



Maine Volunteer River Monitoring Program (VRMP)

Appendix J

Data Management Overview (VRMP QAPP-Section 5.10)

Program-Wide SAP Appendix

From:

Maine DEP (Maine Department of Environmental Protection). 2009. Maine Volunteer River Monitoring Program [VRMP] - Quality Assurance Program Plan. Doc #: DEPLW-0984. Portland, ME. <http://www.maine.gov/dep/blwq/docstream/vrmp/publication.htm>



Section 5.10 – Data Management begins at the bottom of this page.

5.8 Inspection/Acceptance of Supplies and Consumables.

[B8. (Element 17)]

TOPICS

- Inspection/acceptance of supplies and consumables

Supplies and consumables are inspected upon receipt by the individual volunteer groups. In the event that these are found to be unacceptable, they will be returned to the supplier or manufacturer. Once accepted, if an item appears to be damaged or soiled, the item is not used but discarded or returned to the provider, as appropriate. Also, chemical expiration dates will be checked at least annually, and expired chemicals will be disposed of as directed on label instructions (Table 5a). Local laboratories or the VRMP staff may be able to assist volunteer groups with the disposal of these chemicals.

5.9 Non-Direct Measurements / Data Acquisition Requirements.

[B9. (Element 18)]

TOPICS

- Non-direct measurements / data acquisition requirements

The VRMP anticipates using geographic data derived from U.S.G.S. maps and State of Maine Office of GIS (geographic information system) coverages for tasks such as organizing volunteer groups and data by H.U.C watershed, data visualization, map-making tools, learning about local watershed land-uses, etc. The State of Maine GIS data layers typically have metadata (data used to describe other data) associated with them explaining any background information or limitations it may have. Any other types of non-direct-measurements or external-sources, and their limitations, will need to be explained in individual volunteer group's SAPs. VRMP staff knowledgeable in GIS and GPS technology will assist volunteers groups as time allows.

5.10 Data Management.

[B10. (Element 19)]

TOPICS

- Data collection / forms
- Data entry and submission
- Data review by volunteer group managers and VRMP staff
- Data availability via the Internet

Each group will need to include in their SAPs explanations of the details about their handling of datasheets and chain-of-custody/sample-submission forms as well as protocols for datasheet and data file proofing/checking. Examples of datasheets and chain-of-custody/sample-submission forms can be found in Appendix 10. At a minimum, volunteer groups belonging to the VRMP will be required to take some minimum steps to ensure proper datasheet handling and data entry. Those minimum requirements are presented later in this section.



An overview of assistance that will be provided by EGAD staff to the VRMP is listed in Table 5b. (EGAD stands for Maine DEP's Environmental and Geographic Analysis Database.)

Table 5b. EGAD Staff Support for the VRMP.

DEP-EGAD staff will be creating a process to assist VRMP staff with reviewing, validating, and uploading volunteer-group water quality data into the EGAD database.

Key data management and QA/QC support that will be provided to the VRMP by EGAD staff includes:

- i. Design a standardized water quality data import spreadsheet template for VRMP volunteer groups;
ii. Facilitate the importing of volunteer data, via "pre-EDD" (pre-Electronic Data Deliverable) spreadsheets, forwarded to and reviewed by VRMP staff;
iii. Assist VRMP staff with some aspects of electronic QA checks of data being uploaded to EGAD (e.g., data validation procedures utilizing computer database scripts programmed by EGAD staff); and
iv. Create a standardized format and scripts for output of VRMP river data (e.g., tables, graphs, selected analysis), which will reduce the effort from VRMP to complete annual reports for individual rivers and streams.

A. Forms for Data Collection and Chain of Custody (Lab Sample Submission)

VRMP volunteer water quality data will be entered onto standard field forms (which may be updated periodically, especially during initial years of the VRMP) (Appendix 5). (The VRMP has obtained input from some existing volunteer groups to gather feedback so that the VRMP datasheets will meet the needs of as many volunteer groups as possible. Alternatively, groups that insist using their own style of datasheets shall also fill out standardized VRMP forms in addition to any of their own datasheets.) These VRMP datasheet forms will include space for data elements that will be entered into the VRMP database; specific locations for QA/QC data collected by the volunteers; areas to record which parameters were sampled along with specific methods, equipment, and calibration types; and any other important notes or observations.

