

DATA MANAGEMENT

Supplement to the Clean Water Act Section 106 Tribal Guidance

Track 2 CAFO: Chemical Assessment FY 2006

Site 6: 11-T26N-R7E

Type: Cattle

Creek: North Omaha

April 2006 thru September 2006

Parameter	Units	Reference Value	27-Apr	6-Jun	28-Jun	7-Aug	21-Aug	13-Sep
Temperature	° C	N/A	11.00	19.50	18.75	19.25	18.75	13.00
pH	None	6.5-8.5	8.33	8.27	8.49	8.61	8.48	8.14
Spec Cond	mS/cm	N/A	0.649	0.692	0.707	0.730	0.762	0.762
TDS	mg/L	500	422	450	460	474	495	495
Sodium	mg/L	100	18.5	18.5	16.7	16.2	18.7	21.4
Calcium	mg/L	80	92.5	103	95.0	92.2	105	103
Magnesium	mg/L	30	29.4	29.8	27.1	26.9	30.4	30.6
Nitrate	mg/L	10	4.2	5.4	5.9	5.2	4.9	4.8
Sulfate	mg/L	250	36	38	35	33	39	50
Tot Hard	gr/gal	20	20.6	22.2	20.4	20.0	22.7	22.4
Iron	mg/L	0.3	0.75	1.17	3.21	4.27	2.62	2.88
Manganese	mg/L	0.05	0.43	0.53	0.39	0.38	0.28	0.30
Chloride	mg/L	250	7	8	7	7	8	10
Fluoride	mg/L	2	0.3	0.2	0.4	0.3	0.2	0.4
Tot Phos	mg/L	N/A	0.1	0.3	0.3	0.4	0.4	0.4
TKN	mg/L	N/A	1.41	1.65	2.24	1.65	1.34	2.09
Ammonia	mg/L	N/A	<1.10	<0.10	<0.10	<0.10	<0.10	<0.10
Tot Coliform	cfu/100ml	200	2910	>200	19863	48392	24196	24196
Fec Coliform	cfu/100ml	200	290	680	1640	2220	1390	10660
E. coli	cfu/100ml	N/A	248	1722	1291	1352	1060	19863

Note: None - No results available.

N/A - Non applicable

< - Less than

> - Greater than

Table of Contents

INTRODUCTION	1
WHAT IS THE PURPOSE OF THIS TRIBAL DATA MANAGEMENT DOCUMENT?	1
DATA MANAGEMENT	3
I. WHAT IS DATA MANAGEMENT?	3
II. DATA MANAGEMENT PROCEDURES	4
III. DATA TYPES	6
IV. DATA MANAGEMENT FORMATS	9
STORET-COMPATIBLE DATA SUBMISSIONS	12
I. GETTING YOUR DATA INTO A STORET-COMPATIBLE FORMAT	12
II. WAYS TO SUBMIT DATA	16
III. RETRIEVING DATA FROM THE STORET WAREHOUSE.....	19
IV. ONGOING SUPPORT	20
ATTACHMENT A: EXAMPLES OF TRIBAL DATA MANAGEMENT	21
ATTACHMENT B: GLOSSARY	26

Data Management

Supplement to the Clean Water Act Section 106 Tribal Guidance

Introduction

What is the purpose of this tribal data management document?

In October 2006, the United States Environmental Protection Agency (EPA) issued the *Final Guidance on Awards of Grants to Indian Tribes under Section 106 of the Clean Water Act* (Tribal 106 Guidance). The Tribal 106 Guidance provides a flexible framework for tribes and EPA regions to ensure that tribes' water quality program goals are met while being consistent with Section 106 of the Clean Water Act (CWA), the Governmental Performance and Results Act (GPRA), and other federal requirements. The Tribal 106 Guidance outlines new water quality reporting requirements and data expectations for all tribal programs receiving 106 funds. These requirements will help tribes collect and manage critical data and information for effective management of water quality programs, to measure environmental results, and report them to EPA.

As discussed in Chapter 8 and Appendix A of the Tribal 106 Guidance, EPA requires that tribes submit to EPA a tribal Assessment Report that contains information about water quality on tribal land and demonstrates national results for the Section 106 Tribal program. The Assessment Report consists of: (1) a description of the tribe's monitoring strategy; (2) a water quality assessment; and (3) surface water quality data submitted electronically in a STORET-compatible format that includes metadata for at least the nine basic parameters that are being monitored (dissolved oxygen, pH, water temperature, turbidity, total phosphorus, total nitrogen, macroinvertebrates, *E. coli* or enterococci, basic habitat information), as required considering the Tribal 106 Guidance and a tribal water quality program's level of sophistication.

This document is intended to supplement the Tribal 106 Guidance by providing useful suggestions and tips to tribes about how to establish a data management system that reflects tribal water quality goals and objectives, and manage the system so that the data is accessible and reliable. As a result of good data management, tribes are empowered to use the data to determine the physical, chemical, and biological conditions of a water body so water quality trends can be identified and measured to ensure that water quality can be maintained and/or improved. An organized and thorough data management program provides the foundation for measuring and achieving water quality environmental results while helping tribes meet the EPA data reporting requirements. Attachment A provides examples of existing tribal management programs. Attachment B provides a glossary of terms used throughout this document.

This document provides basic data management principles that will be helpful to all tribes and all water quality programs, whether their programs are at the fundamental, intermediate, or mature level as defined by the Tribal 106 Guidance.

Two additional templates have been created for assisting tribes in developing monitoring strategies and performing data analysis and assessment¹. These resources are available to assist tribal programs in meeting the tribal Assessment Report requirements. The data management principles presented in this document are meant to assist tribes with getting data into a usable format in order to begin putting together data assessments (statistics, data summaries, etc). This document is not intended as a tutorial for data submission to EPA, however information about data submission to EPA and support for submission is included in the document.

This document has been reviewed and approved for publication by EPA. Mention of trade names or commercial products or services does not convey, and should not be interpreted as conveying, official EPA approval, endorsement, or recommendation. Any use of words such as “should” or “may” are not intended to be mandatory language, rather, these are statements of suggestion that a tribal program may or may not adopt depending upon available resources.

¹ EPA Tribal Monitoring Strategy and Assessment Workgroup. 2008. *Developing a Tribal Water Monitoring Strategy: Supplement to the Clean Water Act Section 106 Tribal Guidance*.

EPA Tribal Monitoring Strategy and Assessment Workgroup. 2008. *Data Assessment and Reporting: Supplement to the Clean Water Act Section 106 Tribal Guidance*.

