Quality control of STD data (MLI)

GENERAL INFORMATION

The quality control of STD data is based on the algorithms found in the UNESCO manual:

UNESCO (1990) GTSPP Real-Time Quality Control Manual, Intergovernmental Oceanographic Commission, Manuals and Guides 22, SC/90/WS-74, 121 pp.

The thresholds of certain GTSPP tests have been adapted to conditions found in the estuary and Gulf of St. Lawrence. Several additional tests were added to improve the overall quality control of the STD profile.

The current quality control procedure validates the pressure (or depth), temperature, salinity, and sigma-T data as well as the principal metadata concerning the time–space coordinates of the CTD profile. Eventually other variables, such as dissolved oxygen, fluorescence, and light transmission, could also be included.

The quality control procedure is divided into five steps:

Step 1: Tests validating the important metadata such as the time and position
Step 2: Tests comparing data values within a profile
Step 3: Comparison of the profile to a climatology
Step 4: Comparison of the profile to other profiles from the same mission
Step 5: Visual inspection of the cruise track and of the profiles themselves

All metadata can be modified, in particular the time–space coordinates, without making the profile unusable.

No data are modified by the quality control procedure. During the step 2 tests, quality flags are added (see the next section) to qualify the data as good, doubtful, erroneous, or missing. If the data must be modified for some reason, these modifications are made outside the quality control procedure and the quality flags must be adjusted in consequence.

DESCRIPTION OF THE INDIVIDUAL QUALITY FLAGS

The tests performed during step 2 add quality flags to the pressure, depth, temperature, salinity, and sigma-T data. The quality flag is a whole number between 0 and 9. The meanings of the quality flags are given in the following table:

Flag	Meaning
0	no quality control
1	value seems correct
2	value appears inconsistent with other values
3	value seems doubtful
4	value seems erroneous
5	value was modified
6 to 8	reserved for future use
9	value missing

The quality flags of higher value take precedence over those with lower value; e.g., a QC flag of 0 has a lower priority than a flag of 9. As such, if a test judged a data value as doubtful (flag 3) and the following test judged it as erroneous (flag 4), the quality flag 4 would be retained.

The following table shows how the quality flags are set for the dependent data channels. The dependent channels are those calculated from data measured by the CTD. The CTD measures temperature (TE90), conductivity, and pressure (PRES); salinity (PSAL), density (sigma-T [SIGT]), and depth (DEPH) are calculated. Salinity is calculated using measurements of conductivity, temperature, and pressure; density from salinity, temperature, and pressure; and depth from pressure and latitude.

PRES	DEPH	TE90	PSAL	SIGT
[4]	4	-	4	4
[3]	3	-	3	3
-	-	[4]	4	4
-	-	[3]	3	3
-	-	-	[4]	4
-	-	-	[3]	3
-	-	-	-	[3]

[x]: QC flag set as the result of a test

For example, if a temperature recording is judged erroneous by a quality control test (quality flag 4), then the QC flag 4 will be added to the right of that temperature measurement. A QC flag of 4 will also be added to the right of the value for salinity and sigma-T on the same line.

GLOBAL QCFF FLAG

The QCFF flag allows one to determine which test the quality flag results from. It only applies to the step 2 quality control tests. Each test in this step is associated with a number 2^x, where x is a whole positive number. Before running the quality control, a QCFF value of 0 is attributed to each line of data. When a test fails, the value of 2^x that is associated with that test is added to the QCFF. In this way one can easily identify which tests failed by analyzing the QCFF value. If the QC flag of a record is modified by hand, a value of 1 is added to the QCFF.

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

TEST	DESCRIPTION (QCFF)
Test 1.1	GTSPP Platform Identification
Test 1.2	GTSPP Impossible Date/Time
Test 1.3	GTSPP Impossible Location
Test 1.4	GTSPP Position on Land
Test 1.5	GTSPP Impossible Speed
Test 1.6	GTSPP Impossible Sounding
Test 2.0	IML Minimum Descent Rate (2)
Test 2.1	GTSPP Global Impossible Parameter Values (4)
Test 2.2	GTSPP Regional Impossible Parameter Values (8)
Test 2.3	GTSPP Increasing Depth (16)
Test 2.4	GTSPP Profile Envelope (Temperature and Salinity) (32)
Test 2.6	GTSPP Freezing Point (128)',
Test 2.7	GTSPP Spike in Temperature and Salinity (one point) (256)
Test 2.8	GTSPP Top and Bottom Spike in Temperature and Salinity (512)

Test 2.9	GTSPP Gradient in Temperature and Salinity (1024)
Test 2.10	GTSPP Density Inversion (point to point) (2048)
Test 2.11	IML Spike in Pressure, Temperature and Salinity (one point or more) (4096)
Test 2.12	IML Density Inversion (overall profile) (8192)
Test 3.5	IML Petrie Monthly Climatology (Temperature, Salinity and Sigma-T)
Test 4.1	GTSPP Profile Consistency
Test 4.2	IML Annual Deep Water Profile Consistency
Test 5.1	GTSPP Cruise Track Visual Inspection
Test 5.2	GTSPP Profile Visual Inspection

If the salinity of one record fails the test 2.7 then the QCFF value would be 256. If the test 2.0 had already failed, the QCFF value would then be 2+256=258.

