Drifting buoys Data Management 2018

Drifting buoys quality control manual

Version 1.0

September 2018

DOI http://doi.org/10.13155/52040

1 Table of contents

2 Introduction

This document is the Drifter quality control manual. It is derived from Argo floats quality control on trajectories (http://dx.doi.org/10.13155/33951).

Changes from the previous version of the manual are highlighted in yellow.

The Drifter data system has three levels of quality control.

- The first level is the real-time (RT) system that performs a set of agreed automatic checks on all drifter measurements. Real-time data with assigned quality flags are available to users within a few minutes.
- The second level of quality control is the slightly delayed-mode (SDM) system. It includes human quality controls. SDM data are available within a few days.

- The third level of quality control is the delayed mode (DM) system. This can include regional scientific analyses of all drifter data with other available data (the procedures are still to be determined), or statistical analysis, or ….

This document contains the description of the Drifter real-time and SDM procedures.

3 Data sent in accordance with drifter type

The table below lists all the drifter types available at time of writing, and the parameters sent by these drifters.

Parameter	**SVP**	**SVP-B**	**SVP-BS**	**SVP-BSC**	**SVP-BTC**	**SVP-BRST**	**…**
Timestamp	x	x	x	X	x	X	
Latitude	x	x	x	X	x	X	
Longitude	x	x	x	X	x	X	
Air Pressure	x	x	x	X	x	X	
Sea surface temperature	x	x	x	X	x	X	
High resolution sea surface temperature from Reference Sensor for Temperature	-	-	-	-	-	x	
Salinity	-	-	x	x	-	-	
Conductivity	-	-	-	x	-	-	
Submergence	x	x	x	x	x	x	
Battery volatge	x	x	x	x	x	x	
Depth	-	-	-	-	x	-	
Temperature at depth	-	-	-	-	x	-	

4 Drifters quality controls Tests

4.1 Introduction

Because of the requirement for delivering data to users within a few minutes from measurement time, the quality control procedures on the real-time data are limited and automatic. These procedures also include the generation of alerts (so-called backlists) which are analysed in slightly delayed mode before taking action. The test limits are described here.

If data from a drifter fail these tests, those data will not be distributed on the GTS. However, the raw data, including those having failed the tests, should be forwarded to the Global Drifter Servers. The BUFR code form is used to send the drifter data on the GTS.

4.2 RTQC (Real Time Quality Control) tests

When an entry below indicates that an information "*cannot be missing*", it means that otherwise any corresponding data will not be distributed on the GDAC and GTS.

1. Platform identification

This information *cannot be missing*. Every centre handling drifter data and posting them to the GTS needs to prepare a metadata set for each drifter and in this WMO number establishes correspondence to each drifter telecom identifier (for example Iridium IMEI). There is no reason why, except because of a mistake, an unknown drifter ID should appear on the GDAC or GTS. The WMO number must have 7 digits and the first 2 digits must form one of the possible WMO block numbers, namely 11—17 , 21—26, 31—34, 41—48, 51—56, 61—66, or 71—74. The last three digits must be equal or greater than 500.

Action: If the correspondence between the drifter telecom identifier cannot be matched to the correct WMO number, if the WMO is missing or is not a drifting buoy number, none of the data should be distributed on the GTS.

2. Date and time

This information *cannot be missing*. In addition, a test requires that the observation date and time from the drifter be sensible and close to real-time.

- Date greater than 1980
- Observation date/time minus processing date/time between -4 days and +1 hour

Action: If any one of the conditions is failed, the date should be replaced by current timestamp of data reception minus emission delay Iridium (or Argos) if available, so the data are flagged accordingly and may be distributed on the GTS. If Iridium (or Argos) timestamp not available, the data flagged as bad data and none of the data should be distributed on the GTS.

3. Location

This information *cannot be missing*. In addition, the observation latitude and longitude from the drifter be sensible.

Action: If either latitude or longitude fails, the position should be replaced by Iridium (or Argos) position if available, so the data are flagged accordingly and may be distributed on the GTS. If Iridium

(or Argos) position not available, the data flagged as bad data and none of the data from the drifter should go out on the GTS.

- Latitude in range −90 to 90
- Longitude in range −180 to 180

4. Position on land test

The test requires that the observation latitude and longitude from the drifter be located in an ocean.

Use can be made of any file that allows an automatic test to see if data are located on land. We suggest use of at least the 5-minute bathymetry file that is generally available. This is commonly called ETOPO5 / TerrainBase and can be downloaded from http://www.ngdc.noaa.gov/mgg/global/global.html.

Action: If the data cannot be located in an ocean, all the data should be flagged dubious (flag 3).

5. Impossible speed test

Drift speeds for drifters can be generated given the positions and times of the drifters when they are at the surface and between profiles, taking into account the actual dates/times of the positions (which may differ from the dates/times of the geophysical parameters, when for example the GPS fails to make acquisition and an older position is used, which can happen up to 2-3 days). Even if there are strong regional variations in current speeds (high values in the Gulf Stream, for example) we would not expect the drift speed to exceed 3 m s^{-1}. If it does, it means either a position or time is bad data, or a drifter is mislabeled. Using the multiple positions that are normally available for a drifter while at the surface, it is often possible to isolate the one position or time that is in error.

Action: Flag the position, the time, or both at flag 3

6. Global physical range test

This test applies a gross filter on observed physical parameters: air pressure, water temperature, and water salinity(for the buoys which have this sensor). It needs to accommodate all of the expected extremes encountered in the oceans.

