

Chapter 8 – Data Review

All data must be verified prior to its submittal to the SWQMIS database. An example checklist that may assist with data verification is provided [here](#).

Contractors submitting data through TCEQ water programs must use the procedures, checklists, and/or forms required by their contracts (for example, TMDL data review checklists, CRP data summaries, or NPS data review checklists) to document data review.

SWQMIS Data Loading Report

The DM&A Team assists in data verification and validation prior to loading into SWQMIS. The SWQMIS data loader is a tool that performs checks of data flat files. The SWQMIS data loader in the Test environment of SWQMIS can be used by contract data submitters to test their data deliverables prior to submitting the data to their TCEQ project manager. This data loader tool helps the data provider confirm that the dataset is correct in format and complete in content. The data checks also ensure that DM&A can upload the data to the Production environment of SWQMIS without errors. The loading tool also produces a summary report used for further diagnostics of any errors. Project managers can use this report for secondary data review.

Data Dictionary for the SWQMIS Data Loading Validator Report

Submitting and Collecting Entities and Monitoring Types

This section of the report identifies all the combinations of Submitting Entity, Collecting Entity, and Monitoring Type Codes reported in the dataset, and descriptions of the codes are provided.

Frequency of Parameter Occurrence

This section of the report identifies the parameter codes, parameter descriptions, and the number of times the parameter appears in the dataset. Also includes the minimum “less-than” value, maximum “less-than” value, minimum “greater-than” value, maximum “greater-than” value, minimum quantifiable value, maximum quantifiable value, and mean values from the dataset.

Stations in Dataset

This section of the report identifies sampled station IDs. Descriptions are provided for each station, along with the Basin ID and number of sampling events for each station.

Outliers (Requires verification prior to loading.)

This section of the report identifies the tag IDs, station descriptions, end dates, parameter codes, less than/greater than symbols, and values reported in the dataset that fall outside the predefined screening levels. The screening levels are listed in this section of the report as the minimum and maximum. If the minimum and/or maximum screening values need to be revised, complete a Parameter Code Request and submit to DM&A according to the process outlined in Chapter 2 in the SWQM DMRG.

Historical Basin Comparison

This section of the report provides each measurement that does not fall between the historical minimum and maximum value for a parameter in a basin. Dataset values outside the historical data levels for the basin-parameter code combination along with Tag ID, basin ID, station ID, parameter code, less-than/greater-than symbol, and the reported value are retrieved from the provided dataset. Historical minimum value, historical maximum value, historical mean value, and historical number of samples reported for the basin-parameter code combination are calculated using the most recent data (5-year period) currently in SWQMIS in that basin for that parameter.

Historical Station Comparison

This section of the report provides each measurement that does not fall between the historical minimum and maximum value for a parameter at that station. Tag ID, station ID, station description, parameter code, less-than/greater-than symbol, and the reported value are retrieved from the provided dataset. Historical minimum value, historical maximum value, historical mean value, and historical number of samples reported for the station-parameter code combination are calculated using the most recent data (5-year period) currently in SWQMIS at that station for that parameter.

Highest Values per Parameter

This section of the report provides the top ten highest values for each parameter code within the data set. Reported fields include station ID, station description, end date, end time, parameter code, less-than/greater-than symbols, value, and end depth.

Lowest Values per Parameter

This section of the report provides the top ten lowest values for each parameter code within the data set. Reported fields include station ID, station description, end date, end time, parameter code, less-than/greater-than symbols, value, and end depth.

Data Management Review

In addition to the verification checks automatically performed by the SWQMIS data loading tool, TCEQ data managers also perform verification and validation checks using output from the Data Loading Validator Report. Using the report as a guide, data managers compare the quality assurance (QA) document associated with the data load (QAP, QAPP) to the report output. The data manager verifies that the data are intended to be stored in SWQMIS and that the proper signatures appear on the QA document. The data manager then verifies that the sampling dates coincide with the effective date of the QA document. The use of the correct Tag Prefix is verified, as well as the use of proper Submitting Entity, Collecting Entity, and Monitoring Type codes. The data managers also verify all station IDs and parameter codes in the Data Loading Validator Report against stations and parameters described in the project QA document. Finally, the data managers ensure that the data submitter has verified all outliers in the data set. In the event that the data managers find discrepancies between the data set and the quality assurance document, the data managers will contact the TCEQ Project Manager for resolution.

Data Not Adhering to QA Documents

If submitted data is identified as not meeting the project's data quality objectives as stated in the project QAPP, the data manager will return the data set to the TCEQ Project Manager. The data manager will provide a Water Quality Planning Division (WQPD) Data Resubmittal Form to the project manager. It is the project manager's responsibility to complete the form detailing any excursions of the report from the QA document. The project manager will work with the TCEQ QA Specialist to decide on a course of action that addresses the excursion. The project manager will also submit to data management any relevant documentation detailing the excursion from the QA document; often this documentation is recorded directly on the Data Resubmittal Form. When the agreed-to course of action is complete, the project manager, QA specialist, and data manager all sign the form. The project manager resubmits the data set, the Data Resubmittal Form, and any necessary documentation to the data manager. Once the data has been successfully loaded into SWQMIS, the Data Resubmittal Form will also be stored in the database to accompany the project's QA document(s). The Data Resubmittal process may also be initiated by the project manager upon preliminary review of submitted data.