Quality Assurance Quality Control

Overview

ONC has developed and fully implemented a comprehensive process-oriented quality assurance (QA) model in combination with a product-oriented data quality control (QC) model. The QAQC data model systematically intercepts and examines the instruments and the data stream at various stages of the instrument function by different qualified teams to ensure long-term data quality control and assurance.

Prior to deployment, instruments undergo physical and digital testing by ONC's Physical and Digital Infrastructure teams. Appropriate metadata is configured by the Data Stewardship Team and data quality is assessed in the context of the testing environment by the Data Analytics and Quality Team. Healthy instruments are then deployed and go through a post-deployment commissioning phase where data quality in the context of the deployment site is assessed.

Automatic real-time tests determine the initial validity of data during parsing and ingestion into the ONC database. Qualifying the data in real-time ensures that every reading with a QAQC test immediately has an associated **QAQC descriptor or flag**. Additional data quality is achieved through regular instrument monitoring and manual QAQC that is performed on a regular basis by Subject Matter Experts with specific expertise in instrument types and/or sites. ONC scalar data quality control model conveys information about the data quality of individual data by integrating the results from multiple types of automatic tests and manual review. Quality descriptors associated with manual review override automatic quality tests. The overall quality descriptor of the data is given an integer indicator, which are standardized across all ONC data and are based on the Argo quality control flagging system (How to use quality control flags and errors), as well as including some ONC-defined flags. This data quality flagging system is incorporated in data visualization tools and data products available for download. Descriptive annotations are added to further supplement data by providing users information about instrument issues, complex data quality issues, and instrument lifecycle phases.

For further reading and a citable reference, see our Oceans 3.0 paper here: https://www.frontiersin.org/articles/10.3389/fmars.2022.806452/full.

Quality Assurance Quality Control (QAQC) Flags

ONC QAQC Flag	Summary	Detail
0	No quality control on data	No test - not checked
1	Data passed all tests	Passed all automated tests and/or passed manual checks
2	Data probably good	Failed at least one minor test (study area or station range test) and/or flagged by manual checks
3	Data probably bad	Failed at least one specialized test (Temperature-Conductivity test) and/or flagged by manual checks
4	Data bad	Failed at least one major test (instrument, region or global/observatory test) and/or flagged as failed by manual checks
6	Data bad	Insufficient valid data for reliable down-sampling (<70% of expected) (data completeness test) (Not used in ODV spreadsheet txt files, 3 is used instead)
7	Averaged value	Calculated from clean data with flags 0,1,2 (Not used in ODV spreadsheet txt files, worst flag is carried through)
8	Interpolated value	Calculated from clean data with flags 0,1,2
9	Missing data	Raw data gap filled with Not-A-Number (NaN) values

Data Product Filtering

Clean Data: As described in the quality control option data product option, the clean option will filter out poor quality data and will optionally replace it with null values if the fill data gaps option is selected. (In most data products, null values are represented by 'NaN', which stands for *Not-A-Number*.) Poor quality data is qualified as quality control flags 3 and 4 as defined above.

Raw Data: No filtering or modification is done to the data. All the quality control flags are reported.

Development of Quality Assurance Quality Control

One of the problems facing a real-time oceanographic observatory is the ability to provide a fast and accurate assessment of the data quality assurance and quality control (QAQC). Ocean Networks Canada is in the process of implementing measures of real-time quality control on incoming scalar data that meet the guidelines of the Quality Assurance of Real Time Oceanographic Data (QARTOD) group. QARTOD is a US organization tasked with identifying issues involved with incoming real-time data from the U.S Integrated Ocean Observing System (IOOS). A large portion of their agenda is to create guidelines for how the quality of real-time data is to be determined and reported to the scientific community. ONC is striving to adhere to the QARTOD's 'Seven Laws of Data Management' to provide trusted data to the scientific community.

Note that engineering data (e.g. instrument voltages, groundfaults etc.) are excluded from quality controls.

QARTOD's Seven Laws of Data Management:

1. Every real-time observation distributed to the ocean community must be accompanied by a quality descriptor.
2. All observations should be subject to some level of automated real-time quality test.
3. Quality flags and quality test descriptions must be sufficiently described in the accompanying metadata.
4. Observers should independently verify or calibrate a sensor before deployment.
5. Observers should describe their method / calibration in the real-time metadata.
6. Observers should quantify the level of calibration accuracy and the associated expected error bounds.
7. Manual checks on the automated procedures, the real-time data collected and the status of the observing system must be provided by the observer on a timescale appropriate to ensure the integrity of the observing system.

