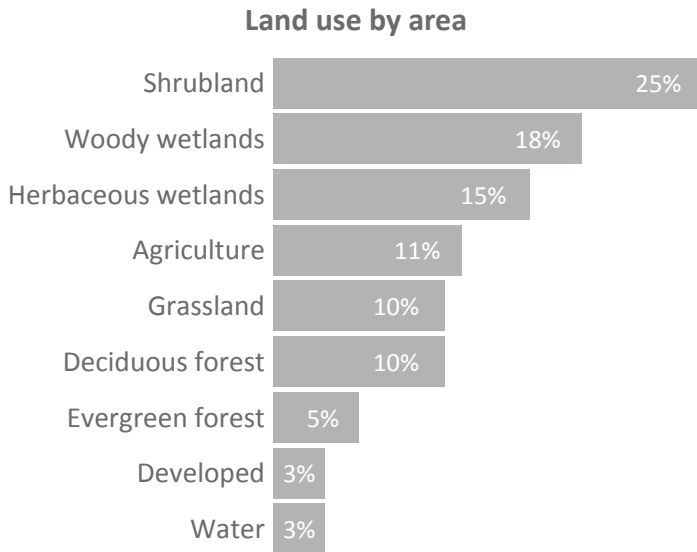
This is a standard pie graph that you often see in presentations and displays. This is particularly poor communication because you cannot tell how large each slice is related to the others, and it is hard to tell what land use type goes with what color. The line around the graph is distracting and the title is non-descriptive.

EXAMPLE 6: AFTER FORMATTING, PIE GRAPH



If you change the pie graph to looking from the top, it is easier to tell how much space each slice takes up. One could take this a step further and add percentages to each land use type. The colors also relate to the land use type and bold type face is used to denote aggregated groups.

EXAMPLE 7: AFTER FORMATTING, PIE GRAPH TURNED INTO BAR GRAPH



Maybe a pie graph really isn't even the correct way to visualize your data. Pie graphs can disorient, and it can be hard to actually tell how the real percentage each land use (or benthic macroinvertebrate type) is occupying in the sample area. Use a stacked bar chart, which ranks the data from largest to smallest land use area.

EXAMPLE 8: SAME DATA IN THREE DIFFERENT TYPES OF GRAPHS

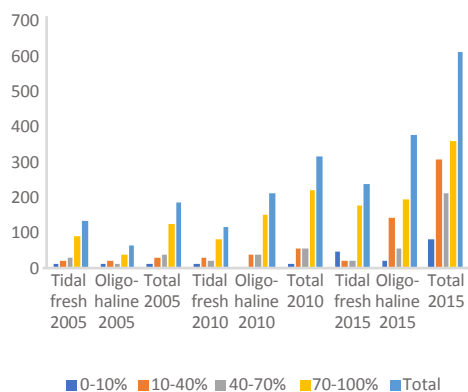
Below is an example of how monitoring data can be used in different types of graphs. These data show the number of hectares of submerged aquatic vegetation that cover the bottom of shallow areas of the York River in the tidal fresh and oligohaline salinity zones of the river. The data is given in a table, with the number of hectares broken out by percent bed density. Below the table are three ways to visualize the data. Picking the most appropriate graph should be determined by what key message you're trying to convey as well as the ease of understanding the figure.

Submerged aquatic vegetation (SAV) area (hectares) in the York River

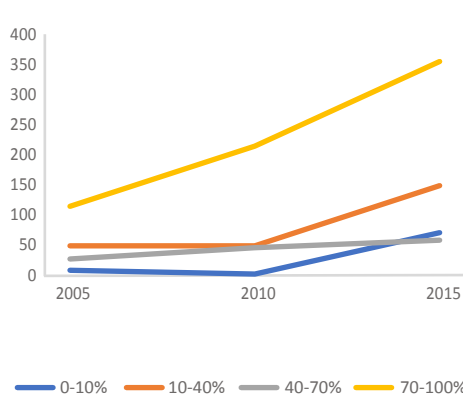
Sample site and year	Bed density				Total
	0-10%	10-40%	40-70%	70-100%	
York tidal fresh 2005	7	16	21	81	125
York oligohaline 2005	1	10	8	34	53
York total 2005	8	50	29	115	178
York tidal fresh 2010	2	22	13	73	110
York oligohaline 2010	0	28	32	143	203
York total 2010	2	50	45	216	313
York tidal fresh 2015	37	14	16	167	234
York oligohaline 2015	15	135	44	188	367
York total 2015	72	149	60	355	601

SAV area in the York River

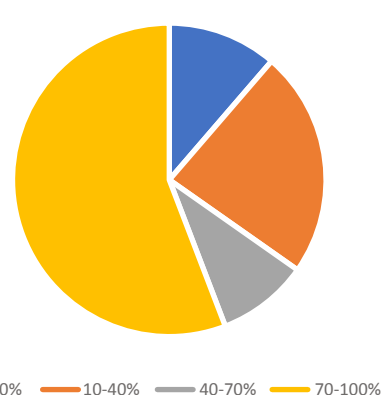
SAV area by sample site



Total SAV area



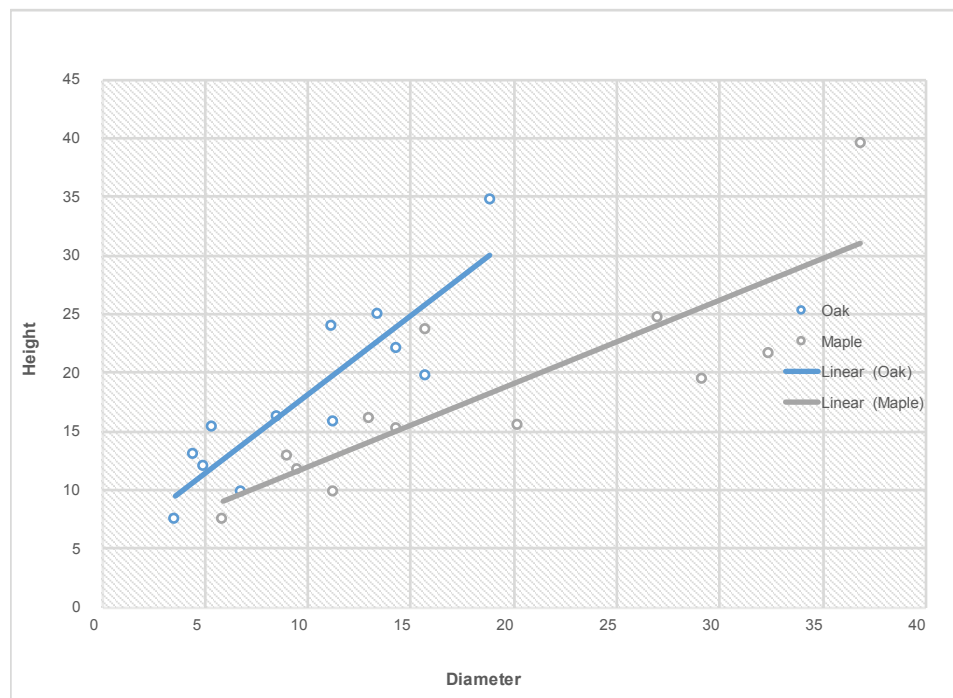
2015 SAV area



This data is for illustration only. This does not reflect actual SAV levels in the York River. This figure shows that different stories can be created from the same dataset depending on what graph or visualization you use to interpret the data.

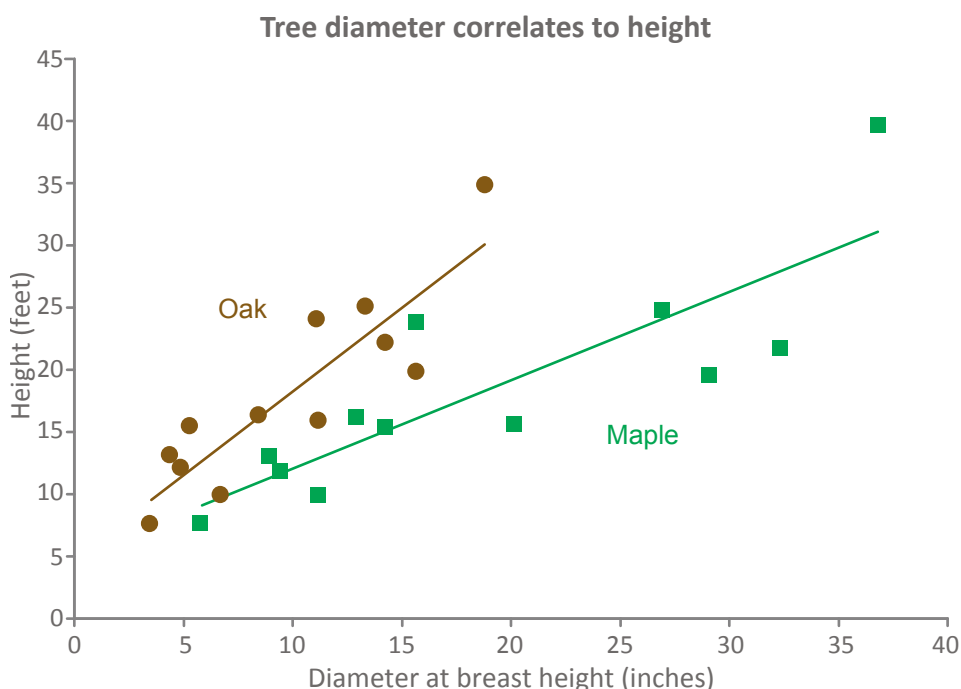
In these graphs, the data are connecting two measured parameters. This is a basic correlation graph, although to determine the actual correlation, you would need to perform basic statistics. But, the x-axis parameter generally relates in some way to the y-axis data, either increasing or decreasing as the y-axis increases. Examples include chlorophyll *a* and Secchi depth, stream water temperature and amount of pollution tolerant macroinvertebrates, and tree diameter with tree height.

EXAMPLE 8: BEFORE FORMATTING, COMPARISON LINE/SCATTER GRAPH



Comparison scatter plots can be visualized in many ways. Here is an example of the wrong way to format a comparison scatter plot. The grid lines are not necessary and the points are two different colors. The colors are not related to the data being presented, and the axis labels are too small and do not include units. There is also no title.

EXAMPLE 9: AFTER FORMATTING, COMPARISON LINE/SCATTER GRAPH



Here is the same scatter plot after cleaning. The colors are not very important, but by changing the points to different shapes and single, solid colors, the scatter of the data really pops. A title also improves communication of the message.

DATA IN FIGURES

Parts of a figure

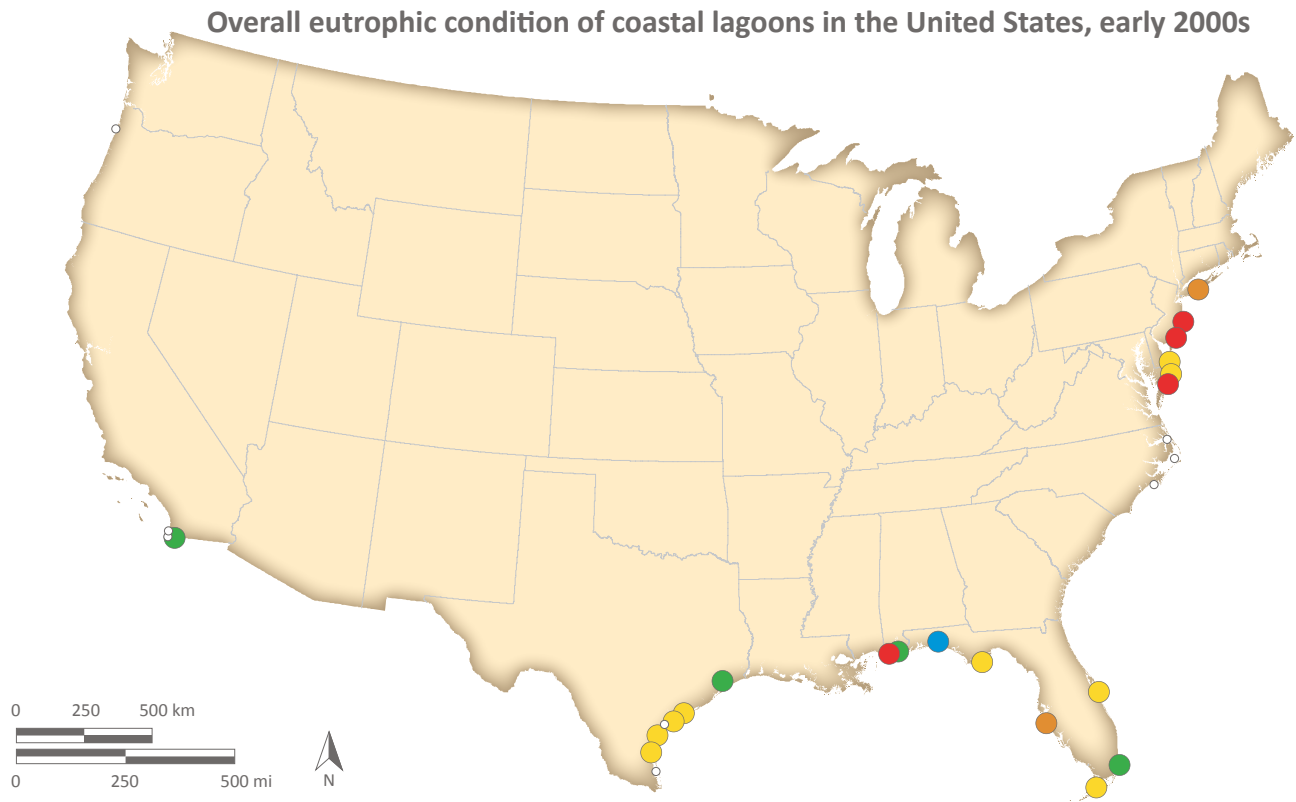
- **Maps**—Maps can be figures themselves or they can be included in a larger, synthetic figure. At a minimum, maps themselves should include a legend, scale bar, and north arrow. A smaller, inset map of the general geographic area may be needed to provide context to the reader.
- **Graphs**—Once a graph is made and designed for an audience, it can be incorporated into larger, synthetic figures.
- **Photos**—Photos are a great visual element for any audience. Incorporating one or more photos into figures can illustrate information, such as a pollution hotspot, different types of benthic species, or the location of a sampling site. Photos have to be cropped and saved at the correct resolution for the presentation or document that they will appear in. For more information, see Thomas et al. 2006.
- **Text**—Text in a figure includes short, descriptive statements, labels, and titles. Text in a figure should help the reader understand the figure without having to refer to the caption or the narrative (if in a document).
- **Caption**—The caption is usually an expanded title or legend that describes the figure and some more description of the figure that isn't included in the visual element (such as background information). Captions are usually used in a document, not in a PowerPoint presentation.
- **Title**—A main title for a figure should be a short description explaining the figure. Within the figure, a title could be used for individual elements, like a map, graph, or photo.

Design of a figure

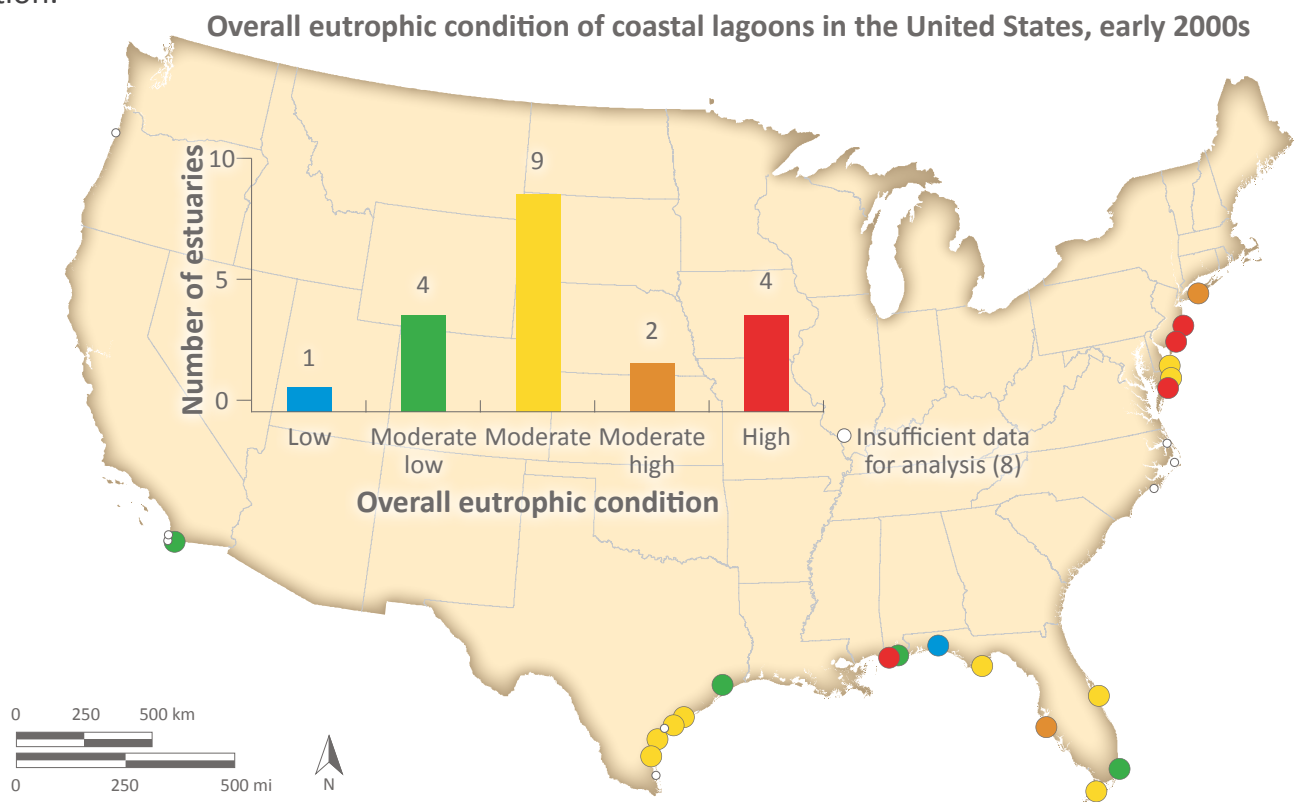
Designing a figure usually starts with a paper and pencil. Ask yourself what message you're trying to convey and what elements do you already have assembled that could be used to support your message. Think of the audience and the type of document (presentation or printed) the figure will appear in. Once you've sketched out the way you want the figure to look, assemble the different pieces in a software program. At this point, you will also want to think about what colors are appropriate for the figure.

FIGURE THAT INCLUDES BOTH DATA AND LOCATION INFORMATION

A basic map with a scale bar, north arrow, and title is the first step to creating more complex figures. Notice that the data on this map does not have a legend yet.



The author has added a legend by using a bar graph to summarize the data by overall eutrophic condition.



The author then adds more information by providing each data point's geographical location.

