
Quality control of plankton data (MLI)

Phytoplankton and Zooplankton

GENERAL INFORMATION

Plankton include zooplankton (animals) and phytoplankton (plants). Zooplankton are generally sampled using a net that filters a volume of water while being towed vertically, horizontally, or obliquely through the water column. Organisms bigger than the net's mesh size are collected and preserved. Phytoplankton are generally sampled directly from a bucket or from a Niskin bottle; a volume of water is collected and preserved.

Plankton data are primarily counts, dry weights, or wet weights. Other data may include individual measurements or length frequency classes.

Plankton data are inherently difficult to quality control with automated tests: there may be very large counts of organisms or very low counts, depending on the natural processes involved. Therefore, only a few automated tests can be applied to sampling positions and times.

Some manual controls of metadata and data are done along with the automated tests.

AUTOMATED QUALITY CONTROL TESTS OF METADATA

The validity of the position and time metadata is checked by applying the Stage 1 tests defined in the *GTSP real-time quality control manual*¹; these tests are performed in the Matlab environment. The same tests are applied to CTD and bottle metadata.

The following is a list of the tests that are currently available:

¹Stage 1: Quality control of position and time metadata.

TEST	DESCRIPTION
Test 1.1	GTSP Platform Identification
Test 1.2	GTSP Impossible Date/Time
Test 1.3	GTSP Impossible Location
Test 1.4	GTSP Position on Land
Test 1.5	GTSP Impossible Speed
Test 1.6	GTSP Impossible Sounding

¹UNESCO, 1990. *GTSP real-time quality control manual. Intergovernmental Oceanographic Commission, Manuals and Guides no. 22.*

Test 1.1: Platform identification

This test verifies that all the mission samples were collected from the same ship.

Test 1.2: Impossible date/time

This test verifies that the date and time of the beginning and end of the sampling fall within the mission dates.

Test 1.3: Impossible location

This test verifies that the sample's position is possible; that is, that the latitude falls between -90 and 90 and the longitude between -180 and 180 .

Test 1.4: Position on land

This test checks whether the sample's position falls within the land polygons describing the area of the St. Lawrence Gulf and estuary. The area covered is from -70 to -56 in longitude and from 45 to 52 in latitude.

Test 1.5: Impossible speed

This test checks the ship's speed between two consecutive samples. Ship speed is calculated from the time-space position at the beginning of the sample and those from the end of the preceding sample. If the end position or date/time of the preceding sample is missing, the test uses the coordinates at the beginning of the preceding sample to determine ship speed. The calculated speed is compared with the ship's cruising speed.

Test 1.6: Impossible sounding

Test 1.6 compares the sounded depth to a 3 km grid bathymetric map of the estuary and Gulf of St. Lawrence to determine its validity. A depth is considered valid if it is within 20 m of the depth noted on the bathymetric grid. The area covered is from -70 to -56 in longitude and from 45 to 52 in latitude.

AUTOMATED QUALITY CONTROL TESTS FOR DATA

Data duplication is systemically checked. Uniqueness is determined by combining the taxonomic name, sex, development stage, and the size of an organism or group of organisms belonging to the same sample.

MANUAL QUALITY CONTROL TESTS FOR METADATA AND DATA

The data manager begins by recording how the sample was handled—from sampling to analysis—to make sure that the submitted data are matched with the correct metadata. While sample handling may be fairly standard in some situations—a monitoring program, for example—it may be very complex in some research projects.

A plankton data set often includes more metadata than data; metadata include details on sampling, preservation, and analysis. There are no systematic quality control tests for these types of metadata and data, so checks are done by validating and looking at data within a spreadsheet.

Sampling metadata:

- ✓ Spatial-temporal information and sounding
- ✓ Sampling gear description and dimensions
- ✓ Sampled depth*
- ✓ Sampling method description
- ✓ Volume of water filtered or sampled*

* For current datasets, sample depths and filtered volumes are verified by comparing different methods of calculation as well as data resulting from other sensors deployed at the same time.

Preservation metadata:

- ✓ Sample splitting
- ✓ Sample preservation

Analysis metadata:

- ✓ Sample splitting and/or dilution prior to analysis*
- ✓ Plankton identification
- ✓ Sex and developmental stage of zooplankton

* If the original data compilation sheet is available, split and dilution information are verified. Otherwise, these are checked according to the species. For example, large organisms would not be identified in a highly diluted subsample while smaller organisms could be.

Most common data types:

- ✓ Plankton counts*
- ✓ Presence/Absence
- ✓ Wet and dry weight
- ✓ Length: total or body component

*Some checks are done on the identified organism itself. Developmental stages, such as copepodites, nauplii, and larvae, are validated with respect to the species' life history. Taxonomic names are checked against a standard species list.

Extreme values of all quantitative data are verified.