

Section 5:

Data management

This Section will show you how to:

- Organize and keep track of the data you collect

Managing data

By now you know that volunteer water monitoring requires attention to detail and precision at every step of the way. If you have been following the Sections in this guide, you have determined the reason you are monitoring (your purpose). You have set your QA/QC objectives and you have designed a plan. You are now ready to start collecting data and recording it in a format that is eventually usable to yourself and others. What, exactly does this entail?



To get the most out of this part of your monitoring program, it is critical that you set up a data management program before you actually begin monitoring. By setting it up in advance:

- You eliminate/minimize errors in recording and transferring data.
- You don't lose data.
- You can go back to the original data sheets if there are problems or questions.
- You can easily access and use data once you've stored it.
- You can format it in a way that will be useful and acceptable to others.

Decide your data management needs

The extent of data management you will need depends upon the purpose for which you are gathering data. If you are collecting information that may be used for enforcement, for example, you will need a more rigorous program than if you are collecting data to determine what organisms are in the wetland in your back yard.



The Golden Rule of data transfer

Keep the number of times data is transcribed to a minimum. The more times you transcribe data (from one sheet to another), the more chance of errors.

Based on your own specific purpose, use the following guidelines to set up your own program: You do not have to follow each one of the guidelines below, nor do you have to follow them sequentially.

- Develop or use already-developed data collection sheets and checklists. This will ensure uniform collection and recording of results in both the field and the laboratory. If you are working with an organization, you will probably use sheets and checklists developed by the organization's program.
- Continually review the data sheets and checklists as you collect data to ensure the information is complete. To do this, you may have a signature line for a sampling team captain or third party to indicate the data sheet was checked or approved. If there are problems, this reviewer will contact any sampler whose field sheets contain significant errors or omissions.
- If using a laboratory, use a Chain of Custody form (or transmittal letter) to document the transmittal of samples. You can get these from most laboratories.
- If using a laboratory, the laboratory manager should review the QA/QC parameters used and include the results with the laboratory report.
- Review field and laboratory QA/QC results and determine if data quality objectives (set in *Section 3*) have been met. Make a decision whether to keep the data or not. Many times, even though the data does not meet QA/QC objectives for a particular purpose, it will meet objectives for another purpose and may still be usable. In such cases, data may be "flagged" to indicate how it did not meet its original QA/QC objectives.
- Enter data that meets the data quality objectives into a spreadsheet or database. If you plan to send data to a central database, such as the MPCA's Water Quality Database, check in with the database

managers to find out how to organize the data for submittal. Then set up your data management system (spreadsheet or database) in a way that is compatible with the database requirements. If you are using a contract laboratory, you can require the lab to provide results in an electronic format that is compatible with database requirements.

- Have a second individual review the entered data.
- Program the database to screen data for errors or review the data manually by checking to see that results are within an acceptable range. For example, pH can only range from 1 to 14 standard units (s.u.). A pH of 16 s.u. is not possible.

Use field data sheets and laboratory reports

Collecting raw data, both field and laboratory, is discussed in detail in *Section 4: Design Your Monitoring Effort*. A few existing data collection sheets are included in *Appendix G*.



Develop a database or spreadsheet

Decide in advance how data sheets will be handled, and how and where they will be stored and then archived.

We suggest you use a computer to store and access data in either a database or spreadsheet. Unless you have a lot of data, spreadsheets will generally be easier to use.

When setting up a spreadsheet, first check the format of the database you may use in the future and pattern your spreadsheet after it by creating similar fields.

When you're entering data, check it, and check it again. Then have an outside individual review it yet again. Make certain that you have a record of individuals who check the results and also record the dates they were reviewed.

Be wary of releasing an electronic form of data to users before the numbers are checked and rechecked or before all data is entered.



Watch out for these data entry errors

- Entering data in the wrong units (entering concentrations as micrograms/liter instead of milligrams per liter)
- Reversing numbers
- Misplacing decimal points
- Entering in the wrong row or column

To prevent multiple versions of a database from being circulated:

1. Wait to release the results until all data has been entered and checked (i.e., resist the temptation to release draft databases).
2. Include a field for dates and initials of the last update or approval so that you can easily tell if you are working with the most current version. Also, to avoid losing your data, make a backup copy and store it at another location.

Make data available to others

Entering data from coordinated programs

If you participate in an organized volunteer monitoring program, the project sponsor may enter your data into one of its databases. Following are some possibilities:

- Data from the Citizens Lake Monitoring Program (CLMP), the Citizens Stream Monitoring Program (CSMP) and the Citizen- Assisted Lake Monitoring Program (CAMP) are all entered into the MPCA Water Quality Database.
- CAMP data is also entered into the Metropolitan Council's Environmental Information Management System (EIMS).
- The Volunteer Stream Monitoring Partnership (VSMP) has an agreement to house its data on the Metropolitan Council's EIMS in the future. (<http://www.vsmf.org>).
- The Department of Natural Resources (DNR) and the State Climatologist maintain a database for volunteer precipitation monitoring collected by volunteers in the Climatological Network (<http://www.climate.umn.edu/doc/historical.htm>).
- The DNR also houses a database for volunteer lake gauge readings, which is online at <http://www.dnr.state.mn.us/lakefind/results.html>.