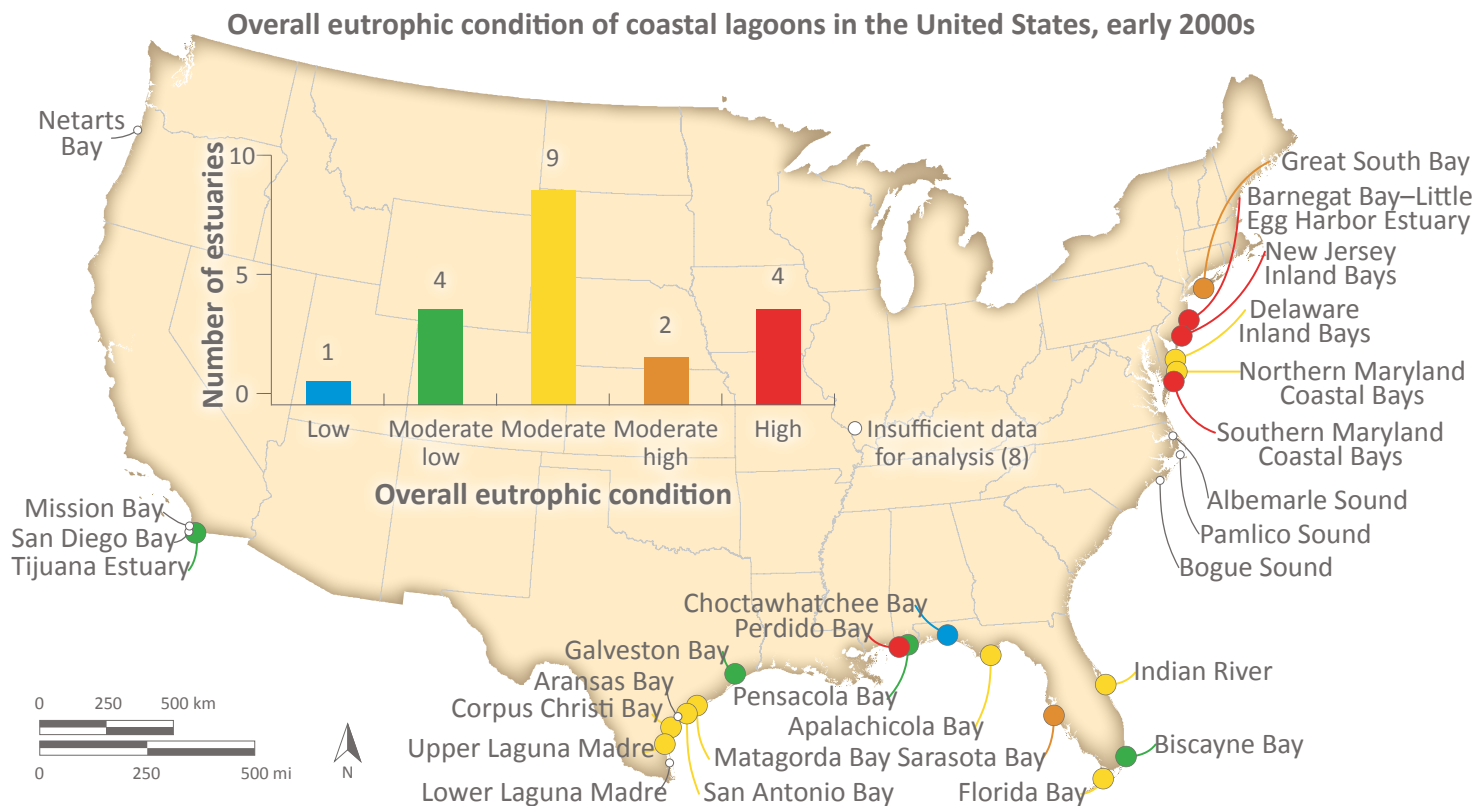
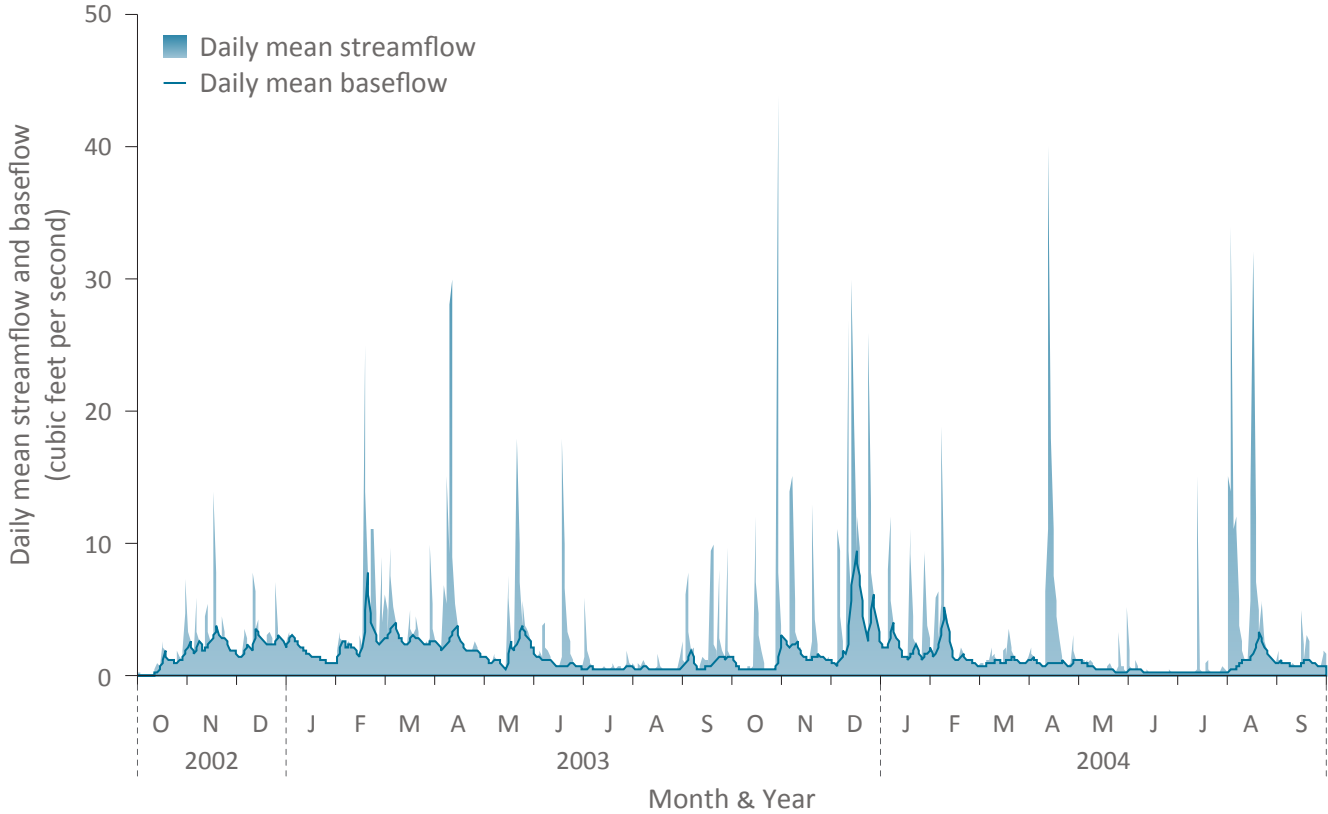
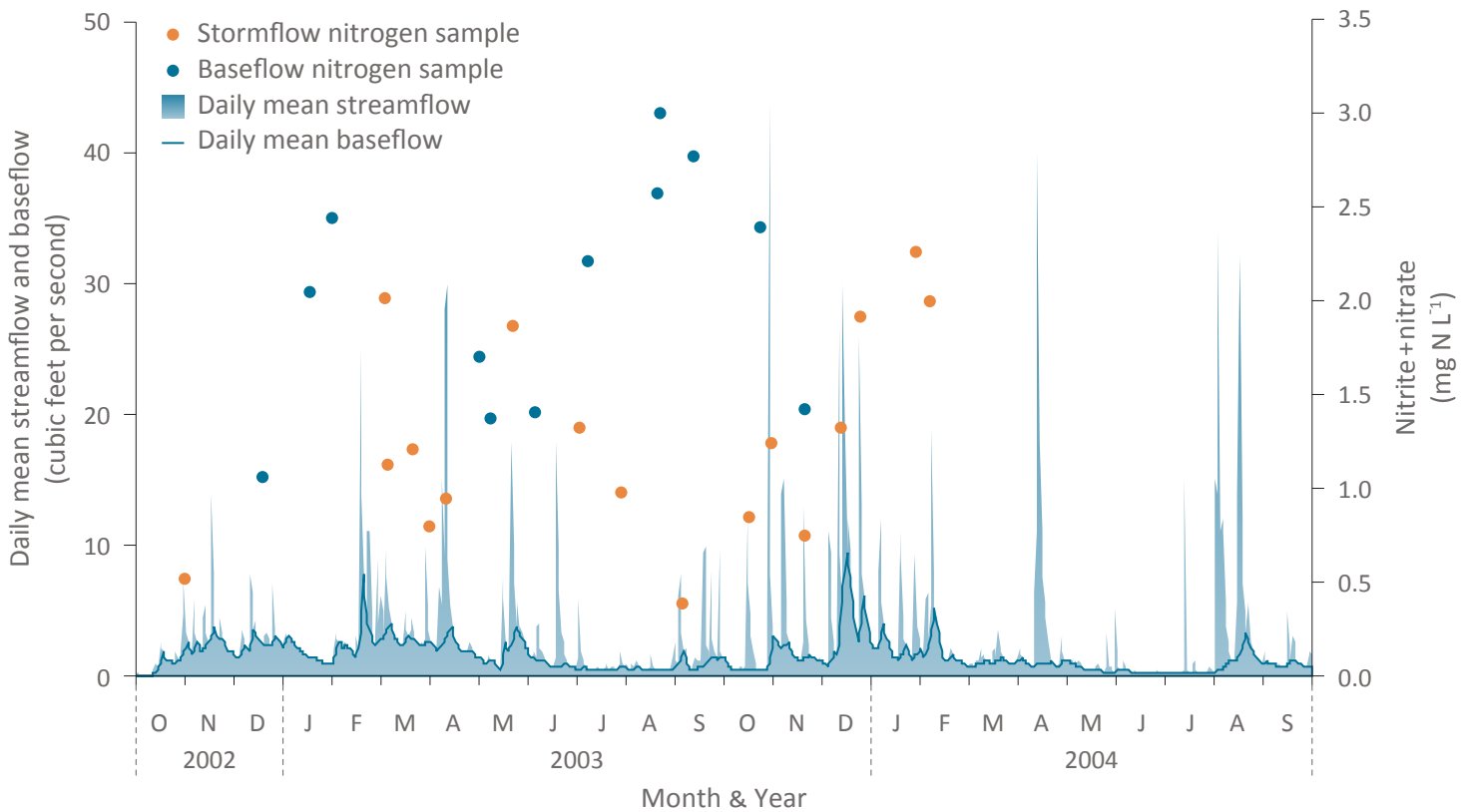


FIGURE THAT INCLUDES MULTIPLE ELEMENTS

This graph is the base from which a more complex figure can be built. The caption can be used to explain the data here, so no title is needed.

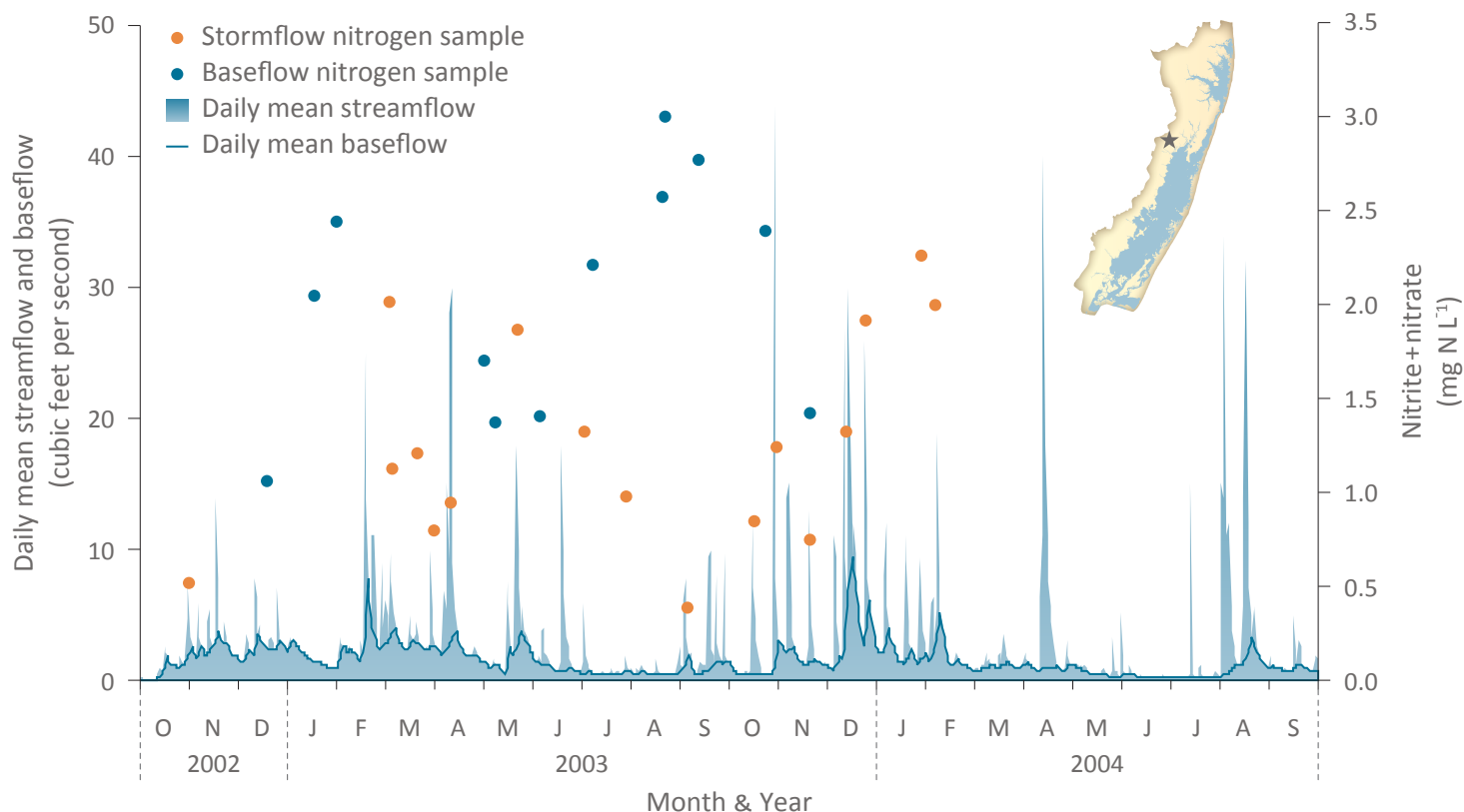


The author has then added a second data series, which goes on the secondary axis on the right of the graph. This allows different types of data to be shown together; an updated caption is necessary.



The finishing touches of this figure are to add a title that easy to understand quickly and an inset map showing the location of the sampling site. Less information would need to be in the caption, but the caption would still need to explain the data.

Flow rate & nitrogen concentrations in Bassett Creek



DATA INTERPRETATION SETS THE STAGE FOR SYNTHESIS

The contents laid out in the data interpretation section you just read provide the foundation for synthesis of your data into science communication products. Understanding how to clean your data, describing your data with statistics, and making graphs, figures, and tables have all been discussed in this section. A final step, which leads into more advanced synthesis and data visualization is combining simple elements into more complex figures, like the one above.

Once you have interpreted your data in these ways and are ready to integrate a variety of basic elements into a communication product, you can jump into the Synthesis section of this manual. Synthesis provides an overview of data visualization techniques and storytelling that you can use to engage and inform your audience about your data. Report cards are tools for communicating your story and also covered in the Synthesis section.

SYNTHESIS

VISUALIZATION AND STORYTELLING HELP SYNTHESIZE DATA INTO INFORMATION

While interpreting data is a critical aspect of any volunteer monitoring or research project and needs to be conducted before synthesis, data contained within spreadsheets or a database cannot be useful until it is properly synthesized into a story that is digestible by your intended audience. In *Houston, we have a narrative*, Randy Olson defines a story as "a series of events that happen along the way in the search for a solution to a problem." More on stories, narrative, and storytelling can be found later in this manual. But, the key point here is that a stated problem and an attempt at a solution is necessary to engage and inform your audience. The following pages will help you learn how to synthesize your interpreted data into key messages and an overarching story. The key messages, story, and report cards are the foundations of your communication strategy and should be given the proper amount of time and effort needed to be effective.



Data visualization is one key aspect of synthesis and includes proper design of graphs, figures, and tables as well as producing conceptual diagrams and infographics. These visual elements will be incorporated into products, such as presentations, newsletters, websites, report cards, and social media content as support for the overarching story that you will be telling.



Storytelling is another key aspect of synthesis and one that most people struggle to articulate and use in their scientific presentations and products. Storytelling includes giving key messages, using narrative, and providing context (Why should I care about your story?). Storytelling and science are not mutually exclusive and anyone producing scientific data should incorporate storytelling into their repertoire to be more effective with their intended audience.



Report cards are a commonly used tool for synthesizing and integrating data into a short, overarching story about ecosystem health. However, that isn't to say that report cards are basic communication tools. The amount of sampling, analysis, and technical skills needed to produce a report card should not be underrated. A special note about report cards: Only those groups that are collecting Tier 2 or 3 level data should be producing report cards.

For more extensive information on synthesis, please see the list of Supporting Materials at the beginning of this manual as well as the reference section at the end of the document. This part of the manual is here to provide an overview of ways to synthesize your data, but there are many topics covered in the resources and references list that will help you tell your story.

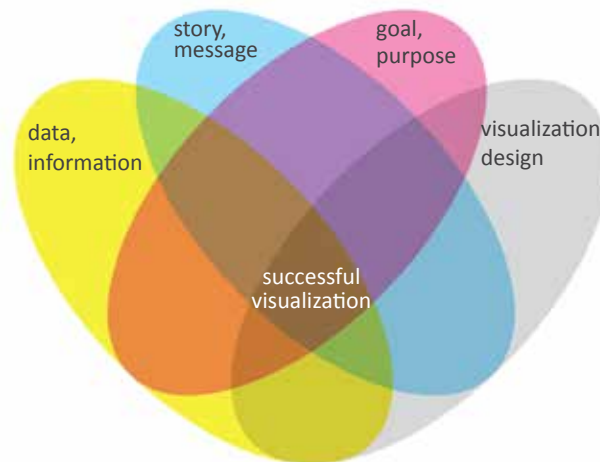
DATA VISUALIZATION

What is data visualization?

Humans are visual creatures. We interpret information by collecting data and presenting it in a visual form so that the information conveyed can be better understood. Effective visualization makes complex data more accessible, understandable, and usable to a general reader. To effectively communicate data to your audience, individual data developed in the Data Interpretation portion of this manual need to be combined into an integrated story about your ecosystem. This step—developing and communicating a data-driven story about your ecosystem that is visually appealing—is where the importance of data visualization comes in.

What makes a good visualization?

The goal of data visualization is to communicate information clearly and efficiently via tables, charts, statistical graphics, plots, and information graphics (info-graphics). Visualizing data helps us identify patterns in the data, and provides a unique, or new, perspective on the dataset. By starting with robust data and information, crafting a story and message, and then creating an effective visualization and design, you will create a successful visualization. But where do you begin with the process of visualizing data, and how do you determine the best visual representation for your data?



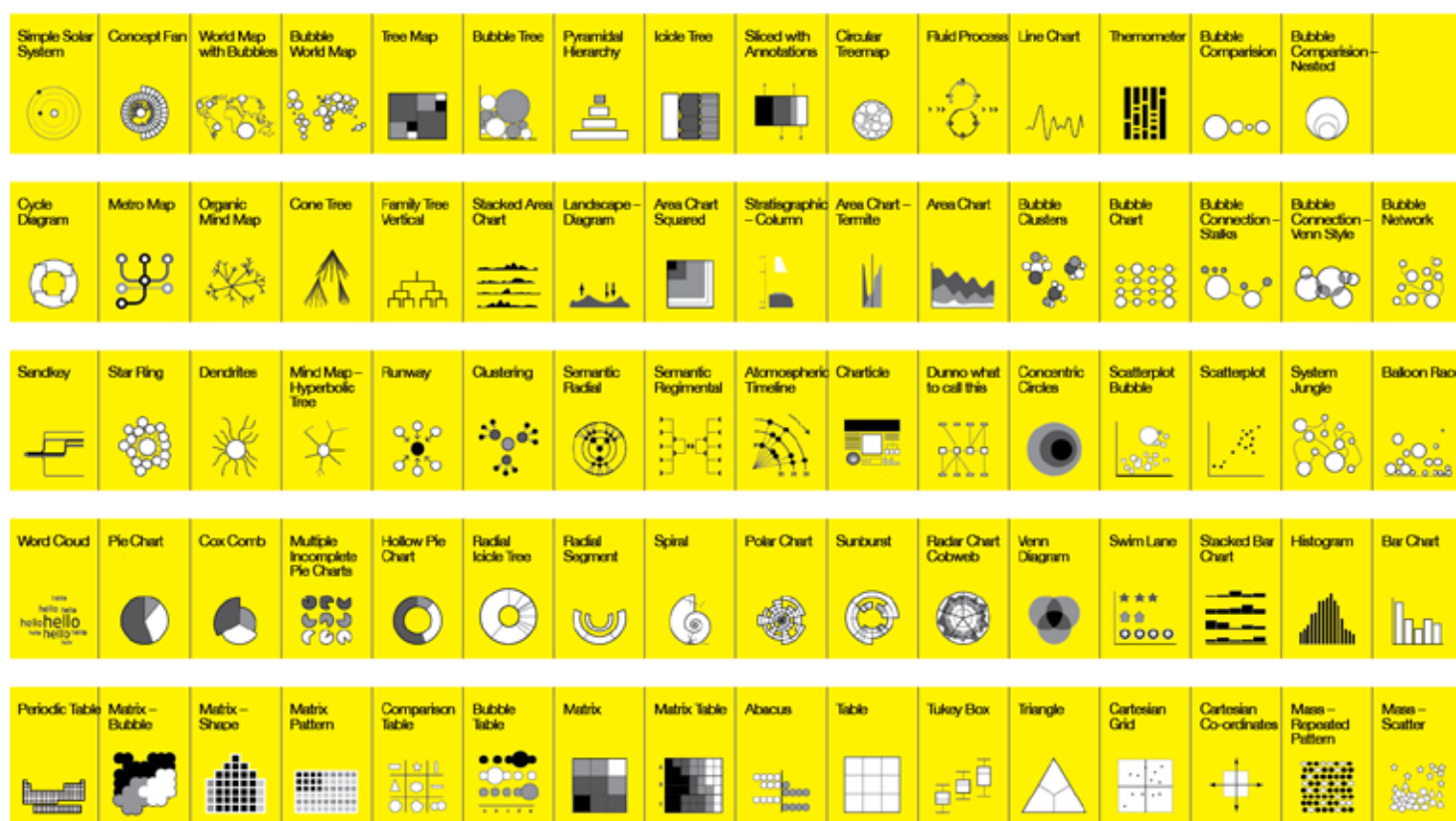
This visual created by David McCandless shows the four main elements (or ingredients, if you will) that go into the creation of any successful visualization.

Data visualization process

The data visualization process has several key steps:

1. **Data acquisition and organization.** The first section of this manual, Data Interpretation, is helpful in tackling this step of the data visualization process. Once data has been collected and organized, move on to thinking about the goal/purpose.
2. **Goal/purpose.** Think about the purpose of the visualization. How will it help the reader/audience to better understand the data? Think about your audience—what context will they need to understand the data? Will culture, prior knowledge, and accessibility limit the audience's ability to understand your data visualization?
3. **Visualization/design.** Chose the right chart/visualization type. There are many different types of data visualizations, and each can be tailored to fit the needs of a dataset. Once you know the goal of your visualization, you'll be able to better cater the type of visualization for your needs. Figure by David McCandless.

Examples of types of data visualizations



4. **Story.** Develop a compelling narrative. Identify key messages of the visualization. We will talk about narratives and their role in data synthesis in the next section of this chapter.

Balancing function with form in data visualizations

Use visual cues to help guide your audience through the data. Every choice—from size, color, and shape of each symbol to the way the symbol is positioned—needs to be made with your communication goals in mind.

- Use colors and symbols to emphasize important elements and help tell your story.
- Symbols can be used as visual aids to directly show your audience what you are talking about, rather than depending on text alone. Symbols are language independent and universal.

Symbols can represent something tangible

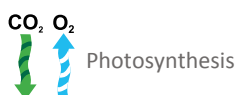


Callinectes sapidus (blue crab)









Vallisneria americana (wild celery)

Symbols can represent something invisible or intangible, like a chemical process (e.g. photosynthesis)



Symbols are consistently recognizable all over the world no matter a person's language, culture, or heritage:

mathematics π , weather , music , religion , corporate branding , signage , and organizations 

- Fonts can be used as a design element and for contrast.

Contrast with
SIZE

Contrast
with **color**

Contrast
with **TYPE**

- Use color, size, scale, shapes, and labels to direct attention to key messages.

Statistician Edward Tufte and Author David McCandless (*Information is Beautiful*) are both practitioners and teachers of data visualization. Tips from their publications on the data visualization process and best practices are included below.

Tips from Edward Tufte

- Integrate words, images, and numbers into one visualization.
- The data should drive the content and final product.
- The layout and design of a visualization should help the reader interpret the data.
- Minimize style variation (have a consistent look and feel) so the reader is focused on the data.
- Tailor the visualization to your reader.

Tips from Information is Beautiful

- Structure the visual in a hierarchy of information.
- Try to add a visual element, like an icon, glyph, or shape wherever possible.
- Less is more. What can you take out?
- Grids are good. Use a grid to harmonize your layout.
- Use colors deliberately. Have a color palette that reflects the topic of the data.
- Make text work harder. Text is a graphical element too.