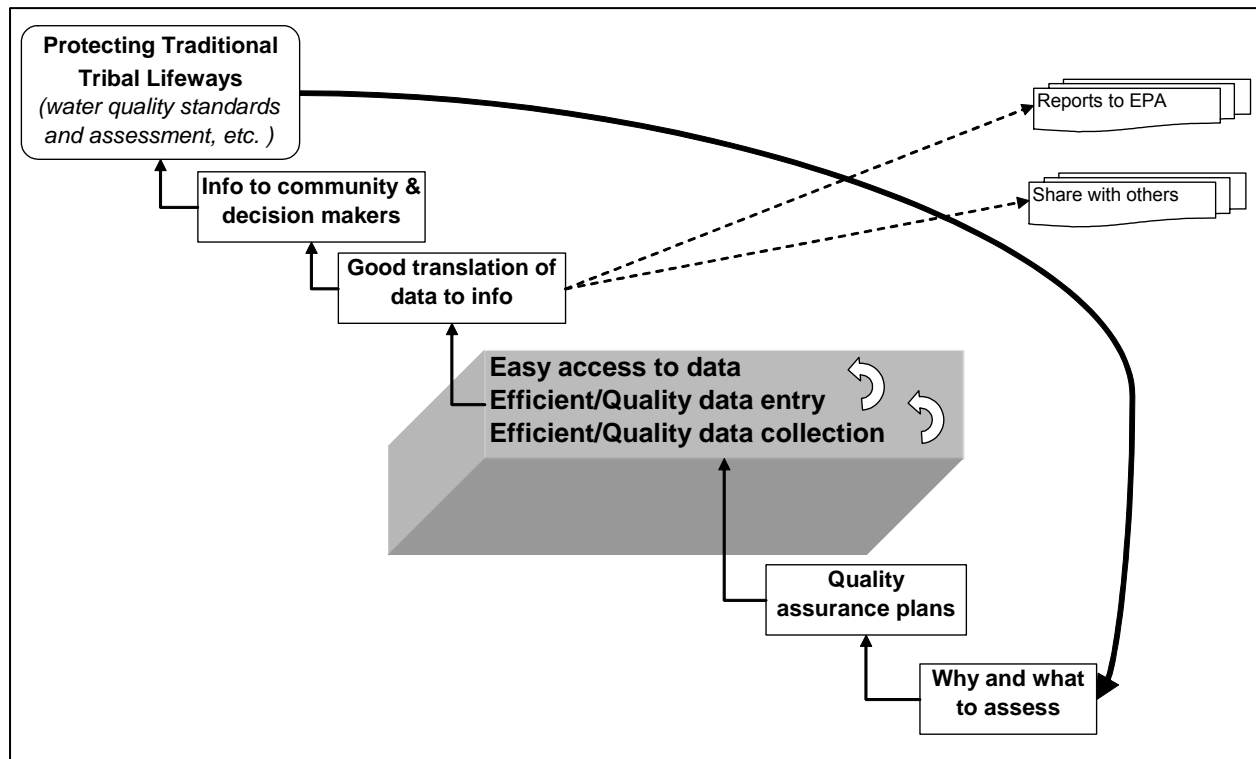
Data Management

I. What Is Data Management?

Data management is the process of reporting and storing the information you collect as part of your water quality monitoring program. It helps you store information in a format that allows you to manipulate, summarize, and analyze your results to make decisions. Managing data properly does take time, but it will save you time in the end and help you maintain and present accurate information about water quality in your area. Data management is a crucial part of a monitoring strategy² and can be a way to integrate your thinking about many aspects of your work.

Figure 1 shows that the ultimate goal of monitoring water quality is to protect traditional tribal lifeways. It also shows that there are many necessary components that are part of accomplishing that goal – of which good data management is one. Without the components in the highlighted box (data management) the steps could not be followed. In this process, tribal data can be shared with others but that does not need to dictate the process of management.

Figure 1. Components Contributing to Protection of Traditional Tribal Lifeways



² EPA Tribal Monitoring Strategy and Assessment Workgroup. 2008. *Developing a Tribal Water Monitoring Strategy: Supplement to the Clean Water Act Section 106 Tribal Guidance*. (Please refer to Appendix A of the Tribal 106 Guidance as well.)

Water quality monitoring data generally consist of the following information:

1. Where a sample or measurement was collected
2. When it was collected
3. What was collected
4. How it was collected
5. Why it was collected
6. Who collected it

Data management ensures that this information is accessible and understandable. The purpose of data management is to convert your monitoring results into a useful set of data to use for making decisions about your water quality monitoring program and meet tribal water quality goals.

II. Data Management Procedures

Data management procedures are processes for storing and sharing your water quality monitoring data in a controlled and standard way. Take a step back to consider the goals of your water quality monitoring program. Any decisions you can make in advance of managing your data can help you develop standard procedures and save time. For example, you may want to have a meeting with others in your group who collect water quality monitoring data, your lab, or your manager to discuss ways to report the data so that information can be documented in similar ways over time. This will avoid confusion when looking at past information, and prevent the need to review procedures to make them consistent. Coordinating with your laboratory will allow it to report data in formats that are helpful and consistent for you. These decisions and meetings will help develop effective data management to help you understand your data over time and will help any new people that are hired in your organization to understand the data.

To determine what data management procedures may be right for you, ask yourself the following questions:

1. What data are currently collected?
 - Some examples may be:
 - Field measurements for temperature, pH, and turbidity on field sheets
 - Continuous data from YSI or other data loggers from lakes and streams monitoring locations
 - Water chemistry sample results from a lab
 - Benthic macro invertebrates counts or fish tissue sample results
2. What data are being recorded?
 - Some examples may be:
 - Results for the core and supplemental water quality indicators (see page 4-10 of the Tribal 106 Guidance) being used to meet monitoring strategy goals and objectives
 - Field sheet result values
 - Metadata for your results, such as the methods you used in collected or analyzing a sample

- Locational information for sampling stations
3. How will field and laboratory data be handled?
- Some examples may be:
 - Field data are entered on a field sheet at the station, then these data are entered in a spreadsheet and the hard copy field sheet is filed away
 - Field data are entered using a portable digital data entry form and then directly uploaded to your database
 - Results from a lab may be sent back to you in a standard formatted spreadsheet
 - These processes may already be documented in your Quality Assurance Project Plan (QAPP) or QAPP Standard Operating Procedures (SOPs)
- Metadata are data that are supplemental to your data, or “data about data”. For example, the result values received from your lab are data, while information like the time the sample was taken and the sample collection procedure are both metadata.*
4. How will data be entered and validated?
- Some examples may be:
 - Field data are entered in a standard data spreadsheet that has consistent headings and naming conventions
 - A staff member enters the field results data from each station visited on a certain day into the program database, and the field results are linked via a unique identifier assigned to them to the sample results that will come back from the lab
 - A second staff member produces a report from the program database and cross checks the data values with those received from the lab
 - These processes may be documented in your QAPP or QAPP SOPs
- Note: An EPA-approved QAPP must be in place before a tribe can take environmental measurements using EPA project funds.*
5. How will data be managed to make them useful for their intended purpose?
- Some examples may be:
 - Data are within a database that multiple staff members can access from their computers using a user name and password
 - Data that are of questionable quality is flagged within a comment field in the spreadsheet so that when data analysis is done, these data can be carefully considered
6. How many people are performing data entry or data management?
- Some examples may be:
 - All staff members need access to a single database that includes monitoring data for all types of monitoring sites
 - Certain staff handle biological and habitat data while others handle lab data so separate spreadsheets are kept
 - A spreadsheet tracks who has entered the data and at what time, so that others know if an entry has been changed.
7. Who needs to know information about the data and how do they want to see the information?
- Some examples may be:

- Tribal council members want a report and a graph showing how water temperature and flow have been changing over the past 2 years
- Staff putting together the tribal Assessment Report need raw continuous pH data to determine if levels have spiked
- Reports to EPA that require certain information in a particular format.

Working through the above questions will help you decide which procedures you will need to create and which tools you may need in order to manage your data. Once you have answered these questions, you may want to develop written procedures for data handling, records management, and documentation. For example, when you consider how field and lab data will be handled, you may decide that you will enter data collected in the field into spreadsheets and keep hard copies of field sheets in a filing cabinet. You may decide that the lab data will be analyzed from an outside lab and that they will send you electronic copies of results that you will copy into spreadsheets. In any case, you could set up a few procedures about how your water quality monitoring program will operate and how you will record and track information. You could come up with a standard format for spreadsheets so that the items are reported in a specific order and names of what you are recording stay the same.

As far as records management, a good practice is to keep hard copies of all lab reports and field sheets in a filing cabinet, and electronic copies of all files on a specific computer with backup files on a CD, DVD, or external drive.

Once you have decided on what your data management procedures are, one way to approach documentation is to write out all the steps of your data management process as if you were explaining them to a new person. You can save and date the document and update it anytime your procedures change. This will help you remember what the procedures are and communicate them to others. For example, if you collect your field data in the spring and summer and produce a report to your tribal council each September, you can describe that process in your documentation, and include a copy of the report you use each year.

Note: It is considered a good practice to keep archived backup files in a physically different building than your main hardcopy or electronic files. This protects your files against an unforeseen disaster situation (e.g., fire, flood, tornado, etc.)

