

## **Protocol P-13-53**

# **Preparation of Algal Data Files and Reports for Submission to the USGS NAWQA Program**

Candia A. Knowles, Frank Acker, and Kathleen Sprouffske

### **1. PURPOSE**

- 1.1. The U.S. Geological Survey's (USGS) National Water-Quality Assessment Program (NAWQA) collects algae samples from rivers throughout the United States. The Phycology Section of the Patrick Center for Environmental Research is responsible for the identification and enumeration of algae in these samples. For each group of NAWQA algal samples, data files containing results of these analyses are generated and transmitted to the NAWQA Biological Transactional Database (BioTDB). These files form the basis of a data set used to assess regional and national water quality conditions.
- 1.2. The purpose of this protocol is to describe the procedures necessary to perform checks on data quality and completeness. These procedures are followed to ensure that the data files submitted to the NAWQA BioTDB are complete and accurate.

### **2. SCOPE**

- 2.1. This protocol is applicable to the data verification stage of sample analysis. It details the steps necessary to review the data before generating a complete set of data files.
- 2.2. Personnel responsible for the procedures in this protocol include data entry personnel, sample analysts and the Phycology Section Project Manager.

### **3. REFERENCES**

- 3.1. Cotter, P. 2002. Tabulator Installation and User's Guide. Version 3.51. ANSP, PCER.
- 3.2. PCER, ANSP. 2002. Analysis of Diatoms on Microscope Slides Prepared from USGS NAWQA Program Algae Samples. Protocol No. P-13-39.
- 3.3. PCER, ANSP. 2002. Log-In Procedures for USGS NAWQA Program Algal Samples. Protocol No. P-13-47.
- 3.4. PCER, ANSP. 2002. Subsampling Procedures for USGS NAWQA Program Periphyton Samples. Protocol No. P-13-48.
- 3.5. PCER, ANSP. 2002. Preparation of Diatom Slides Using NAPHRAX™ Mounting Medium. Procedure No. P-13-49.
- 3.6. PCER, ANSP. 2002. Analysis of Soft-Algae in USGS NAWQA Program Qualitative Multihabitat (QMH) Samples. Protocol No. P-13-51.
- 3.7. PCER, ANSP. 2002. Archiving Algal Samples, Diatom Slides and Images. Protocol No. P-13-56.

- 3.8. PCER, ANSP. 2002. Analysis of Soft-Algae and Enumeration of Total Number of Diatoms in USGS NAWQA Program Quantitative Targeted Habitat (RTH and DTH) Samples. Protocol No. P-13-63.
- 3.9. United States Geological Survey, National Water Quality Assessment Program. 1997. Procedures for Processing NAWQA Algal Samples. Draft Manuscript. February 1997.

#### 4. METHODS

- 4.1. Prior to file verification and report generation, obtain the three paper folders of documentation for each subproject (defined by NAWQA study unit, the year of collection, and whether the samples are periphyton or phytoplankton): “Sample Tracking and Subsampling”, “Diatom Analyses” and “Soft-Algae Analyses” (see Protocol No. P-13-58). These file folders are started by personnel involved in the sample log-in and preparation phase. They should be circulated with their corresponding samples and subsamples. All pertinent printouts, bench sheets or other documentation should be added to these folders as they are completed (see section 4.4. for details).
- 4.2. **Data Entry Review.** Confirm that the needed data have been created (i.e., analyses are complete) and have been entered into the database. The verification of data cannot start until all data are entered. Listed below are the NAWQAdat database tables containing necessary data, and salient points concerning the verification required:
  - 4.2.1. **Sample Identification.** Should be complete and verified by this point. For each sample, there must be one record with complete data for the fields: “Sample ID”, “Site Location ID”, “Client Sample ID”, “Subproject ID”, “Collection Date/Date1”, NAWQA Sample ID, “result\_code”, and “Sample Auto ID”.
  - 4.2.2. **Sample Volumes/Areas.** This table should have been completed during the log-in and subsampling procedures. In some cases, data entered at time of log-in were incomplete or inaccurate. This table is only necessary for quantitative data files and has one record for each sample. Check the following critical fields for completeness and accuracy (these fields cannot be zero or null) and run the queries “Data Entry Check: Field Volumes/Areas” and “Data Entry Check: ANS Sample Volume” in PHYCLGY to make sure all fields required for biovolume calculations are present. When data are complete, print out the results of this query and place them in the “Sample Tracking and Subsampling” folder.
    - 4.2.2.1. Area Sampled. Verify that these values were properly transferred to the Academy’s database.
    - 4.2.2.2. Client DCF Factor (usually = 1). This factor can vary from “1” if there was subsampling and dilution or concentration prior to receipt at the Academy. Since the Academy measures the sample volume, this factor remains “1” unless a portion of the sample was lost prior to receipt at the Academy.
    - 4.2.2.3. **Sample Volume (ANS).**
    - 4.2.2.4. Original Sample Volume. This critical volume relates the amount of sample to the area that was sampled. In most cases there was no subsampling and this value is the same as the “Sample Volume (ANS)” field (the assumption being that a laboratory measurement is more accurate than a field measurement). Where