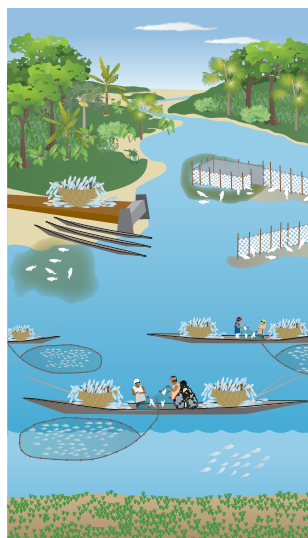
Tools for data visualization

Conceptual Diagrams






Conceptual diagrams are an essential component of environmental problem solving and a useful tool in your science communication toolbox. They can accompany your organized and evaluated data to provide synthesis, visualization, and context. A conceptual diagram is essentially a ‘thought drawing’ that can be used to communicate complicated processes and environmental concepts in a visually pleasing and understandable format. It can take many letters, words, sentences, and paragraphs to describe the processes that make up an ecosystem. Conceptual diagrams can take our current understanding of science, help to identify priorities and environmental values, and create a shared vision.

There are six important reasons to use conceptual diagrams:

1. To help communicate key messages and visualize scientific data.
2. To provide a better understanding of the ‘big picture’ that can’t be captured in a single photograph.
3. To span cultural boundaries and language barriers.
4. To better define words that are ambiguous by committing the image to the message being portrayed.
5. To better define scales of processes and linkages within habitats and communities.
6. To identify gaps in knowledge and solicit an agreed synthesis.








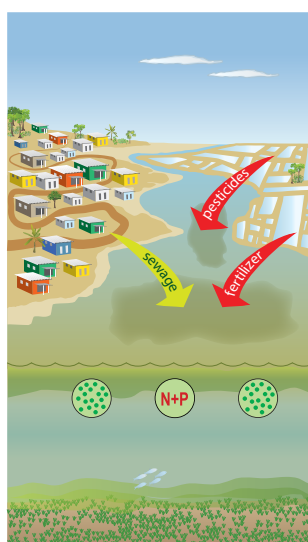
Pollution

The recent abundance of fish stocks is not sustainable with overfishing  and so many fishers . At landings, dead fish thrown back into the water contaminate the lake . Active shrimp pens and abandoned nets trap sediment  and kill juvenile fish .









Sedimentation

During monsoon season, an excess of sediment  is deposited in the Lake, mostly from Mahanadi River tributaries, nearby settlements , and agricultural lands . As the Lake becomes more shallow and its sea outlets fill in with sediment  increased flooding occurs .



Fishing and Aquaculture

As land use changes from forest to settlements  and paddy agriculture , sewage , and fertilizer and pesticides  runoff increases into the Lake. Algae blooms  that float and sit on the bottom are the result of that extra nutrient input .

This series of diagrams show some of the pressures facing a Lake Ecosystem in India. The symbols are embedded into the text to help explain what is happening in each scenario.

Visualizations provide synthesis and context

We use symbols and conceptual diagrams to provide a synthesis of scientific understanding, but in a graphic visualization that frames the message and provides context.

Steps to create a conceptual diagram

1. Identify the audience. Who is the target audience and what is the medium of the final product (print, digital)?
2. Develop an issue statement. Write one simple declarative sentence describing the conceptual diagram.
3. Prioritize key features. List the features needed to describe your message and prioritize to eliminate any more than seven.
4. Prioritize the major drivers. List the drivers or threats and prioritize.
5. Make a short list of symbols. List the symbols that will enhance the audience's ability to understand the concept. Don't forget a legend.
6. Hand draw a sketch with a base and symbols. What is the available space and appropriate context (2D, 3D).
7. Create with software. Some resources for creating successful, science based conceptual diagrams include:

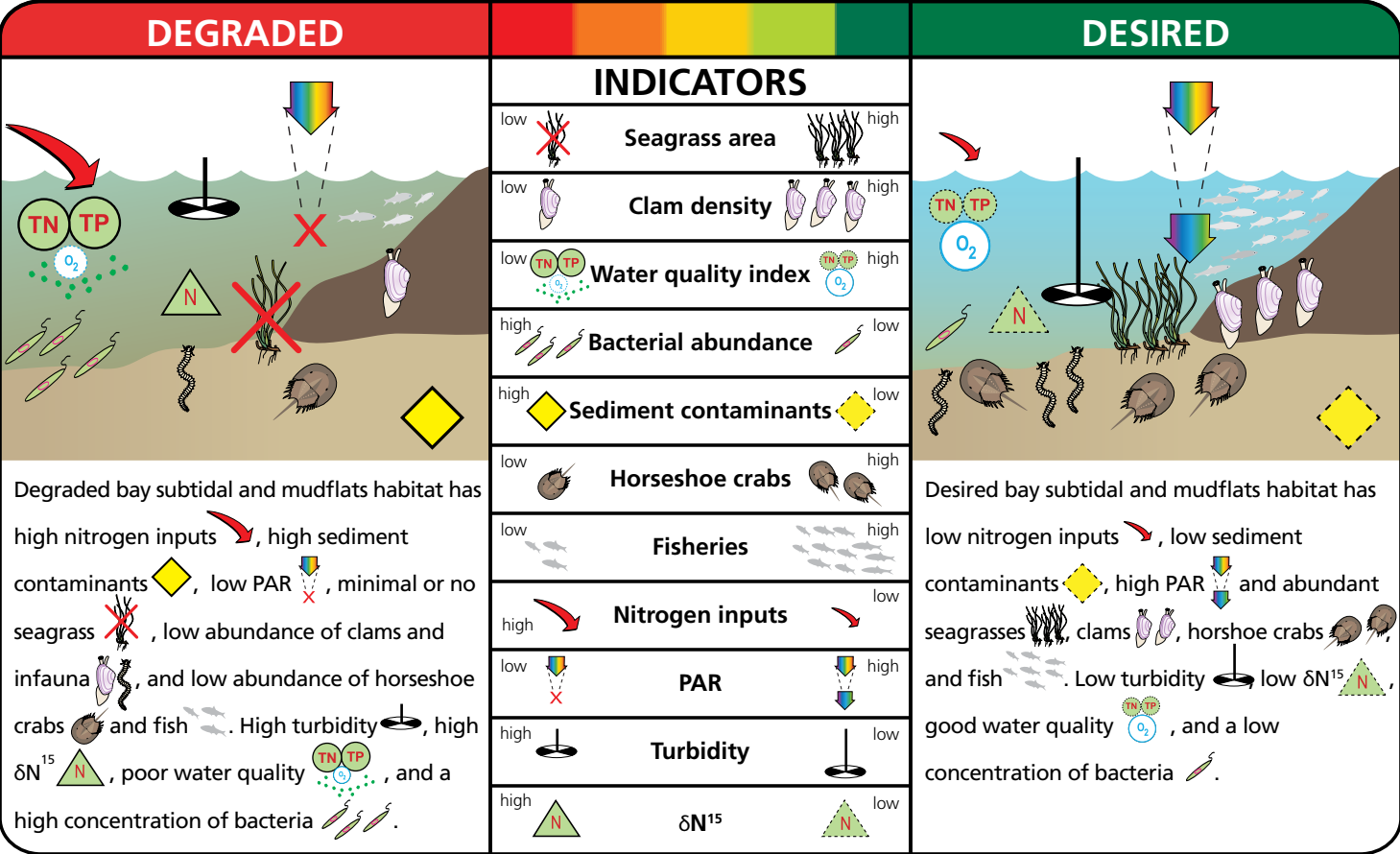
Software—Adobe Illustrator, SVG Edit, or Inkspace

Symbols and bases—The Integration and Application Network (IAN) has produced a series of scientific symbol libraries. The libraries contain over 2,800 custom-made symbols designed specifically for enhancing science communication. The IAN symbol libraries are available cost- and royalty-free with appropriate attribution (ian.umces.edu/symbols), and allow for the creation of diagrams and other data visualizations without being an expert graphic designer or artist.

Visualizations should be compelling. The diagram to the right was combined with Chesapeake Bay climate change projections to provide the context of sea-level rise to a general public audience.



BAY SUBTIDAL AND MUDFLATS

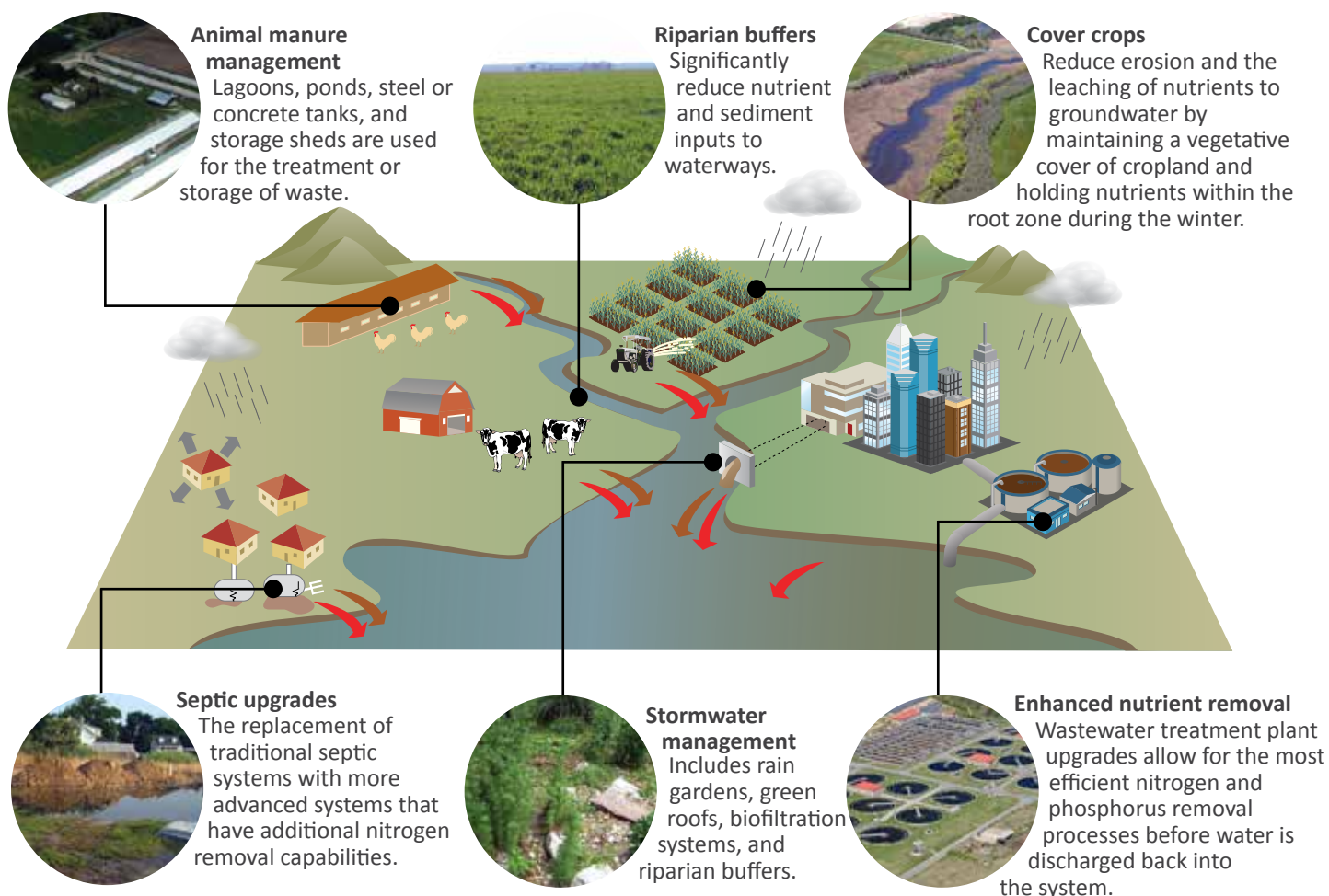


In the diagram above, symbols are incorporated into a conceptual diagram illustrating degraded vs. desired conditions in the sub-tidal region of the Maryland Coastal Bays. The size, number, and color of symbols are used to portray additional information.

SYNTHESIS

Infographics

Visual elements can be combined to provide unique information. For example, a combination of a photo and a conceptual diagram can effectively orient the audience to your study site, or explain methodology. Photos and graphs together can help with the visualization of your results. Results can be overlaid on to maps, which helps the audience envision the overall context of your results.

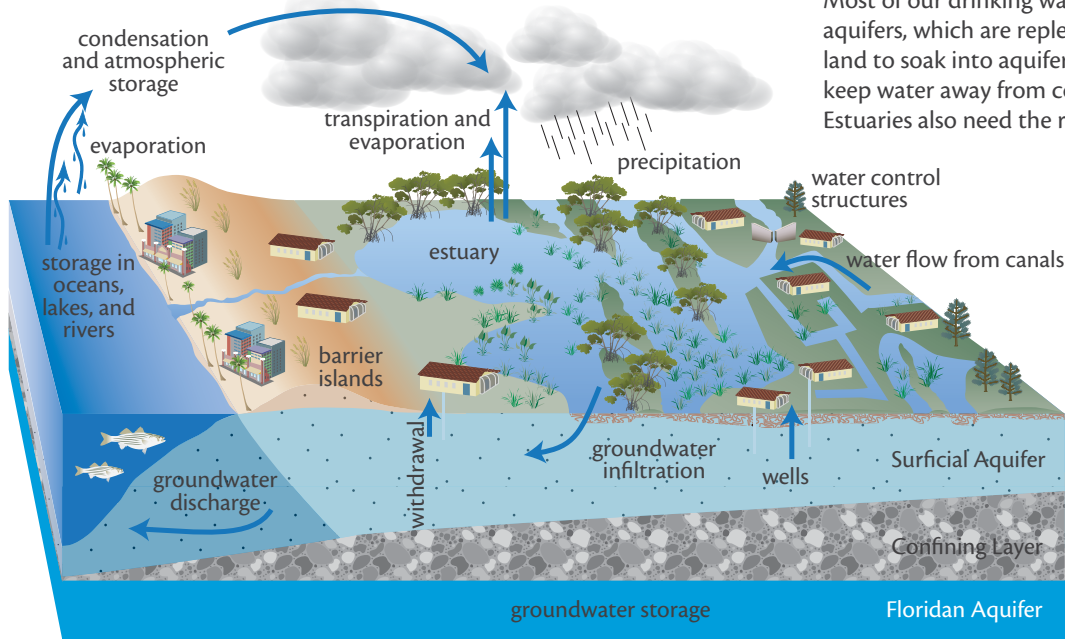


The diagram above illustrates an array of best management practices (BMPs) that when used together, target the numerous sources and delivery pathways of excess nutrient loads to a body of water. Photographs of each BMP were included with the conceptual diagram to give the reader an image in their head of each practice, so they can relate them to their own experience or interaction with each practice.

The infographic on the right was created as part of a four-page newsletter highlighting the features and threats to Rookery Bay National Estuarine Research Reserve in southwest Florida. Three conceptual diagrams were combined with symbols and data to highlight the importance of an estuary, and how Rookery Bay contributes to the economy, and the natural features such as system provides.

The Water Balancing Act

Maintaining **aquifers**, preventing **flooding**, and protecting **estuaries**

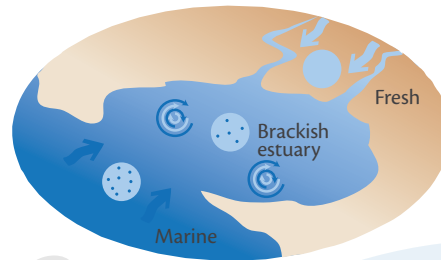


Most of our drinking water comes from underground aquifers, which are replenished by rain. Holding water on land to soak into aquifers is important, as is the need to keep water away from communities to prevent flooding. Estuaries also need the right amount of fresh water from land to stay healthy. The biggest challenge to water managers is to maintain estuarine health despite the need to recharge aquifers and prevent flooding.

49 million
49 million gallons are pumped from the Surficial Aquifer every day.

Estuaries are the **heart** of Rookery Bay

Estuaries are areas where **fresh water** from streams, rivers, or canals mixes with **salt water** from the ocean, resulting in **low salinity** (brackish) water.



70%
70% of rainfall occurs from May through October, while peak water usage by seasonal visitors and agriculture occurs from January through March, which is during the dry season.

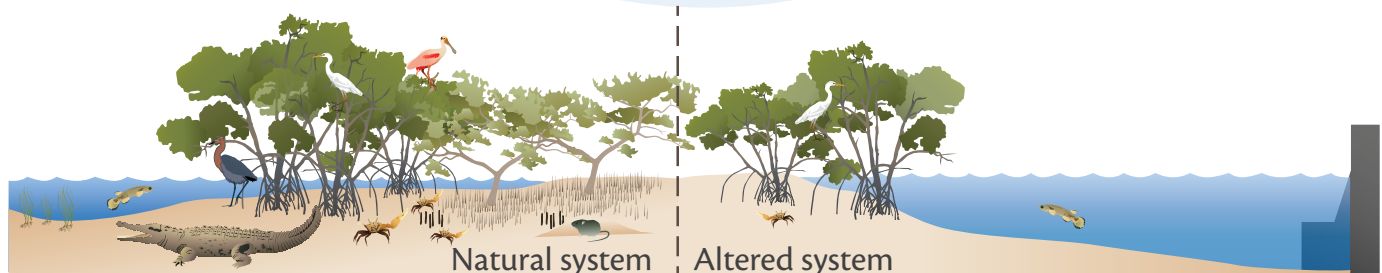
\$1.2
Tourism generates \$1.2 billion each year in Collier County.

Estuary wetlands help protect communities from **flooding** and **storms**, while removing pollutants from the waters heading to the Gulf of Mexico.

500,000
500,000 pounds of stone crab claws are harvested annually in Collier County, which includes Rookery Bay Reserve.

750,000
750,000 pounds of other seafood is harvested in Collier County annually.

\$23 billion
Coastal wetlands provide storm protection that saves \$23 billion per year.



STORYTELLING

Introduction

To better engage your audience, combining narrative and visuals for science communication is essential in generating an illustrated, factual story. Stories are more interesting than a string of facts and data, and when combined with compelling visuals, they improve knowledge retention and shorten reading time. It is easier to remember good stories and effective visual elements, and the audience/reader can follow a storyline and view visual elements quickly. Science narratives should focus on informing, not persuading, the audience.

The history of narrative and visual communication

Storytelling is an ancient form of human communication. People have gathered around campfires and in caves since the early evolution of humans. They would regale each other with stories, and in many cultures, these stories are what constitute their shared oral history.

Stories are both easier to recall and easier to remember than facts alone, so storytelling is an effective way to communicate important aspects of culture. Michael Dahlstrom, a professor of journalism and communication at Iowa State University, has identified three key elements of good storytelling: causality (relating cause-effect relationships), temporality (a beginning, middle, and end), and character (telling the story in terms of people).

Both narrative structure and visual processing are hard-wired into the human brain. Brain scans of people watching a story unfold on video show their brains are highly activated and they display the same activation as a group, compared to people watching videos devoid of a story (people walking through a public place). Brain scans of people viewing different color hues show a higher activation in some colors (red) than other colors (green). These observations show that our neurobiology makes us particularly attuned to stories and visual cues like color.



Stone age painting depicting a man riding on a horse. This painting was found in the Bhimbetka rock shelters in India. Source: Photo by LRBurdak / CC BY-SA 3.0.



Albert Anker's painting, *Der Grossvater erzählt eine Geschichte*, 1884 (*The Grandfather tells a story*). Source: Wikimedia Commons.

What makes a good story?

Narrative stories have similar essential elements: character, plot, setting, point of view, and style. These can be broken into three main themes:

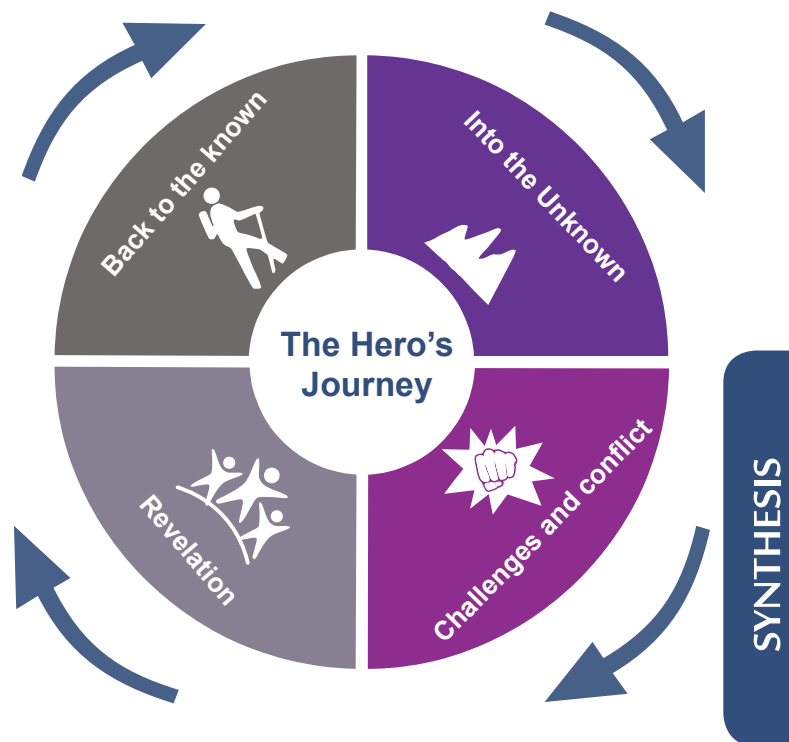
1. Causality: cause and effect relationships
2. Temporality: timeline with past, present, future.
3. Character: they involve people.

The classic storytelling structure-the Hero's Journey, was identified and outlined by Joseph Campbell, an American mythological researcher. In this structure, which is used in many modern movies and books, the Hero, or main character is called to action, leaving a safe, known space, for the unknown. During this time in the unknown, conflicts will arise, he/she will face challenges and temptations, and a revelation is made, allowing a return to the known.

Telling your story

While scientific narratives may not include the classical components of the Hero's Journey, scientific research is linked, as it contains many of the same elements as a narrative: a call to action, an adventure into the unknown, and an eventual resolution or transformation of thought.

Science narratives should focus on informing, not persuading the audience. The intent of narratives in science communication is important. The focus should be on providing basic understanding and knowledge to be used in decision-making or for developing informed opinions. The use of science to persuade the audience becomes advocacy, rather than serving as an honest broker. Informing the audience gives them information which can be used to persuade them to think differently. Persuading the audience gets something from them and sells them something.



The Hero's Journey is a classic storytelling structure described by Joseph Campbell. This structure can be applied to scientific narrative.

"It is one of the novelties of human language that it encourages storytelling and thus creative imagination. Scientific discovery is akin to explanatory storytelling, to myth making and to poetic imagination."

- Karl Popper, Philosopher

While persuasion can inform the audience, it is a slanted viewpoint if the intent is simply persuasion and not education.

Provide context: Why should the audience care?