III. Data Types

Data types describe the way organizations, such as tribes, generally categorize water quality monitoring information. Your organization may not have these exact names for the same types of data, but these are some examples of the concepts that you are probably already capturing or should capture. These types of data influence the way you might organize your data within a data management system. Here are some examples of the types of data related to water quality. Some of these can be considered metadata, which contextualize monitoring data results, such as who, why, where, and how.

1. Who you are
 - Examples:
 - An organization that collects water quality monitoring information

- Name of staff member who was doing field measurements and taking water samples
 - Name of lab staff person who was analyzing a sample
2. Why you are collecting the data - An organization may have one or more projects to collect data.
 - Examples:
 - Lakes profile study
 - Wadeable stream condition study
 - Little Northeastern Creek condition study
 3. Where you are collecting data - Stations are places where monitoring activities occur.
 - Examples:
 - Atlantic Beach Near Gordan's Pond
 - Collets Run, above confluence with Entry Run
 - Chatfield In-Reservoir Near Dam
 - Alluvial Well – Flying C Ranch
 - Stations or monitoring locations have metadata associated with them. This may include using a GPS device to record the latitude/longitude of the station, recording directions to the station, or other overall observations of the station that may impact sampling. Metadata are an important component of your data, as they provides context for your monitoring activities and results.
 4. How you collect data - Monitoring activities are conducted by an organization's staff.
 - Examples:
 - Field measurement (parameters measured in field using instruments, e.g., temperature or Secchi disk depth)
 - Field observation (parameters observed in the field using the eye or ear, e.g., stream bank erosion)
 - Routine sample (media collected in the field for analysis by the lab, such as a water sample)
 - Quality control (QC) sample (samples used to ensure that no contamination is introduced during collection, transport or analysis of routine samples, – e.g., field and lab blanks - or samples used to assess the variability in collection or analysis methods, e.g., duplicates and splits)
 - Integrated vertical profile sample (discrete sample derived from a continuous vertical segment at your station, e.g., temperature measured at intervals from surface to bottom of waterbody)
 - Activities will have metadata associated with them. For example, the sampling collection method used may be an important piece of information. Your QAPP or SOPs will often contain this information.
 5. What data you are collecting - Results are the data collected or produced from monitoring activities.
 - Examples:
 - pH
 - Total nitrogen

- Fecal coliform
 - Count of mayflies
 - Undercut bank presence (% Rch)
 - Results will have metadata associated with them. For example, the analytical method that was used to attain a result value is an important piece of information that should accompany results. Your QAPP will often contain this information.
6. What field data you are collecting - Field data are results that come from assessing onsite water quality characteristics
 - Examples:
 - Water temperature
 - pH
 - Existing streambank erosion
 7. What lab data you are collecting - Laboratory data are results that must be analyzed by a lab.
 - Examples:
 - Total phosphorous
 - Total nitrogen
 - *E. Coli*
 8. How you check the data you are collecting – quality assurance (QA) or QC results are another type of result data that are very important to capture, not only for overall data quality, but for later data analysis. You may decide to track these types of results separately from the regular results, while linking them via a common monitoring location (see page 4-20 of the Tribal 106 Guidance for additional information on the QA/QC process and tribal responsibilities).
 9. What additional data do you collect – Assessment data, geographic information systems (GIS) data, wildlife surveys, etc.

These data types may already drive how you organize and manage data overall. For example, from year to year, your organization name will likely stay the same, although staff members may change. Similarly, the project will likely stay the same from year to year. Many tribes organize their information under one project. However, some tribes want to set up separate projects to track information related to specific studies or funding sources. Stations or monitoring locations often stay the same over time, but sometimes additional stations and water quality indicators will be added as a tribal water quality environmental program becomes more sophisticated. The monitoring activities and results are the data types that will change each time water quality data are collected.

As you begin or continue to implement data management for these data types within spreadsheets or a database, a few principles are important to keep in mind:

1. Required or optional data – within databases and spreadsheets, it is important to agree to what types of data you will track and how much of these data will be required. For example, you might always require the entry of the unit of measure for any result value.

Or, you might say that whenever a result value is not detected, a reason for the non-detect must be given in a separate field. In a database, enforcing these rules is easy once you have set them up. Spreadsheets may be documented so that all users can adhere to the rules that have been defined by your program.

2. Common identifying conventions – it is an important practice to begin to use common naming conventions, especially in tracking what types of water quality indicators or parameters you are sampling. For example, if you are tracking dissolved oxygen by a consistent name, it will be much easier to find, manipulate, and analyze all the measurement values associated with this parameter. Another example may be making sure a consistent naming scheme is in place for identifying all of your monitoring stations. In a database, drop down lists based on valid values can make it very easy to implement common naming conventions. You can also set up common values lists within spreadsheets.
3. Standard data formats – similar to common identifying conventions, having standard formats for certain data elements will also make it easier to manipulate data once you have begun to analyze it. For example specifying that all dates need to have a format of YYYY-MM-DD or that all latitude and longitude coordinates should be in decimal degrees will make the data useful for all staff members. Be sure to know the exact capabilities of the software you are using before assuming that the data format you define is absolute.

You may also be managing assessment data or GIS data. Regardless of the types of data you currently collect or plan to collect, you should list the data types and determine how you will manage information for each data type. The same data management principles apply whether you are managing water quality monitoring data, assessment data, or GIS data.

Water quality monitoring data consists of field data or lab results associated with sampling events at a monitoring location.

Assessment data are decisions about water quality that are made based on comparisons of water quality monitoring data against water quality standards.

IV. Data Management Formats

Figure 2 is a road map for you to determine your current level of data management and what recommendations you can consider in fully implementing data management as you move forward (e.g., Recommendation A, etc). There are many ways to manage your data, so you should choose the one that works best for your organization now and in the long term. Each format has advantages and disadvantages. But don't underestimate the potential for what seems like less work now making far more work for you in the future as your program develops.

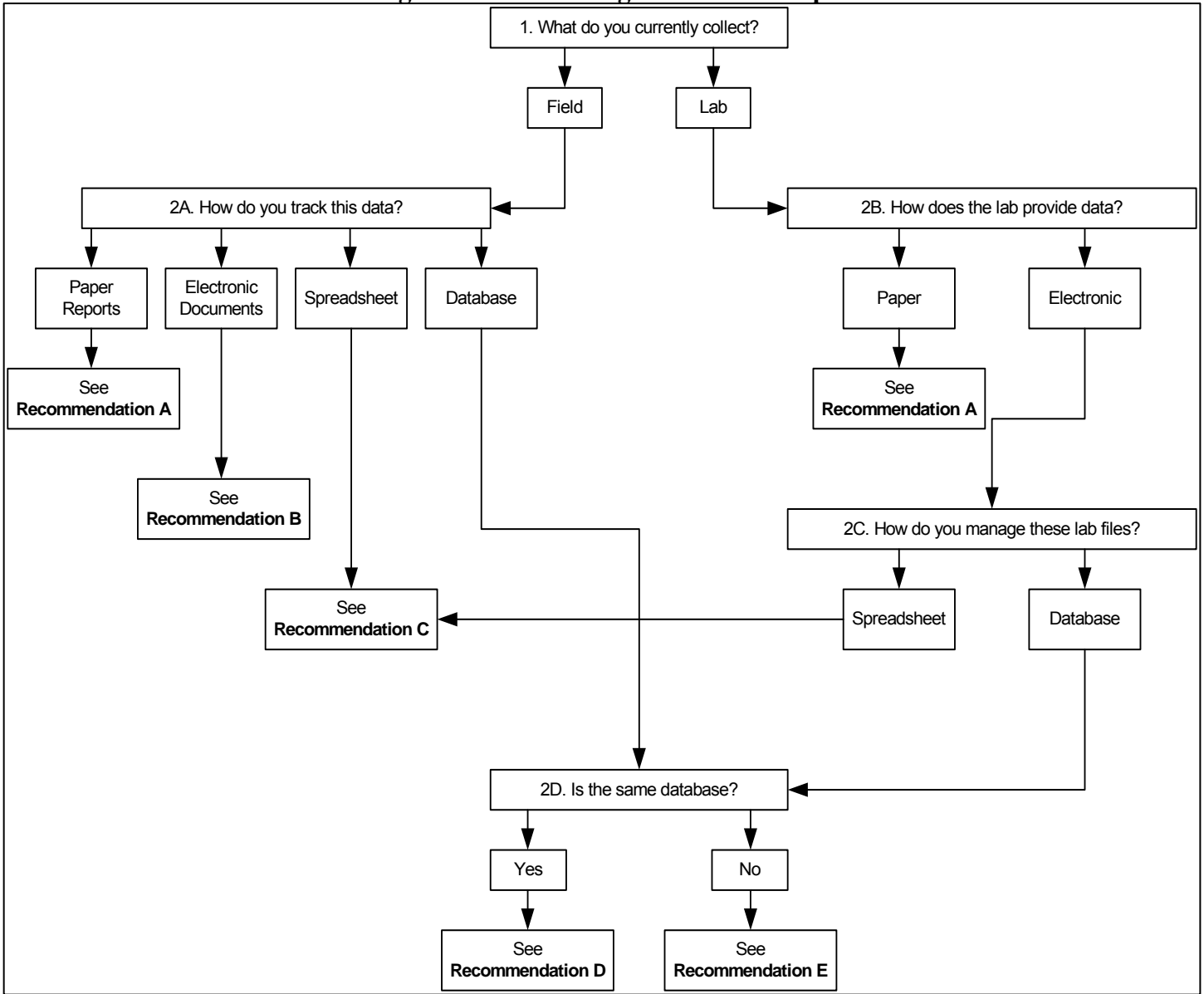
Some formats you may be currently using include:

1. Paper reports
2. Electronic documents, such as word processing or scanned documents

- 3. Spreadsheets
- 4. Databases
- 5. GIS

While GIS is a common way of analyzing and presenting data, it is supplemental to most datasets and is not a core way of managing data.

Figure 2. Data Management Road Map



Recommendation A. Some tribes keep track of information in paper reports that are filed in filing cabinets. Keeping papers can be relatively easy, especially if that is what you have been doing previously and/or if labs mail you paper reports. It is also smart to keep paper copies of important data and reports as a back-up system to data stored electronically. However, keeping track of all of your data on paper makes it very difficult to analyze the data, make environmental

decisions, and share the information with your community. You also run a greater risk of losing the data over time. You could also contact your lab and request that they provide data to you in an electronic format specified by your environmental program. Once you get data in an electronic format, you can expand your data management capabilities and use the data collected to meet the water quality goals and objective of your environmental program.

Recommendation B. Other tribes keep track of information in electronic documents such as Microsoft (MS) Word, Word Perfect, or PDF documents. The advantage of having electronic documents is that they are easy to print and revise. Although electronic documents are helpful, they are not very effective methods for managing data. It is difficult to find, manipulate, sort, and analyze data when it is in a document format. If you maintain electronic documents with your data, you may consider pasting or entering the data into a spreadsheet or database. Please note that if you do transfer data from electronic documents to spreadsheets or a database, it is important to do a QC check to ensure that the data were transferred correctly.

Recommendation C. Spreadsheets organize data in tables (rows and columns) that you manually construct and can directly rearrange, sort, search, and use to perform many calculations and statistical and graphical analyses. Several spreadsheet programs are available, such as MS Excel. If you have all of your data in spreadsheets, your options begin to open up a little bit more. You can easily copy and paste data, rearrange data for specific reports, and even create charts to present the data. If you are using a spreadsheet format, it is helpful to define a consistent standard format for your spreadsheets (i.e., same columns, same column order). You may also develop a standard way of reporting certain items, like the parameter names and units of measurement. For example, if you sometimes record “temp” and other times record “temperature,” or sometimes record “degrees” and other times “degC,” you should choose the name and units you want to use consistently. Unlike hard copy or electronic word processing documents, more formal business rules (e.g., required or conditionally required fields) can be incorporated into a spreadsheet. Because of the inherent lack of a centralized location, it is frequently found that the process of manually rearranging data is quite time consuming and can generate many files and versions. Consequently, it becomes even more important to institute a good file management plan. You can develop a storage plan for how to name and track your spreadsheets, and to ensure that you keep one master version of the spreadsheet to avoid confusion about what is the most up to date, correct spreadsheet. It is very important to keep separate read-only copies of spreadsheets somewhere in case you make a mistake as you change a spreadsheet, and to prevent others from inadvertently changing data in them.

Recommendation D. Databases (such as MS Access, Oracle, or MS SQL databases) organize data in tables, but you typically interact with them through data entry screens. Although databases take a considerable amount of preparation time, once they are completed, it is easy to enter, store, retrieve, perform relational queries, and report the data. Database systems can vary in complexity, from simple “Out of the Box” programs like MS Access, to more complex, customized systems, like Oracle or SQL Server. Databases are the best way to ensure data consistency and that the most correct version of the data is maintained. They can, however, represent a fairly significant technical hurdle to those who have not worked with them. Therefore, it is recommended that you identify a qualified person who has primary responsibility for database management and maintenance. If you have a database, you are most likely already meeting your need for managing data.

Recommendation E. If you have multiple databases to manage your field and lab water quality data, it is recommended that you combine your data into one database. Part of the power of having a database is that you can store multiple types of monitoring data together, such as lab and field data, in an easily accessible format. A database will allow these data to be related to each other for a given station, on a specific date. A database also eliminates the need to store data more than once, such as station information. Raw monitoring data should be stored separately from assessment data because this allows for future assessments to be made based on the same data. For example, the result for 8° C for water temperature would be kept in a separate database from the assessment conclusion that the stream is determined to be meeting the Tribe's threshold for supporting a cold water fish habitat.

GIS software gives you another way to store and display your data that combines simple database and graphing functions with the power of putting data directly onto a map of the watershed. GIS software allows you to visually analyze, manipulate and report on your data based on the location of your sampling point in the river or watershed. Most GIS software also easily imports other types of data formats into a table to which the spatial information can be linked.

STORET-Compatible Data Submissions

I. Getting Your Data Into a STORET-Compatible Format

The STORET (short for STOrage and RETrieval) Data Warehouse (or STORET Warehouse) is a repository for water quality, biological, and physical data and is used by tribal environmental organizations, state environmental agencies, EPA, other federal agencies, universities, private citizens, and many others. Formatting your monitoring data to be STORET-compatible, not only benefits the EPA, it also benefits tribes and other agencies. By putting data in a STORET-compatible format, you are formatting the data you are managing (with Excel or Access, etc.) into a standard set of data elements, formal business rules and allowable values that comply with the STORET Warehouse, allowing the data to be placed in the STORET Warehouse. Once data are in the STORET Warehouse, they can be retrieved for use in a variety of studies and analyses. For instance, your tribe can acquire your data and data from surrounding jurisdictions to identify upstream issues that might be of concern to your reservation. The STORET Warehouse can also be useful in the event of an internal data loss.

As part of the required tribal Assessment Report, tribes are required to submit monitoring data to EPA electronically in a STORET-compatible format for each assessed surface water body. This section describes what it means to put data into a STORET-compatible format.

There are different ways to submit data to the STORET Warehouse, which will be discussed in the next section. It is important to note that the Water Quality Exchange (WQX) is the new way to place data into the STORET Warehouse and will be discussed in the next section. The STORET-compatible format as outlined below is very similar to the WQX format, and any efforts in putting data in a STORET-compatible format are well worth it for any type of data submission.

For each data type, STORET requires that certain information be reported in a very specific way. These data types answer the questions defined in Section III: Who, Why, Where, How, and What. Tables 1 through 3 show information needed to submit data to STORET Warehouse and where to get the information.

- *Field Name* is the data parameter that can be sent to STORET. You will likely have similar names in your own data sets.
- *Description* explains what the field name means and other helpful hints.
- *Source* provides information on where to get the data from.