QAQC Implementation

The quality control testing is split into 3 separate categories. The first category is in real-time and tests the data before the data are parsed into the database. The second category is delayed-mode testing where archived data are subject to testing after a certain period of time. The third category is manual quality control by an ONC data expert.

Real-time data quality testing at ONC includes tests designed to catch instrument failures and major spikes or data dropouts before the data is made available to the user. Real-time quality tests include meeting instrument manufacturer's standards and overall observatory/site ranges determined from previous year's data. We have also designed dual-sensor tests for CTDs to catch conductivity cell plugs which cause a sudden drop in conductivity.

Delayed-mode tests include spike, gradient and stuck value tests. Spike and gradient tests are delayed by 2 sample periods, while stuck value results normally require 3 hours of delay. However, if the data comes in out of order, reprocessing may take up to one day to complete.

All resampled data is checked for completeness at the time of its calculation from raw data with the exception of engineering data, data from irregular or scheduled sampling (samples per sampling period of zero), data from externally-derived sensors. If a resample period does not have at least 70% of the expected data, it is flagged with a QAQC flag value of 6.

Manual quality control consists of a review of the data by a ONC data expert. This is done periodically; some primary data (CTDs in particular) are reviewed daily, others are reviewed in intervals up to every 6 months. (We are continuing to develop and improve our QAQC process and coverage). QAQC values manually entered by ONC data experts override all other tests (see below).

In addition to quality control testing, data may be *annotated*. Annotations are general purpose comments by ONC data experts, scientists and operations staff to note and explain times of service interruption or data issues. Annotations are compiled into the metadata reports that accompany all data products. Prior to or independent of requesting a data product, users may also search the annotation database directly using the Annotations search page, however, that method is not as effective as clicking on the links in step 2 of Data Search.

Test Categories

Major Test: A test that sets gross limits on the incoming data such as instrument manufacturer's specifications or climatological values. If failed, we recommend that the flagged data not be used.

Minor Test: A test that sets local limits on the incoming data such as site-level values based on statistics and dual-sensor testing to catch conductivity cell plugs. If failed, the data are considered suspect and require further attention by the user to decide whether or not to include this data in their analysis.

How does ONC determine the final quality control flag?

All data are passed through different levels of testing to create a quality control vector containing the output for each test (i.e. multiple QC flags per datum). The levels of testing include:

- Major tests set gross limits on the incoming data such as instrument tests (based on the manufacturer's specifications), global, observatory and region tests (based on climatological values). If data fails a major test, it is an indication of a serious issue such as an instrument misbehaving or a conductivity-cell plug. Data failing the major tests are flagged as 4 and we recommend that the flagged data not be used.
- Data points that pass major tests are then evaluated by minor tests. Minor tests included study area, location and station tests, and other tests such as T-C curve test. Failure to pass those tests results in a flag or 2 (station tests) or 3 (T-C curve test) and could indicate an unusual event (different water mass) but not necessarily an instrument issue. Flagged data should be considered suspect and requires further attention by the user to decide whether to include this data in their analysis.

The overall output quality control flag is determined from the set of QC flags for each datum as follows:

- If passed all tests, the final output flag assigned is 1 (Data passed all tests).
- If there are real-time automatic test failures, the most severe flag is assigned.
- If there are manual quality control flags they override the automatic tests with the most severe manual flag applying (there should not be any cases of multiple manual flags).

Note that automatic QAQC flags can be inherited from other sensors. If a given sensor value is derived from other sensors, then it will inherit all QAQC flags from Quality Control Tests on those sensors. For example, if a Density sensor is derived from Conductivity, Pressure, and Temperature, then the QAQC flags from those sensors will all be applied to Density. In some cases inherited flags can be excluded if they are redundant. Flags from an inherited test will be integrated into the final flags according to the level of the inherited test.

How do you determine which tests have been applied to the data you downloaded?

The now deprecated FGDC Metadata Product used to contain a summary of the applied QAQC tests. The Sensor Maintenance pages do have this information and are available without logging in, for example: https://data.oceannetworks.ca/SensorListing?DeviceId=23355&SensorId=13663#applied_qaqc_tab To find these information find the Device Details page, click on the Sensor tab and then click on a sensorID. To find a Device Details page, click on the device links in Data Search, Plotting Utility, et. One could also edit the example URL with the deviceID and sensorIDs, if known (available in CSV data product headers for instance). A dedicated QAQC detail product with all the flags for each reading is also a possibility for a future improvement, contact us if you're interested. For now, the time series scalar MAT file data product does contain all of the contributing tests and individual test results, including ways of determining which flag contributed to the final flag.