DESCRIPTION OF THE QUALITY CONTROL TESTS

Test 1.1: Platform Identification

This test verifies that all the mission profiles were sampled 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 profile fall within the mission dates.

Test 1.3: Impossible Location

This test verifies that the profile'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 profile'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 speed between two consecutive profiles. The ship speed is calculated from the time–space position at the beginning of the profile and those from the end of the preceding profile. If the end position or date/time of the preceding profile are missing, the test uses the coordinates at the beginning of the preceding profile 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 St. Lawrence Gulf and estuary 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.

Test 2.0: Minimum Descent Rate

Test 2.0 is only executed if the DPDT variable (descent rate) was saved or if it is possible to calculate it. If the descent rate is below the minimum indicated in the processing history (usually 0.1 m/s), a flag of 3 is added to the salinity record.

Test 2.1: Global Impossible Parameter Values

Test 2.1 checks if the pressure, temperature, and salinity data are globally possible based on the criteria in the table below. If a data value is judged impossible and thus erroneous, its QC flag is replaced by 4.

Code	Variable	Units	Minimum value	Maximum value
BATH	Sounding	m	0	10000
TEMP	Water temperature	^{o}C	-2.5	35
PSAL	Salinity	(psu)	0	50
PRES	Pressure	db	0	10000

Test 2.2: Regional Impossible Parameter Values

Test 2.2 checks if the pressure, temperature, and salinity are regionally possible based on the criteria in the table below. If a data value is judged impossible and thus erroneous, its QC flag is replaced by 4.

Code	Variable	Units	Minimum value	Maximum value
BATH	Sounding	m	0	600
TEMP	Water temperature	^{o}C	-2.5	35
PSAL	Salinity	(psu)	0	35
PRES	Pressure	db	0	600

The region is defined by the following coordinates (longitude, latitude):
(-56.0, 52.0), (-73.0, 49.5), (-73.0, 46.0), (-64.5, 46.0), (-62.3, 45.2), (-56.0, 48.2), (-56.0, 52.0)

Test 2.3: Increasing Depth

A given pressure is considered doubtful when it is shallower than the preceding record and its QC flag is set to 3. Erroneous or missing pressure values are not considered.

Test 2.4: Profile Envelope (Temperature and Salinity)

Test 2.4 checks whether the temperature and salinity data are within the envelope of permitted limits by depth range (see the table below). The data value is judged doubtful if it does not fall within the permitted interval and its QC flag is set to 3. Erroneous or missing temperature or salinity values are not considered.

Code	Depth interval (m)	Units	Minimum value	Maximum value
TEMP	0 – 50	^{o}C	-2.5	35
TEMP	50 – 100	^{o}C	-2.5	30
TEMP	100 – 400	^{o}C	-2.5	28
TEMP	400 – 1100	^{o}C	-2	28
PSAL	0 – 50	(psu)	0	35
PSAL	50 – 100	(psu)	1	35
PSAL	100 – 400	(psu)	3	35
PSAL	400 – 1100	(psu)	10	35

Test 2.6: Freezing point

The freezing point is calculated from the salinity and the pressure. A temperature value lower than the corresponding freezing point is judged erroneous and its flag is set to 4. Temperature data previously judged erroneous or missing are not considered.

Test 2.7: Spike in Temperature and Salinity (not including the first and last records)

A spike is detected by comparing a data value (V2) with the previous (V1) and next (V3) values. If (| (V2 - (V3+V1))/2 | - | (V1-V3) /2 |) is greater than the threshold value, then V2 fails the test. The threshold values for temperature and salinity are noted in the table below. The data that fail this test are doubtful and their QC flags are set to 3. The temperature and salinity data previously judged erroneous or missing are not considered.

Code	Units	Threshold value
PRES	m	5
TEMP	^{o}C	2
PSAL	(psu)	0.3

Test 2.8: Top and Bottom Spike in Temperature and Salinity

A spike at the surface is detected by comparing the value of the tested data point (V1) to the value of the next point (V2) so that if | (V1 - V2 | is less than the threshold value no spike is detected. A spike at the bottom is detected by comparing the value of the tested data point (V2) with the value of the preceding data point (V1) so that if | (V2 – V1) | is less that the threshold value no spike is detected. If a spike is identified, then the QC flag 3 is assigned. Temperature or salinity records previously judged erroneous or missing are not considered.