Please note Tables 1 through 3 come from the STORET Data Template, an MS Excel file that is available from the EPA STORET Web site (http://www.epa.gov/storet/national_websim.html). The Template also contains other worksheets that can be very helpful to understanding STORET. The template was created as a result of consultation with tribes throughout the country. It can be used to actually manage data or simply as a guide to see how data can be organized for easy import into the STORET Warehouse.

Table 1. Project Information in STORET

<i>Field Name</i>	<i>Description</i>	<i>Source</i>
Project ID	A unique ID for the Project to which data will be assigned	User Created
Name	A unique Name for the Project	User Created
Start Date	Date on which the Project started (MM/DD/YYYY format)	Date which monitoring work began (may cover multiple years)
Duration	Planned duration of the Project	Either planned end date of monitoring for this purpose or "Ongoing"
Purpose	Project goals, expectations, or why these data were collected	User Created

Table 2. Station Information in STORET

<i>Field Name</i>	<i>Description</i>	<i>Source</i>
Station ID	Unique identifier for the Station	User Created
Station Name	A geographically descriptive Station name	User Created
Primary Type	Station's primary type	Listed in template allowable values
Latitude	Latitude. See available format options. (DD.DDDDDDD recommended)	GPS reading, Map Interpolation, etc. – based on methods listed in template
Longitude	Longitude. See available format options. (DDD-DDDDDDDD recommended)	GPS reading, Map Interpolation, etc. – based on methods listed in template
Geopositioning Method	Method used to determine Lat/Long coordinates	Listed in template allowable values
Geopositioning Datum	Datum used to determine Lat/Long	Listed in template allowable values
State	Postal abbreviation for Station's primary state, not required for Great Lake and Ocean Stations	The state containing the Station
County	Name or FIPS code for Station's primary county	The county containing the Station
Station	Station Description	User Created

Field Name	Description	Source
Description		

Table 3. Results Information in STORET

Field Name	Description	Source
Activity ID	A unique identifier for each activity (an activity can be a unique lab ID, a suite of on-site field measurements, an individual sample, etc)	User-created
Medium	Medium in which this activity occurs	Listed in template allowable values
Activity Type	Type of activity	Listed in template allowable values
Activity Category	Category of activity	Listed in template allowable values
Activity Start Date	Date activity began (MM/DD/YYYY)	Date samples or measurements taken
Sample Collection Procedure ID	Valid STORET sample collection procedure, required if activity type is sample	User-created based on QAPP
Characteristic Name	Valid STORET characteristic name	Listed in template allowable values
Result Value	Measured value of characteristic	From lab results or field sheet
Result Value Units	Units for characteristic Result	From lab Results or field sheet
Sample Fraction	Fraction for characteristic, required for certain characteristics (see Non-Taxa Characteristics for details)	Lab Results
Field/Lab Procedure	The field or lab procedure used to analyze the sample, required for certain characteristics (see Non-Taxa Characteristics for details)	Listed in template allowable values or created by organization staff based on QAPP
Field/Lab Procedure Source	Source acronym for field/lab procedure, required if field/lab procedure entered	Listed in template allowable values or use organization ID

To manage and report data in STORET-compatible format, you will need to provide at least the data fields listed in Tables 1 through 3. The STORET template provides detailed guidance on the requirements for each field that you may need to provide. For example, you will need to make sure that:

- Field length of data in a data field is not longer than what STORET will accept
- You provide information in the defined formats or select the appropriate allowable value from the STORET lists, as defined in the template
- You provide the unique identifier that STORET requires for all projects, stations, and activities within your organization.

If you are implementing this within your own data management system, meeting these STORET requirements will not be difficult. This helps to make the use of these data easier for everyone involved by aiding data standardization, which is essential for data sharing at the national level.

Metadata, or data about your data, are very important for telling potential data users how data were gathered and why certain data may be unreliable. For this reason, it is encouraged that all relevant QC data be entered into STORET, as well as specific sample collection procedures and field/lab procedures.

Data standards are consistent ways to report data in formats so they are easy to share. When reporting Results, you will need to use the appropriate allowable value for characteristic names. The STORET characteristic names are based on an EPA data standard. Table 4 shows the list of the parameters defined in the Tribal 106 Guidance and the corresponding STORET characteristics. While you are only required to manage and report data on particular characteristics listed in the Tribal 106 Guidance, many tribes have decided that it is easier to manage and report data on all the characteristics they monitor. One of the first activities you will want to do if you plan to submit characteristics to STORET that are not listed in Table 4 is to match the characteristic names you use to those in STORET. You can look up a long list of allowable values for characteristics on the STORET template.

Table 4. Water Quality Parameters (allowable values in STORET)

<i>Parameter as Defined in the Tribal 106 Guidance (page A-4)</i>	<i>Corresponding STORET Characteristic</i>
Dissolved Oxygen	Dissolved oxygen (DO)
Water Temperature	Temperature, water
Turbidity	Turbidity
pH	pH
Phosphorous	Phosphorous ^a
Total Nitrogen	Nitrogen
<i>E. Coli</i>	Escherichia coli
Enterococci	Enterococcus Group Bacteria

^a Total phosphorous is calculated as the sum of both orthophosphate and phosphorous, depending on your methodology, you may need to include both of these characteristics to accurately describe total phosphorous. Alternatively, you can use the characteristic of 'Phosphorous' combined with the sample fraction of 'Total' to describe total phosphorous.

Because biological and habitat data are handled differently than physical/chemical characteristics, you should contact your region or the STORET Helpdesk at 1-800-424-9067 or at storet@epa.gov to discuss this electronic format.

There are also standard ways to report the procedures you use when collecting data in the field and that the laboratory uses to evaluate samples. For example, when you report a result to STORET you will need to select a Field/Laboratory Procedure source, such as the National Institute for Occupational Safety and Health or the EPA. You will also select the Procedure ID, such as 150.1 for pH or 150.2 for pH by continuous monitoring. All of these allowable values are listed on the STORET template. If you use procedures that are unique, you can ask that they be added to STORET by contacting the STORET Helpdesk. One of the first activities you will want to do if you plan to submit data to STORET is to identify the sources and IDs for the field and laboratory procedures you use for each characteristic. You can work with your laboratory to

find out its procedures for analyzing certain parameters. Tribal QAPPs, SOPs, and even data logger manuals are also good resources for procedural IDs.

Overall, the following steps will help you get your data into a STORET-compatible format. At any point in this process, contact the Tribal STORET support line at 1-800-844-0638 or owsupport@rti.org if you have any questions about how to organize your data, what are the allowable values, or other topics. For these steps below, the STORET template available on the EPA Web site will show specifics such as allowable and required values: http://www.epa.gov/storet/national_websim.html.

1. Ensure that you are collecting all the required data fields listed in Tables 1 to 3, above, and as shown in the template available on the EPA Web site.
2. Figure out what values STORET will allow for your parameters and field and laboratory procedures. Refer to the STORET template for help.
3. Get your data formatted according to STORET's specific required fields, field length, and field type. You might want to use the STORET template, which lays out these items.
4. Before you can send your data to STORET, you need to request an Organization ID by contacting the STORET Helpdesk at 1-800-424-9067 or at storet@epa.gov.

