Converting Data to Information

This section will show you how to:

◆ Work with graphs and tables to interpret your data.

◆ Work with graphs, tables and charts to report your data.

◆ Use common assessment methods, benchmarks and indices for lakes, streams and rivers, and wetlands.

The payoff for all your hard work

The purpose of your monitoring program may be to make the results available to fellow volunteers, the community where you are monitoring or regulatory agencies. In order for you and others to make sense of the numbers you have generated, the numbers will need to be transferred into a format and context that is coherent and easy to understand.

Some data are quite simple to interpret. For example, Secchi disk readings are easy to correlate with chlorophyll-*a* data to determine whether or not algae is the primary factor affecting water transparency. Other numbers will require more expertise. You may be able to do some of the work yourself, especially if you have some background in science or the patience to learn. Otherwise, you may decide to work with an agency or organization that will interpret the numbers for you.

Interpreting data

This Section will introduce you to the basics of data interpretation and reporting. Think of data interpretation as a process in which you ask a series of questions that lead you to **findings** and **conclusions**.

**Findings** are objective observations about your data. **Conclusions** are how you explain why the

data look the way they do. For example, if you are monitoring a lake to determine its trophic state, your **findings** can, for example, indicate nutrient concentrations or the relationships between chlorophyll-*a* concentrations, total phosphorus, and Secchi transparency. Based on those findings, you can draw **conclusions** as to whether or not excessive nutrients are causing algae blooms and high chlorophyll-*a* concentrations, and in turn problems with water clarity. Conclusions are drawn by statistically testing hypotheses developed from your findings.

Long-Term Water Quality Data Tells a Lake's "Story"

Since 1979, Lay Monitors on Lake St. Catherine, located in the towns of Poultney and Wells, have collected weekly summer samples of total phosphorus and chlorophyll-*a* and taken weekly Secchi water clarity readings. These volunteers have helped document baseline water quality conditions for their lake. Their data show a decline in conditions, especially noted by the drop in water clarity annual means since 1998 as seen on the graph below. This decrease in clarity has been attributed to the arrival of the invasive alewife fish, *Alosa pseudoharengus.* Alewives are "top grazers," who feed on zooplankton (microscopic animals). Fewer zooplankton are not able to keep the phytoplankton (algae) levels down and algae populations explode, lowering the clarity.

Lake St. Catherine
Data show a decrease in water clarity since the discovery of the invasive alewife fish, whose feeding habits have affected the lake conditions

---
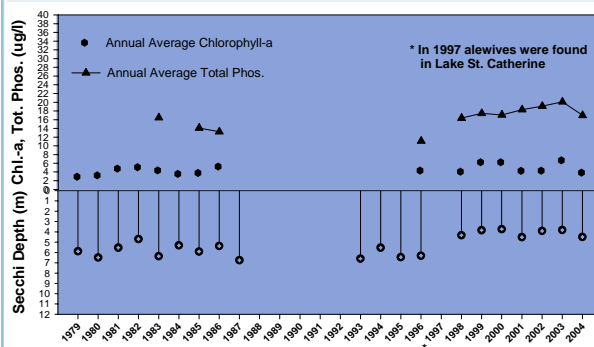
Determining findings

Asking some of these questions can help you arrive at findings:

◆ Which sites consistently did not meet the water quality goals? By how much?

◆ Are there seasonal differences in results?

◆ Did flow or rainfall affect results?

◆ Do results change in a consistent manner upstream or down?

◆ Do changes in one parameter coincide with changes in another? For example, is there an inverse relationship between Secchi transparency readings and chlorophyll-*a* measurements?

Creating graphs

To assess findings, first graph your data to visually display results. This will help you compare parameters. Table 7-1 lists graphs you can consider creating. (More on creating graphs is covered later in this Section).

Reaching conclusions

Once you have organized your data into findings, you can start to assess whether or not you can answer your monitoring question(s), address your study purpose and make conclusions. Then, once you develop conclusions, you can organize them into a presentation. Good presentation of information is essential to effectively communicate and gain credibility for your results.

In reducing your data down to usable information, the key is to make conclusions that your

Table 7-1: Graphs and Comparisons to Consider When Assessing Data

Graph	Comment
Flow vs. any parameter	May show nonpoint source pollution effects or dilution of dissolved parameters at high flows
Date vs. observed values/concentrations	May show trends or seasonal variation
Precipitation vs. any parameter	May show how parameters respond to rainfall and/or nonpoint source pollution effects
Secchi transparency readings vs. chlorophyll-*a* measurements	May show that algae blooms are the primary factor affecting transparency, or suggest that non-algae turbidity or color (organic acids) is affecting transparency
Chlorophyll-*a* measurements vs. total phosphorus	May indicate that phosphorus is the controlling factor for algae growth
Secchi transparency readings vs. total phosphorus	Shows relationship between primary nutrient and water clarity
Dissolved oxygen and temperature depth profiles	May show stratification or mixing status in lakes
Parameters vs. numerical standards/criteria	May indicate problem areas
Bacteria vs. total suspended solids or turbidity	May indicate that bacteria are associated with solids, and reductions in bacteria could be achieved by controlling or trapping solids
Chemical values or biological assessments (biometrics) vs. river mile or sampling station	May show trends by location or points/locations where major changes are noticeable (to see upstream-to-downstream trends)

data support. One conclusion may be that additional data are needed. That is an acceptable conclusion. You may arrive at a conclusion that others disagree with. Following these steps will put you in a strong position to defend your conclusions:

◆ Follow a logical process that has a scientific basis.

◆ Get help from other knowledgeable people. Most professionals and scientists enjoy reviewing and assessing datasets.

◆ Document your assumptions and your assessment process.

Take your conclusions back to your "why" question, on which you based your monitoring plan. If you can answer the question(s), your work is done.

It is likely, however, that you only will answer part of your question(s), or find some additional questions. For example, you may need to assess whether findings and conclusions can be explained by natural conditions, human alterations, and/or errors in sampling or analysis.

Natural conditions or human alterations may affect findings and conclusions

Consider some of the following questions to help you decide if human alterations or natural conditions can explain your results.

◆ Might natural upstream-to-downstream changes in the river account for your results? Your benthic macroinvertebrate results might be explained by natural shifts in the macroinvertebrate community composition from headwaters to mouth.