- Air pressure must be within the range 850 to 1055 hPa.
- Water temperature must be within the range range −2.5 to 40.0°C
- For ice buoys temperature range may be lower.
- Salinity must be within the range 2 to 41.0 PSU
- Air temperature must be winthin the range -60 to 42°C

Action: If a value fails, it should be flagged as bad data and only that value need to be removed from distribution on the GTS.

7. Blacklist alerts generation

Automated blacklists, available on QCTools website (http://esurfmar.meteo.fr/qctools/), are generated for main physical parameters (SLP, SST, …) and position.

“Data quality” as indicated here, for physical parameters, is appreciated by comparing to other sources such as model outputs (various models must be considered).

8. Date out of deployment

If the observation date is outside the deployment interval, position and parameters are flagged bad (qc = 4). The deployment interval is set by the 2 global attributes:

- deployment_date
- recovery_date

4.3 SDQC (Slightly Delayed Quality Control) tests

1. Visual QC

Subjective visual inspection of drifter values by an operator.

Because incurring delays in the data dissemination are not an option, this inspection cannot be done before real-time distribution. However, it must be performed at regular intervals (every working day) for values that are disseminated on the GTS and critical to operational applications, such as air pressure and water temperature. This implies, in order to decide to flag the data and/or stop transmission, to investigate the following cases:

- Beached drifters (whether the data quality remains acceptable, and the data would make a difference because located away for main observation networks;

- Intermittent transmission (whether the data quality remains acceptable);

- Abnormal speeds (whether the buoy has been picked up at sea and has travelled by ship for a while);

- Data quality that is suboptimal (increasing biases and/or noise, whether the barometer or thermometer are failing);

- Data position that proves jittery (appreciated by looking at trajectories; in such case the Iridium position may be used in place of the GPS position, flagging the data accordingly)

- Drogue loss :
  When the submergence count is below 3 (maximum value allowed: 100), the drifter data are flagged accordingly.
  If the submergence information is missing, or outside the expected range (0 to 100), the flags indicate that the drogue status is unknown.
  In addition, and of importance for the operational use of real-time-current estimates, a visual inspection must review changes in submergence or strain gauge counts, as well as GPS Time To First Fix, to detect potential drogue loss, and flag the data accordingly.

Note, most cases above can be detected automatically, but they cannot be resolved automatically.

2. Grey list

This list is updated to reflect when the operator, alerted by a blacklist, has taken action to stop real-time dissemination on the GTS of measurements from a faulty sensor.

The grey list contains the following 7 items:

- Drifter WMO Id
- Parameter: name of the grey listed parameter
- Start date: from that date, all measurements for this parameter are flagged as bad or probably bad
- End date: from that date, measurements are not flagged as bad or probably bad
- Flag: value of the flag to be applied to all measurements of the parameter
- Comment: comment from the PI on the problem
- DAC: data assembly center for this float

Example:

Drifter WMO Id	Parameter	Start date	End date	Flag	Comment	DAC
1900206	SLP	20030925		3		JMA

Each DAC manages a grey list, sent to the GDACs.

The merged grey list is available from the GDACs.

- Grey list format: ascii csv (comma separated values)
- Naming convention: xxx_greylist.csv
  xxx: DAC name (e.g.: aoml_greylist.csv, coriolis_greylist.csv, jma_greylist.csv)
- PLATFORM, PARAMETER, START_DATE, END_DATE, QC, COMMENT, DAC
  e.g. 4900228, SST, 20030909, , 3, , AOML
  e.g 1900206, SLP, 20030925, , 3, , JMA

The decision to insert a drifter parameter in the grey list comes from the PI or the delayed-mode operator. A drifter parameter should be put in the grey list when sensor drift is too big to be corrected adequately in real-time process, or when the sensor is judged to be not working correctly.

The grey list only concerns real-time files (RT-files) or slightly delayed files. When an anomalous drifter is dead or has been adjusted in delayed-mode, it should not appear in the grey list. When an anomalous drifter is active and has been partially adjusted in delayed-mode, it should remain in the grey list only if real-time adjustment is not adequate.

5 Delayed-mode quality controls

The delayed-mode quality controls build on the real-time controls, only to add the benefit of long time series to better qualify and appreciate abrupt or slow changes in data quality.

This section will be refined as experience is gained reprocessing entire time-series of data from drifters that have ceased to function.

6 Appendix

6.1 Reference Table 2: drifter quality control flag scale

This table describes the drifter qc flag scales. Please note that this table is used for all measured parameters.

Flag	Meaning	Real-time comment	Delayed-mode comment
0	No QC was performed	No QC was performed	No QC was performed
1	Good data	All real-time QC tests passed.	The adjusted value is statistically consistent and a statistical error estimate is supplied.
2	Probably good data	Probably good data	Probably good data
3	Probably bad data that are potentially correctable	Tests SDQC1 or SDQC2 failed and all other real-time QC tests passed. These data are not to be used without scientific correction. A flag '3' may be assigned by an operator during additional visual QC for bad data that may be corrected in delayed-mode.	An adjustment has been applied, but the value may still be bad.
4	Bad data	Data have failed one or more of the real-time QC tests. A flag '4' may be assigned by an operator during additional visual QC for bad data that are uncorrectable.	Bad data. Not adjustable. Data replaced by FillValue.
5	Value changed	Value changed	Value changed
6	Not used	Not used	Not used
7	Not used	Not used	Not used
8	Interpolated value	Interpolated value	Interpolated value
9	Missing value	Missing value	Missing value