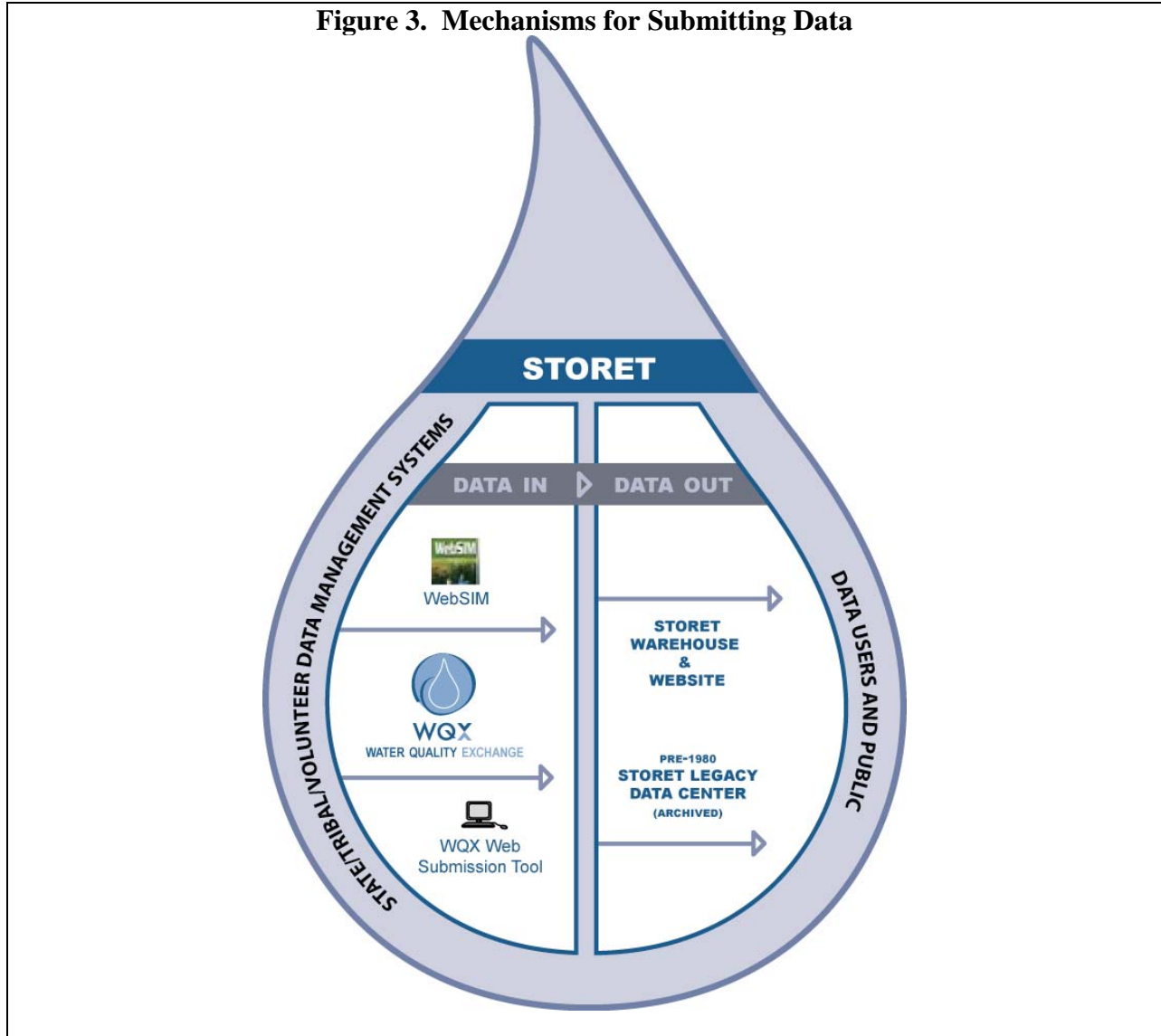
There are other applications that have been created by members of the user community that may be used to help you put your data into a STORET-compatible format. You may find some of these tools on the STORET Web site (<http://www.epa.gov/STORET/otherapps.html>).

II. Ways to Submit Data

As you may know, the STORET Warehouse contains raw surface and ground water biological, chemical, and physical data (from 1999 and beyond) collected by federal, state and local agencies, Indian Tribes, volunteer groups, academics, and others. The STORET Warehouse will continue to be the repository for all STORET data.

Section 5 talked about how to put your data into a format that is accessible by the STORET Warehouse. Once your data are in this format, the next step is to submit the data electronically to the STORET Warehouse using the different tools and mechanisms available. Figure 3 shows the different mechanisms for submitting data to the STORET Warehouse.

Figure 3. Mechanisms for Submitting Data



EPA expects that once data are in a STORET-compatible format, they will be submitted to the STORET Warehouse. EPA recognizes that due to resource limitations, certain tribes need EPA support in completing the data submission. In these cases, submitting the data electronically to your project officer (via e-mail, CD, DVD, etc.) may be sufficient for meeting the data submission requirement. This option is not discussed below, and should be discussed further between you and your project officer.

To submit your data electronically into the STORET Warehouse, determine what mechanism is best for your tribe. Each approach has advantages and disadvantages, but all are acceptable means for meeting the data submission requirements set forth in the Tribal 106 Guidance. Here are some options to consider (described in more detail below):

1. Submit data through the Web STORET Import Module (National WebSIM)

2. Submit data through WQX using an Exchange Network Node or Node Client
3. Submit data through the WQX Web Submission Tool (WQXWeb)

This document does not include detailed instructions on how to submit data using these tools. Please see the resources for support at the end of this document for more information about the steps to be taken to utilize these tools.

1. Submit data through the Web STORET Import Module (National WebSIM)

National WebSIM is a Web-based user interface for uploading text files of Project, Station, and Results data. More information about this option, including a data management template and National WebSIM tutorials can be found at http://www.epa.gov/storet/national_websim.html. In addition, the National WebSIM tutorials on the Web site help you understand how it works and how to submit electronic files to STORET.

Please note that National WebSIM will be phased out as it is replaced with WQXWeb. WQXWeb will be similar to National WebSIM, so any effort made with National WebSIM will not be wasted. EPA plans to continue to support National WebSIM as WQXWeb is released so that tribes can transition to the tool in a complete and comfortable fashion. National WebSIM will remain available until September 2009. Any progress that you make now towards data management, including submitting data to STORET through National WebSIM, will help you organize your data and prepare for the future.

2. Submit data through WQX

WQX is another way to submit your data into the STORET Data Warehouse. WQX is a framework that makes it easier for tribes, states, and others to submit and share water quality monitoring data over the Internet. WQX uses EPA's Exchange Network (<http://www.exchangenetwork.net/exchanges/water/wqx.htm>) to flow water quality monitoring data to STORET from a tribe's data system via Exchange Network nodes or node client. Tribes that have been awarded Exchange Network grants use WQX to submit files. Most tribes submitting data through WQX directly have database systems for managing data. Some tribes have begun working together toward utilizing WQX for submitting data. If you are considering this route, you should stay aware of activities going on in your region through your Exchange Network grant representative, 106 grant representative, and monitoring coordinator. More information about the WQX can be found at <http://www.epa.gov/storet/wqx.html>.

EPA's Exchange Network is a way to exchange environmental information securely over the internet, using an approach based on standard technology and protocols.

Exchange Network Grants are awarded to states, territories, and federally recognized tribes to assist in the development and implementation of Exchange Network Nodes, exchanges, tools, and related backend systems.

3. Submit data through WQXWeb

WQXWeb is a tool that will eventually replace the National WebSIM, but it will work in much the same way. Like WebSIM, WQXWeb allows the uploading of text files, but also allows the creation of a WQX file and the submittal of this file to the STORET Warehouse via WQX. Standardized templates, as well as tutorials, will continue to be available to assist tribes in formatting data in the correct format for data submission. The tool also allows tribes to correct data errors upon submission. Tribes will be able to easily transition from the National WebSIM to the new WQX web submission tool because they will be familiar with formatting and preparing data for submission using a Web-based tool.

Once data is submitted to the STORET Warehouse, it is available for use by your tribe, EPA, and the general public who have an interest in the water quality data in your area. As a part of a tribal Assessment Report, the data submitted to EPA will enable water quality managers to have a better picture of how water quality is improving in Indian Country. If you have privacy concerns about the availability of monitoring data through the STORET Warehouse, please see Appendix C (Letter to Tribal Leaders on Information Access) of the Tribal 106 Guidance, and consult with your project officer.

Some tribes may find that they have errors in data that they have submitted to the STORET Warehouse. All the tools for submitting data described above allow a tribe to change any submitted data. For example, if a set of results taken during a certain sampling season was taken using equipment that was later found to be defective, the tribe can use any of the submission tools to delete the erroneous data and resubmit the new data.