◆ Does weather appear to influence your results? For example, do problem levels coincide with intense rainstorms? Might elevated temperature levels be caused by unusually hot weather?

◆ Do problem levels coincide with rising flow? For example, are elevated bacteria counts only present during storm flows, which would indicate nonpoint runoff sources? Or are they only present during low flows, which might suggest point discharge sources?

◆ Does the presence of specific sources explain your results? For example, can you attribute increased bacteria levels to a wastewater treatment plant or a failing septic system?

◆ Do changes in one parameter appear to explain changes in another? For example, could low dissolved oxygen be explained by high temperature?

◆ Do your visual observations explain any of your results? Did your volunteers report any strange pipes, eroding banks or dry weather seeps from storm drains? Did volunteers see evidence of pollution (e.g., tires, trash, oil slicks)?

◆ For multiple years of data, are there overall trends that coincide with changes in land use or habitat? For example, did the macroinvertebrate community improve over time following streambank stabilization work?

◆ If you are monitoring the impact of a pollution source, are there other upstream impacts that might be influencing and confusing your results? For example, if there is no riparian vegetation for shade upstream of an outfall, it might be difficult to figure out which factor or combination of factors is causing elevated temperatures.

Sampling and analysis factors affecting findings and conclusions

Your results may also be explained by the way you collected and analyzed samples, rather than by changes in the resource itself. To determine if this is the case, consider the following questions:

♦ Could flaws in your field and/or laboratory techniques explain your results? Could high concentrations be due to contamination or sampling error? Double check your QA/QC sample results to confirm data quality.

♦ Was your sampling representative of the resource and range of conditions observed? For example, was your sampling primarily conducted when river flows were low? Did you catch storm-related runoff or just base flow? Plot sample times against continuous stream flow records, if available, to check which parts of the flow regime were sampled.

♦ Was your analytical method sensitive enough to detect levels of concern?

♦ Did the time of day you sampled affect your results? For example, dissolved oxygen is typically lowest in the early morning and highest in the afternoon.

Understanding variability

Variability happens. Even with rigorous sampling methods and QA/QC protocol, all monitoring data will have variability. Natural systems are inherently variable, and through sample handling and analysis, we introduce additional variability. Uncertainty, in turn, compounds variability.

Uncertainty arises because there is no such thing as a truly exact measurement, and samples cannot be collected continuously, forever.

Instead, samples are collected *periodically* to represent an environment that is *continually changing over time and space.* These periodic samples are analyzed using methods that have limits in resolution, precision and accuracy.

Information on variability is taken with permission from the article "Variability Happens: Basic Descriptive Statistics for Volunteer Programs" by Julie Rector: *The Volunteer Monitor,* Vol. 7, No.1, Spring 1995.

Working with statistics

Statistics is the science of making decisions in the face of uncertainty. We cannot eliminate uncertainty and variability, but we can use statistics to estimate their contribution to our observed results and make informed decisions based on the data. Statistical methods and a number of assessment methods and indices have been developed to help with water quality data interpretation. Volunteer monitoring programs generally use statistics for three main purposes:

♦ To summarize and report monitoring findings.

♦ To evaluate QA/QC data.

♦ To help interpret data and draw conclusions.

The most frequently used descriptive statistics are those that *describe central tendency* and those that *describe the distribution or variability*. The following examples illustrate these processes as they are used for lakes, streams and wetlands monitoring. Other statistical analyses, such as *trend analysis,* can also be completed. However, they may require years of data and/or more advanced statistical techniques.

If you are interested in more advanced statistical techniques, see references such as *Statistical Methods for Environmental Pollution Monitoring* by R.O. Gilbert, 1987, Van Mostand Reinhold Co., New York.
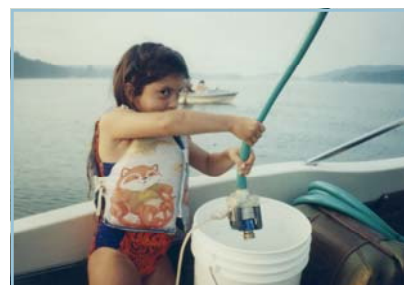
Common measures

Common measures used to analyze data include averages (i.e., arithmetic means), geometric means and medians.

♦ **Average** is calculated by adding all the values and dividing by the number of values. Averages are representative or typical of all the sample observations. A problem with averaging can occur when you have a few very high or very low numbers that distort results. The term "mean" also refers to the average.

♦ **Geometric mean** reduces the influence of very high and very low numbers on the average of a dataset. The geometric mean is commonly used to summarize bacteria data, since the values can fluctuate from single digits into the thousands. There are several ways to calculate a geometric mean. This is only one method:

   1. Multiply all the values in the dataset.

   2. Raise that number to the power of 1 ÷ the number of values in the dataset .

♦ **Median** is the value that divides the distribution of the data into two halves. In other words, 50% of the values are above the median and 50% are below. Medians are meant to be a value representative or typical of the dataset. The median is not affected by outliers (values either extremely high or low) and is frequently more representative of data than the average. This is particularly true when the dataset contains only a few very high or very low numbers.

In general, it is appropriate to use the *average* when datasets are normally distributed (with most values clustered around the average) with no outliers (values follow a bell-shaped curve when graphed). It is better to use the *median* if the dataset is skewed and/or if there are outliers. The only time volunteers typically use the *geometric mean* is for *E. coli* and total phosphorus data.

Average, Geometric Mean and Median Calculation Examples

Consider the following set of *E. coli* data (cfu/100mL = colony forming units per 100 milliliters of water):

Date 1: 22 cfu/100mL   Date 5: 188 cfu/100mL
Date 2: 234 cfu/100mL   Date 6: 77 cfu/100mL
Date 3: 17 cfu/100mL   Date 7: 89 cfu/100mL
Date 4: 36 cfu/100mL

The average, also called "the mean," (95 cfu/100mL) is the sum of all the values divided by the number of values.