Real-time Quality Control Tests

Instrument Level: Tests at this level are to determine whether the data meet manufacturer's range specifications for each sensor. Failure of this test level is considered major and is likely due to sensor failure or a loss of calibration. We are currently working on entering calibration and configuration to the database so that it can be returned in the metadata.

Station Level: Similar to Study Area tests the tests at this level put the data through more stringent limits based on the previous data from each station such as Saanich Inlet VENUS Instrument Platform or Folger Deep.

Region Level: Similar to Observatory level tests the tests at this level help to eliminate extreme values in water properties that are not associated with the overall region. Failure of this test is considered major and could be due to sensor drift, bio fouling etc. Minimum/maximum values are chosen based on years of data at various sites that are considered to be of good quality.

Observatory Level: This level was active for data prior to 2015-11-05, it is no longer being used, replace by Region level. Tests at this level are to eliminate extreme values in water properties that are not associated with the overall observatory region. Failure of this test is considered major and could be due to sensor drift, bio fouling etc. Minimum/maximum values are chosen based on years of data at various sites that are considered to be of good quality.

Study Area Level: This level was active for data prior to 2015-11-05, it is no longer being used, replaced by Station level. Testing at this level puts the data through more stringent limits based on the previous data from each study area / location, such as Barkley Canyon or Saanich Inlet.

Single-Sensor Testing

Range: Minimum/maximum values for this level of testing originate in the statistics of the previous years of data. The limits are set as +/- 3 standard deviations about the mean without considering seasonal effects. Failure of this test minor as it could mean a little seen water mass or short term bio-fouling such as a plugged conductivity cell. Further testing is provided to determine whether a failure is a plugged conductivity cell.

Dual-Sensor Testing

Temperature-Conductivity Testing: This specialized test is designed to catch a dropout in conductivity that are not necessarily outside the site level range given by the single-sensor testing. This is a dual-sensor test that uses both the temperature and conductivity sensors of a single device to determine whether the conductivity data are good. Dropouts in conductivity are very apparent in a T vs C plot and can be flagged relatively easily using a simple equation. Failure of this test is minor. All derived sensors (salinity, density and sigma-T) inherit the quality flag from this test.

Delayed-Mode Quality Control Tests

Delayed-mode tests involve the determination of flags for values using subsequent data points. Examples include:

Spike Testing: Designed to identify implausible spikes in scalar data. (Applies most specifically to CTD data). Requires 3 consecutive values, so test results are *slightly* delayed and may be reprocessed if the data come in out of order.

Gradient Testing: Designed to identify whether the difference between adjacent measurements is too steep. Requires 3 consecutive values, so test results are *slightly* delayed and may be reprocessed if the data come in out of order.

Stuck Value Testing: Flags data that remains identically the same after predetermined length of time. Results will be delayed by the tolerance time, may be reprocessed if the data come in out of order.

Data Product Testing

In the generation of data products, processing steps that modify the data need to propagate the QAQC test results on the raw data through to the final result.

Data Completeness Test: Also know as the "70% rule". All resampling regularizes the data into an array of timestamps, each one representing the centre of a resample time period. If the amount of data falling in that time period is insufficient to compute a reliable resampled value (averaging, min/max, interpolation), then the final QAQC flag will be modified. The following processes and flags can be assigned:

- averaging:
    - if the amount of 'good' data (flag 1 or 2) in the period is greater than or equal to 70% of the total for the given time period, the data are averaged and QAQC flag 7 is assigned
    - if the amount of 'good' data in the period is less than 70% and more than 0%, the min/max QAQC flag is assigned 6
    - if all the data in the period is 'bad' (flag of 3 or 4, and the clean option has been selected), the min/max QAQC flag is assigned 4
- min/max resampling:
    - if the amount of 'good' data in the period is greater than or equal to 70% of the total for the given time period, the QAQC flag of the min /max value is kept
    - if the amount of 'good' data in the period is less than 70% and more than 0%, the min/max QAQC flag is assigned 6
    - if all the data in the period is 'bad' (flag of 3 or 4, and the clean option has been selected), the min/max QAQC flag is assigned 4
- interpolation:

- when data are linearly interpolated for unavailable time periods, then the QAQC flag 8 is assigned (e.g. latitude and longitude data are returned in data products interpolated to the time stamps of the request device, as a positioning reference)

Manual Quality Control

All ONC data undergo some form of manual quality control to assure the user that an expert is regularly checking the data. If real-time or delayed tests do not pick up an entire episode of ‘bad’ data, manual quality control will. This overrides all tests.