The STORET Warehouse does not provide QA/QC services (i.e., does not check data for “bad” values), but rather, trusts data providers to perform QA/QC before the data is submitted. It is generally unnecessary to report questionable data to the EPA, unless the data are used as part of your water quality assessment. Data with questionable data quality can be flagged within STORET if shared.

III. Retrieving Data from the STORET Warehouse

You may wish to perform analyses on data that has been loaded to the STORET Warehouse by you or other organizations having jurisdiction in your area. In order to do this, you will need to retrieve the necessary data from the STORET Warehouse. You may download data from the STORET Data Warehouse by going to the STORET Web site (<http://www.epa.gov/storet/index.html>) and clicking on the “Download Data” link. You will have the choice between accessing Legacy STORET data or Modernized STORET data. All data that was loaded during or after 1999 will be found in the Modernized STORET data. The STORET Warehouse offers several query options. Once you have selected a query option, a form will help you filter your data down to only the data that are needed. Once you have requested the data, you will need to select what data fields that you want and provide your e-mail address so that you can receive a web-link telling you where to download your data set. Detailed instructions for downloading data from the STORET Data Warehouse are available from http://www.epa.gov/storet/Downloading_STORET_Data.pdf.

A second option for retrieving data from the STORET Data Warehouse is through Web Services. Web Services offer a way to allow computers to retrieve data directly from the STORET Data

Warehouse in predefined formats using a computer to computer protocol for directly sharing information. You may find more information on STORET Web Services at: http://www.epa.gov/STORET/web_services.html.

IV. Ongoing Support

If you have questions about getting your data into a STORET-compatible format or how to submit data to STORET, there are ways to seek help. Table 5 lists some resources that may be helpful to you.

Table 5. STORET Resources for Tribes

Item	Location
STORET Web site	http://www.epa.gov/storet/index.html
Information on National WebSIM	http://www.epa.gov/storet/national_websim.html
Information on WQX	http://www.epa.gov/storet/wqx.html
STORET Tribal data management template	http://www.epa.gov/storet/national_websim.html
National WebSIM Tutorials and Example Files	http://www.epa.gov/storet/national_websim.html
Data Management Applications	http://www.epa.gov/STORET/otherapps.html
Information on Exchange Network	http://www.exchangenetwork.net/
STORET Helpdesk	1-800-424-9067 or storet@epa.gov
Tribal STORET/WQX Support Line	1-800-844-0638 or at owsupport@rti.org

The STORET helpdesk and tribal STORET/WQX support line are useful resources to assist you with any issues that arise with STORET/WQX. EPA's STORET helpdesk can set up Organization IDs and enter field/lab procedures into the system. The tribal support line can answer questions about data management and reporting, or help you troubleshoot problems with STORET/WQX.

As part of submitting data, you should determine how you plan to submit data and how often you should submit it to EPA. That way, you can reserve the staff hours necessary to submit data on a routine basis. The more often you submit data, the more familiar you will become with the process. It will get easier every time. Once you have decided on a plan for submitting data, you should document it and revise the plan as necessary.

ATTACHMENT A: Examples of Tribal Data Management

This section provides examples of tribal data management systems that were provided by tribes. Please note that the Fond du Lac, Penobscot Indian Nation, and Yurok Tribe Environmental Program are considered to have mature data management programs because they use databases to manage their water quality environmental data and provide data to EPA in a STORET-compatible format.

Fond du Lac

Fond du Lac's water quality monitoring program includes physical, chemical, and biological data from reservation lakes and streams, going back to 1998. In 2008, we will begin collecting similar data from a targeted group of wetlands. Physical field data include a Hydrolab profile in lakes (temperature, dissolved oxygen, pH, specific conductance, turbidity and depth), Secchi transparency; and in streams, continuous temp loggers, flow/discharge measurements, and habitat assessments. We have begun to use BioSonics integrated hydroacoustic/GPS equipment to collect data that will generate accurate bathymetric maps, including substrate type and vegetation mapping. Chemistries include nutrients (TKN, nitrite-nitrate, ammonia, total phosphorus, ortho P), total suspended solids, total hardness, alkalinity, color (correlation verified one season with dissolved organic carbon), chloride and metals (arsenic, cadmium, chromium, copper, lead, nickel, selenium, zinc), chlorophyll a (plus ash free dry mass for stream periphyton). Biological data include algal community and zooplankton from lakes (will add fish this year), and benthic macroinvertebrates and fish in streams. Additionally, we have sampled fish tissue for mercury and other contaminants of concern, and have done two sediment studies on reservation lakes and the St. Louis River, looking at physical characteristics (grain size, moisture, organic content), mercury, lead, PCBs, and toxicity tests.

Our water quality data is integrated into an Access database, which was organized and designed to facilitate STORET transfer. We record field data and lab data in separate, hard copy notebooks, which are kept as a permanent record after the information is entered into the database. Contract labs submit both electronic files (generally Excel spreadsheets) and hard copies to us; our database has some forms designed specifically for ease of import of these files (including our Hydrolab output). Our sediment data is in a separate Access database, designed to be similar to the National Sediment Quality Database.

The contract lab that analyzes most of our chemistry data formats the electronic files so that they can be easily imported to our Access database. Since the database was designed to be similar to STORET, our intent was to make data migration as simple as possible.

We have used National WebSIM to migrate our water quality data to STORET. After the training we received, we successfully uploaded all of our water chemistry data from 1998-2006 into STORET. Since we have had staff turnover in the interim, we have not yet entered the 2007 data; will be training new staff to carry on this duty.

Our Access database has provided us with a comprehensive electronic "parking lot" for all of our water quality monitoring data. Because of the way it is designed, the data is in a secure place

with only two key staff able to enter or edit the data, and it is backed up nightly. We have been able to extract data—either specific parameters or specific waterbodies—to be able to conduct analyses, do assessment and reporting, as well as being able to easily share data with other resource management agencies in a format they can use.

Penobscot Indian Nation

Penobscot Indian Nation (PIN) collects general water quality monitoring data from late May to early October. Table 1 describes the different locations and types of monitoring we do.

Table A-1. PIN Water Monitoring Program

Location or Type	Number of Stations	Characteristics monitored
Penobscot River	84	temperature, pH ^a , dissolved oxygen, conductivity ^a , BOD ^a , <i>E. coli</i> bacteria ^a , color ^a , total suspended solids ^a , turbidity ^a , secchi disc visibility, foam (index). Additional parameters monitored less frequently include: phosphorous, nitrogen, chlorine, closed cell pH, alkalinity, tannin and lignin, chlorophyll <i>a</i> , organic carbon
Penobscot River Tributaries	30	temperature, dissolved oxygen, conductivity, <i>E. coli</i> bacteria, color, total suspended solids, and turbidity
Trust Land Lakes and Ponds		temperature, pH ^a , dissolved oxygen, conductivity ^a , <i>E. coli</i> bacteria ^a , color ^a , total suspended solids ^a , turbidity ^a , secchi disc visibility, foam (index), alkalinity, develop trophic status index
Biomonitoring	4-6	This method involves examining communities of aquatic insects to determine water quality conditions
Ambient Toxics Monitoring		Contaminant monitoring primarily focuses on aquatic species that are used by Tribal members or that have ecological importance. Species examined include: Smallmouth bass, white suckers, muskrat, snapping turtles, freshwater mussels, and fiddlehead ferns. Toxics examined include: dioxins, furans, PCBs, chlorophenols, and trace elements (including Mercury). Regularly collect Penobscot River fish samples for ME DEP programs including Dioxin Monitoring Program and Surface Waters Ambient Toxics Program. Provide assistance to University of Maine researchers (i.e. bioassay development, histeopathology, endocrine disruption)
Other Water Quality Projects	13 ambient sites 4 point source sites	Monitoring Project to document and determine the cause of algae blooms during summer months on the Penobscot River. Sampling within the watershed for temperature, nitrogen, phosphorous, chlorophyll <i>a</i> , total suspended solids, dissolved organic carbon, algal taxonomy, presence of toxins and physical/visual characteristics. Preliminary study to determine presence and characteristics of sediments in the Penobscot River. Assist USGS with occurrence & distribution study of contaminated sediments. Continuous temperature monitoring Project to examine daily water temperature patterns at numerous sites within the mainstem and tributaries of the Penobscot River. Assessment & management plan of non-point source pollution on Trust Lands and Indian Island Spring thaw sampling on Lakes/Ponds to determine effects of acidity on fish Cooperative waste load allocation study in 1997, 1998 and 2001 with Maine Department of Environmental Protection, EPA, and several dischargers to collect data for dissolved oxygen modeling of the Penobscot and Piscataquis Rivers.
<p>^a Because the Penobscot Nation Water Resources Program has a Water Quality Monitoring Cooperative Agreement with ME DEP, collected data are made available for use in Maine's 305 (b) Water Quality Assessment report to Congress. The classifications of over 400 miles of Penobscot River watershed were upgraded in the last reclassification effort by DEP as the result of Penobscot Indian Nation water quality data.</p>		

