

QARTOD III

Scripps Institution of Oceanography
La Jolla, CA

Quality Control, Quality Assurance,
and
Quality Flags

Mark Bushnell, NOAA/NOS/CO-OPS
November 2, 2005



QUALITY ASSURANCE

Processes employed *a priori* to support the generation of high quality data

- Sufficiently accurate, precise, reliable sensor w/ adequate resolution
- Calibration, calibration checks, and/or *in-situ* verification
- Proper deployment considerations (rugged, anti-foulant, corrosion, etc.)
- Robust data communications
- Maintenance intervals
- Creation of QC process
-



QUALITY CONTROL

Follow on steps supporting the delivery of high quality data

- Format, checksum
- Timely arrival
- Threshold checks (min, max, rate of change)
- Neighbor checks, climatology checks, model comparisons
- Signal/noise ratios
- Verification of user satisfaction
- Generation of data flags



QARTOD I – BREAKOUT GROUP FLAG STATEMENT

It is important that an **Aggregate Quality Level** be delivered with each real-time data Record collected. While many scientific researchers will want to use the test-specific quality information, more common users (the workshop group postulated that their number would be about 85% of the total anticipated users of IOOS-type systems) of the real-time data systems may only be interested in a general or aggregate quality level. The group of “casual” users would probably want a **very simple quality level of only a few choices**. The data provider would map the test-specific flags to the aggregate quality level, using product- and system-specific mapping which is adequately described in the metadata. The group recommended a 4-level aggregate quality level designation.



QARTOD I – Summary

The following minimum standards for QA/QC were agreed upon by the workshop participants for real-time ocean observations:

1. Every real-time observation that is distributed to the ocean community ***must*** be accompanied by a quality descriptor.
2. All observations should be subject to automated real-time quality tests.
3. The real-time quality is best described by an aggregate quality flag (a simple overall descriptor) and a detailed quality test record (indicative of the results of any individual quality tests applied) to suit both the common user and the real-time scientist. The Aggregate quality levels were recommended to be few (although the order is arbitrary):
 - **-9 = missing value**
 - **0 = quality not evaluated**
 - **1 = bad**
 - **2 = questionable/suspect**
 - **3 = good**
4. The quality flags and quality test descriptions ***must*** be sufficiently described in the accompanying metadata.

QARTOD I

Examples of such flagging schemes abound, including these sources:

CSIRO XBT QC flags : www.nodc.noaa.gov/GTSP/document/codetbls/gtsppcode.htm

WOCE QC flags: www.coaps.fsu.edu/RVSMDC/woce/qccodes_NetCDF.shtml

ARGO, IUGG also cited



QARTOD II – In Situ Currents breakout group

QUESTION 2: What categories of real-time quality descriptor flags should be applied?

The group agreed the categories of real-time quality descriptor flags should follow the QARTOD I recommendation:

- 9 = missing value;
- 0 = quality not evaluated;
- 1 = bad;
- 2 = questionable/suspect;
- 3 = good.



QARTOD II – Remote Currents breakout group

QUESTION 2: What categories of real-time quality descriptor flags should be applied?

The quality descriptor flags defined in QARTOD I are appropriate; however, an additional flag is required to identify interpolated (radial or total vector) data.

“Interpolation” refers to filling values within Doppler spectra or within a range cell (at the radial level). This data quality flag does not apply to interpolations of Level-3 data products. Rather it indicates an interpolation technique was used to create the radial or total vector since this information may influence the utility or application of the data.

-9 = missing value

0 = quality not evaluated

1 = bad

2 = questionable/suspect

3 = good

4 = interpolated data

QARTOD II – In Situ Waves breakout group

The group generally concurred with the descriptor flags suggested by QARTOD I, with one modification (shown in bold):

-9 = missing value

0 = quality not evaluated

1 = bad

2 = questionable/suspect

3 = **passed real time QA/QC** (*as opposed to the suggested flag of “good”*)

Data with a flag of “3” should not be considered final since it has only passed real time QA/QC. We may need an additional flag for final QA'd data to be included with metadata, although this additional flag would not be considered real time.

It should be noted that Datawell has its own set of descriptor flags, where 0 = good, 1 through 8 are different quality levels. The user can define a further set (e.g. 9+).

However, these are not incompatible with Q1 standards and can be mapped to those flags for standardization.

QARTOD III - PREPARATION

Global Temperature & Salinity Profile Program (GTSP) flags

Data Quality Flags

A detailed explanation of the data quality flags used in the UOT and GTSP is provided in the CSIRO Cookbook. Table 1 shown below is a shortened version.

Table 1. Data Quality Classes

Class	Quality	Description
Flag 0	No QC done	Flag 0 data are the level at which all data enter the working archive. They have not yet been quality controlled.
Flag 1	Good data	Flag 1 data are top quality data in which no malfunctions have been identified and all real features have been verified during the quality control process.
Flag 2	"Probably" good data	Flag 2 data are good data in which some features (probably real) are present but these are unconfirmed. Flag 2 data are also data in which minor malfunctions may be present but these errors are small and/or can be successfully corrected without seriously affecting the overall quality of the data.
Flag 3	"Probably" bad data	Flag 3 data are suspect data in which unusual, and probably erroneous features are observed.
Flag 4	Bad data	Flag 4 data are those in which obviously erroneous values are observed.
Flag 5	Changed	Flag 5 data have been altered by a QC Centre, with original values (before the change) preserved in the history record of the profile.

International Standard Flag Set?

The whole set of GTSPP flags?

Code	Meaning
0	No QC was performed
1	Good data
2	Probably good data
3	Bad data that are potentially correctable
4	Bad data
5	Value changed
6	Not used
7	Not used
8	Interpolated value
9	Missing value

Too many flags for real time operations?

NOAA/NOS/CO-OPS PORTS®

PORTS® UNIFORM FLAT FILE FORMAT (PUFFF)

Each file, besides the basic observed data, has several fields containing information regarding the quality of the data as determined on a real-time, single sample basis. These fields are the Data Quality Assurance (DQA) bit mask, Data Quality Class Code (DQCC), and Data Quality Action Codes (DQAC).

The DQA bit mask immediately follows the data fields on line 7. It consists of 32 digits, each digit either a zero or a one (0 or 1). Bit 0 (zero) is the first character, bit 31 is the last character. If the character is '1', this means it failed a particular test. The meaning of the bits varies according to the data type.

The DQCC and the DQAC follow the DQA bit mask. The DQCC is a three digit code. The first digit is '3' or '4'. If it is '3', there are no failure codes relating to real-time use of the data. If it is '4', the data failed the DQA in some way for real-time use. The next two digits are the number of DQAC's following the DQCC. The DQAC's are defined in Appendix IV.

Maximum influence

Death & taxes

Gravity & corrosion



QARTOD recommendations

Subtle suggestions

Noise

Hints & allegations

The Life Aquatic with Steve Zissou – Team Zissou