$22 + 234 + 17 + 36 + 188 + 77 + 89 = 663$

$663 \div 7 = 95$ cfu/100mL (rounded off)

The geometric mean (63.27 cfu/100mL) can be calculated using the following equation:

$$GM = (r1 \times r2 \times r3 \dots rN)^{1/N}$$

Where:

r = the value for samples 1, 2, 3 through the Nth sample

N = the total number of samples collected

Using our example data, the calculation looks like this:

$GM = (22 \times 234 \times 17 \times 36 \times 188 \times 77 \times 89) = 4,059,088,697,664$

$(4,059,088,697,664)^{1/7} = 63$ cfu/100mL (rounded off)

It may be easier to convert the fraction (in this case 1/7) into a decimal for calculations.

The median (77 cfu/100mL) is the number in the middle when values are ranked in order from lowest to highest. In this example, 50% of values are above and 50% below the number 77.

In this case, the median and geometric mean are more representative of the dataset than the average, as the average is greater than all but two of the measured values.

Converting Data to Information

Measures of distribution

Commonly used measures of distribution include range, quartiles, standard deviation and confidence intervals.

- **Range** is defined as the difference between the maximum and minimum values of your dataset. If you have a wide range, it means there is a lot of variation in your data. A small range indicates low variability and, therefore, greater likelihood that the average (i.e., arithmetic mean) is representative of the dataset.

- **Quartiles** are the values below which lie 25%, 50% and 75% of the values in a dataset. Another way to look at the quartiles is that 50% of your data, or the interquartile range, lies between the 25% and 75% quartiles. If these quartiles are far apart, it means there is a lot of variability in your data. If they are close together, it means your dataset is relatively consistent and is clustered about the median.

- **Standard deviation** describes the variability of the datapoints around the average. For a normally distributed population, the average plus or minus one standard deviation represents a 66% confidence interval. Confidence intervals and standard deviations will be larger when there is a lot of variability. Most scientific calculators have a function for calculating standard deviation, and some will perform confidence intervals.

- **Confidence interval** is a group of continuous values that tends to include the true value a predetermined portion of the time. For example, if we say that the 95% confidence interval for parameter "y" is 6 to 26, that means we are confident that 95% of the time the true value of parameter "y" is between 6 and 26. You may not be able to establish accurate confidence intervals until several years of data have been accumulated.

Deciding which measure to use depends upon the type of data you are summarizing. In general, Table 7-2 (page 60) provides suggestions for the different parameters, but you should check with your data user or consult historical datasets to see how they have been summarized.

Measures of Distribution Calculation Examples

Consider the following set of total phosphorus data from Maidstone Lake Lay Monitors in 2004:

June 5:	7.0 µg/L	July 18:	17 µg/L
June 12:	5.0 µg/L	July 25:	5.2 µg/L
June 20:	6.0 µg/L	Aug. 2:	10 µg/L
June 27:	5.0 µg/L	Aug. 8:	5.9 µg/L
July 4:	5.0 µg/L	Aug. 14:	5.8 µg/L
July 11:	5.9 µg/L	Aug. 28:	5.2 µg/L

The range of the data is 12 µg/L
(maximum value - minimum value)
17 - 5.0 = 12

The quartiles can be determined by arranging the values in ascending order and dividing the values into four equal groups. The median (5.9 µg/L) marks the 50% quartile (for a dataset with an even number of values, the median is the average of the two middle values). The 25% quartile is values of 5.0 or less and the 75% quartile is values of 7.0 or less.

$$
\begin{array}{l}
17 \\
10 \\
75\% \;\; \dfrac{7.0}{6.0} \\
5.9 \\
50\% \;\; \dfrac{5.9}{5.8} \\
5.2 \\
25\% \;\; \dfrac{5.2}{5.0} \\
5.0 \\
5.0
\end{array}
$$

Standard deviation indicates the range of variation in the measurements taken. It is calculated using the following equation:

X = a measured value
$\overline{X}$ = the average of the values
n = the number of values
Σ = the sum of the calculations for each measured value

$$ S = \sqrt{\frac{\Sigma\,[x - \overline{x}]^2}{n - 1}} $$

First, figure out the average of the sample measurements.
7.0 + 5.0 + 6.0 + 5.0 + 5.0 + 5.9 + 17 + 5.2 + 10 + 5.9 + 5.8 + 5.2 = 77.9 ÷ 12 = 6.9 µg/L

Then, for each measured value, calculate the next part of the formula. Using the first sample value as an example, the calculation would look like this:

$(7.0 - 6.9)^2 = (0.1)^2 = 0.01$

After performing that calculation for each value in the dataset, add the results together, divide the result by 11 (12-1) and take the square root of that quotient.

0.01 + 3.6 + 0.8 + 3.6 + 3.6 + 1.0 + 102 + 2.9 + 9.6 + 1.0 + 1.2 + 2.9 = 132 ÷ 11 = 12

$\sqrt{12}\; = 3.5$

Confidence intervals require an advanced statistical analysis and can provide a dependable range of values to expect for a given parameter. Consult the World Wide Web for more information.

The Lake Champlain Basin

Quebec

Richford

Chazy

St. Albans

Plattsburgh

Burlington

Lake Placid

Lake Champlain

Montpelier

Port Henry

Middlebury

New York   Vermont

Ticonderoga

Whitehall

Map provided by the Lake Champlain Basin Program

Specific considerations for water monitoring statistics

When calculating lake parameters, it is general practice to calculate growing season average. In Vermont, the growing season is loosely defined as mid-May through mid-September.

Central tendencies for pollutants in runoff to tributaries, streams and rivers are frequently summarized as flow-weighted average concentrations. Flow-weighted concentrations take into account the fact that concentrations of some parameters vary with flow. For example, concentrations of particulate pollutants (TSS, TP) may be higher at higher flows, which have more energy to suspend and transport particles. This higher concentration, combined with the higher flow, means that a disproportionate amount of the load of that particulate pollutant is transported during high flow events.

Finally, you should have at least five data points to calculate averages, geometric means, medians and quartiles.