there is an indication of subsampling with a subsequent dilution of concentration, or when loss of sample is noted, this field or the “Client DCF Factor” field must be adjusted.

- 4.2.3. **Subsample Information.** There will be two records for each sample if only one set of subsamples was made: a diatom (DT1) and a soft-algae (PR1) subsample. There will be more records if more than one subsample was made (e.g., DT2, PR2). For quantitative samples, verify the “Subsample Volume” and “Dilution/Concentration Factor” fields. If subsample volumes do not equal 20 ml, confirm that the volume was adjusted to 20 ml (by dilution or concentration).
- 4.2.4. **Slide Information.** There will be one record for each slide that was made. There should therefore be two records for each sample (unless a 3<sup>rd</sup> or 4<sup>th</sup> slide was produced). Verify the “Volume Cleaned Material,” “Dilution/Concentration Factor” and “Volume of Sample on Slide” fields, especially for quantitative samples. Similar to the “Subsample Information” table, check the “Dilution/Concentration Factor” field if the “Volume Cleaned Material” does not equal 20 ml.
- 4.2.5. **Palmer-Maloney Fractions.** For quantitative samples only, there are one to three records for each sample, depending on the number of dilution or concentration steps needed. It is important to verify that there is a record for the Palmer-Maloney fraction that was used in the analysis (in addition verifying that the fraction is from a periphyton subsample [designated with “PR” at the beginning of the subsample code]). Verify the “FractionVolume” and “Dilution/Concentration Factor” fields. It is important to verify that the amount of dilution or concentration relates to the original subsample, and not necessarily to the fraction from which it was diluted.
- 4.2.6. Verification is not necessary for several tables including **Diatom Hierarchy, Non-Diatom Hierarchy, Diatom Taxa** and **Non-Diatom Taxa**. The “...Taxa” tables are child tables of the “...Hierarchy” tables; the “...Count tables are child tables of the “...Taxa” tables. It is good practice to ensure that referential integrity between these tables and related tables is in force when the reports are generated.
- 4.2.7. Verify that **Diatom Count Info** is complete. If diatom analyses were performed using the “Tabulator” program, simply verify that the record for each sample is complete. For diatom analyses where bench sheets were used, verify data entry. Run the “Data Entry Check: Diatom Count Info” query in the PHYCLGY database to produce a table of database records. Print it and check it against data on the bench sheets. Similarly, verify the **Non Diatom Count Information** table by checking that data entry was correct. Run the query “Data Entry Check: Non Diatom Count Information” in the PHYCLGY database to produce a table of database records. Print it and check it against the bench sheets.
- 4.2.8. **Diatom Count.** For analyses using the “Tabulator” program, verify that “Diatom Count” table records have been properly saved with corresponding documentation (count reports signed and dated with all changes and corrections entered and documented). For “Diatom Count” records entered from bench sheets, verify that the data entry has been checked (the database query “Data Entry Check: Diatom Count” in the PHYCLGY database produces a table that is printed out and used to check against the bench sheets).