The most important question to address is “so what?”. The audience needs to care enough to try to understand what is being communicated. Thus, it is important to provide both the scientific context and the societal context of the science being communicated. This context setting should be covered early in the science communication, rather than waiting until near the end. Starting out with explaining ‘why’, and then following up with ‘how’ and ‘what’ allows the science communicator to engage with the audience early, create interest and the follow up with more explanation. Once the audience understands why you are using science to tackle a question or issue, they are far more likely to want to know more about how you are going about the research.

Defining your message

Often times, the documents we are creating must be short and concise, which makes it unrealistic to include every aspect of your research or results. In Randy Olson's book, *Houston We Have A Narrative*, he introduces a myriad of tools that can assist scientists and communicators in identifying key messages, and telling their stories.

Randy Olson has an activity he termed the “The Dobzhansky template”, named after the geneticist that used it first, that highlights key words and allows the author to hone in on the general theme of a narrative.

The template reads "Nothing in _____ makes sense except in the light of _____."

“Scientific theories... begin as imaginative constructions. They begin, if you like, as stories, and the purpose of the critical or rectifying episode in scientific reasoning is precisely to find out whether or not these stories are stories about real life.”
- Peter Medawar, Biologist



Randy Olson's book, 'Houston, We Have A Narrative', is a helpful resource when crafting your science story.

The point of focusing on key words is that it provides a clear focus on the most important aspects of a science communication product. Distilling the message into key words provides an impetus to collect the photographs or other visuals to illustrate these attributes elucidated by keyword selection.

Using “And...but...therefore”

ABT is shorthand for a sentence structured with “. . . and. . . but. . . therefore. . .”, and can be used to outline writing short paragraphs, and even applied to full text documents. The ABT format sets up an issue with the statements linked by ‘and’, in which tension is created using the ‘but’ statement, and is followed by the resolution using the ‘therefore’ statement. The format guides the author into a strong narrative structure. A contrast to the ABT is AAA format of “. . . and. . . and. . . and. . .” in which repetitive statements or strings of facts become monotonous and boring. Another contrast is the DHY format “. . . despite. . . yet. . . however. . .” in which the reader is taken on a circuitous journey with no resolution; often resulting in a series of caveats.

Here is an example of the AAA statement transformed into an ABT statement:

*The Chesapeake Bay is the largest estuary in the United States, **and** it is an incredibly complex ecosystem, **and** its rivers, wetlands, and forests provide homes, food, and protection for countless animals and plants.*

*The Chesapeake Bay is the largest estuary in the United States, **and** it is an incredibly complex ecosystem, providing habitat for countless animals and plants. **But**, the health of the Bay is degraded by pollution. **Therefore**, it is important to continue monitoring efforts.*

Keyword

ian.umces.edu

Nothing in _____
makes sense, except in
light of _____.
(keyword)

Keyword

ian.umces.edu

Nothing in conservation
makes sense, except in
light of people.
(keyword)

Science communication students were given a blank card (above) with The Dobzhansky template, and instructed to use the statement and blank space to identify their own keywords (bottom).



Students at a science communication course work to structure their writing in an ABT framework. Photo: Jamie Currie.

Randy Olson has developed the ABT concept and an Index of Narrative Strength. The narrative index looks at the use of 'and' and 'but' within a given length of text. Based on the ABT framework, text with a greater frequency and use of 'but' relative to 'and', the more narrative strength there is to the content being presented. The Narrative Index analysis has shown that people using narrative structure are much more effective communicators than those who do not.

ABT Statement ian.umces.edu

_____ **and** _____

_____ **but** _____

_____, **therefore** _____

ABT Statement ian.umces.edu

Cities are home to over 50% of people and provide considerable ecosystem services but are often ignored in research, therefore we need to pay more attention to urban nature.

Science communication students were given a blank card (left) with the ABT structure, and instructed to use the space to develop their own ABT statement (right).

Applying the Hero's Journey for narrative structure

A longer publication or communications product can be developed using the ABT structure embedded into the Hero's Journey. This allows for the same progression, tension, and resolution, but for the entirety of the scientific narrative. Randy Olson supplies the template below for developing a complete narrative, where each of these nine steps would be comprehensive sections of the narrative:

"In an *ordinary world* _____

a *flawed character* _____

has a *catalytic event* _____

which upends his/her world. But after *taking stock* _____

the character decides to *take action*, _____

but when the *stakes get raised* _____

the character must *learn the lesson* _____

in order to *overcome the opposition* _____

and *achieve the goal* _____."

Once each of these nine steps are identified, the author can go ahead and apply data visualizations to support each section of text.

Active titles enhance comprehension.

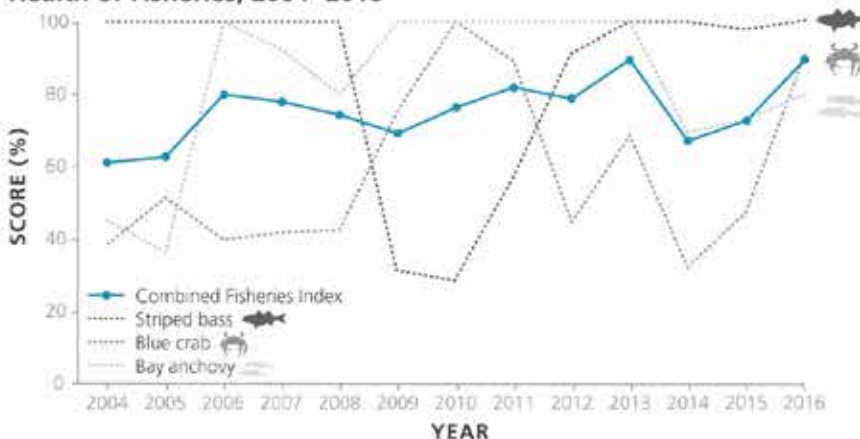
The use of titles that are declarative statements provides the reader with the take home point up front. Active titles allow the reader to understand the intent of the supporting text or visual elements. Passive titles like 'Results' or 'Comparisons' only provide readers with guideposts for the text, but active titles are more engaging.

More gains for fish populations



A striped bass is handled by Captain Pete Ide at Fishing Creek in Chesapeake Beach, MD. Photo by Chesapeake Bay Program.

Health of Fisheries, 2004–2016



Four indicators of fisheries health showed widespread score improvements in 2016.

Striped bass, bay anchovy, and blue crab are ecologically, economically, and socially important fish species in Chesapeake Bay, and are incorporated into the Combined Fisheries Index. While the Index is variable over time, the 2016 score of 90% was a large improvement over the previous year's score of 73%. This is the highest the fisheries score has been since the start of the report card. The Index is an average of all three species scores. Striped bass held steady with a 100% score, while both blue crab and bay anchovy improved. Blue crab scored a 90% and bay anchovy scored an 80%. Fisheries indicators can vary greatly from year to year since these species move throughout the Bay and beyond.

The use of the active title "More gains for fish populations" succinctly summarizes the content found in three different elements of this page—a data figure, a photograph, and accompanying text. Before reading further down the page, the audience automatically knows what is happening to fish populations.

The importance of combining narrative and visuals

When determining the best way to synthesize and interpret data, and communicate your science, it is vital to combine narrative and visuals to create a compelling story. Narrative and visuals engage the audience. Stories are more interesting than a string of facts and good visuals are compelling to view. Narrative and visuals are memorable. It is easier to remember a good story than disassociated vignettes and good visuals are easy to recall. Narrative and visuals shorten reading time. Readers/viewers can follow a storyline more readily and visuals can be quickly scanned.

REPORT CARDS

Report cards are a popular and effective tool for promoting numerous issues, ranging from bacteria levels at swimming beaches to the ecosystem health of freshwater streams. They serve as an important outreach method for generating community interest and increasing citizen understanding of ecosystem health, water quality, and watershed issues. Typically, after a report card is 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 an increase in applications by citizens for new systems.

Report cards can be created for freshwater or saltwater systems, or in this document, tidal and nontidal. Different indicators can be used depending on the system. While some indicators are the same for tidal and nontidal waterbodies, others are different. The following sections go into detail about tidal and nontidal indicators, report card analysis and scoring, and communication in a report card document.

TIDAL INDICATORS

This and the following sections 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 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 Resources and References).

Core indicators

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.

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.

Sampling and data analysis

An overview and methods for sampling and data analysis are provided for each indicator in the following sections. A summary table of preferred and minimum recommendations is provided here. 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 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 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.

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 <i>a</i>	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










Thresholds

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. 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.

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

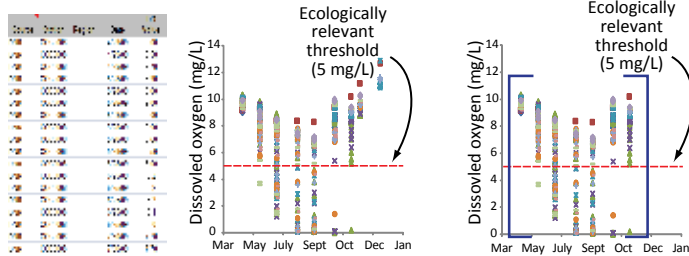
The core indicators used in this protocol and examples of threshold values used to compare observed data to the reference community.

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.

1. Sort data by station and date

2. Plot data to see temporal pattern in order to determine months to assess

3. Separate out data for determined months



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 the figure below, 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·l⁻¹, it fails because the threshold is 5.0 mg·l⁻¹. However, it is much closer to passing than a value of 1.0 mg·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.

1. Sort data by station

2. Calculate score for each data point
Ex: If DO ≥ 5 mg l⁻¹, then Score = Pass

3. Calculate the score for each station
Ex. ((Total # of scores = Pass)/(Total # of scores for that station))*100 = % total

Source	Station	Region	Date	DO
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00
DNR	0000008	Mid-Missouri	4/29/09	9.00

Source	Station	Region	Date	Depth	DO Value	Threshold	Score
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P

Source	Station	Region	Date	Depth	DO Value	Threshold	Score	Station Score
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	100
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	
DNR	0000008	Mid-Missouri	4/29/09	1.0	9.00	5.00	P	

5. Calculate region scores by averaging all station scores per region.

To calculate overall lake or watershed score, sum the region values weighted by % of total area.

Source	Station	Region	Date	DO (mg/L)	pH	%	Station Average	Region	Score	Area (sqm)	Percent	Weighted score
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000
DNR	0000008	Mid-Missouri	4/29/09	9.00	p	100	100	Mid-Missouri	100	1000000000	100	1000000000

4. Station scores for each station

Source	Station	Region	Station Score
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100
DNR	0000008	Mid-Missouri	100

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

Multiple thresholds

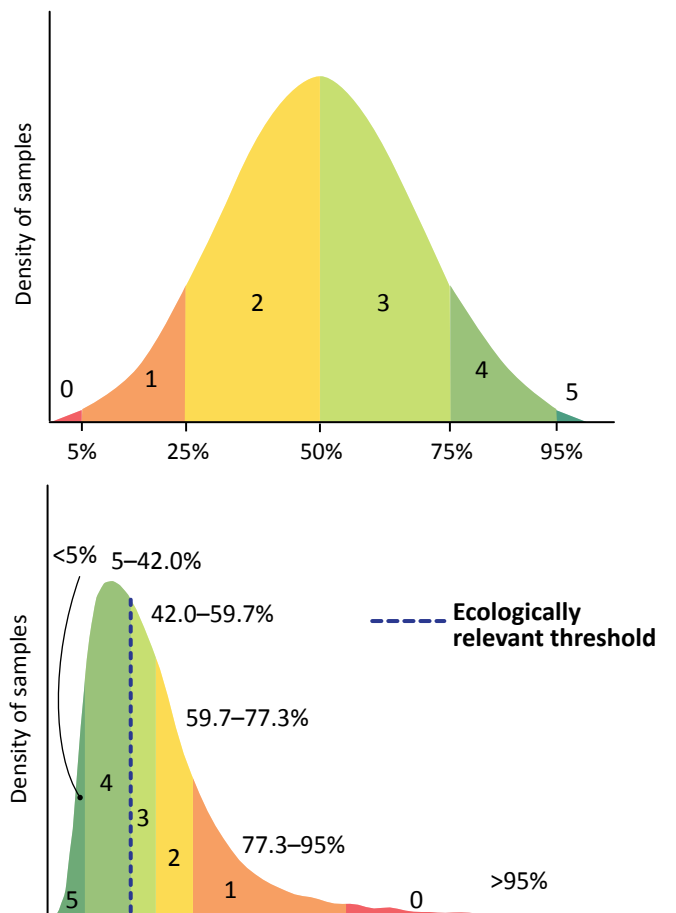
Multiple thresholds are used to score indicators based on a gradient of healthy to unhealthy conditions. For example, total phosphorus is an indicator of the amount of phosphorus in the water system. However, the amount of phosphorus, from acceptable levels, to just a little bit too much, to a truly 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 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. 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. 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.

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		

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



Example frequency distribution (top)—scores are divided equally among percentiles. Example frequency distribution (bottom)—scores are anchored by an ecologically relevant threshold, then divided equally among percentiles.

Scores are standardized

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 sections), 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. 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.

Summary

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

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

A grade and description are assigned based on the score that the indicator or subregion achieves.



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.

Descriptions of ecological health that correspond with each grade.

Indicator data analysis

Tidal: Dissolved Oxygen

Dissolved Oxygen (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. 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.

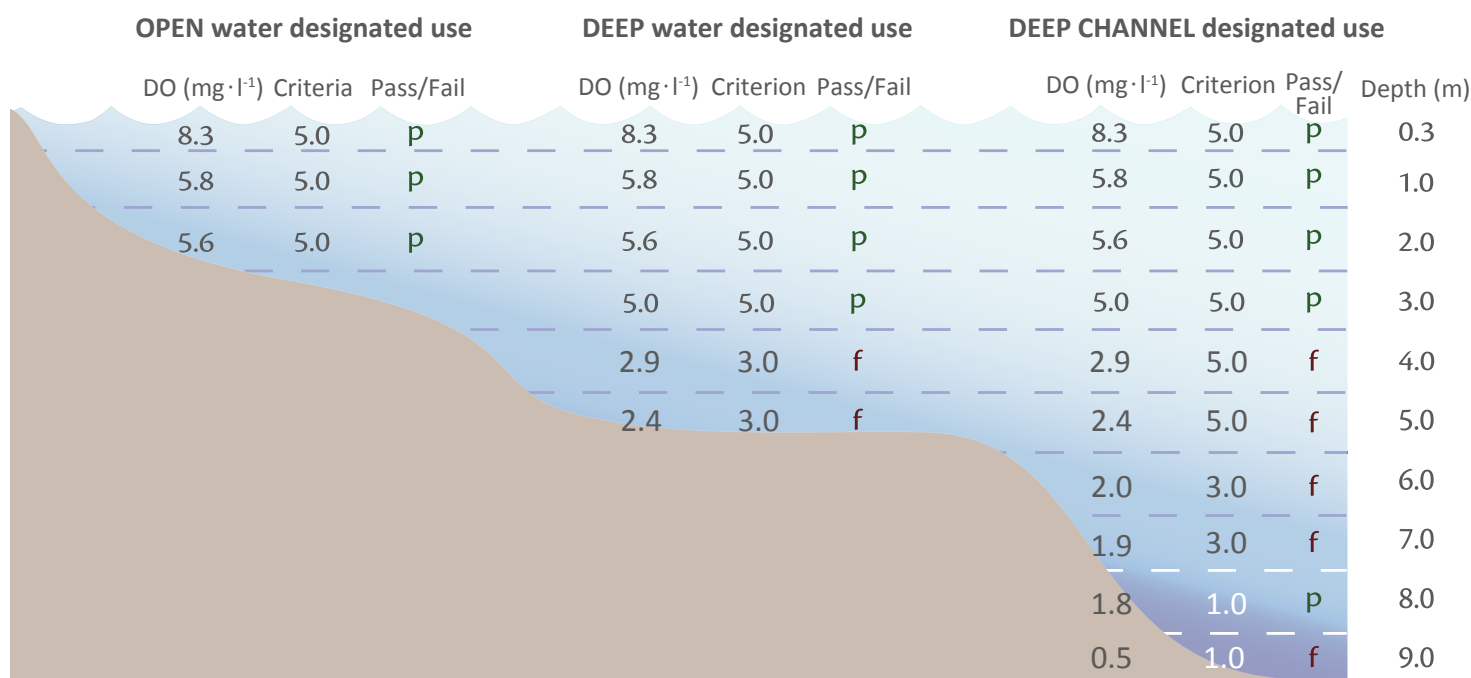
Stratification

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. 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 $0.1 \text{ kg}\cdot\text{m}^{-3}\cdot\text{m}^{-1}$ for the upper boundary and deepest change of $0.2 \text{ kg}\cdot\text{m}^{-3}\cdot\text{m}^{-1}$ 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.

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. Please contact the Chesapeake Bay Program for the latest pycnocline and designated use areas.

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. Sites where no stratification is expected are open water designated use areas, and all measurements in those areas use a $5.0 \text{ mg}\cdot\text{l}^{-1}$ criterion.



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

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 $5.0 \text{ mg}\cdot\text{l}^{-1}$.

For example, where a single stratification layer is evident (deep water designated use areas), the $5.0 \text{ mg}\cdot\text{l}^{-1}$ criterion will apply to samples above the pycnocline, and a $3.0 \text{ mg}\cdot\text{l}^{-1}$ 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 $1.0 \text{ mg}\cdot\text{l}^{-1}$ 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 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 subregion 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 ($5 \text{ mg}\cdot\text{l}^{-1}$ for open water, $3 \text{ mg}\cdot\text{l}^{-1}$ for deep water, or $1 \text{ mg}\cdot\text{l}^{-1}$ 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.
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.

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.

Tidal – Chlorophyll *a*

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 (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.

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).
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 = $3.8/5.0 = 0.76 \times 100 = 76\%$).
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.

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.

8. 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 km² out of a total 20 km² = 0.25, 76% x 0.25 = 19%).
9. Sum the resulting sub-region scores into an overall score.
10. 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.

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

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

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

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

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

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

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

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

Tidal – Water Clarity

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. The four major salinity regimes are tidal fresh (0–0.5ppt), 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 = $3.8/5.0 = 0.76 \times 100 = 76\%$).
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.

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 km² out of a total 20 km² = 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.

Tidal – Nutrients

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. The thresholds for total nitrogen and total phosphorus are different, so make sure the appropriate threshold is 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 = $3.8/5.0 = 0.76 \times 100 = 76\%$).
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.

Total nitrogen ($\text{mg}\cdot\text{l}^{-1}$) 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	$>0.6-\leq 0.9$	$>0.6-\leq 0.9$	$>0.5-\leq 0.6$	$>0.4-\leq 0.5$
3	$>0.9-\leq 1.3$	$>0.9-\leq 1.2$	$>0.6-\leq 0.8$	$>0.5-\leq 0.6$
2	$>1.3-\leq 1.8$	$>1.2-\leq 1.6$	$>0.8-\leq 1.0$	$>0.6-\leq 0.8$
1	$>1.8-\leq 2.8$	$>1.6-\leq 2.8$	$>1.0-\leq 1.5$	$>0.8-\leq 1.2$
0	>2.8	>2.8	>1.5	>1.2

Total phosphorus (mg·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-≤0.14
0	>0.23	>0.28	>0.15	>0.14

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 km² out of a total 20 km² = 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.

Tidal – Aquatic grasses

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. Examples of these goals for some river systems are provided below.

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

While aquatic grass hectares are calculated by the Virginia Institute of Marine Science's (VIMS) group, the total number of acres for your waterbody need to be 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 subregion 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.

Tidal – Bacteria

The EPA threshold for enterococci in swimming and contact recreation areas is 104 MPN 100 ml⁻¹ when using IDEXX (USEPA 1986) and 104 cfu 100 ml⁻¹ when using membrane filtration. Based on EPA guidelines, risk for swimming-associated illness is too high when the criterion (104 MPN 100 ml⁻¹) is exceeded. The EPA threshold for *E. coli* is 235 organisms 100 ml⁻¹ for any single water sample.

Data provided from the laboratory are analyzed to calculate a percent of samples below the appropriate (enterococci or *E. coli*) threshold. The percent of samples in a sampling season (Memorial Day to Labor Day) that a sample was below the appropriate threshold is the percent passing (score). Station scores are then averaged into a sub-region score. If your sampling is targeted to a specific area, such as beaches, simply average the samples for an overall score. If your sampling is either random, or if you have reference site samples included, calculate an average sub-region score. Then calculate an overall score based on area-weighted sub-region scores. A summary of steps for calculating bacteria scores is:

1. Make sure the data used for analysis are from the relevant months. For bacteria, the minimum sampling period is Memorial Day to Labor Day with twice monthly sampling.
2. Make sure the appropriate threshold for enterococci or *E. coli* is used.
3. Calculate the percent of samples that were below the threshold for a station score. Do not average the individual station values before calculating the percent. Compare each station value directly to the threshold to see if it meets the threshold value. (For example: a data value of 200 MPN 100 ml⁻¹ is above the 104 threshold, therefore it scores a zero. A data value of 100 MPN 100 ml⁻¹ is below the threshold and therefore it scores a one. Take the average of the ones and zeros to find the percent of samples that are below the threshold.)
4. Calculate sub-region scores by averaging the scores of the stations in each sub-region.
5. Assign a grade to each sub-region score.

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

If your sampling is targeted to a specific area, simply average the samples for an overall score. If your sampling is either random, or if you have reference site samples included, calculate an overall score based on area-weighted sub-region scores:

6. Calculate the area of each sub-region and area-weight the sub-region average before calculating the average bacteria score for the entire waterbody (e.g., total bacteria = 76% for sub-region 1, sub-region area = 5 km² out of a total 20 km² = 0.25, 76% x 0.25 = 19%).
7. Sum the resulting sub-region scores into an overall score.
8. Based on the overall score, assign a grade for the entire waterbody.