Common assessment methods, benchmarks and indices

In addition to descriptive statistics, there are some fairly common assessment methods, benchmarks and indices used by scientists that tell us a lot about surface water quality. This subsection provides a general overview of some of these common assessment methods and indices. Additional information can be found in the

"Load" in Water Monitoring

Load refers to the total amount of a parameter delivered to a point per unit of time (e.g., pounds per year) such as the amount of phosphorus delivered by a stream to a lake each year. Loads are important to consider for waterbodies such as lakes and wetlands that are sensitive to longer-term inflow or recycling of pollutants. In monitoring programs, sampling occurs at a mix of high and low flow events, but the average (arithmetic mean) concentration will not represent the relationship between flow and concentration. Consider the following example:

Event	Total phosphorus µg/L	Flow cfs
1	33.0	2
2	29.0	7
3	45.0	16
4	35.0	4
5	55.0	25

The average for these samples is 39.4 µg/L, and the median is 35.0 µg/L total phosphorus. The flow-weighted average as shown below using a very simple approach is:

Total phosphorus µg/L x Flow cfs
33.0 x 2 = 66
29.0 x 7 = 203
45.0 x 16 = 720
35.0 x 4 = 140
55.0 x 25 = 1375
Add the products (66 + 203 + 720 + 140 + 1375 = 2504)
Add the flow values (2 + 7 + 16 + 4 + 25 = 54)
Divide the sum of the products by the sum of the flow values to get the flow–weighted average (2504 ÷ 54 = 46.4 µg/L). In this case it is much higher than the arithmetic average and the median.

When to Consider Flow-Weighted Averages
Flow-weighted averages are important to consider when doing loading studies such as determining the magnitude of pollutant loads discharged by a lake's tributaries. Calculation of flow-weighted averages are generally more complex than presented in the example because you also need to consider flow occurring between the sampled events and how concentrations can be represented for this unmonitored flow.

In Vermont, you will also need to consider the influence of snowmelt runoff in tributaries, streams and rivers. Snowmelt runoff can be significantly different with respect to pollutant concentrations than other runoff events because pollutants that accumulate over the winter are mobilized with the snowmelt. If snowmelt concentrations are high compared to other events, we suggest calculating your statistics with and without the snowmelt values to test the sensitivity of the result. Median may also be a better measure of central tendency than average when considering snowmelt. Other parameters may vary over other continuous periods, such as ice-free periods when the waterbody stratifies. In any case, you must be sure that you are comparing datasets that are for the same period, seasonal or otherwise.

Converting Data to Information

Table 7-2: Suggested Statistical Summaries for General Chemical and Physical Parameters

Table 7-2 is adapted from *Data to Information: A Guide Book for Coastal Volunteer Water Quality Monitoring Groups in New Hampshire and Maine,* by Dates and Schloss, (University of Maine Cooperative Extension and University of New Hampshire/Maine Sea Grant Extension, 1998).

Parameter	Statistical Summary	Parameter	Statistical Summary
Total suspended solids	Average Median Flow-weighted average[1] Range Quartiles Confidence intervals or standard deviation	pH	Median or average[3] Quartiles Minimum
Temperature (water or air)	Seasonal average Seasonal median Maximum Range Quartiles	Alkalinity	Median Quartiles Minimum
		Chlorophyll-*a*	Seasonal average[2] Range Maximum and minimum Median Quartiles Confidence intervals or standard deviation
Dissolved oxygen	Seasonal median Minimum Quartiles		
Turbidity	Median Maximum Quartiles	Flow	Average Maximum and minimum Median Quartiles
Nutrients (e.g. nitrite plus nitrate or total phosphorus)	Seasonal average[2] Flow-weighted average[1] Median Quartiles Confidence intervals or standard deviation	Water clarity/transparency	Seasonal average[2] Seasonal median Maximum and minimum Range Quartiles Confidence intervals or standard deviation
Conductivity	Average Median Quartiles	Bacteria (*E. coli*)	Geometric mean Quartiles Maximum

[1]Flow-weighted averages are used for stream or river monitoring to represent concentrations weighted by flow. Flow-weighted averages account for concentration-flow relationships.

[2]For lakes typically presented as growing season (loosely defined as mid-May through mid-September in Vermont) average.

[3]The average is acceptable in well-buffered systems where fluctuations are not extreme. It is also acceptable if you measure pH to the nearest 0.1 unit. If you measure to the nearest 1.0 unit, then use the median.

many manuals cited throughout this Guide. Specific assessment methods, benchmarks and indices described in this Section include:

♦ Determining the mixing status of a lake.

♦ Determining the trophic state of a lake.

♦ Comparing to water quality standards.

♦ Using biometrics for assessing streams and rivers.

♦ Using habitat indices for streams and rivers.

Determining the mixing status of a lake

Mixing status refers to the frequency of vertical (i.e., top to bottom) mixing of water in lakes. Mixing can be characterized as:

♦ Dimictic: mixes spring and fall.

♦ Intermictic: mixes intermittently during the summer with periods of thermal stratification.

♦ Meromictic: does not mix, always stratified.

♦ Polymictic: mixes from top to bottom throughout the summer.

♦ Monomictic: mixes once a year.

These characteristics can significantly influence the conditions of a lake. For example, in some lakes that stratify where sediments release significant amounts of phosphorus (internal phosphorus loading), concentrations of phosphorus in bottom waters can become very high. In dimictic lakes, where mixing only occurs in the spring and fall, these bottom phosphorus-rich waters are not brought to the surface during summer months. However, in intermictic lakes, this mixing of bottom water can be a significant source of phosphorus. Most lakes in Vermont are dimictic or polymictic.

Vertical mixing is controlled by the presence or absence of thermal stratification. Thermal stratification occurs when layers of water with different temperatures form a thermal density gradient that resists the energy of wind and makes it more difficult for waters to mix.

To assess mixing and stratification, temperature measurements are taken by lowering a probe to

specified depths (typically every meter from the surface of the lake to the bottom) and recording the temperature at each depth. These measurements are frequently complemented with dissolved oxygen data to characterize oxygen gradients from the surface to the bottom. Conducting this sampling regularly (monthly or weekly) from spring through fall should allow for a characterization of the mixing status of the lake.

Analysis of mixing is best done visually with graphs. You can complete the analysis by graphing each sample date separately as shown in Figure 7-1. With a series of these graphs covering the monitoring season, you can determine when the lake was well-mixed vertically, as in Plot A, versus Plot B, where temperature drops significantly in five meters, indicating the lake is thermally stratified.