Entering data independently collected

If you have developed your own local program, you may want to deposit your data into one of the state or regional databases. Submitting data to a state or regional database increases the chances that federal, state and local agencies and organizations are aware of, have access to and use the data you've collected. To do so, follow the protocols established by the database manager(s). Some databases to consider to house your data are:

- MPCA Water Quality Database (formerly known as STORET). Administered by the USEPA and coordinated by MPCA, this database is linked to a web-based access system that is online, effective July 2003 (www.pca.state.mn.us/data/eda/index.html).
- Met Council Environmental Information Management System (EIMS) is designed to house data from the Twin Cities metropolitan area. Web-based access is under development. For more information, call 651-602-1056.
- The Minnesota River Basin Data Center (MRBDC) at Mankato State University houses a lot of data, including water quality information on the Minnesota River and its tributaries (<http://mrbdc.mankato.msus.edu>). Contact is MRBDC, Mankato State University, 184 Trafton Science Center South, Mankato, MN 56001; 507-389-5492; mrbdc@mnsu.edu.
- Local governments may also house water quality databases. Contact Soil and Water Conservation Districts, Watershed Management Organizations and County Water Planners for more information. Online directories for these organizations are provided in *Section 2*.

The following tables identify the information that must be submitted for data to be loaded into the Water Quality Database. Keep in mind that it is important to contact the MPCA as you are setting up your data management system (i.e., before you begin entering data into your spreadsheet or database system) to ensure it is compatible with the database. This will minimize steps needed to load your data into the database.



Send complete information

Note that it is generally not sufficient to simply send in the monitoring data for inclusion in a database. You must also send information about the data, such as the monitoring location, field sampling procedures, equipment used, analytical laboratory, lab methods, etc. This information is known as meta-data.

Appendix F includes meta-data descriptions and requirements for the MPCA Data Quality Database.

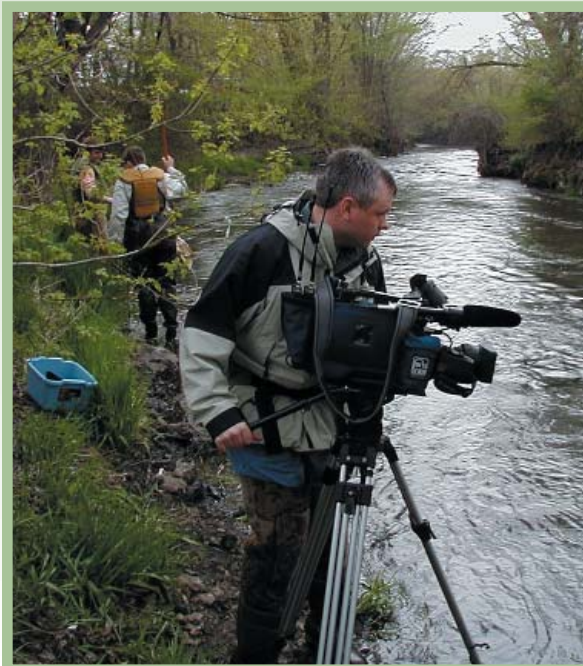


Exhibit 5-1 Information that must be submitted for MPCA Water Quality Database

Meta-data element	Required?	Notes
Project Information		
Name Purpose Start date & duration Lead organization Contact information Laboratory info. (if one is used) Sampling methods & equipment Sample medium Sampling station information	Yes, for all monitoring efforts	Only needs to be supplied once, when the project data is first submitted for inclusion in the database
QAPP summary or citation (i.e. where to find project QAPP)	For data submitted for 305(b) or 303(d) use (in addition to info. required for all projects)	
Laboratory Information		
Name Contact information Analyses and methods Parameter name and reporting units Lab certified for parameter? Comparable Standard Method	Yes, for all monitoring efforts where a lab is used	Only needs to be supplied once, when the project data is first submitted for inclusion in the database (or if the lab changes)
Detection limit	For data submitted for 305(b) or 303(d) use (in addition to info. required for all projects)	
Monitoring Station Information		
Name Type Description Ecoregion (optional) Travel directions Latitude-longitude or UTM Method and reference (datum) for determining lat.-long. or UTM HUC code; RF1 reach (optional)	Yes, for all monitoring efforts (except as noted)	Only needs to be supplied once, the first time data is provided for a particular station
Data (sampling results) Information		
Project ID Station ID Date and time Lab ID (as applicable) Depth Methods QA sample type (as applicable) Measurement (i.e., result) and units Project personnel Remarks (as needed)	Yes, for all monitoring efforts	Required every time data is submitted to the database
Lab sample temp. (as applicable) Time of sample	For data submitted for 305(b) or 303(d) use (in addition to info. required or all projects)	