Suggested narrative

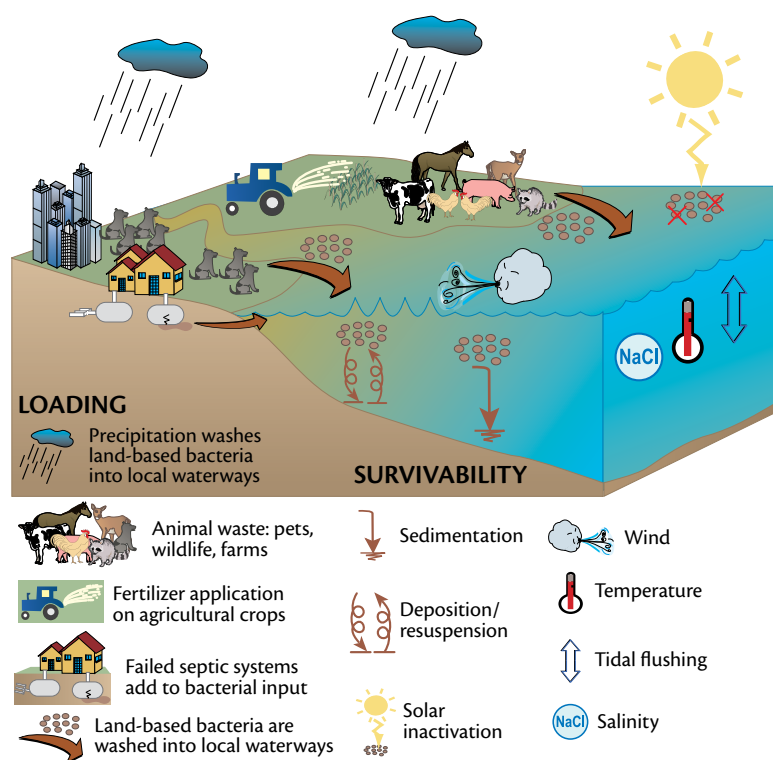
Bacteria indicators differ from other ecosystem health indicators in that they include both targeted (samples are taken at fixed locations designed to evaluate swimming illness risk) and random (samples are taken at randomly assigned locations to represent all potential locations) sampling. Indicator bacteria are useful to evaluate how safe water is for swimming, but are not easily used to describe ecosystem health. There is no clear link between more traditional measures of ecosystem health (core indicators) and bacteria concentration. For that reason, it is not recommended that bacteria scores be integrated with other ecosystem health indicators.

There are also many factors that can affect bacteria concentration and therefore, the interpretation of results. The suggested list below describes some of these important factors and should be included in a narrative statement within the bacteria section of your report card to help provide context and interpretation of results. This narrative can also be in a separate document that's referenced in the report card.

- *Rainfall and dry weather data.* Perhaps the most important transport mechanism for fecal bacteria to coastal waters is by rainfall runoff. Bacteria are transported from animal feces by stormwater, and measurements of fecal indicator bacteria may often be high following rain events. The annual score for bacteria in a report card is designed to represent the percent of days when it is safe to swim (as recommended by EPA guidelines), so these analyses should include sampling during or immediately following significant rainfall. To help interpret the score, report the number of sampling days on which rainfall was a factor. However, when comparing among regions or time series, it is useful to remove the rainfall data from analysis so that comparisons are performed using similar conditions. With any comparison that has different numbers of rain dates, you will need to drop the data from rain dates, or else the results are biased toward higher values in the dataset with more rain dates. This allows direct comparison of results from other tributaries and at individual swimming areas in different years.
- *Potential sources.* Fecal bacteria (and pathogens) can come from a variety of animal sources including humans, wildlife, pets, and even soils. It is mostly assumed that fecal pollution from human sources presents higher risk to humans, but this is difficult to prove; U.S. EPA recommends that fecal indicator bacteria thresholds be applied regardless of the likely bacteria source. It is very difficult to determine the source of bacteria found in the water. Even so, in reporting bacteria scores, it is useful to discuss the potential sources of the bacteria to provide context and interpret results.
- *Scoring.* Currently, there are single thresholds for both enterococci and *E. coli* bacteria for full contact recreational use. The use of a single threshold indicator, while helpful, does not show the resolution that a multiple threshold indicator does.
- *Limitations of indicator bacteria.* When fecal indicator bacteria are present, pathogens are more likely to be present, but they may not always be there. The likelihood of getting sick from swimming is therefore not perfectly correlated with indicator bacteria concentration. Still, these

indicators are the current, best information to predict illness risk, and EPA guidelines say that the risk of illness from swimming is too high when bacteria concentrations exceed the guidelines. Due to the difficulty in assigning risk from different sources, and because rainfall is a major contributor of fecal pollution, the Maryland Department of the Environment recommends that people do not swim in the 48-hour period immediately following rainfall greater than one inch.

- *Health implications.* To improve the linkages between illness and swimming, we recommend that gastrointestinal illnesses following swimming are reported to the health department and other public health databases.
- *Homework/tips.* Including information in the report card about what citizens can do to decrease bacteria is always helpful.
- *Site specific details.* Site specific details help citizens identify locations of high bacteria concentrations and raise awareness of where bacteria concentrations are a problem in the ecosystem.



This diagram illustrates the sources and fates of bacteria in an ecosystem. This is part of a suggested narrative and suggested visualization for this tidal indicator.

NONTIDAL INDICATORS

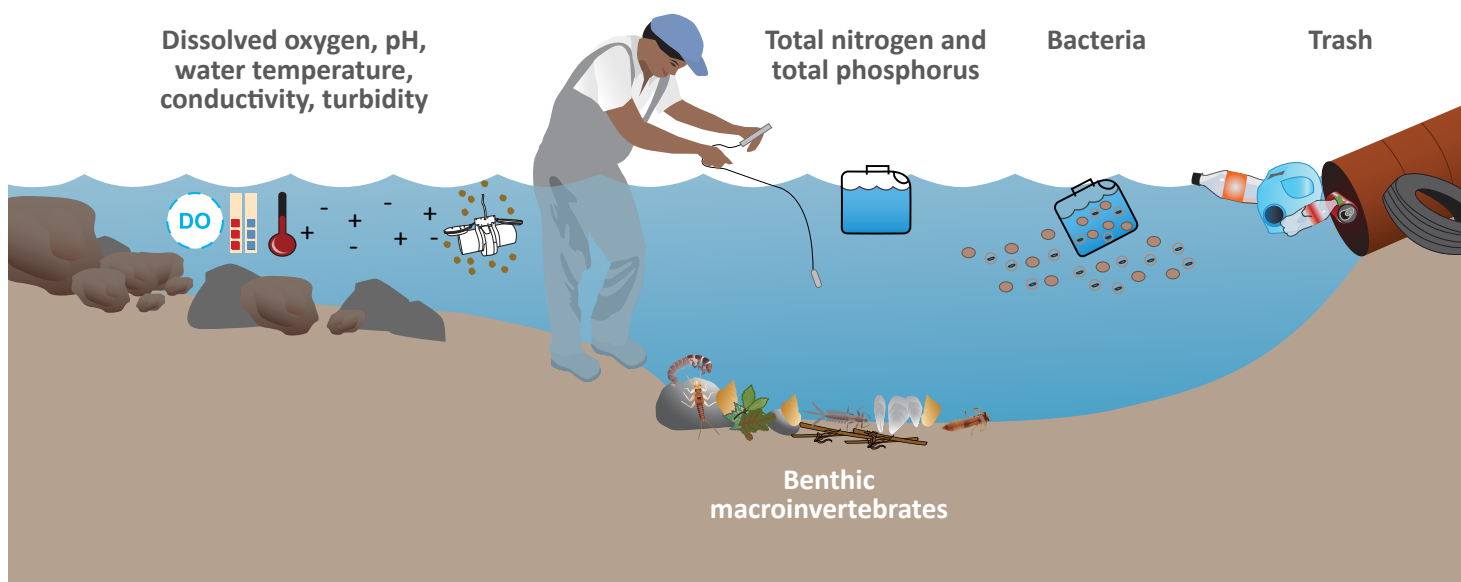
This section discusses in detail how to sample and analyze the indicators that should be incorporated into a nontidal (watershed) report card.

For nontidal systems, there are several types of indicators that are needed to assess health. These include water quality (e.g., nutrients and conductivity), habitat (e.g., stream bank erosion, benthic macroinvertebrates), and biological (e.g., fish) indicators. In this document, we divide them into annually and periodically assessed indicators. Annual indicators include water quality and pollution indicators. Periodical (multi-year, decadal) indicators include habitat and biological indicators.

Core indicators

All the indicators in this protocol document were chosen by the Mid-Atlantic Tributary Assessment Coalition (MTAC) to be used by report card-producing organizations in the Mid-Atlantic region for watershed assessments. The indicators and the methods for evaluation are specifically chosen for nontidal rivers and streams. They were chosen due to their ease of collection and communication, low cost, 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 indicators are:

- total nitrogen and phosphorus
- conductivity
- turbidity
- dissolved oxygen, pH, and water temperature
- bacteria
- trash
- benthic macroinvertebrates



This conceptual diagram illustrates the indicators discussed in this document. They include water quality (e.g. dissolved oxygen, pH, temperature, conductivity, turbidity, and nutrients), pollution (e.g. bacteria, trash), and habitat (e.g. macroinvertebrates).

These indicators should be measured and analyzed by all monitoring programs that wish to compare the health of their watersheds with adjacent watersheds.

Elective indicators

The indicators discussed in this document provide a consistent base for data comparisons among water systems. However, elective indicators, such as stream bank erosion, bottom habitat, and toxic contaminants, may also be measured if organizations have a particular interest in them.

Sampling and data analysis

An overview and methods for data analysis are provided for each indicator in the following chapters. A summary table of preferred and minimum recommendations is provided in the table below.

Summary of preferred and minimum sampling recommendations for all indicators.

Indicator	Preferred sampling period	Preferred sampling resolution	Minimum sampling period (needed for data analysis)	Minimum sampling resolution	Salinity regime (needed for data analysis)
Total nitrogen, total phosphorus, conductivity, DO, pH, water temperature	Year round	Twice monthly	April-November	Once monthly	Once quarterly
Turbidity	Year round	Twice monthly	April-November	Once monthly	Once quarterly, twice in winter and spring preferred
Benthic macroinvertebrates	Year round	Once quarterly	N/A	N/A	Not necessary
Bacteria	Year round	Twice monthly	April-November	Once monthly	Once quarterly
Trash	Year round	Twice monthly	April-November	Once monthly	Once quarterly












The recommended amount of samples taken is twice monthly, or higher. The more samples taken, the more accurate the assessment of the watershed. However, if groups do not have enough resources to measure twice or more a month, once monthly sampling may be conducted so that there are enough samples from which to calculate an average. Additionally, some samples should be taken during or right after storm events, since rainfall and run off affect these indicators. Sampling should occur for at least one storm event (1 inch of rainfall in 24 hours) per quarter.

Due to funding and time constraints of watershed organizations, it is understood that a group may only have enough capacity to sample just once per month (the minimum recommended amount). Therefore, this protocol also provides a minimum sampling effort that is required to adequately assess and score the indicators.

Thresholds

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. The thresholds are significant because they represent the point where prolonged exposure to unhealthy conditions leads to a negative ecosystem response (Longstaff et al. 2010). Thresholds described in this document were derived from peer-reviewed scientific articles and consultation with Chesapeake Bay nontidal analysts (ICPRB 2011; US EPA 2000 a; US EPA 2000 b; US EPA 2000 c).

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 throughout the document.

Health indicator	Example threshold value	Comparison of data to threshold
 Total nitrogen	$<0.64 \text{ mg}\cdot\text{l}^{-1}$	 = Proportion of data that meets threshold values for each indicator
 Total phosphorus	$<0.01 \text{ mg}\cdot\text{l}^{-1}$	
 Conductivity	$\leq 42 \text{ }\mu\text{siemens}\cdot\text{cm}^{-1}$	
 Turbidity	$<3 \text{ NTUs}$	
 Dissolved oxygen	$>5.0 \text{ mg}\cdot\text{l}^{-1}$	
 pH	$>6.5 \text{ \& } <8.5$	
 Water temperature	$<68^{\circ}\text{F } (20^{\circ}\text{C})$	
 Bacteria	$\leq 235 \text{ organisms}\cdot\text{100 ml}^{-1}$	
 Trash	N/A	
 Benthic community	IBI = 3.0	

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

Scoring of data

The recommended time period for nontidal data is year-round. All samples collected within a calendar year should be included in data analyses. Once thresholds 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.

Pass/Fail scoring method

A pass/fail scoring method is a simple method used to calculate indicator scores based on the percent of measurements that met an ecologically relevant threshold. The process is outlined below, 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 it doesn't describe how close a failing value is to passing. For example, if a dissolved oxygen measurement is 4.9 mg·l⁻¹, it fails because the threshold is 5.0 mg·l⁻¹. However, it is much closer to passing than a value of 1.0 mg·l⁻¹.

1. Sort data by station

Source	Station	Region	Date	DO Value
DNR	0000008	Mid-Michigan	4/29/09	9.00
DNR	0000000	Mid-Michigan	4/29/09	9.00
DNR	0000008	Mid-Michigan	4/29/09	9.00
DNR	0000000	Mid-Michigan	4/29/09	9.00
DNR	0000000	Mid-Michigan	5/20/09	9.00
DNR	0000008	Mid-Michigan	5/20/09	9.00
DNR	0000000	Mid-Michigan	5/20/09	9.00

2. Calculate score for each data point Ex: If DO ≥ 5 mg l⁻¹, then Score = Pass

Source	Station	Region	Date	Depth	DO Value	Threshold	Score
DNR	0000008	Mid-Michigan	4/29/09	0.0	9.00	5.00	P
DNR	0000000	Mid-Michigan	4/29/09	0.0	9.00	5.00	P
DNR	0000008	Mid-Michigan	4/29/09	0.0	9.00	5.00	P
DNR	0000000	Mid-Michigan	4/29/09	0.0	9.00	5.00	P
DNR	0000000	Mid-Michigan	5/20/09	0.0	9.00	5.00	P
DNR	0000008	Mid-Michigan	5/20/09	0.0	9.00	5.00	P
DNR	0000000	Mid-Michigan	5/20/09	0.0	9.00	5.00	P

3. Calculate the score for each station Ex. ((Total # of scores = Pass)/(Total # of scores for that station))*100 = % total

Source	Station	Region	Date	Depth	DO Value	Threshold	Score	Station Score
DNR	0000008	Mid-Michigan	4/29/09	0.0	9.00	5.00	P	100
DNR	0000000	Mid-Michigan	4/29/09	0.0	9.00	5.00	P	
DNR	0000008	Mid-Michigan	4/29/09	0.0	9.00	5.00	P	
DNR	0000000	Mid-Michigan	4/29/09	0.0	9.00	5.00	P	
DNR	0000000	Mid-Michigan	5/20/09	0.0	9.00	5.00	P	
DNR	0000008	Mid-Michigan	5/20/09	0.0	9.00	5.00	P	
DNR	0000000	Mid-Michigan	5/20/09	0.0	9.00	5.00	P	

5. Calculate region scores by averaging all station scores per region.

To calculate overall lake or watershed score, sum the region values weighted by % of total area.

Source	Station	Region	Date	DO (mg/L)	pH	%	Station Average	Region	Score	Area (sqm)	Percent	Weighted score
DNR	0000008	Mid-Michigan	4/29/09	9.00	p	100	100	north lake	100.0	1000000.00	7.14285714	7.00000000
DNR	0000000	Mid-Michigan	4/29/09	9.00	p	100		mid lake	100.0	1000000.00	7.14285714	7.00000000
DNR	0000008	Mid-Michigan	4/29/09	9.00	p	100		south lake	100.0	1000000.00	7.14285714	7.00000000
DNR	0000000	Mid-Michigan	4/29/09	9.00	p	100		Lake	100.0	1000000.00	7.14285714	7.00000000
DNR	0000000	Mid-Michigan	5/20/09	9.00	p	100						
DNR	0000008	Mid-Michigan	5/20/09	9.00	p	100						
DNR	0000000	Mid-Michigan	5/20/09	9.00	p	100						

4. Station scores for each station

Source	Station	Region	Station Score
DNR	0000008	Mid-Michigan	100
DNR	0000000	Mid-Michigan	100
DNR	0000008	Mid-Michigan	100
DNR	0000000	Mid-Michigan	100
DNR	0000000	Mid-Michigan	100
DNR	0000008	Mid-Michigan	100
DNR	0000000	Mid-Michigan	100

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

Multiple thresholds






Multiple thresholds are used to score indicators based on a gradient of healthy to unhealthy conditions. 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 high, medium, or low. 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 can be assessed with greater precision than using a pass/fail method.

Scores are standardized

In order to integrate individual indicator scores into a more encompassing index (e.g., a water quality index), 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., system-wide) score can be calculated as the area weighted average of regional scores.

Core indicators discussed in this document that are measured against multiple thresholds include nitrogen, total phosphorus, conductivity, and turbidity. Measurements are compared 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		

A grade and description are assigned based on the score that the indicator or subregion 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

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 sections), a grade can be assigned. For the ecological indicators in this protocol, the grading scale follows the Chesapeake Bay report card scale of 0–100%, with equal interval breaks. This was determined through consensus meetings with the Chesapeake Bay Program. Grades are equally divided to provide a clearer picture of health. Following the typical school grading scale (<60% = F, 60–70% = D, etc) would result in consistently failing grades, which does not provide information about small improvements or declines in ecosystem health. The equally divided grading scale and multiple thresholds both allow evaluation of small changes in ecosystem health, even in the very poor, poor, and moderately poor ranges. A narrative description of the major categories are provided, which relate the grade to ecological health.

For the bacteria indicator in this protocol, the grading scale does not follow the overall watershed report card scale, but rather follows the traditional 10-point intervals. Since bacteria is a human health indicator, a stricter grading scale was needed to ensure that bacteria scores were communicated properly to the public.

Summary

This overview of the core indicators and thresholds should provide a general understanding of this protocol. The following chapters provide more detail and step-by-step instructions for analysis and assessment of each indicator.



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.

Descriptions of ecological health that correspond with each grade.

Score	Narrative
100	Excellent
90 - <100	Good
80 - <90	Moderate
70 - <80	Moderately Poor
60 - <70	Poor
<60	Very Poor

Scoring and description for bacteria indicator.

Data analysis

Nontidal – Nutrients

Once samples have been analyzed in the lab, a spreadsheet of data will be provided. The thresholds for total nitrogen and total phosphorus are different, so make sure the appropriate thresholds are being used. A set of multiple thresholds has been determined for nitrogen and phosphorus. These threshold levels are based upon how benthic organisms are affected by increasing nutrient levels. For total nitrogen and total phosphorus, each measurement is separated into ecoregion and compared to a corresponding set of thresholds. The five most relevant ecoregions are Mid-Atlantic Coastal Plain, Southeastern Plain, Piedmont, Ridges, and Valleys. For nitrogen and phosphorus analysis, these five ecoregions are combined into two groups. The two major groups are 1. Piedmont, Valleys, and Ridges; and 2. Coastal Plain (Mid-Atlantic Coastal Plain, Southeastern Plain; ICPRB 2011).

Applying thresholds to individual sites allows determination of total nutrient condition. Each data point is compared to the thresholds in the appropriate table and scored from 0–5. Each measurement score (0–5) is averaged into a station score for the entire year. Then, station scores are averaged into a sub-watershed score. Once the sub-watershed score is calculated, calculate the total overall score by area-weighting each sub-watershed score and averaging them for an overall watershed score. For example, we can consider an example Site x, located in the Piedmont ecoregion. The total nitrogen measured at Site x was $1.70 \text{ mg}\cdot\text{l}^{-1}$. So when looking at the table, we can compare to the threshold levels to see which range the measurement falls into. For total nitrogen, Site x is greater than 1.65 but less than 2.15, so it scores a 3.

Ecologically relevant multiple thresholds for total nitrogen by ecoregion.

Score	Piedmont, Valleys, & Ridges ($\text{mg}\cdot\text{l}^{-1}$)	Coastal Plain ($\text{mg}\cdot\text{l}^{-1}$)
5	<0.64	<0.82
4	≥ 0.64 - <1.65	≥ 0.82 - <1.52
3	≥ 1.65 - <2.15	≥ 1.52 - <2.22
2	≥ 2.15 - <2.65	≥ 2.22 - <2.66
1	≥ 2.65 - <3.66	≥ 2.66 - <3.61
0	≥ 3.66	≥ 3.61

Ecologically relevant multiple thresholds for total phosphorus by ecoregion.

Score	Piedmont, Valleys, & Ridges ($\text{mg}\cdot\text{l}^{-1}$)	Coastal Plain ($\text{mg}\cdot\text{l}^{-1}$)
5	<0.01	<0.02
4	≥ 0.01 - <0.03	≥ 0.02 - <0.06
3	≥ 0.03 - <0.05	≥ 0.06 - <0.09
2	≥ 0.05 - <0.06	≥ 0.09 - <0.12
1	≥ 0.06 - <0.09	≥ 0.12 - <0.17
0	≥ 0.09	≥ 0.17

A summary of steps for calculating total nitrogen and total phosphorus scores is:

1. For nitrogen and phosphorus, the sampling period is year-round with once-a-month sampling.
2. Make sure the appropriate thresholds for total nitrogen and total phosphorus are used.
3. Assign scores of 0–5 to each sampling value.
4. Average the 0–5 scores for a station score.
5. Calculate sub-watershed scores by averaging the scores of the stations in each sub-watershed. Remember that 10 sampling sites are needed for each sub-watershed, so average to the sub-watershed level if possible, but otherwise average to the watershed level if not.
6. Assign a grade to each sub-watershed or watershed score.

If you have a score for each sub-watershed, you can determine the average % score and grade for the overall watershed by area-weighting the sub-watershed scores.

7. Determine the percent-area for each sub-watershed. For example: sub-watershed 1 area = 5 km², divide by the total watershed area of 20 km² = 0.25.
8. Multiply the sub-watershed proportion (0.25) by the sub-region score (76%) to equal 19%.
9. Sum the resulting sub-watershed scores into an overall watershed score.
10. Based on the overall score, assign a grade for the entire watershed.

For health assessments, it is recommended that 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 12 measurements all year and another site has 5, the site with 12 measurements would have more influence on the final average nitrogen or phosphorus score than the site with 5 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.

Nontidal – Conductivity

First, temperature affects conductivity. If the conductivity meter automatically compensates for temperature, use the data directly. However, if it does not or if the samples were brought back to the office, the conductivity data needs to be adjusted using the water temperature data collected at the same time as the conductivity data.

Data from the laboratory results are analyzed to calculate a percent of samples below the appropriate threshold. A set of multiple thresholds has been determined for conductivity. These threshold levels are based upon how conductivity levels impact organisms in the environment. For conductivity, each measurement is separated by ecoregion and compared to a corresponding set of thresholds. The four most relevant ecoregions are Piedmont, Valleys, Ridges, and Coastal Plain (which includes Mid-Atlantic Coastal Plain, and Southeastern Plain; ICPRB 2011).

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

1. For conductivity, the sampling period is year-round with once-a-month sampling.
2. Make sure the appropriate threshold for conductivity is used. Assign scores of 0–5 to each sampling value.
3. Average the 0–5 scores for a station score.
4. Calculate sub-watershed scores by averaging the scores of the stations in each sub-watershed. Remember that 10 sampling sites are needed for each sub-watershed, so average to the sub-watershed level if possible, but otherwise average to the watershed level if not.
5. Assign a grade to each sub-watershed or watershed score.

If you have a score for each sub-watershed, you can determine the average % score and grade for the overall watershed by area-weighting the sub-watershed scores.