At volunteer training sessions (discussed elsewhere), volunteers will be encouraged to take the time to carefully enter and double-check data on the forms because this information could play a role in waterbody assessment and management actions. The Maine Lakes Program (MDEP and VLMP) has found that most individuals are more likely to double-check their form prior to submitting if their signature is required, thus VRMP forms will include a required signature line at the bottom. Volunteer groups will be queried by the VRMP if a form is not signed, and may possibly exclude data from inclusion into the VRMP database if a reasonable explanation for lack of signature is not provided.



Original datasheets, along with chain-of-custody or sample-submission forms for (approved) laboratory analyses (if applicable), will be completed by members of the individual volunteer groups and will eventually reside as archives within the group's project files that are maintained by the group's leader or designee. Additionally, copies of these forms will be sent to and maintained by the VRMP at VRMP headquarters.

B-1. Data Input and Review – Volunteer Group Phase

- Data will be input by the volunteer group's data management designee ("data manager") into a standardized, EGAD/VRMP-created water quality data import spreadsheet template called a "pre-EDD" (pre-Electronic Data Deliverable) spreadsheet. These simple spreadsheet templates will be in MS Excel format. The use of a template by VRMP groups will facilitate, speed up, and standardize the importing of volunteer data into EGAD. (*Note:* VRMP/EGAD may need to take special steps to create specially-tailored data import spreadsheets for a few of the already-existing volunteer groups who have been entering and storing data in a certain manner for many years. EGAD staff can create database scripts that can import data contained in these different style spreadsheets. It is, however, in the best interest of EGAD staff time to minimize the number of groups who receive this custom treatment. Existing groups will be encouraged to migrate towards the standardized spreadsheets by VRMP staff.)
- Prior to a volunteer group's submission of electronic water quality data spreadsheet to the VRMP (as described in item B-2), another member of the volunteer group (e.g., the group's leader or another member, but not the "data manager") will compare 100% of the group's original field data sheets to the data entered electronically into the "pre-EDD" datasheet template, to see if the records match. A review of field duplicate data and laboratory quality assurance information will also be conducted (this step can be done by the "data manager"). Any easily-fixed problems (e.g., "typos") will be corrected in the electronic files. More complicated problems will be documented in the "pre-EDD" datasheet template file in a location designated on the spreadsheet.
 - This volunteer group review of the data requires at least 2 people – the data manager and another person.
- Data identified as problematic or incorrect will either be corrected (based upon writings on the data sheet), deleted, or flagged as "problematic" on the data spreadsheet so that VRMP staff can delete those specific data (or take other corrective steps) before the spreadsheet is uploaded into EGAD.
- The volunteer group's data manager will then sign the Chain of Custody portion of the VRMP field data sheet (Appendix 5) indicating that the data (quality assurance) review has been completed before passing on the data spreadsheet to the VRMP. Volunteer groups will be encouraged to review and submit water quality data as a single "batch" (package) once per year to the VRMP to maximize efficiency. Data spreadsheets may either be emailed or mailed to the appropriate recipient.



- Clean, readable copies of the original datasheet hardcopies will be passed along to the VRMP (item B-2) to accompany the electronic files. These hardcopies will eventually be archived in a VRMP file cabinet at VRMP headquarters.

--- Continued on the next page ---



OPTIONAL – ITEM FOR FUTURE CONSIDERATION

B-1-opt. Data Review – Regional Data Coordinator or Volunteer Group “Data Swap” Phase

NOTE:

- *This step currently is considered optional under this version of the VRMP QAPP because it is expected in the early years of the VRMP that few, if any, people will volunteer to be regional data coordinators for the VRMP. Volunteer watershed groups and their members are expected to be the only volunteers participating in the VRMP during its early years.*
 - *This data review step (B-1b) may become implemented as the VRMP grows beyond 10-12 groups, in order to keep data review by VRMP staff (described in item B-2) manageable.*
 - *This data review step could potentially substitute for task B-1 for some volunteer groups in the future.*
-
- Prior to a volunteer group’s submission of electronic water quality data spreadsheet to the VRMP, either a (volunteer) regional data coordinator or another volunteer group’s data management designee will compare the group’s data sheets and electronic data files to the best of their ability, as documented in their SAP, to see if the records match. A review of field duplicate data and laboratory quality assurance information will also be conducted. The data reviewer will compare 100% of the group’s original data sheets to the data entered electronically into the “pre-EDD” datasheet template, to see if the records match.
 - Data identified as problematic or incorrect will either be corrected (based upon writings on the data sheet), deleted, or flagged as “problematic” on the data spreadsheet so that VRMP staff can delete those specific data (or take other corrective steps) before the spreadsheet is uploaded into EGAD. This reviewer will document any data deletions and why they occurred, and then pass that documentation on to the VRMP.
 - This data reviewer will then sign the Chain of Custody portion of the VRMP field data sheet (Appendix 5) indicating that the data review has been completed before submitting the data spreadsheet to the VRMP.
 - Copies of the original datasheet hardcopies will be passed along to the VRMP to accompany the electronic files. These hardcopies will be archived in a VRMP file cabinet at VRMP headquarters. Data spreadsheets may either be emailed or mailed to the appropriate VRMP staff recipient (generally the program director).



B-2. Data Review – DEP (VRMP, EGAD) Phase

[[Some of the procedures related to data validation are also mentioned in sections 7.1 and 7.2.]]

- After a water quality datasheet has gone through the data review processes outlined in items B-1, the VRMP will then review the data. VRMP staff will compare the group's data sheets and electronic data files to see if the records match. A review of field duplicate data and laboratory quality assurance information will also be conducted. The data reviewer will compare 100% of the group's original data sheets to the data entered electronically into the "pre-EDD" datasheet template, to see if the records match. (As the VRMP grows beyond 10-12 groups, it may begin using the steps outlined in item B-1b instead, and then perhaps scale back the proportion of hardcopy and electronic data records that it reviews.)
- A summary of techniques that VRMP staff will use to identify possible data problems, as well as protocols for resolving problematic ("flagged") data, are provided in Table 5c.
- Data identified as problematic or incorrect will either be corrected (based upon writings on the data sheet and after checking with the volunteer group data manager) or excluded from uploading into EGAD. Details about dealing with problematic data are provided in section 7.2.
 - The VRMP will document any data exclusions and why they occurred, and then save that documentation in an electronic file dedicated to archiving the VRMP data review processes each year.
- VRMP staff will then sign the Chain of Custody portion of the VRMP field data sheet (Appendix 5) (and archive it in VRMP files) indicating that the data review has been completed.
- Copies of the original datasheet hardcopies will be archived in a VRMP file cabinet at VRMP headquarters. The paper files will be ordered by Hydrological Unit Code (HUC) watershed, then alphabetically by volunteer group name, and finally, within each volunteer group folder, by year.

C. Data Review Tracking

- The Chain of Custody (COC) portion of the VRMP field data sheet (Appendix 5) will be the process used to track the various levels of data review discussed above. These COCs/field-sheets will be archived in a VRMP file cabinet.



Table 5c. Techniques VRMP staff will use to identify and deal with possibly problematic or “flagged” data problems. *See also section 7.2.*

- VRMP staff will exclude data records or data sets that fail to meet QA/QC criteria (e.g., precision, accuracy) for the various parameters outlined in Table 3a.
- Automated data validation techniques (described in sections 7.1 and 7.2) also will be employed by VRMP staff to identify problematic data and reduce the chance of human error. For example, EGAD staff will write scripts to identify data that fall outside of expected ranges for particular water quality parameters. Data that fall a great distance outside of expected ranges may have had data entered with incorrect placement of decimal points.
- Best professional judgment will be used to decide whether to exclude data from the dataset. Notes and local knowledge about conditions will be considered.
- VRMP staff will contact the volunteer groups to gather more background information and other details that might help resolve the status of questionable data. Field visits to potential problem sites may also occur.

D. Data Uploading Into EGAD

Once data has been subjected to the various review steps mentioned in section 5.10 (items B-1 through B-4) and data validation steps described in sections 7.1 and 7.2, it will be emailed to the appropriate EGAD staff person and then uploaded into EGAD.

E. Data Availability Via the Internet

The water quality data submitted by VRMP staff to EGAD are anticipated to be periodically uploaded to Internet-accessible databases (e.g., possibly DEP’s own website, USEPA’s WQX, University of Maine’s “PEARL”, USEPA-STORET, Bowdoin College’s Watershed Web) by DEP EGAD staff.