- 4.2.9. **Diatom Count Qualitative.** Verify qualitative records following a procedure similar to the one for verifying diatom count records (if bench sheets were used, the “Data Entry Check: Diatom Count Qualitative” query in the PHYCLGY database produces the table utilized in data checking).
- 4.2.10. Verify the **Non Diatom Count** and **Non Diatom Count Qualitative** tables by ensuring that the data entry from bench sheets has been checked (the queries “Data Entry Check: Non Diatom Count” and “Data Entry Check: Non Diatom Qualitative” in the PHYCLGY database produce tables to check against the bench sheets).
- 4.2.11. Check the **Microscope Lenses** table to ensure that the lens entered in the Diatom Count and Non-Diatom Count tables has the following data: lenses used in the non-diatom analyses (usually 40-50x magnification) should have values for Palmer-Maloney volume (in ml); lenses used for diatom analyses (usually 90-100x) and phytoplankton analyses (usually 15-50x) should have values for field diameter (in  $\mu\text{m}$ ). Additionally, all lenses must have the correct factor to convert from units to micrometers.
- 4.2.12. Verify the **Biovolume Measurements** table using the application “frmBiovolumeVerification” (a form in the PHYCLGY database). After the subproject is entered, check both the soft-algae and diatom components to see that the proper number of measurements were made. If fewer than five measurements are recorded for each abundant taxa, require the analyst to measure additional specimens.
- 4.3. **Final Preparatory Work for Data File Generation.** These are the final steps to take before moving on to Procedure No. P-13-55 to use the BioTDB Export application to generate data files.
  - 4.3.1. Download the most recent data from the BioTDB using the NAWQA ASR Download application in the NAWQAApp database (for more information on the NAWQA ASR Download application, see Procedure No. P-13-47). To do this, click the “Download” button on the first tab sheet, entitled “Download”.
  - 4.3.2. Determine if the data used in the biovolume calculation match the most recent data from the BioTDB by running the query “kms\_OriginalVolume\_0\_updates\_to\_subproject\_from\_BioTdb.” from the NAWQAApp database. This query compares key values from the BioTDB to those in the ANSP databases. If any values do not match, update them in the NAWQAApp database. Variables compared are Client Subsample Volume, Sample Volume (Client), Preservative Volume, After Decant Volume, and Area Sampled.
  - 4.3.3. Run the next set of queries from the NAWQAApp database. They check the “original volume” data in NADED against the most recent volume data from the BioTDB and update values as necessary.
    - 4.3.3.1. Run the “kms\_OriginalVolume\_1\_FlagAfterDecantVolume” query. This query returns all Sample Volumes/Areas records that have a value for “AfterDecantVol.” All records returned from this query are problematic and must be dealt with on a case-by-case basis to determine the correct values for “original volume” and “dilution/concentration factor.”

- 4.3.3.2. Run the “kms\_OriginalVolume\_2\_NotSubsampled” update query. Updates the Original Volume and Client DC Factor fields in the NADED database for samples that were not subsampled in the field.
- 4.3.3.3. Run the “kms\_OriginalVolume\_3\_Subsampled\_NotPhytoplankton” update query. Updates the Original Volume and Client DC Factor fields in the NADED database for non-phytoplankton samples that were subsampled in the field.
- 4.3.3.4. Run the “kms\_OriginalVolume\_4\_Phytoplankton” update query. Updates the Original Volume and Client DC Factor fields in the NADED database for phytoplankton samples that were subsampled in the field.
- 4.4. **Final Archived Files.** As noted earlier, there will be a set of three files (paper folders) used to archive the data from a subproject. The items to be included in each of these files are listed in Protocol No. P-13-58.
- 4.5. Once all data are verified and reported, reference specimens will be archived in the ANSP Diatom Herbarium. See protocol P-13-56 for detailed archiving procedures.