Figure 7-1: Temperature Plots

Plot A: Temperature Profile May 23, 2000
Well-mixed vertically
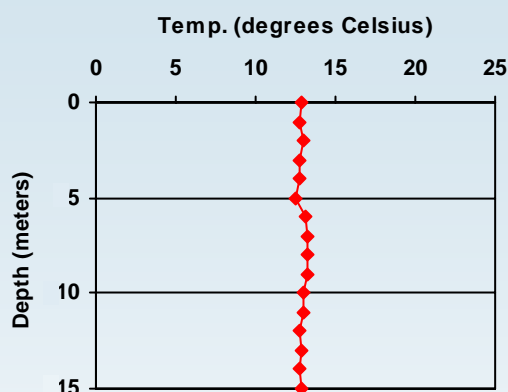
Plot B: Temperature Profile July 16, 2000
Stratified

Converting Data to Information
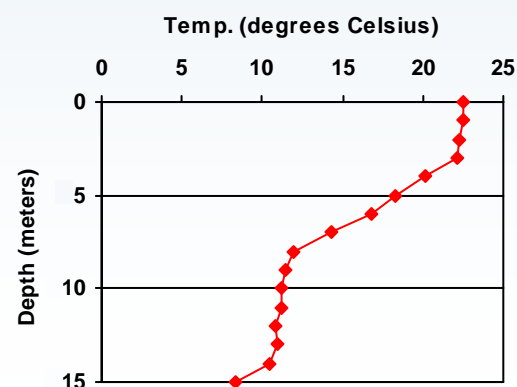
Assessing a Lake's Trophic State

Most lakes naturally contain aquatic plants and algae. The amount of plant and algae life a lake can support is referred to as the lake's "productivity." Plants and algae require nutrients for growth; the more nutrients in a lake, the more plants and algae it supports. Nutrient concentrations increase over thousands of years through a natural aging process called eutrophication. Eutrophication is a continuous process, which can be divided into three broad phases, or "trophic states"- oligotrophic, mesotrophic and eutrophic.
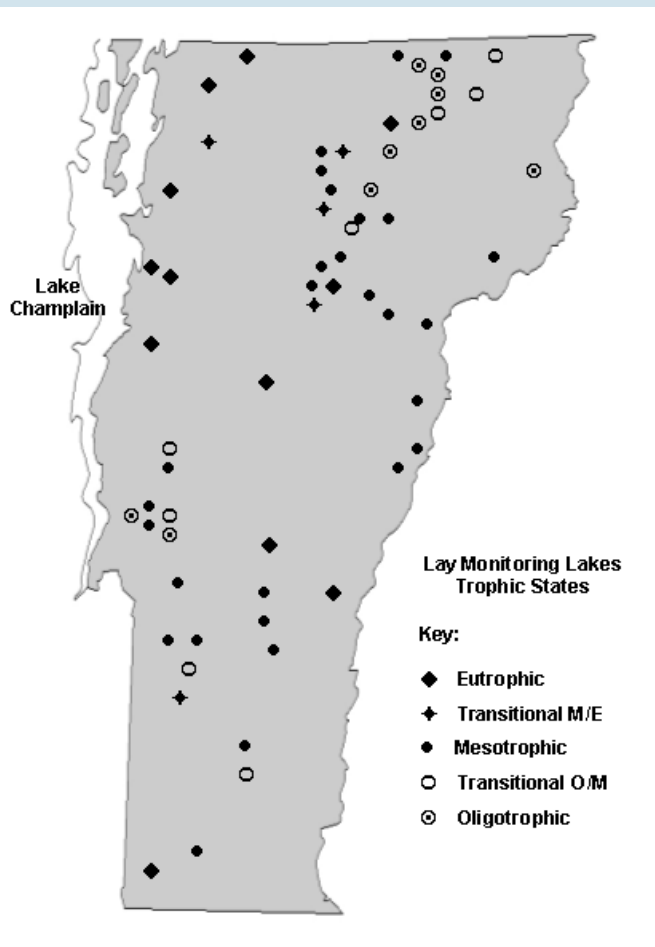
Although eutrophication is a natural process, human land use activities can greatly accelerate the process by contributing excessive nutrients to waterbodies through failing septic systems, shoreline erosion, fertilizer and roadway runoff, farming and logging practices and many other point and nonpoint sources. Establishing a lake's trophic state through monitoring helps document if and when land use practices impact the lake's water quality. Monitoring programs also work to identify potential pollution sources to help protect the lake's water quality. The map on the right shows the trophic states of Vermont lakes monitored under the Lay Monitoring Program.

Trophic States

Oligotrophic- Referred to as "young" lakes, characterized by deep, clear water; low nutrient enrichment; little algae growth (low productivity); few aquatic plants; bare sand or rock along most of the shoreline (little mud); and often supporting cold-water fish species.

Mesotrophic- Referred to as "Intermediate" lakes, characterized by moderate nutrient enrichment; moderate algae growth; moderate aquatic plant growth; moderate sediment accumulation over the lake bottom; and usually supporting warm water fish species.

Eutrophic- Referred to as "old" lakes, characterized by high nutrient enrichment; abundant algae growth (high productivity); extensive aquatic plant beds; extensive sediment accumulation on the lake bottom; and supporting exclusively warmwater fish species.

Lake Champlain

Lay Monitoring Lakes Trophic States

Key:

- ◆ Eutrophic
- ✦ Transitional M/E
- ● Mesotrophic
- ○ Transitional O/M
- ⊙ Oligotrophic

Index Ranges for Parameters to Assess Trophic State as Determined by the Vermont Lay Monitoring Program*

Trophic State	Secchi disk water clarity (meters)	Chlorophyll-*a* concentration (µg/L)	Total phosphorus (µg/L)
Oligotrophic (sparsely enriched)	> 5.5	0-3.5	0-7.0
Mesotrophic (moderately enriched)	3.0-5.5	3.5-7.0	7.0-14
Eutrophic (very enriched)	0-3.0	> 7.0	> 14

*Ranges in this table are those established by the Vermont Lay Monitoring Program and used to assess trophic state. Index ranges vary from state to state.

Thanks to Lay Monitors, like the Maidstone Lake crew to the left, enough years of data have been provided to document the trophic state of many Vermont lakes. Maidstone Lake is an oligotrophic lake located in Maidstone (north eastern Vermont).

