What Do CBIBS QC Codes Mean?

CBIBS data are run through two sets of QARTOD (Quality Assurance of Real Time Oceanographic Data) tests. These quality control (QC) tests flag data that likely are not accurate representations of actual conditions at the buoy. The tests generate a code value that describes the data.

The first test is done on data in real time as it is collected. This test applies QC in real time for immediate QC values. This is the data that is stored in the data files available in the “Data Download” section of the CBIBS website (https://www.buoybay.noaa.gov/data/data-download). This has some drawbacks, because it is applied directly to the data stream. If there are any communication delays or other processing anomalies, this style of analysis may be incomplete.

A second, more thorough, post-processing algorithm is applied to the data that are exported to the National Centers for Environmental Information (NCEI) datastore. The NCEI datastore is uploaded yearly. A user can choose the datasource depending on their needs; CBIBS data are available here. Data available through the Data Download section at the CBIBS website are processed and uploaded daily.

The QC values for both processes are described in the table below. The individual test results are aggregated, and the highest value is applied to the data as a primary flag. For example, all other tests may pass, but if the location test fails with a “BAD” flag, that flag will be applied to the data. The post-processing algorithm includes flags performed by a manual review. This flags values that may be within acceptable limits for an automated system, but are known to be bad or suspect from a human review.

QC Code	Value	Description
1	GOOD Value	Data have passed critical real-time quality control tests and are deemed adequate for use as preliminary data.
2	Not Evaluated, Not Available, Unknown.	Data have not been QC-tested, or the information on quality is not available.
3	Questionable or Suspect.	Data are considered to be either suspect or of high interest to data providers and users. They are flagged suspect to draw further attention to them by operators.

QC Code	Value	Description
4	BAD	Data are considered to have failed one or more critical real-time QC checks. If they are disseminated at all, it should be readily apparent that they are not of acceptable quality.
9	MISSING	Data are missing; used as a placeholder.

These automated tests are run in JAVA for real-time data and Python for post-processing:

Automated Tests	Description
Rate of Change	This test checks if the data is changing faster than the standard deviation plus a factor.
Climatology	This test uses historical climate data to determine if data is within acceptable ranges
Gap	This test checks the expected number of measurements over a defined time interval.
Gross Range	This test aims to identify data that fall outside either the sensor measurement range or is a statistical outlier
Location	This test uses location data to determine if the data is within an acceptable range.
Flat Line	This test checks if the data is flat over three indices Flat can mean the same value or repeating values within a threshold
Syntax Test	This test checks for null values