6. Determine the percent-area for each sub-watershed. For example: sub-watershed 1 area = 5 km², divide by the total watershed area of 20 km² = 0.25.
7. Multiply the sub-watershed proportion (0.25) by the sub-region score (76%) to equal 19%. Sum the resulting sub-watershed scores into an overall watershed score.
8. Based on the overall score, assign a grade for the entire watershed.

For health assessments, it is recommended that 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 12 measurements all year and another site has 5, the site with 12 measurements would have more influence on the final average conductivity score than the site with 5 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.

Ecologically relevant multiple thresholds for conductivity for the Piedmont ecoregion.

Conductivity ($\mu\text{siemens cm}^{-1}$)	Score
≤ 42	5
$>42 - \leq 100$	4
$>100 - \leq 158$	3
$>158 - \leq 249$	2
$>249 - <544$	1
≥ 544	0

Ecologically relevant multiple thresholds for conductivity for the Valleys ecoregion.

Conductivity ($\mu\text{siemens cm}^{-1}$)	Score
≤ 49	5
$>49 - \leq 137$	4
$>137 - \leq 267$	3
$>267 - \leq 430$	2
$>430 - <626$	1
≥ 626	0

Ecologically relevant multiple thresholds for conductivity for the Ridges ecoregion.

Conductivity ($\mu\text{siemens cm}^{-1}$)	Score
≤ 21	5
$>21 - \leq 66$	4
$>66 - \leq 130$	3
$>130 - \leq 214$	2
$>214 - <521$	1
≥ 521	0

Ecologically relevant multiple thresholds for conductivity for the Coastal Plain (Mid-Atlantic and Southeastern) ecoregion.

Conductivity ($\mu\text{siemens cm}^{-1}$)	Score
≤ 56	5
$>56 - \leq 108$	4
$>108 - \leq 182$	3
$>182 - \leq 257$	2
$>257 - <526$	1
≥ 526	0

Nontidal – Turbidity

Data from field sampling are analyzed to calculate a percent of samples below the appropriate threshold. A set of multiple thresholds has been determined for turbidity. These threshold levels are based upon how turbidity levels impact organisms in the environment. For turbidity, each measurement is compared to a corresponding set of thresholds.

Applying these thresholds to individual sites allow determination of total turbidity condition. Each data point is compared to the thresholds and scored from 0–5. Each measurement score (0–5) is averaged into a station score for the entire year. Then, station scores are averaged into a sub-watershed score. Once the sub-region score is calculated, calculate the total overall score by area-weighting each sub-watershed score and averaging them for an overall score. A summary of steps for calculating turbidity scores is:

Ecologically relevant multiple thresholds for turbidity.

Turbidity (NTUs)	Score
<3	5
≥3 - <4.75	4
≥4.75 - <6.5	3
≥6.5 - <8.25	2
≥8.25 - <10	1
≥10	0

1. For turbidity, the sampling period is year-round with once-a-month sampling.
2. Make sure the appropriate threshold for turbidity is used.
3. Assign scores of 0–5 to each sampling value.
4. Average the 0–5 scores for a station score.
5. Calculate sub-watershed scores by averaging the scores of the stations in each sub-watershed. Remember that 10 sampling sites are needed for each sub-watershed, so average to the sub-watershed level if possible, but otherwise average to the watershed level if not.
6. Assign a grade to each sub-watershed or watershed score.

If you have a score for each sub-watershed, you can determine the average % score and grade for the overall watershed by area-weighting the sub-watershed scores.

7. Determine the percent-area for each sub-watershed. For example: sub-watershed 1 area = 5 km², divide by the total watershed area of 20 km² = 0.25.
8. Multiply the sub-watershed proportion (0.25) by the sub-region score (76%) to equal 19%. Sum the resulting sub-watershed scores into an overall watershed score.
9. Based on the overall score, assign a grade for the entire watershed.

For health assessments, it is recommended that 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 12 measurements all year and another site has 5, the site with 12 measurements would have more influence on the final average turbidity score than the site with 5 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.

Nontidal – Vital Signs

Field sampling measurements should be marked on a field data sheet, then entered in a spreadsheet or database. Data are compared against ecologically relevant criteria and assigned as passing or failing. For analysis, each data observation is compared to a corresponding threshold.

Comparison to criteria

For nontidal streams, DO, pH, and water temperature thresholds are defined based on a designated use set by state agencies. Designated uses include water contact recreation, support of marine life, support of shellfish harvesting and public water supply. You will need to determine if your stream has a designated use of warmwater or coldwater.

Dissolved oxygen

For coldwater nontidal streams, the dissolved oxygen concentration may not be less than 5.0 mg·l⁻¹ at any time, with a minimum daily average of not less than 6.0 mg·l⁻¹. For warmwater nontidal streams, the dissolved oxygen concentration may not be less than 5.0 mg·l⁻¹ at any time. Each individual data point is compared to this criterion and scored as pass or fail.

pH

Both coldwater nontidal streams and warmwater nontidal streams must have a pH measurement between 6.5 and 8.5. Each individual data point is compared to this criterion and scored as pass or fail.

Water temperature

For coldwater nontidal streams, the temperature must not exceed 68°F (20°C). For warmwater nontidal streams the temperature must not exceed 90°F (32°C). Each individual data point is compared to this criterion and scored as pass or fail.

Passing scores for DO, pH, and temperature fall into the following thresholds from warm and cold water regions.

Stream type	DO	pH	Temperature
warmwater	>5.0 mg·l ⁻¹	6.5-8.5	<90°F (32°C)
coldwater - instantaneous concentration	>5.0 mg·l ⁻¹	6.5-8.5	<68°F (20°C)
coldwater - minimum daily average	>6.0 mg·l ⁻¹	6.5-8.5	<68°F (20°C)

Scoring

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-watershed score. An overall watershed score is calculated as an area-weighted average of the sub-watershed scores. A summary of the data analysis steps for vital sign indicators is listed below:

1. For vital sign indicators, the sampling period is year-round with once a month sampling.
2. Make sure the appropriate threshold for do, pH, and water temperature is used.
3. Compare measured value to the threshold and assign it pass/fail.
4. For each pass value, assign it a 100 (one hundred), and for a fail, a 0 (zero). Average the 100s and 0s (zeroes) for each station. This is the average % passing, and therefore the station score.
5. Calculate sub-watershed scores by averaging the scores of the stations in each sub-watershed. Remember that 10 sampling sites are needed for each sub-watershed, so average to the sub-watershed level if possible, but otherwise average to the watershed level if not.
6. Assign a grade to each sub-watershed or watershed score.

If you have a score for each sub-watershed, you can determine the average % score and grade for the overall watershed by area-weighting the sub-watershed scores.

7. Determine the percent-area for each sub-watershed. For example: sub-watershed 1 area = 5 km², divide by the total watershed area of 20 km² = 0.25.
8. Multiply the sub-watershed proportion (0.25) by the sub-region score (76%) to equal 19%.
9. Sum the resulting sub-watershed scores into an overall watershed score.
10. Based on the overall score, assign a grade for the entire watershed.

For health assessments, it is recommended that 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 12 measurements all year and another site has 5, the site with 12 measurements would have more influence on the final average do, pH, or water temperature score than the site with 5 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.

Bacteria

The EPA threshold for *E. coli* is 235 organisms·100 ml⁻¹ for any single water sample. The laboratory cultures the water sample and then counts how many bacteria organism colonies are on the plate which is called the colony forming units (cfu). Some labs use a slightly different method where that reports the number of organisms as the most probable number (MPN) in the dish. Data provided from the laboratory are analyzed to calculate a percent of samples below the *E. coli* threshold. The percent of samples in a sampling season (Memorial Day to Labor Day) that were below the threshold is the percent passing (score) for each station. A summary of steps for calculating bacteria scores is:

1. Make sure the data used for analysis are from the relevant months. For bacteria, the minimum sampling period is Memorial Day to Labor Day with twice monthly sampling.
2. Make sure the appropriate threshold for *E. coli* is used.
3. Calculate the percent of samples that were below the threshold for a station score. Do not average the individual station values before calculating the percent. Compare each station value directly to the threshold to see if it meets the threshold value. (For example: a data value of 300 organisms·100 ml⁻¹ is above the 235 threshold, therefore it scores a zero. A data value of 100 organisms·100 ml⁻¹ is below the threshold and therefore it scores a one. Take the average of the ones and zeros to find the percent of samples that are below the threshold for each station.)

4. For this protocol, we do not recommend calculating an overall grade for the sub-watersheds or overall watershed because bacteria data are so variable.

Scoring and description for bacteria indicator.

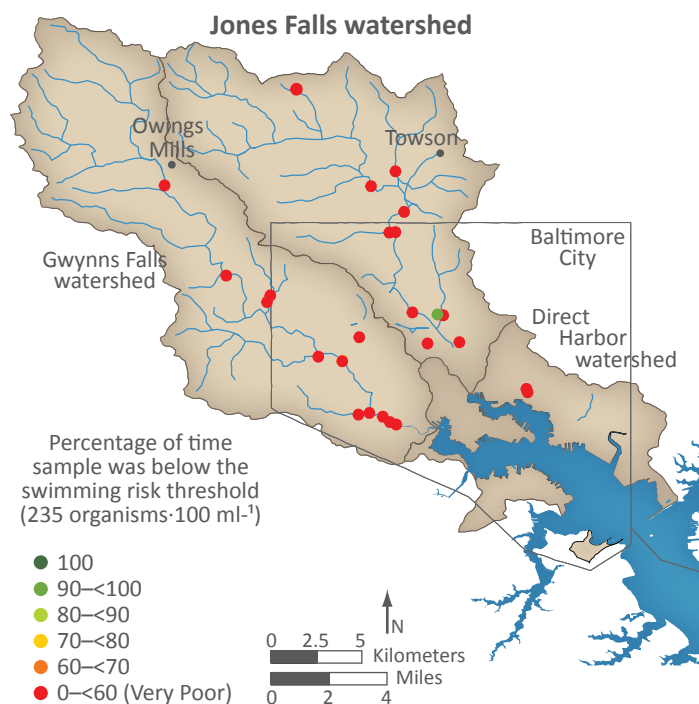
Score	Narrative
100	Excellent
90 - <100	Good
80 - <90	Moderate
70 - <80	Moderately Poor
60 - <70	Poor
<60	Very Poor

Communicating bacteria score results

Since bacteria is a human health indicator, it is communicated differently than ecological indicators. For bacteria, station average scores are calculated, then presented on a 10-point scale (not the 20-point scale used by ecological indicators). Furthermore, due to the variability of bacteria scores within small areas, a map of station average scores should be presented along with the overall sub-region or region information. To interpret scores correctly, scores on the map and the associated text should be described as the "Percentage of time samples were below the swimming risk threshold." For an overall sub-region average, "fire danger" symbols or dials can be used to illustrate relative risk of becoming sick from swimming. This is provided by calculating an overall sub-region score, with Low risk = 100% passing and High risk = 60% passing. To calculate the sub-region score, station scores are averaged into a sub-region score.

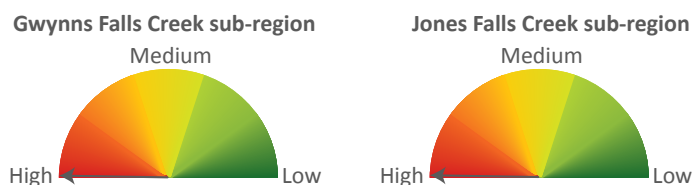
When creating the "fire danger" symbol, use the following steps to calculate the angle of the arrow. This angle is the proportion the score takes up out of 180°.

5. Take the sub-region score and subtract it from 100. For example: sub-region score = 75%. $100 - 75 = 15\%$.
6. Next determine what percent the resulting number (15 in our example) is out of 40. 40 covers the range between 60 and 100. So, the angle will be equal to $15/40$ multiplied by 180. Using this example, the angle is 67.5° .



Baltimore Harbor watershed map for E. coli scores. Note the 10-point ranges for this indicator compared to the 20-point ranges for ecological indicators.

Annual relative risk of humans becoming sick from swimming



A dial or "fire danger" symbol can be used to illustrate the relative risk of becoming sick from swimming.

Suggested narrative

Bacteria indicators differ from other ecosystem health indicators in that they include both targeted (samples are taken at fixed locations designed to evaluate swimming illness risk) and random (samples are taken at randomly assigned locations to represent all potential locations) sampling. Indicator bacteria are useful to evaluate how safe water is for swimming, but are not easily used to describe ecosystem health. There is no clear link between more traditional measures of ecosystem health (e.g. TN, TP, etc) and bacteria concentration. For that reason, it is not recommended that bacteria scores be integrated with other ecosystem health indicators.

There are also many factors that can affect bacteria concentration and the interpretation of results. The suggested list below describes some of these important topics, which could be described in a narrative statement within the bacteria section of the report card. This discussion can also be in a separate document that is referenced in the report card.

- *Rainfall and dry weather data.* The most important transport mechanism for fecal bacteria to streams is often rainfall runoff. Bacteria are transported from animal feces by stormwater, and measurements of fecal indicator bacteria may often be high following rain events. To help interpret the score, report the number of sampling days on which rainfall was a factor. However, when comparing among regions or time series, it is useful to remove the rainfall data from analysis so that comparisons are performed using similar conditions. With any comparison that has different numbers of rain dates, drop the data from rain dates, to reduce bias toward values in the dataset with more rain dates. Analysis of dry weather data only allows direct comparison of results from other tributaries and at individual swimming areas in different years.
- *Potential sources.* Fecal bacteria (and pathogens) can come from a variety of animal sources including humans, wildlife, pets, and even soils. It is mostly assumed that fecal pollution from human sources presents higher risk to humans, but this is difficult to prove; US EPA recommends that fecal indicator bacteria thresholds be applied regardless of the likely bacteria source. It is very difficult to determine the source of bacteria found in the water. Even so, in reporting bacteria scores, it is useful to discuss the potential sources of the bacteria to provide context and interpret results. For example, interpolation of high bacteria concentrations might be different if there are large numbers of geese in an area or if there are many residences with failing septic systems.
- *Scoring.* Currently, there are single thresholds for *E. coli* bacteria for full contact recreational use. The use of a single threshold indicator, while helpful, does not show the resolution that a multiple threshold indicator does.
- *Limitations of indicator bacteria.* When fecal indicator bacteria are present, pathogens are more likely to be present, but they may not always be there. The likelihood of getting sick from swimming is therefore not perfectly correlated with indicator bacteria concentration. Still, these indicators are the current, best information to predict illness risk, and EPA guidelines say that the risk of illness from swimming is too high when bacteria concentrations exceed the guidelines. Due to the difficulty in assigning risk from different sources, and because rainfall is a major contributor of fecal pollution, Maryland Department of the Environment recommends that people do not swim in the 48-hour period immediately following rainfall greater than one inch.

-
- *Health implications.* To improve the linkages between illness and swimming, we recommend that gastrointestinal illnesses following swimming are reported to the health department and other public health databases.
 - *Homework/tips.* Including information in the report card about what citizens can do to decrease bacteria is always helpful (e.g. pick up pet waste, maintain septic systems, etc.).
 - *Site specific details.* Site specific details help citizens identify locations of high bacteria concentrations and raise awareness of where bacteria concentrations are a problem in the ecosystem.

Nontidal – Benthic Community (Benthic macroinvertebrates)

Benthic macroinvertebrates are a critical indicator for nontidal streams because they integrate a variety of environmental conditions that will be reflected in the individual invertebrate and in the diversity of benthic macroinvertebrates at a site. The benthic macroinvertebrate indicator for incorporation into a watershed (nontidal) report card is still under development. A few key points should be noted here:

- There will be different methods for Tier 1 versus Tier 2 benthic macroinvertebrate data. While most report cards require Tier 2 data collection, we do acknowledge there are many sites with Tier 1 data that should be incorporated into a report card. We will distinguish between the two tiers and the certainty in the results based on those tiers.
- In the past, a benthic macroinvertebrate indicator was developed for the Mid-Atlantic Tributary Assessment Coalition (MTAC) which can still be of use by groups. The methods for determining that indicator are below and should be used until the new indicator can be fully developed (over the calendar year 2018).
- There are several distinctions that need to be made for the new indicator.
 - Tentatively, Tier 1 data consists of Order level collection by non-traditional groups around the entire Chesapeake Bay watershed. Tier 2 data consists of Family level collection by non-traditional groups around the entire Chesapeake Bay watershed. Tier 2 data may or may not need analysis for certified professionals in a lab. This is To Be Determined.
 - Tier 1 data cannot be used for trends information.
 - The intention is to develop a Tier 2 indicator that can be used for trends. A site needs to be re-visited over time to create trends. A process to determine which sites have Tier 2 data and should be re-visited by non-traditional groups will be determined over the 2018 calendar year.

Methods for determining benthic macroinvertebrate indicator from MTAC protocol

Data analysis of the benthic macroinvertebrate indicator consists of averaging individual station scores over the watershed for an average watershed score. Unlike water quality indicators, this indicator uses the six most recent years' worth of data to determine current condition. Data from the six most recent years provides a good assessment of current health conditions. Data from 6 to 10 years old should be used with caution or flagged. Data that is older than 10 years should not be used to evaluate current conditions.

Benthic macroinvertebrate sampling includes both targeted and random sites. Targeted sites are used to focus on potential issues within a stream reach. Using both targeted and random sites in your assessment provides more data within the specified time-frame. However, if you want to roll up the data into an overall score, only use random sites. Using only random sites is necessary for averaging because it ensures unbiased sampling results to be included in the assessment.

To roll up the individual sampling site scores into an overall watershed score, a minimum of 5 sampling sites are needed for small (HUC-12) watersheds and a minimum of 10 sampling sites are needed for larger (HUC-8) watersheds.

HUCs, or hydrologic unit codes, are the subdivisions of watersheds in the United States. For more information about HUCs, please visit the USGS website: <http://water.usgs.gov/GIS/huc.html>. HUC-12 and HUC-8 watersheds show very detailed information, which is important for benthic macroinvertebrate sampling. Unfortunately, there are few state programs that cover these small areas enough to average into a watershed score. Pennsylvania, for example, only has sporadic random sampling sites throughout the state. Some small watersheds can be evaluated, but not all. In these cases, it is best to provide just the targeted and sampling site scores on a map, rather than provide an overall average score.

The following bullet points provide a step-by-step process for data analysis.

- Become familiar with the CBP Interactive Mapping website: http://www.chesapeakebay.net/indicators/indicator/health_of_freshwater_streams_in_the_chesapeake_bay_watershed. This will help determine how many sampling sites are within your watershed, which helps to determine if you can average the scores or just provide a map of individual sampling sites.
- Benthic Index of Biotic Integrity scores are available for download from the Chesapeake Bay Program's database: http://www.chesapeakebay.net/data/downloads/watershed_wide_benthic_invertebrate_database. Benthic IBI scores and ratings are provided in a downloadable Excel spreadsheet.
- Determine which HUCs are within your watershed boundary before downloading benthic data from the CBP website.
 - Download the data from the website.
 - Download the most recent six years worth of data from the database. The most recent six years of data provides a good assessment of current condition.
 - See Database information below for more detailed instructions.
- Using the spreadsheet, determine the average score for each of the HUCs within your watershed. Otherwise, you can use the individual sampling site scores from the Random and Targeted tab to display the individual sampling site scores, as discussed above.
- The last step is to average each HUC average score into an overall watershed score. Do not area-weight because you are already using randomly sampled sites.

IBI Score (%)	Rating	Grade
≥ 67	Excellent	A
50 – <67	Good	B
30 – <50	Fair	C
17 – <30	Poor	D
<17	Very Poor	F

Benthic IBI scores and ratings are provided in the downloadable Excel spreadsheet. This table helps to determine the overall watershed average score, rating, and therefore the grade.

Database information

Benthic Index of Biotic Integrity scores are available for download from the Chesapeake Bay Program's database: http://www.chesapeakebay.net/data/downloads/watershed_wide_benthic_invertebrate_database.

- Download the most recent six years worth of data from the database. The most recent six years of data provides a good assessment of current condition.
- Click download the data.
- The Data Source should be nontidal benthic data.
- Select indicators and calculated metrics below Data Source drop-down menu.
- Select indicators and calculated metrics again under Data Type.
- Click continue.
- Choose state as the attribute.
- Enter the date range, which the website allows a maximum of five years. You can download three years at a time, since the last six years of data is needed.
- Type in your email address.
- Download data.

CLEANING YOUR DATA TO ANALYZE IN A REPORT CARD

Now you're ready to analyze your data for report card scoring!

- Start by choosing one indicator. Let's choose dissolved oxygen.
 1. Start a new excel spreadsheet called "Dissolved Oxygen Data Analysis".
 2. Name the first tab "All DO data".
 3. Copy all the data from your cleaned spreadsheet tab into the new spreadsheet.
 4. Next, delete the data you don't need in this spreadsheet. Keep information like the Site, Latitude, Longitude, Date, Depth, O₂ (mg/l), Notes/Comments.

	A	B	C	D	E	F	G	H	I	J
	Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/l)	NOTES/COMMENTS			
1										
2	E15	40.762900	-73.849110	1/10/11	3	11.58				
3	FLC1	40.761980	-73.836660	1/10/11	3	10.90				
4	FLC2	40.765100	-73.843100	1/10/11	4	11.44				
5	FB1	40.772620	-73.853080	1/10/11	3	11.28				
6	E6	40.785840	-73.861200	1/10/11	3	11.08				
7	E14	40.800480	-73.864330	1/10/11	4	11.02				
8	E7	40.802800	-73.819960	1/10/11	3	11.24				
9	E4	40.782300	-73.921120	1/10/11	3	11.41				
10	E8	40.801560	-73.780740	1/10/11	3	11.45				
11	E10	40.843460	-73.764080	1/10/11	3	11.73				
12	E11	40.790280	-73.765620	1/10/11	3	11.96				
13	LN1	40.779840	-73.761040	1/11/11	3	11.98				
14	E15	40.763080	-73.849270	2/7/11	3	13.34				
15	FLC1	40.762080	-73.836760	2/7/11	3	11.14				
16	FLC2	40.765080	-73.842840	2/7/11	3	12.59				
17	E14	40.800780	-73.864960	2/7/11	3	12.08				
18	FB1	40.772820	-73.853280	2/7/11	3	13.25				
19	FB1	40.772820	-73.853280	2/7/11	3	13.25				
20	E6	40.786340	-73.861580	2/7/11	3	13.59				
21	E7	40.803270	-73.819840	2/7/11	3	13.23				
22	E4	40.781900	-73.922040	2/7/11	3	11.56				
23	E8	40.801980	-73.781000	2/7/11	3	13.68				
24	E10	40.843420	-73.764580	2/7/11	3	14.08				

5. In this example, there are no notes or comments in that column, so you can remove it as well.
6. Now copy the entire sheet and paste it on a new tab called "DO data analysis".
7. At this point, it is helpful to organize the spreadsheet using a Custom Sort by Site and then by Date.

Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/l)
BR5	40.813700	-73.871500	5/16/11	3	6.95
BR5	40.813940	-73.871720	5/31/11	3	6.53
BR5	40.814020	-73.871820	6/6/11	3	7.21
BR5	40.813940	-73.871470	6/13/11	3	4.95
BR5	40.814100	-73.871940	6/20/11	3	5.33
BR5	40.813540	-73.870680	6/28/11	3	4.31
BR5	40.813800	-73.871400	7/5/11	3	4.98
BR5	40.813740	-73.871280	7/11/11	3	11.38
BR5	40.813920	-73.871520	7/18/11	3	8.96
BR5	40.813900	-73.871500	7/25/11	4	8.45
BR5	40.813940	-73.871720	8/1/11	3	4.40
BR5	40.813920	-73.871500	8/9/11	3	6.93
BR5	40.800460	-73.865130	8/15/11	3	3.17
BR5	40.813750	-73.871440	8/22/11	3	4.60
BR5	40.814020	-73.871720	8/30/11	5	5.52
BR5	40.813920	-73.871460	9/6/11	3	3.22
BR5	40.813640	-73.871180	9/12/11	3	4.70
BR5	40.814040	-73.871780	9/19/11	3	5.39
BR5	40.813940	-73.871620	9/26/11	3	3.53
BR5	40.813900	-73.871720	10/3/11	3	4.36
BR5	40.814000	-73.871720	11/14/11	3	7.73
BR5	40.800580	-73.864830	12/13/11	3	8.15
E10	40.843460	-73.764060	1/10/11	3	11.73
E10	40.843420	-73.764580	2/7/11	3	14.08
E10	40.842550	-73.765230	5/16/11	3	8.70
E10	40.843780	-73.765920	5/31/11	3	9.96

8. Next, add a threshold column and input the threshold you are using. For this example we are going to use a strict single threshold of 5 mg/l.

Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/l)	Threshold (mg/l)
BR5	40.813700	-73.871500	5/16/11	3	6.95	5.00
BR5	40.813940	-73.871720	5/31/11	3	6.53	5.00
BR5	40.814020	-73.871820	6/6/11	3	7.21	5.00
BR5	40.813940	-73.871470	6/13/11	3	4.95	5.00
BR5	40.814100	-73.871940	6/20/11	3	5.33	5.00
BR5	40.813540	-73.870680	6/28/11	3	4.31	5.00
BR5	40.813800	-73.871400	7/5/11	3	4.98	5.00
BR5	40.813740	-73.871280	7/11/11	3	11.38	5.00
BR5	40.813920	-73.871520	7/18/11	3	8.96	5.00
BR5	40.813900	-73.871500	7/25/11	4	8.45	5.00
BR5	40.813940	-73.871720	8/1/11	3	4.40	5.00
BR5	40.813920	-73.871500	8/9/11	3	6.93	5.00
BR5	40.800460	-73.865130	8/15/11	3	3.17	5.00
BR5	40.813750	-73.871440	8/22/11	3	4.60	5.00
BR5	40.814020	-73.871720	8/30/11	5	5.52	5.00
BR5	40.813920	-73.871460	9/6/11	3	3.22	5.00
BR5	40.813640	-73.871180	9/12/11	3	4.70	5.00
BR5	40.814040	-73.871780	9/19/11	3	5.39	5.00
BR5	40.813940	-73.871620	9/26/11	3	3.53	5.00
BR5	40.813900	-73.871720	10/3/11	3	4.36	5.00
BR5	40.814000	-73.871720	11/14/11	3	7.73	5.00
BR5	40.800580	-73.864830	12/13/11	3	8.15	5.00
E10	40.843460	-73.764060	1/10/11	3	11.73	5.00
E10	40.843420	-73.764580	2/7/11	3	14.08	5.00
E10	40.842550	-73.765230	5/16/11	3	8.70	5.00
E10	40.843780	-73.765920	5/31/11	3	9.96	5.00

9. Now, we can use the threshold to score each sample. By using a formula you can make sure the scoring is done accurately. The formula we're using in this case is =IF(F2<G2,0,100). Which in our spreadsheet is saying If F2 (dissolved oxygen) is less than G2 (the threshold or 5 mg/l) give it a 0 score, or if not give it a 100 score. Copy this formula all the way down so that all the samples are scored.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/l)	Threshold (mg/l)	Score					
2	BR5	40.813700	-73.871500	5/16/11	3	6.95	5.00	100.00					
3	BR5	40.813940	-73.871720	5/31/11	3	6.53	5.00	100.00					
4	BR5	40.814020	-73.871820	6/6/11	3	7.21	5.00	100.00					
5	BR5	40.813940	-73.871470	6/13/11	3	4.95	5.00	0.00					
6	BR5	40.814100	-73.871940	6/20/11	3	5.33	5.00	100.00					
7	BR5	40.813540	-73.870680	6/28/11	3	4.31	5.00	0.00					
8	BR5	40.813800	-73.871400	7/5/11	3	4.98	5.00	0.00					
9	BR5	40.813740	-73.871280	7/11/11	3	11.38	5.00	100.00					
10	BR5	40.813920	-73.871520	7/18/11	3	8.96	5.00	100.00					
11	BR5	40.813900	-73.871500	7/25/11	4	8.45	5.00	100.00					
12	BR5	40.813940	-73.871720	8/1/11	3	4.40	5.00	0.00					
13	BR5	40.813920	-73.871500	8/9/11	3	6.93	5.00	100.00					
14	BR5	40.800460	-73.865130	8/15/11	3	3.17	5.00	0.00					
15	BR5	40.813750	-73.871440	8/22/11	3	4.80	5.00	0.00					
16	BR5	40.814020	-73.871720	8/30/11	5	5.52	5.00	100.00					
17	BR5	40.813920	-73.871460	9/6/11	3	3.22	5.00	0.00					
18	BR5	40.813640	-73.871180	9/12/11	3	4.70	5.00	0.00					
19	BR5	40.814040	-73.871780	9/19/11	3	5.39	5.00	100.00					
20	BR5	40.813940	-73.871620	9/26/11	3	3.53	5.00	0.00					
21	BR5	40.813900	-73.871720	10/3/11	3	4.36	5.00	0.00					
22	BR5	40.814000	-73.871720	11/14/11	3	7.73	5.00	100.00					
23	BR5	40.800580	-73.864830	12/13/11	3	8.15	5.00	100.00					
24	BR5	40.813940	-73.871720	1/10/11	3	11.23	5.00	100.00					

10. These are the scores for each sample, but we also want to know the scores for each station. Copy the sheet and paste it into a new sheet called "DO station scores".

11. Add a new column called Station Score.

12. Average the all the sample scores for one station together.

Dissolved Oxygen Data Analysis

Home Insert Page Layout Formulas Data Review View

Cut Copy Paste Format

Wrap Text

Number

Conditional Formatting Format as Table Cell Styles

$=AVERAGE(H2:H23)$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
	Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/L)	Threshold (mg/L)	Score	Station Score										
2	BR15	40.813750	-73.871930	5/18/11	3	8.95	5.00	100.00	94.55										
3	BR15	40.813940	-73.871720	5/21/11	3	8.53	5.00	100.00											
4	BR15	40.814200	-73.871930	6/6/11	3	7.21	5.00	100.00											
5	BR15	40.813640	-73.871470	8/13/11	3	4.96	5.00	0.00											
6	BR15	40.814100	-73.871940	6/20/11	3	5.33	5.00	100.00											
7	BR15	40.813540	-73.870680	6/28/11	3	4.31	5.00	0.00											
8	BR15	40.813800	-73.871400	7/5/11	3	4.98	5.00	0.00											
9	BR15	40.813740	-73.871260	7/11/11	3	11.38	5.00	100.00											
10	BR15	40.813820	-73.871520	7/18/11	3	8.96	5.00	100.00											
11	BR15	40.813800	-73.871500	7/25/11	4	8.45	5.00	100.00											
12	BR15	40.813640	-73.871720	8/1/11	3	4.40	5.00	0.00											
13	BR15	40.813620	-73.871500	8/9/11	3	8.93	5.00	100.00											
14	BR15	40.800490	-73.865130	8/15/11	3	3.17	5.00	0.00											
15	BR15	40.813750	-73.871440	8/22/11	3	4.80	5.00	0.00											
16	BR15	40.814020	-73.871720	8/30/11	5	5.52	5.00	100.00											
17	BR15	40.813620	-73.871460	9/6/11	3	3.22	5.00	0.00											
18	BR15	40.813640	-73.871180	9/12/11	3	4.70	5.00	0.00											
19	BR15	40.814040	-73.871780	9/19/11	3	5.39	5.00	100.00											
20	BR16	40.813640	-73.871620	9/26/11	3	3.53	5.00	0.00	96.96										
21	BR16	40.813600	-73.871720	9/26/11	3	3.96	5.00	0.00											
22	BR16	40.814020	-73.871720	11/14/11	3	7.73	5.00	100.00											
23	BR16	40.800580	-73.864830	12/13/11	3	6.16	5.00	100.00											
24	BR16	40.843480	-73.764060	1/10/11	3	11.73	5.00	100.00											
25	BR16	40.843420	-73.764580	2/7/11	3	14.08	5.00	100.00											
26	BR16	40.842050	-73.765230	9/18/11	3	8.70	5.00	100.00											
27	BR16	40.843780	-73.765920	5/31/11	3	8.96	5.00	100.00											
28	BR16	40.844080	-73.764700	6/6/11	3	8.17	5.00	100.00											
29	BR16	40.844000	-73.764160	8/13/11	3	8.02	5.00	100.00											
30	BR16	40.843240	-73.765040	6/20/11	3	8.23	5.00	100.00											
31	BR16	40.843800	-73.765790	6/28/11	3	7.08	5.00	100.00											
32	BR16	40.844400	-73.764820	7/5/11	3	5.58	5.00	100.00											
33	BR16	40.843560	-73.765040	7/11/11	3	8.97	5.00	100.00											

All DO data DO data analysis DO station scores +

13. At this point, it is good reorganize the station scores, and paste them as values instead of the formulas that are currently there. Copy this sheet, and paste as “values and source formatting” on a new tab called “DO overall scores”.
14. Organize the spreadsheet by the Station Score.

Site	Latitude	Longitude	Date	Depth (ft)	O ₂ (mg/l)	Threshold (mg/l)	Score	Station Score
BR5	40.813700	-73.871500	5/16/11	3	6.95	5.00	100.00	54.55
BR5	40.813940	-73.871720	5/31/11	3	6.53	5.00	100.00	
BR5	40.814020	-73.871820	6/6/11	3	7.21	5.00	100.00	
BR5	40.813940	-73.871470	6/13/11	3	4.95	5.00	0.00	
BR5	40.814100	-73.871940	6/20/11	3	5.33	5.00	100.00	
BR5	40.813540	-73.870680	6/28/11	3	4.31	5.00	0.00	
BR5	40.813800	-73.871400	7/5/11	3	4.98	5.00	0.00	
BR5	40.813740	-73.871280	7/11/11	3	11.38	5.00	100.00	
BR5	40.813920	-73.871520	7/18/11	3	8.95	5.00	100.00	
BR5	40.813900	-73.871500	7/25/11	4	8.45	5.00	100.00	
BR5	40.813940	-73.871720	8/1/11	3	4.40	5.00	0.00	
BR5	40.813920	-73.871500	8/9/11	3	6.03	5.00	100.00	
BR5	40.800460	-73.865130	8/15/11	3	3.17	5.00	0.00	
BR5	40.813750	-73.871440	8/22/11	3	4.80	5.00	0.00	
BR5	40.814020	-73.871720	8/30/11	5	5.52	5.00	100.00	
BR5	40.813920	-73.871460	9/6/11	3	3.22	5.00	0.00	
BR5	40.813640	-73.871180	9/12/11	3	4.70	5.00	0.00	
BR5	40.814040	-73.871780	9/19/11	3	5.39	5.00	100.00	
BR5	40.813940	-73.871620	9/26/11	3	3.53	5.00	0.00	
BR5	40.813900	-73.871720	10/3/11	3	4.36	5.00	0.00	
BR5	40.814000	-73.871720	11/14/11	3	7.73	5.00	100.00	
BR5	40.800580	-73.864830	12/13/11	3	6.15	5.00	100.00	
E10	40.843460	-73.764060	1/10/11	3	11.73	5.00	100.00	86.96
E10	40.843420	-73.764560	2/7/11	3	14.08	5.00	100.00	
E10	40.842550	-73.765230	5/16/11	3	8.70	5.00	100.00	

15. Delete all the rows that do not have station scores. In the example, they are row 15 and down.
16. Delete the Date column, Depth column, O₂ column, Threshold column, and Score column (D through H in the example). Add Overall Score in the next column.
17. Calculate the overall score by averaging all the station scores together.

Site	Latitude	Longitude	Station Score	Overall score
E14	40.800480	-73.864330	37.50	59.66
FLC1	40.761980	-73.836660	37.50	
FLC2	40.765100	-73.843100	45.83	
E6	40.785840	-73.861200	50.00	
E4	40.782300	-73.921120	54.17	
E7	40.802800	-73.819960	54.17	
BR5	40.813700	-73.871500	54.55	
E15	40.762900	-73.849110	58.33	
E8	40.801560	-73.780740	62.50	
FB1	40.772620	-73.853060	68.00	
E11	40.790280	-73.765620	79.17	
E10	40.843460	-73.764060	86.96	
LN1	40.779840	-73.761040	86.96	

This is a simple example, but helps to give an idea of how to score your data. Good luck!

Synthesizing for report cards

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?”

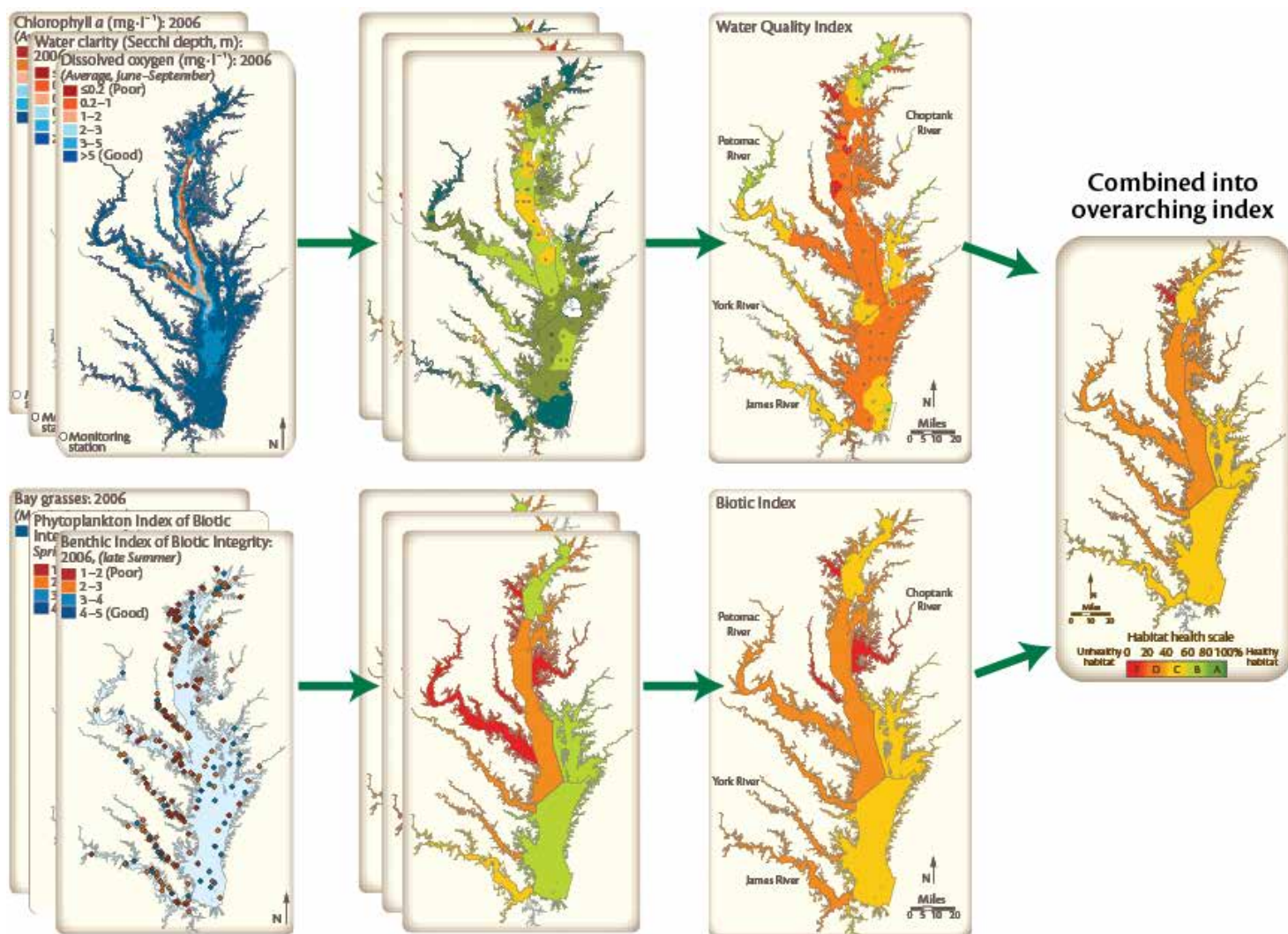
One way to synthesize data is to roll up individual indicators into an overarching index. An index can combine similar types of indicators (e.g., chemical, physical, biological) into one index, or it can be an average of all measured indicators. Overarching indices give a much better integrated assessment (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.

Tidal - How to synthesize

Each of indicators 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.

Selecting reporting regions

Sub-regions of your system may have already been determined to help clarify where to assign sampling sites. 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).

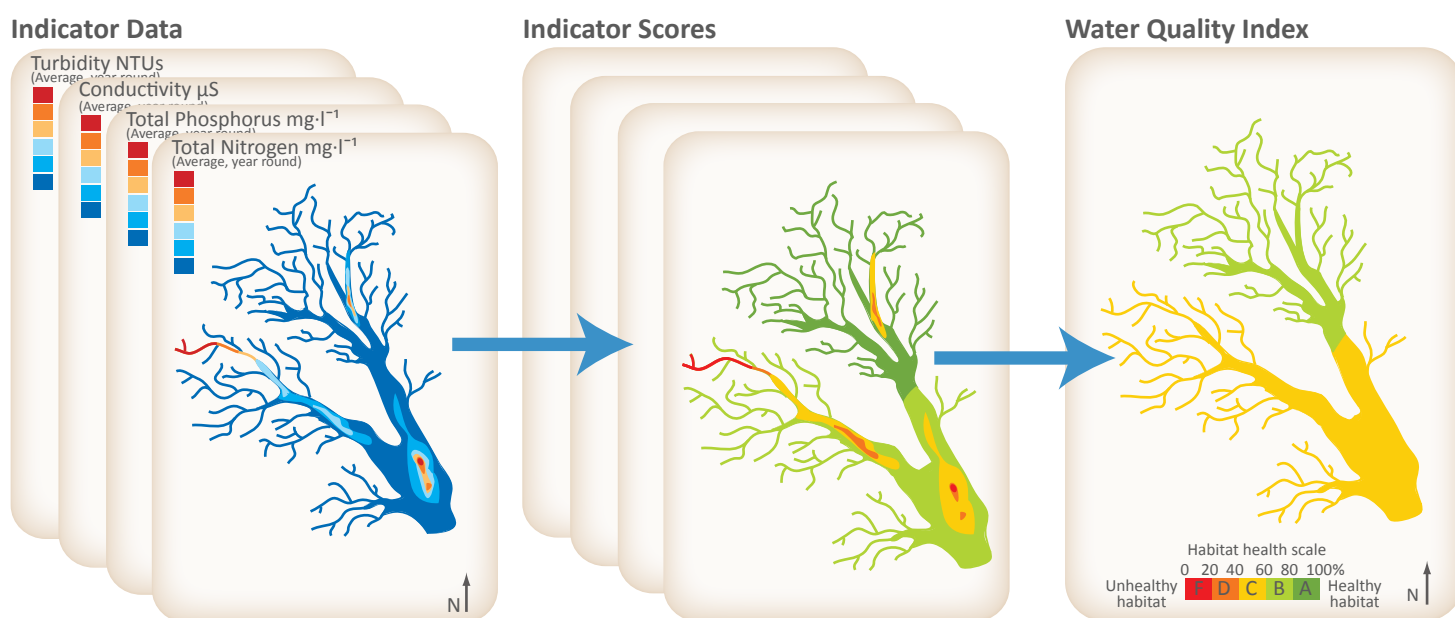


The Chesapeake Bay-wide report card: three water quality indicators and three biotic indicators are evaluated against threshold values. The water quality indicators are then averaged into a water quality index, and the biotic indicators are averaged into a Biotic Index. These two sub-indices are then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

Nontidal - How to synthesize

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 river or stream 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 river or stream?” The audience does not necessarily want to see each measurement that goes into a year-long monitoring program’s database. Rather, they want to know the ultimate outcome of those measurements, or what the data collected mean. Synthesizing data also allows for better communications products that the audience, many times the general public, is able to understand. After synthesizing data and determining the grade of the river or stream, then this information is disseminated through a newsletter or report card.