Determining the trophic state of a lake

Total phosphorus, Secchi transparency and chlorophyll-*a* measurements are parameters used to characterize the "trophic state" of a lake. To figure out a lake's trophic state, calculate the season's average for each of these parameters (with at least eight weeks of data and at least one data point from each month) and compare them to the index ranges determined by the Water Quality Division (see side box, left). An assessment of a lake's trophic state can be made with one season of data, however, baseline conditions are considered substantiated after several years of data are accumulated.

There are three main trophic categories: oligotrophic, mesotrophic and eutrophic. For lakes sampled by the Vermont LMP, oligotrophic lakes mostly occur in the Northeast Kingdom, mesotrophic lakes are generally found along the Green Mountains, and eutrophic lakes are primarily in the Champlain Valley region, an area with fertile soils and over 200 years of agricultural history.

Comparing to water quality standards

You may want to compare your data to Vermont's Water Quality Standards, the fundamental benchmarks by which the quality of surface waters are measured. Water quality standards are used to determine impairment and assess whether a waterbody is meeting its designated uses (see Table 7-3: Designated Uses for Water Classifications). However, keep in mind that failure to meet standards does not automatically mean there is an impairment that will immediately place the waterbody on the state's List of Impaired Waters (303 (d) list).

Assessing compliance with water quality standards and impairment is very specific since there is a regulatory component. This subsection provides a brief overview of Vermont's Water Quality Standards. The complete document of *Vermont Water Quality Standards* is available online from the Water Resources Panel of the Natural Resources Board at **www.nrb.state.vt.us/wrp/rules.htm**. The methods VTDEC uses to assess compliance with the Vermont Water Quality Standards is described in Vermont's Water Quality Assessment and Listing Methodology.

Classes of Vermont surface waters

All surface waters in Vermont are classified as either Class A or Class B and eventually will be designated as a water management type.

Class A waters

Class A waters are managed to maintain the highest quality standards. Within Class A, waters can be designated as Management Type 1 or Type 2. Class A(1) waters are ecologically significant waters, providing significant wildlife and aquatic habitat and managed to have biota "as naturally occurs." Class A(1) waters are managed to maintain an essentially natural condition (e.g., Sterling Pond on Sterling Mountain).

Designated Uses	Class A(1) Ecological Waters	Class A(2) Public Water Supplies	Class B Waters
Aquatic biota, wildlife, & aquatic habitat	√	√	√
Aesthetics	√	√	√
Swimming & other primary contact recreation	√		√
Boating, fishing, & other recreational uses	√		√
Public water supplies		√	√
Irrigation of crops & other agricultural uses		√	√

Table 7-3: Designated Uses for Water Classification

Vermont Water Quality Standards Established for Designated Uses

Numeric water quality standards may be different for waters in Vermont with different designated uses. For example, the numeric *E. coli* standard for:

Class A(1) Ecological waters- Not to exceed a geometric mean based on at least three samples obtained over a 30-day period of 18 organisms per 100mL, no single sample above 33 organisms per 100mL. None attributed to the discharge of wastes.

Class A(2) Public drinking water supplies- Not to exceed a geometric mean based on at least three samples over a 30-day period of 18 organisms per 100mL, no single sample above 33 organisms per 100mL (before filtration). None attributed to the discharge of wastes.

Class B All other waters- Not to exceed 77 organisms per 100mL.

Sterling Pond is one of the highest lakes in Vermont, located at about 3,600 feet on the slopes of Sterling Mountain in Cambridge. It is a natural pond and a Class A(1) waterbody.

It should be noted that Vermont's *E. coli* standards are stricter than the national standards established by the U.S. EPA. The EPA's standards are not to exceed a geometric mean based on at least 5 samples over a 30-day period of 126 organisms per 100mL for recreational waters (analogous to Vermont's Class B waters).

Waters designated as Class A(2) are public drinking water supplies. These waters are managed to maintain water quality suitable for public consumption (with appropriate filtration and disinfection), as well as for aquatic biota, wildlife, aquatic habitat and aesthetics (e.g., Springfield Reservoir). Only about 3% of surface waters in Vermont are presently Class A.

Class B waters

Most of Vermont's surface waters (97%) are Class B. These waters are managed to maintain a level of quality that supports swimming, fishing, boating, aquatic habitat and biota. During the VTDEC's basin planning process, all Class B waters will be recommended to be typed as Water Management Type 1, Type 2, or Type 3 based on public opinion and professional judgments, desired management objectives, attainable uses and present water quality conditions. Waters ultimately will be typed by Vermont's Natural Resources Board after additional public input.

The main difference between management types is the amount of change from a reference condition allowable to the aquatic biota, wildlife and aquatic habitat, as well as standards for aesthetic values. Management Type B(1) allows minor changes in aquatic biota and wildlife, minimal changes in aquatic habitat, and maintenance of consistently excellent aesthetic values.

Management Type B(2) allows moderate changes in aquatic biota and wildlife, minor changes in aquatic habitat, and maintenance of consistently very good aesthetic values. Management Type B(3) allows moderate changes in aquatic biota and wildlife, moderate changes in aquatic habitat and requires achievement of good aesthetic values (seasonal and temporal variability may be allowed).

The management goal for all Vermont surface waters is protection of the aquatic environment for sustainable, healthy, diverse and successfully reproducing populations of aquatic organisms and wildlife, including macroinvertebrates, fish, waterfowl and other organisms that depend on the waterbody for survival.

Numeric water quality standards

A numeric water quality standard is an acceptable concentration of a pollutant in water, associated with a designated use. Numeric standards are associated with each water classification. Specific standards for numerous parameters (pH, phosphorus, temperature, etc.) can be found in the aforementioned Vermont Water Quality Standards document.

Narrative water quality standards

A narrative water quality standard is a statement that defines the acceptable conditions in or on

the water, such as visible oil film, algae blooms, or exotic/invasive species. Narrative standards are sometimes called "free froms" because they keep surface waters free from fundamental forms of water pollution. More specifically, these standards also protect surface waters and aquatic biota from:

♦ Accelerated eutrophication (nutrient enrichment from point and nonpoint sources).

♦ Impairment of the biological community.

♦ Impairment of fish for human consumption.