We manage our data in a MS Access database called Penobscot Indian Nation Environmental (PINE). Field data on physical/chemical parameters are collected on mobile digital assistants

using forms created in Visual CE (software from Syware) based on tables in the PINE database. This data is synchronized with a field database file. The field database has forms for making final edits, checking in the samples and documenting bacteria sample preparation. The relevant field information is moved to a laboratory and archive database file. The laboratory database file has forms for data entry on all in-house analyses done by PIN including: bacteria, total suspended solids, biological oxygen demand, alkalinity, conductivity, turbidity, color, and pH. Laboratory data also get sent to the common archive database file. An uploading tool copies both field and laboratory data into the main PINE database file.

The process for real-time electronic entry of field and laboratory data has been amended to accommodate entering data from paper reports from previous years. Current efforts include making data processing more efficient, incorporating more quality-assurance checks, and producing a good data entry form for information from external laboratories. In addition, we are working to add continuous data from data loggers/data sondes into the database.

PIN is currently matching the PINE database fields to the required WQX elements. We plan to develop a data entry process for the external labs to use so that they can easily import the data when it is received. Once the matching is done, we will run a series of database queries and exporting the Results to an XML file for submission to EPA.

At this stage of developing the PINE database, we are just beginning to be able to look at all of the data for one sampling season at the end of it. Our first Assessment Report to EPA included a summary of the average, minimum and maximum of each characteristic they analyzed in the PIN lab for each site they visited. We also plan to calculate a trophic state index for each of the lakes they monitor. As we begin entering more past data, we plan to look at trends over time.

Benefits of data management for our program include:

- Improved ability to protect traditional tribal lifeways
- Integration of all aspects of monitoring so each team member has a better understanding of how their contribution fits into the effort as a whole
- Improved quality of data entry
- More timely assessment of problems with water bodies, quality-control sampling/analysis efforts, and the ability to produce reports on problems and analysis
- Better and more timely information to decision-makers
- Time savings in the long run

Yurok Tribe Environmental Program

The Yurok Tribe Environment Program (YTEP) collects continuous water quality data (temperature, pH, conductivity, Dissolved Oxygen), bacteria, nutrients, and chemicals. We also measure stage height and flow via continuous monitors. In order to manage our data, we built an SQL-based data warehouse. The database includes an XML translator to generate files and submit them to WQX. Water quality monitoring data is pushed from the database into WQX on a regular basis, as data packages are entered, validated, quality assured (QA'd), and approved. YTEP used Exchange Network funding to enhance their data management system. (For more information on these grants, see the Exchange Network Grants web site at: <http://www.exchangenetwork.net/grants/index.htm>)

YTEP prepares an annual report summarizing water quality data collected. We also prepare a report summarizing the hydrological data collected and presenting the rating curves generated by staff based on data. We do not do much analysis beyond minimum/maximum/average charting, seven-day moving averages, percent exceedance for temperature, and identification of measurements which violate water quality standards.

Benefits of data management for the YTEP include enhanced data quality, improved quality assurance/quality control (QA/QC), streamlined data handling, and increase in efficiency (reduced staff time demand for data handling).

ATTACHMENT B: Glossary

Allowable Values – Acceptable data values that can be input for a given data element. These values provide information on the exact spelling and capitalization needed for acceptance into the STORET Warehouse.

Assessment Data - Decisions about water quality that are made based on comparisons of water quality monitoring data against water quality standards.

Characteristic Names – Standard names used to describe measured or analyzed parameters.

Common Values Lists –List of acceptable data values that are allowed to be input into the STORET Warehouse. Common values ensure consistency between data entered by different organizations.

Conditionally Required Fields – Data fields in STORET/WQX that are required only when certain conditions are met in other data fields. For example, a time zone must be entered if you enter the time of an Activity, but is otherwise not required.

Data Elements – A basic unit of data within a data flow or transfer. It is often called a data field in a database.

Data Field – A column in a database that can hold one specific data parameter. For example, the station ID for all stations would be contained within one data field.

Data Standards – Consistent ways to report data so that they are easily understood and shared between data users.

Distributed STORET – Desktop software that allows users to enter and upload data directly to the STORET Warehouse. Distributed STORET will not be supported after September 2009.

Exchange Network Node – Software that facilitates the transfer of data between an organization's database and other databases on EPA's Exchange Network. Nodes use standards based web services and eXtensible Markup Language (XML) to allow for the seamless exchange of data between organizations regardless of hardware, operating system, or programming environment.

Extensible Markup Language (XML) – An open standard language that describes data using rigid syntax rules. XML allows data to be transferred between different data systems, regardless of computer system or platform. XML overcomes system incompatibility by translating information into a common data structure and format.

Field length – The maximum number of characters or numbers that can be placed in a data field.

Field name – The name of a data field, which provides a common reference for a data field.

Field type – Term referring to whether a data field can contain freely formatted text data, data of a defined format, or must conform to a list of allowable values.

Formal Business Rules – Rules that ensure completeness and consistent structure of data input into the STORET Warehouse.

GIS Data – Data of a spatial nature, such as a map associated with information about your waterbodies, produced using a Geographic Information System.

Metadata – Data that are supplemental to and add context to your data, or “data about data.” For example, the Result values received from your lab are data, while information like the time the sample was taken and the sample collection procedure are both metadata.

National WebSIM – Web-based tool used to upload water quality monitoring data to STORET Data Warehouse from spreadsheets and databases. Will be replaced by WQX Web Submission Tool (late 2008 or 2009).

Node Client – Software that can be downloaded on to individuals’ computers and can initiate Exchange Network requests for data or a submission of data to other nodes. Unlike nodes, node clients can not respond to requests for data.

Quality Assurance (QA)/Quality Control (QC) – Processes, samples, or results that help validate and check the quality of monitoring data and the processes used to acquire results from monitoring activities.

Relational Queries – Queries used to access data in a database that relate data in one table to data in another table by the use of matching values in key fields.

Required Fields – Data fields in STORET/WQX that must be included in a data submission under all circumstances.

Spatial information – Information that allows data to be associated with a location on a map.

STORET Data Template – Spreadsheet tool used to facilitate the entry of water quality data for eventual upload to the STORET Warehouse. Will be called STORET/WQX data template in the future.

STORET Data Warehouse – The repository for water quality data, a warehouse for both STorage and RETrieval of water quality data from across the U.S.

Water Quality Exchange (WQX) – The new data transmission standard for upload to the STORET Warehouse. WQX uses the XML programming language to format data.

Water Quality Monitoring Data – Field data or lab results associated with sampling events at a monitoring location.

Wildlife Surveys – Studies designed to census and measure the number of species in order to assess the ecological health of an area.

WQX Web Submission Tool (WQXWeb) – The replacement for WebSIM that will convert data (from spreadsheets and other flat file formats) to the XML format and migrate the data to the STORET Warehouse. This tool will fully replace the current function of WebSIM.