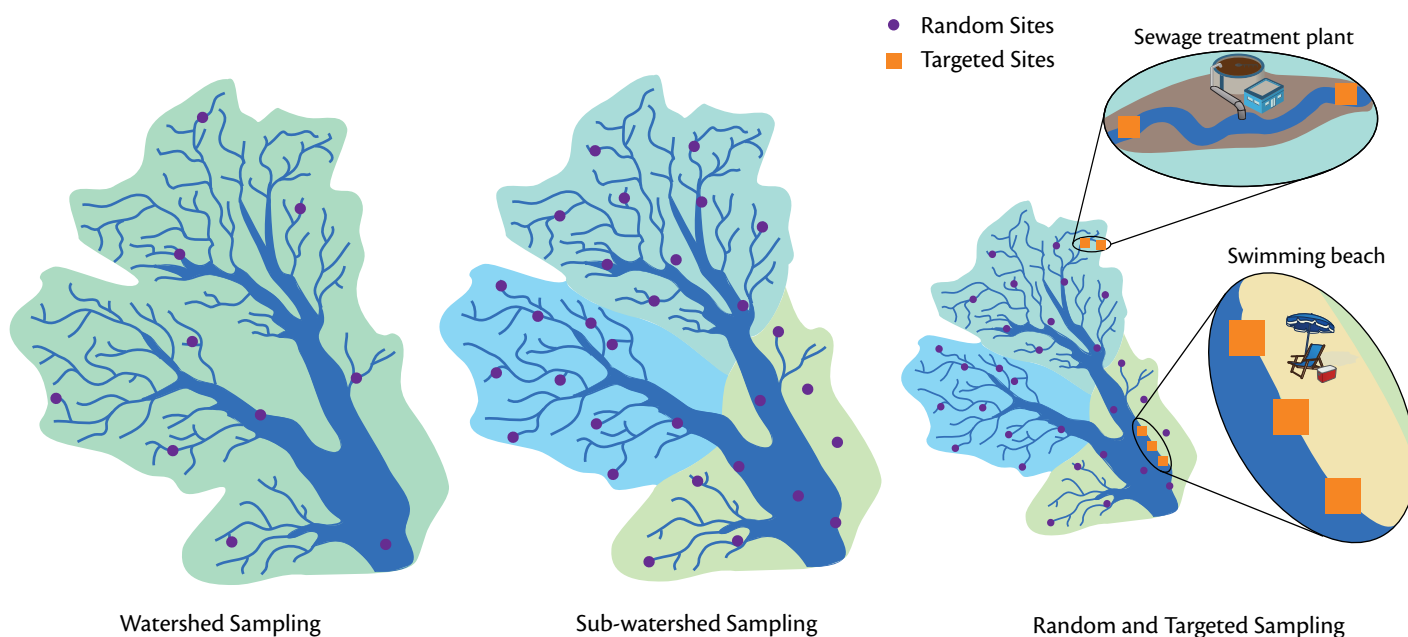
One way to synthesize data is to “roll up” individual indicators into an overarching index. An index can combine similar types of indicators (e.g., chemical, physical, biological) into one index (Figure x), or it can be an average of all measured indicators. Overarching indices give a better integrated 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.



In this example of a nontidal report card, four water quality indicators are evaluated against threshold values. The water quality indicators are then averaged into a Water Quality Index, which gives information on the health of the river or stream.

Selecting reporting sub-watersheds

Sub-watersheds of your system may have already been determined to help clarify where to assign sampling sites. 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 (at least 10 are recommended) in a sub-watershed to provide a representative and accurate score for each indicator. The boundaries of the sub-watersheds are defined by topography, but when delineating sub-watersheds, consider the land use, population, and contribution of the sub-watershed to the entire watershed.



A minimum of 10 sites should be chosen at either the watershed scale, or for each sub-watershed. Site locations should be randomly selected for an unbiased condition assessment. Targeted sampling sites can also be selected to study specific issues, like the safety of swimming beaches or the effectiveness of a sewage treatment plant.


Water quality index

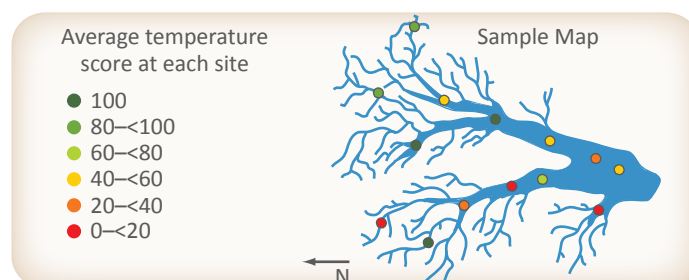
The four core indicators used in this protocol, total nitrogen, total phosphorus, conductivity, and turbidity, can be averaged together for a water quality score for a sub-watershed. Then, scores for each sub-watershed are area-weighted (i.e., the area of the sub-watershed divided by the total area of the watershed) and averaged for one watershed score. Each monitoring program will need to decide if it wants to provide sub-watershed scores, or if it wants to average the indicators into one Water Quality Index (WQI) for the entire river or stream. It is recommended that quantitative grades are given for each indicator and the Water Quality Index. When giving an overall grade for the watershed, all of the indicators will be wrapped together into a single overarching score.

Vital signs indicators


Dissolved oxygen, pH, and water temperature are considered to be vital signs of the watershed, and give the pulse of the system. Vital signs are generally not directly reported on in communication products such as report cards. These indicators are usually “Very good” unless there is a site-specific reason for them to be poor, such as an area of the stream where chemicals have been dumped, causing poor pH levels. Vital signs should be measured and monitored for sudden changes, and can be reported if they are strongly influencing the health of the waterway, such as consistent low dissolved oxygen levels.

When one or more vital sign indicator is showing low scores, then the vital signs indicators should be further examined and evaluated for what may be going on in the river or stream. If one vital signs indicator scores lower than 80%, this low

Indicator	Score	Visual
Dissolved oxygen	95%	
pH	100%	
Water temperature	70%	



When any of these indicators scores are lower than 80%, a thumbs down symbol and a map of the sampling sites with the specific scores should be provided.

Indicator	Score	Visual
Dissolved oxygen	99%	
pH	100%	
Water temperature	95%	

If all vital signs indicators score 80% or higher, then a thumbs up symbol can be used to express the health of the ecosystem.

grade should be expressed with a red thumbs down symbol. Accompanying the symbol should be a map of the watershed displaying each sampling site and showing where the indicator is doing poorly. If all of the indicators are doing well, a green thumbs up symbol can be used, and a map of the sampling sites is not necessary.

Since vital signs indicators should not have much variability and should score well, wrapping them up with the other indicators would skew the grades. For this reason, these vital signs indicators are not wrapped up with the four core indicators or with each other.

Bacteria

Bacteria is a human health indicator, which is communicated differently than ecological indicators. If desired, however, bacteria scores can be incorporated into an overall grade calculation. Bacteria, calculated on a 10-point scale (not the 20-point scale used by ecological indicators), are presented as station averages.

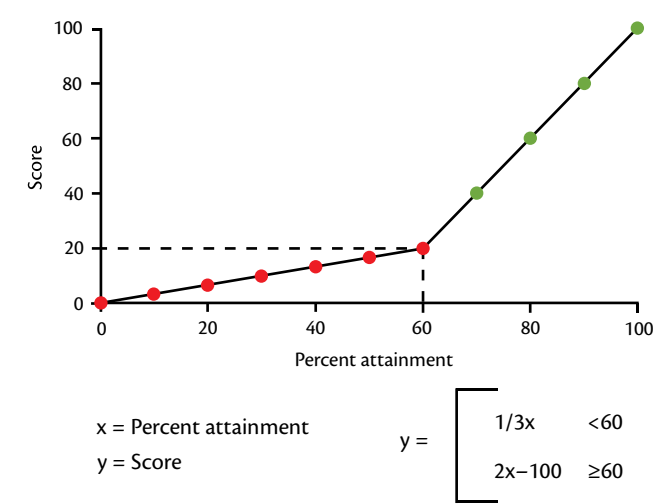
Furthermore, due to the variability of bacteria scores between sampling sites, a map of station average scores should be presented with the specific grade per sampling station. When expressing bacteria data alone, an overall bacteria grade for the watershed should not be expressed. For bacteria, specific sites can have high variability, and averaging all of the sites into a single score loses the resolution of the data. For instance, one site could be consistently poor, but would be averaged out if not expressed individually.

To communicate the data, annually averaged results for each bacteria sampling site should be displayed spatially and accompanied by a “fire danger” diagram indicating the “annual relative risk of humans becoming sick from swimming” in those locations.

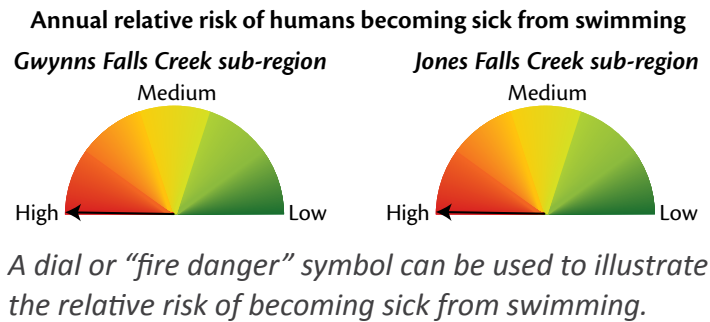
Bacteria station scores can be averaged together for an overall bacteria watershed score. In order to integrate bacteria with the other indicators, it has to be standardized to a 20-point scale.

The final bacteria score can be averaged with other human health indicators into an index. If other human health indicators are available, the bacteria score or Human Health Index can be evenly averaged with the Water Quality Index, benthic community, and trash. Additional human health indicators (such as toxins, like heavy metals, or carcinogens) are not directly addressed in this protocol.

Remember, although we recommend averaging the data to wrap up bacteria with the other indicators, we do not recommend expressing the overall bacteria score in communication materials.



This graph shows the conversion of the 10-point scale used by bacteria, to the 20-point scale used by all other indicators.



Benthic community

Benthic community can be communicated separately, as well as wrapped up with the Water Quality Index, and bacteria indicators into an overall watershed health grade.

To roll up the individual sampling site scores into an overall watershed score, a minimum of 5 sampling sites are needed for small (HUC-12) watersheds and a minimum of 10 sampling sites are needed for larger (HUC-8) watersheds. While these data are quantitative and may be spatially integrated, state data reporting is only completed every other year. The average benthic community score for the watershed will be reported for two years in a row in your report card. This letter grade is the score that will be wrapped in with the rest of the indicators.

Overall grade

The overall grade of the river or stream integrates the Water Quality Index (total nitrogen, total phosphorus, conductivity, and turbidity), bacteria, and benthic community results. These four items can be considered indexes that are integrated into an overall score. In addition to showing the overall score, another method for communicating the results that shows more detail is to generate a grid of the indicators that may be spatially averaged (i.e. all but bacteria) along one axis and the list of sub-watersheds on the other axis. The color in each grid cell would then indicate the grade for that indicator in that location.

Sub-regions	Dissolved Oxygen	pH	Temperature	Total nitrogen	Total phosphorus	Turbidity	Conductivity	Benthics
1								
2								
3								
Average watershed score per indicator								

Using a grid of scores is another method for communicating the results that shows more detail about what is happening with each indicator in each sub-region and the overall watershed.

Communication strategy

A well-rounded communications strategy outlines key messages (what you want to convey), identifies target audiences (with whom you want to communicate), helps choose a spokesperson, and determines communication vehicles (the documents or techniques through which you communicate). 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.

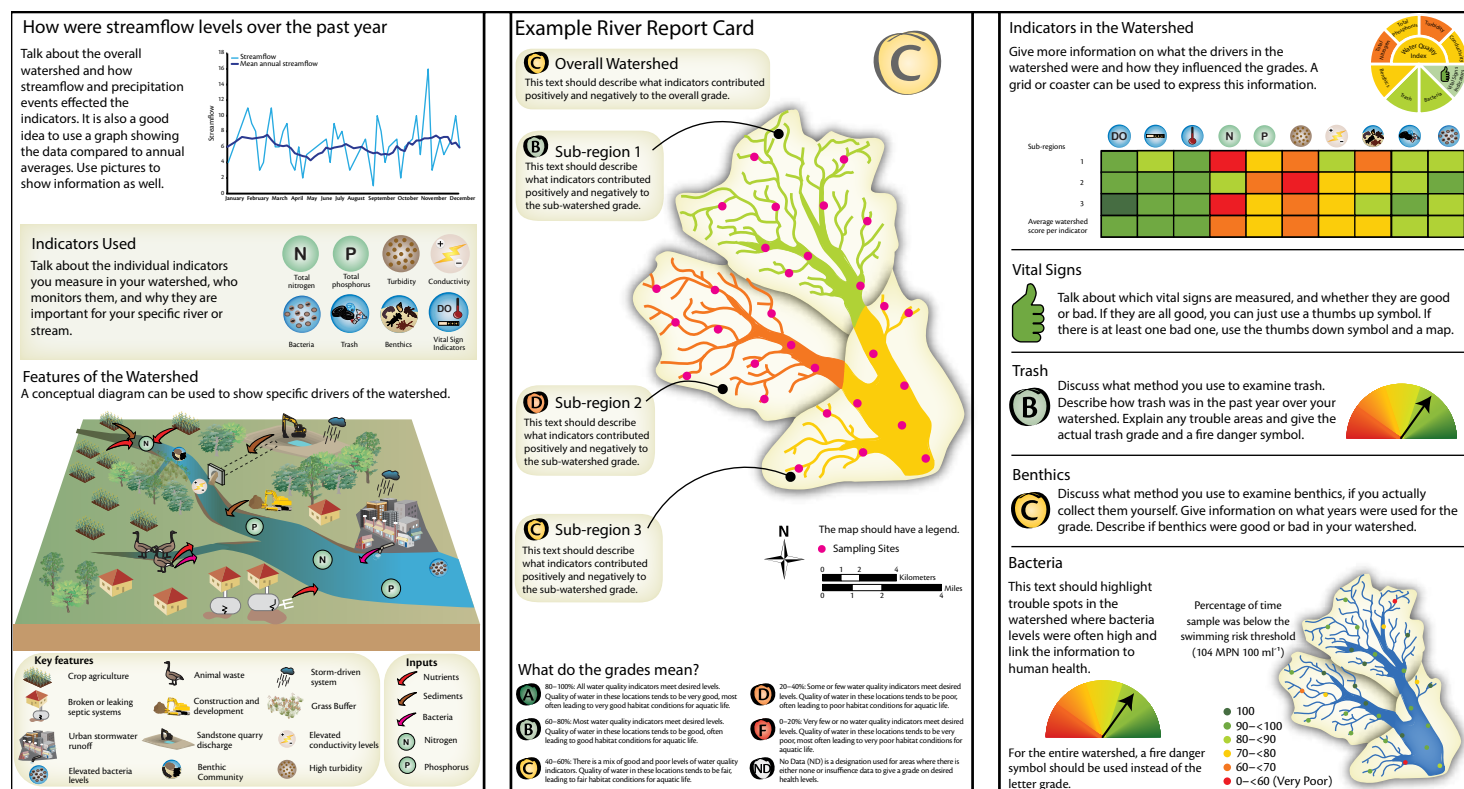
A website is now considered an essential science communication tool. It allows the widest possible audience to be reached in the most timely manner, without the normal delays of print media. The constant ability to edit and refine a website is one of the key features that makes them effective for science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



Examples of different report card products. Top to bottom: 2011 Chester River report card (4-page newsletter), 2011 State of the Anacostia River (8-page brochure), 2012 West & Rhode River Report Card (12-page brochure), and 2012 South River Report Card (17-page pamphlet).

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.



This sample spread of a report card shows how to integrate all of the elements to create a visually appealing and informative communication product.

COMMUNICATING SCIENCE EFFECTIVELY

At the start of this workbook, the interpretation section took you through the very basics of compiling and understanding the data you collect. This allowed you to ensure the validity and accuracy of your data and to format it into usable graphics that accurately represents your measurements in the field.

The synthesis section then moved on to more advanced and creative ways of interpreting your (now) well-organized and compiled data. One thing that should be noted is that although there are many rules to help you be a better scientific communicator and data visualizer, an overarching tip from this section is to simply be more thoughtful about what your data means, and who you are communicating to. Giving more thought yourself or bringing together a room full of people to tackle a concept can really bring it to life and make it much more understandable. The specific tools, such as visualization design tips, ABT, narrative structure, and active titles, provided in this section can help to bring your data to life.

However, these sections do not stand alone sections. Both pieces—interpretation and synthesis – heavily rely on one another in order for you to successfully communicate the data that your organization collects. These are building blocks of effective science communication.

This handbook is best used in tandem with the supporting materials listed at the beginning of this document and best understood when combined with a workshop as you have done over the past two days. We hope that you keep this handbook and use it to help your colleagues understand the data they are collecting and communicate it succinctly.

Thank you for working with us over the past two days.

RESOURCES AND REFERENCES

- Buchanan, C., Foreman, K., Johnson, J., & Griggs, A. 2011 Development of a Basin-wide Benthic Index of Biotic Integrity for Non-tidal Streams and Wadable Rivers in the Chesapeake Bay Watershed: Final Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. Interstate Commission on the Potomac River Basin. Rockville, MD. <http://www.potomacriver.org/2012/publicationspdf/ICPRB11-01.pdf>
- Burgan, B., Carpenter, D., Gould, D., Keeler, B., McGovern, C., & Miller, S. 2007. Indicator Development for Estuaries. (EPA Publication No. 842-B-07-004). Government Printing Office, Washington, D.C.
- Chesapeake Bay Program. (n.d.) Health of Freshwater Streams in the Chesapeake Bay Watershed. Annapolis, MD. http://www.chesapeakebay.net/indicators/indicator/health_of_freshwater_streams_in_the_chesapeake_bay_watershed
- Dates, G. and J. Schoen. (n.d.) Data Interpretation. River Watch Network and Massachusetts Water Watch Partnership at the University of Massachusetts. Amherst, Massachusetts. <http://www.umass.edu/mwwp/pdf/intmanl.pdf>
- Interstate Commission on the Potomac River Basin (ICPRB). 2011. Data Analysis to Support Development of Nutrient Criteria for Maryland Free-Flowing Waters. Maryland Department of the Environment. Baltimore, MD.
- Jackson, L.E., Kurtz, J.C., & Fisher, W.S. Eds. 2000. Evaluation Guidelines for Ecological Indicators. (EPA Publication No. 620-R-99-005). U.S. Government Printing Office, Research Triangle Park, NC. 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. CRC Press, Danvers, MA.
- Longstaff, B.J., Carruthers T.J.B., Dennison W.C., Lookingbill T.R., Hawkey J.M., Thomas J.E., Wicks E.C., & Woerner J. Eds. 2010. Integrating and applying science: A practical handbook for effective coastal ecosystem assessment. IAN Press, Cambridge, Maryland, USA. 244pp. http://ian.umces.edu/press/books/publication/259/integrating_and_applying_science_a_handbook_for_effective_coastal_ecosystem_assessment_2010-05-10/
- 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. 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. mddnr.chesapeakebay.net/eyesonthebay/documents/QAPP_MainTrib_2010-2011_Draft1.pdf

-
- Massachusetts Water Watch Partnership. 1994. Volunteer Water Quality Monitoring Manual. Amherst, MA.
- McCandless, D. 2012. Information is Beautiful. New Edition, revised, recalculated and reimagined. Collins, London, England.
- McCandless, D. 2014. Knowledge Is Beautiful: Impossible Ideas, Invisible Patterns, Hidden Connections--Visualized. Harper Design, New York, New York.
- McKenzie, D.H., Hyatt, D.E., & McDonald, V.J. 1992. Ecological indicators: Volumes 1 and 2. Proceedings of the International Symposium on Ecological Indicators. Elsevier Science Publishers, Essex, England.
- 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).
- Olson, R. 2015. Houston, we have a narrative: Why science needs story. University of Chicago Press, Chicago, Illinois.
- Schoen, J., Walk, M-F., & Tremblay, M.L. 1999. Ready, Set, Present! a data presentation manual for volunteer water quality monitoring groups. Massachusetts Water Watch Partnership. University of Massachusetts, Amherst, Massachusetts. <http://www.umass.edu/mwwp/pdf/ReadySetPresent.pdf>
- Schoen, J. and M-F. Walk. 2002. Data Interpretation Manual for Volunteer Monitors. Massachusetts Water Watch Partnership at the University of Massachusetts. Amherst, Massachusetts. <http://www.umass.edu/mwwp/pdf/data-interp-2002.pdf>
- 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.
- Sisense. 2017. How to choose the right data visualization. Sisense.com. <https://www.sisense.com/whitepapers/choose-right-data-visualization/>
- 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. IWA Publishing, London, England. 132pp. Free PDF download: <https://explorer.ian.umces.edu/public/7604ed>. Buy online: http://ian.umces.edu/press/books/publication/65/communicating_science_effectively_a_practical_handbook_for_integrating_visual_elements_2006-07-01/
- Townend, J. 2002. Practical statistics for environmental and biological scientists. John Wiley & Sons Ltd., West Sussex, England.
- Tufte, E. 1990. Envisioning Information. Graphics Press LLC, Cheshire, Connecticut.

-
- Tufte, E. 1997. Visual Explanations: Images and quantities, evidence, and narrative. Graphics Press LLC, Cheshire, Connecticut.
- Tufte, E. 2001. The Visual Display of Quantitative Information, 2nd edition. Graphics Press LLC, Cheshire, Connecticut.
- Tufte, E. 2006. Beautiful evidence. Graphics Press LLC, Cheshire, Connecticut.
- U.S. Environmental Protection Agency. 1991. Volunteer Lake Monitoring: A Methods Manual (2nd ed.). (EPA Publication No. 440-4-91-002). U.S. Government Printing Office, Washington, DC.
- U.S. Environmental Protection Agency. 1997. Volunteer Stream Monitoring: A Methods Manual. (EPA Publication No. 841-B-97-003). U.S. Government Printing Office, Washington, DC. http://water.epa.gov/type/rsl/monitoring/stream_index.cfm
- U.S. Environmental Protection Agency. 2000 a. Ambient Water Quality Recommendations for Rivers and Streams in Nutrient Ecoregion IX. (EPA Publication No. 822-B-00-019). Washington, DC: U.S. Government Printing Office.
- U.S. Environmental Protection Agency. 2000 b. Ambient Water Quality Recommendations for Rivers and Streams in Nutrient Ecoregion XI. (EPA Publication No. 822-B-00-020). Washington, DC: U.S. Government Printing Office.
- U.S. Environmental Protection Agency. 2000 c. Ambient Water Quality Recommendations for Rivers and Streams in Nutrient Ecoregion XIV. (EPA Publication No. 822-B-00-022). 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). U.S. Government Printing Office, Washington, DC.
- U.S. Environmental Protection Agency. 2004. Technical Support Document for Identification of Chesapeake Bay Designated Uses and Attainability, 2004 Addendum. U.S. Governmental Printing Office, Washington, D.C.
- U.S. Environmental Protection Agency. 2006. Volunteer Estuary Monitoring: A Methods Manual (2nd ed.). (EPA Publication No. 842-B-06-003). U.S. Government Printing Office, Washington, DC. water.epa.gov/type/oceb/nep/monitor_index.cfm
- 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. www.chesapeakebay.net/content/publications/cbp_27849.pdf
- U.S. Environmental Protection Agency. 2012. 5.1 Stream Flow. from Volunteer Stream Monitoring: A Methods Manual. (EPA Publication No. 841-B-97-003). Washington, DC: U.S. Government Printing Office. Retrieved August 8, 2012, from <http://water.epa.gov/type/rsl/monitoring/vms51.cfm>

Valiela, I. 2001. *Doing Science: design, analysis, and communication of scientific research*. Second Edition. Boston University Marine Program, Marine Biological Laboratory. Oxford University Press, Woods Hole, Massachusetts.

Virginia Institute of Marine Science. 2010. Submerged aquatic Vegetation in Chesapeake Bay. www.vims.edu/bio/sav

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.

Other Resources

Science Visualization semester-long course <https://sciencevisualizationclass.wordpress.com/>