The association between the water quality standards and designated uses is less well-defined for narrative standards than it is for numerical standards; however, most narrative standards are written to protect aesthetics or aquatic life. Since narrative standards are not quantitative, the determination that one has been exceeded typically requires a "weight of evidence" approach to data analysis showing a consistent pattern of impacts to uses.

Using biometrics for assessment

Biometrics are used to analyze and interpret biological data by grouping organisms into mean-

ingful biological assemblages. Biometrics represent various aspects of the biological community and typically are chosen to express meaningful biological characteristics, such as species diversity, trophic structure and tolerance or intolerance of various forms of human disturbance.

Biometrics for benthic macroinvertebrates have been used by the VTDEC's Biomonitoring and Aquatic Studies Section (BASS) for lake, river and stream assessments, and are currently being explored to incorporate into wetland assessments. The family level Hilsenhoff Biotic Index (HBI) is the most commonly used tool for assessments of the macroinvertebrate community by volunteers.

The HBI is a measure of the macroinvertebrate community's tolerance toward nutrient enrichment. While it is a useful tool, volunteer groups that use the HBI as a primary means of assessing stream health should be aware that this is only one way of looking at data, and that other metrics, which are intended to reveal other types of changes to stream ecosystems should also be considered.

The Intensive Stream Biosurvey Method 4.3 in the U.S. EPA manual *Volunteer Stream Monitoring: A Methods Manual* (EPA, Nov. 1997) recommends the use of four basic metrics described below. These metrics have been commonly used by monitoring agencies throughout the country and are considered robust measures of stream health. Using multiple metrics is recommended and will allow for more in-depth assessment.

♦ Number of taxa (taxa richness)- a count of the number of taxa (e.g., orders, families, species) found in the sample.

♦ Number of EPT taxa (EPT richness)- a count of the number of taxa in each of three generally pollution-sensitive orders: Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies).

♦ Percent dominance- the percentage of the entire sample that the individuals of the most abundant family make up. It indicates how dominant a single family is at a particular site.

♦ Sensitive taxa index (modified Hilsenhoff Biotic Index)- calculated by multiplying the

---

Calculating the Family Level HBI (Hilsenhoff Biotic Index)

The HBI measures the macroinvertebrate community's tolerance toward pollution and nutrient enrichment. This measure is calculated by multiplying the number of organisms in each family by the family's pollution tolerance value and then adding these for all families represented in the sample and dividing by the total number of individuals in the sample. The tolerance value is a number from 0 (intolerant) to 10 (tolerant) that is assigned by the VTDEC that correlates with the family's tolerance of pollution. The calculation looks like this:

$$HBI = \frac{\sum (X_i\, t)}{n}$$

Where:

X_i = the number of individuals in a family

t = the tolerance value for that family

n = the total number of individuals in the sample

$\sum$ = the summation of $X_i\, t$ for each family in the sample

A low HBI value for a stream or river indicates better water quality.

---

number of organisms in each taxon by the pollution tolerance value assigned to each taxon, adding these for all taxa represented in the sample and dividing by the total number of taxa in the sample.

Additional metrics used by VTDEC's BASS include:

⬥ EPT/EPT + Chironomidae (indicator of taxonomic structure and tolerance/intolerance)- a measure of the ratio of the abundance of the intolerant EPT orders to the generally tolerant Diptera family Chironomidae.

⬥ Percent Oligochaeta (indicator of tolerance/intolerance)- a measure of the percent of the entire sample that is composed of individuals in the order Oligochaeta.

⬥ Percent Model Affinity of Orders (PMA-O)- a measure of order level similarity to a model based on reference streams.

⬥ Pinkham-Pearson Coefficient of Similarity–Functional Groups (PPCS-F)- is a measure of functional feeding group similarity to a model based on reference streams.

Information on calculating metrics for macroinvertebrate data is available from:

⬥ VTDEC BASS at (802) 241-3777 or online at **www.vtwaterquality.org/bass/htm/bs_macro.htm**.

⬥ River Network at (802) 223-3840 or online at **www.rivernetwork.org**.

Using habitat indices for streams and rivers

Completion of some form of habitat assessment is recommended to complement stream biosurveys. A quantitative method for habitat assessment is included in the U.S. EPA manual *Volunteer Stream Monitoring: A Methods Manual*, (EPA, Nov. 1997). This quantitative evaluation method is available for rocky and muddy bottom sampling sites. It consists of a scoring system from 0 (poor) to 20 (optimal) for the following habitat parameters:

Rocky Bottom	Muddy Bottom
1. Attachment sites for macroinvertebrates	Shelter for fish and macroinvertebrates
2. Embeddedness	Poor substrate characterization
3. Shelter for fish	Pool variability
4. Channel alteration	Channel alteration
5. Sediment deposition	Sediment deposition
6. Stream velocity and depth combination	Stream sinuosity
7. Channel flow status	Channel flow status
8. Bank vegetative protection	Bank vegetative protection
9. Condition of banks	Condition of banks
10. Riparian zone width	Riparian zone width

Total scores are summed to get the quantitative assessment. The total value and the individual parameter values can be compared to biosurvey results and biometrics. This will help identify causes of impairments shown by the biometrics.

For example, if the percent dominance metric shows a very high value indicating dominance by one or two taxa, but the quantitative habitat evaluation shows optimal conditions for all parameters, a likely conclusion is that water quality, rather than habitat, may be stressing the aquatic community. Future studies should perhaps focus on water quality parameters.

Using tables and graphs

This subsection explains how your results can be displayed in tables and graphs to help visualize and interpret them. In reports, only include the graphs that help tell your story. Otherwise, raw data tables can be included in appendices.

Tables

Sometimes a table is not considered "exciting," but it is an important tool for organizing data and can present information more precisely than graphs. Use tables sparingly in presentations because they are difficult for the audience to read unless they are very simple.

Graphs and Charts

Pie charts, bar graphs and line graphs (including scatter plots) are the three main types of graphs you will use. You can create these types of graphs with most spreadsheet programs.

Pie charts (and stacked column charts) are different ways to display data and show data as proportions of a whole. They are easy for the general public to understand, but can only be used for data that can be expressed in terms of proportions, or percentages of a whole. For example, they can show the percent each source contributes to the phosphorus load to a lake (Figure 7-2) or the percent composition of taxonomic groups (Figure 7-3).