Code	Units	Threshold value
PRES	m	25
TEMP	^{o}C	10
PSAL	(psu)	5

Test 2.9: Gradient in Temperature and Salinity

Vertical gradients of temperature and salinity are calculated to determine if they exceed the limits specified in the table below. If | V2 - (V1 + V3) / 2 | / (change in pressure) is greater than the gradient limit then V2 fails the test. V1, V2, and V3 are three successive values of temperature or salinity. A QC flag of 3 is then assigned to V2. Temperature or salinity records previously judged doubtful, erroneous, or missing are not considered.

Code	Units	Gradient limit
TEMP	^{o}C /m	10
PSAL	(psu)/m	5

Test 2.10: Density Inversion (1 point)

This test compares the density of one data point (V2) with that of the preceding data point (V1).
If V2-V1 is less than -0.05, the test fails and the density is considered doubtful (QC flag 3). Density values previously judged doubtful, erroneous, or missing are not considered.

Test 2.11: Abnormal spikes (several points)

Test 2.11 detects abnormal spikes in pressure, temperature, and salinity from several consecutive values by comparing the difference between a profile and the same profile that has had a median filter applied to it. The filter's window is determined by an algorithm that increases the filter's window until a profile is obtained where the density inversions are less than 0.01 kg/m^{3}. The window obtained in this way is used to filter the pressure, temperature, and salinity data. When the difference between the measurement of the filtered and

the non-filtered variable exceeds the value indicated in the table below, the data is judged doubtful (QC flag 3). Pressure, temperature, or salinity data that were previously judged doubtful, erroneous, or missing are not considered.

Code	Units	Allowed difference[1]
PRES	m	0.2
TEMP	^{o}C	0.2
PSAL	(psu)	0.015

[1] Allowed differences are sometimes modified according to the sampling interval

Test 2.12: Density inversion (several points)

Test 2.12 detects decreases in density of 0.05 kg/m^3 per decibar. If the test detects an inversion, the density identified as erroneous is replaced by the preceding value and the test is run again until there is no longer an inversion. Pressure and sigma-T data already judged doubtful, erroneous, or missing are not considered. It is preferable to run the test 2.11 first since a positive density spike associated with a doubtful increase in salinity, for example, would then not be considered in this test. The QC flag 3 is assigned to density inversions.

Test 3.5: Petrie's monthly climatology

This climatology was compiled by Petrie et al. (1996) for the Gulf of St. Lawrence. The average and standard deviations of temperatures, salinities, and densities at fixed depths from 21 regions of the gulf were calculated for each month. This test uses the climatology to determine the validity of observations from a mission. If the difference between the observations and the climatology exceeds three standard deviations, then a warning is given. It is then the user's responsibility to determine whether he wants to reject a set of observations or add quality indicators to some observations. The problem with this test is that a data point cannot be rejected simply because it fails the test; it is possible that the profile reflects a particular event, but it is also possible that the instrument was not functioning properly and the data are incorrect. A warning is given if the difference between two profiles is judged as too high. No QC flag is modified as a result of this test.

Petrie, B., K. Drinkwater, A. Sandström, R. Pettipas, D. Gregory, D. Gilbert and P. Sekhon (1996) Temperature, salinity and sigmat-t atlas for the Gulf of St. Lawrence, Can. Tech. Rep. Hydrogr. Ocean Sci., 178: v+256 pp.

Test 4.1: Comparison among profiles from the same mission

Test 4.1 compares the temperature and salinity values from profiles made within a certain spatial and temporal distance. The differences in temperature and salinity from two profiles within 20 km of each other and within an interval of two days should be below the limits presented in the following table for data below 200 m of depth. A warning is given if the difference between two profiles is judged as too high. No QC flag is modified as a result of this test.

Code	Units	Tolerated deviation
TEMP	^{o}C	0.5
PSAL	(psu)	0.3

Test 4.2: Comparison of deep-water temperature and salinity within an interval of one year

Test 4.2 is based on the short- and medium-term (maximum one year) conservation of properties of deep water masses. A database of temperature and salinity is compiled from data 300 m that had been previously validated (with reversing thermometers and salinometer readings). The temperature and salinity

profiles from a mission are compared to the values in the database if they are within a radius of 10 km. The maximum deviations should not exceed those noted in the table below. A warning is given if the deviation between two profiles is considered too great and the average of the deviation is calculated to indicate the magnitude of the difference. No QC flag is modified as a result of this test.

Code	Units	Tolerated deviation
TEMP	^{o}C	0.5
PSAL	(psu)	0.3

Test 5.1: Visualization of the cruise track

This test plots the cruise track, allowing the identification of gross position errors.

Test 5.2: Visualization of profiles

This step is of the utmost importance. It allows one to see and compare the original STD profile with the one resulting from the quality control procedure (in which the data marked as doubtful, erroneous, or missing are not shown). The validity of the data set is determined at this step in the quality control.