Bar graphs put more emphasis on the individual points or summary statistics. They are useful for comparing biosurvey results, the level of a pollutant at one station over time or at several stations at one time and for displaying summarized data. Figure 7-4 shows a bar graph generated by the Huntington Conservation Commission to display the geometric mean for each site in their *E. coli* monitoring program.

Line graphs are good for displaying relationships between points. A line graph displays the data points as points on the graph connected by a line. They often illustrate trends in data; time or space is usually displayed along the x-axis (horizontal) and water quality parameters along the y-axis (vertical). Figure 7-5 shows a line graph used by the Upper Otter Creek Watershed Council to display their *E. coli* monitoring results. The graph compares their results to VTDEC and EPA standards.

Converting Data to Information

Figure 7-2: Example Pie Chart
Sources of phosphorus contributions to Round Lake, 1996

Watershed 64%
Septic 11%
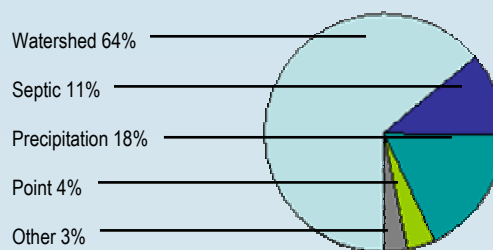Precipitation 18%
Point 4%
Other 3%

Figure 7-3: Example Stacked Column Chart
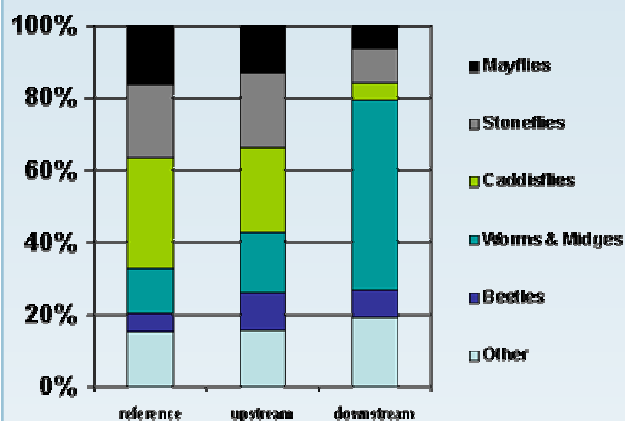Composition of selected macroinvertebrate groups

Figure 7-4: Example Bar Graph
Geometric mean of *E. coli* levels at sampling sites on the Huntington River (data from 2003/2004 sampling seasons)
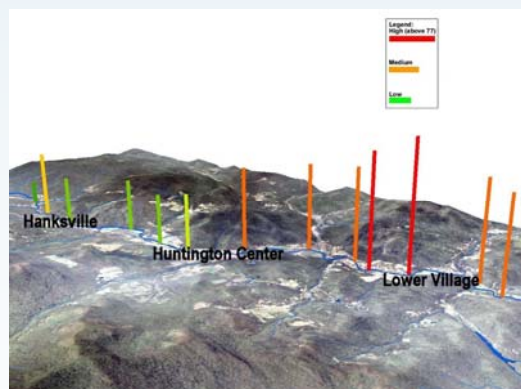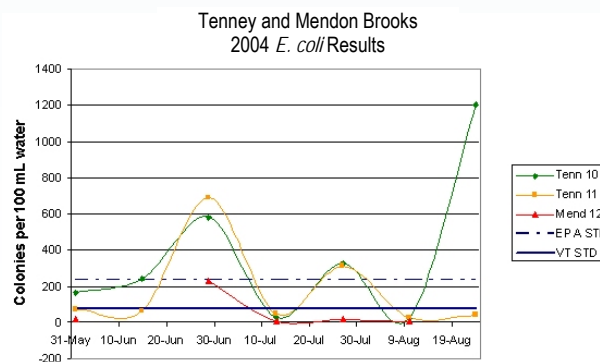

Figure 7-5: Example Line Graph
E. coli colonies per 100 mL

When using a line graph, you must be careful that you have enough datapoints so that the trend implied is valid. This may or may not be the case depending on the variability of the data. For example, if graphing dissolved oxygen concentrations against location (mileage) along a river, it may be appropriate to connect a line through several points that are only a short distance apart and taken at about the same time of day. It would not be appropriate for sites miles apart or where readings were taken at different times of the day.

Reporting your information

If you have spent the time to collect data, you will probably want to share your experience and the data you have collected with others. At the very least, produce a written report that summarizes your work and the results for your most rigorous audience. Once you have this report prepared, you can prepare presentations for different audiences. A presentation you make to the town selectboard, for example, may be very different from a report you make to your staff.

Making an annual report

In your report, summarize your monitoring activities and results, state your findings and conclusions and make recommendations for actions to address problems or changes to your sampling program, if needed. You may produce an annual "state of the waterbody" report that highlights trends, cleanup progress, new trouble spots, etc.

Here is a generic format to follow:

1. Introduction- describe the area and your specific program, your "why" question and your monitoring purpose, include maps of your monitoring location(s).

2. Project description- summarize your design, parameters monitored, sampling methods.

3. Results- describe how data were analyzed, findings, conclusions, recommendations, include charts and graphs.

4. Acknowledgements- give credit to volunteers, professional contacts, anyone who helped in the planning and implementation of your program.

5. References- list information sources you used to prepare your report.

6. Appendices– present any other information you wish to include but that would detract from your narrative report.

Once you have your basic report prepared, share your experience and data with others by:

♦ Participating in the distribution of information to and with other agencies.

♦ Writing and distributing technical reports describing what you learned- current water quality conditions; suspected or identified pollution sources; effects of contaminants on humans and ecosystems.

♦ Communicating with multiple audiences by writing reports or executive summaries for nontechnical audiences.

♦ Writing articles for local weekly newspapers and magazines.

♦ Presenting lessons to peers, school classes, after-school clubs or other organizations and volunteer groups.

♦ Creating a display or booth.

♦ Making presentations to your natural resource conservation district, town selectboard, or Regional Planning Commission to assist the public in understanding the significance of your results.

♦ Providing basic data for other data users.

Now that you have finished reading *Section 7*, return to the Worksheet on pages 5-8 to answer the corresponding questions.