

GUI based interactive system for Visual Quality Control of Argo data

TVS Udaya Bhaskar¹, R Venkat Seshu¹, E Pattabhi Rama Rao¹ & R Devender¹

¹Indian National Centre for Ocean Information Services (INCOIS), Ocean Valley, Pragathinagar (BO),
Nizampet (SO), Hyderabad 500 090, India

[E-Mail: udaya@incois.gov.in]

Received 15 May 2012; revised 7 August 2012

Argo program is aimed at maintaining an array of 3000 free drifting floats to measure temperature and salinity (T/S). Present study consists a PC-based system developed for visualization and quality control of T/S profiles obtained from Argo floats. The system, coded in Java, is user interactive and runs on Windows platform. Default the Argo T/S profiles pass through 19 automatic checks and quality flags are assigned. Using the system, T/S profiles that failed the automatic Argo tests undergo visual review. This visual review is done to determine whether automatic Argo tests were excessively flagging good measurements as bad or vice-versa, to motivate modifications to automatic Argo tests and to determine whether additional tests were necessary to catch problems that could not be detected by the existing tests. Visual review is done by comparing with 1° X 1° monthly climatologies from WOA01. Profile records deviating beyond 2 standard deviations from the mean are flagged as bad. Provision is also given to compare individual T/S profiles with previous 5 profiles via a waterfall plot.

[**Keywords:** Argo floats, Temperature, Salinity, Visual quality control, Climatology, Java.]

Introduction

Argo is an internationally coordinated program directed at deploying the global ocean with 3000 profiling temperature and salinity floats. Profiling float sinks after launch to a prescribed pressure level, typically 2000 dbar and after a preset time (typically 10 days) the float returns to the surface, collecting temperature and salinity profiles¹⁻². On the surface the float transmits the data to ARGOS satellite which are received at the ground stations and analyzed. Primary objectives of Argo are to 1) provide a quantitative characterization of upper-ocean properties 2) use the float observations to improve the satellite altimetry data, 3) initialize ocean and coupled forecast models, and 4) provide input to other global ocean analyses³.

Indian National Centre for Ocean Information Services (INCOIS) being the National Data Centre (DAC) for Indian and Argo Regional Centre (ARC) for the Indian Ocean, archives Argo T/S data from Indian Ocean (20° – 140° E and 70° S – 30° N). The spatial distribution of 243 Argo float Indian array in December 2011 is shown in Fig. 1.

Table 1 list Indian float deployments by year and 10° latitude band. The Indian Argo program used mainly two types of floats: the Autonomous Profiling Explorer (APEX) floats built by Webb Research

Corporation and the PROVOR built by the Metocean. APEX floats forms 90% of the total floats deployed by India. Two time scales have been defined for Argo data processing: Real Time and Delayed Mode.

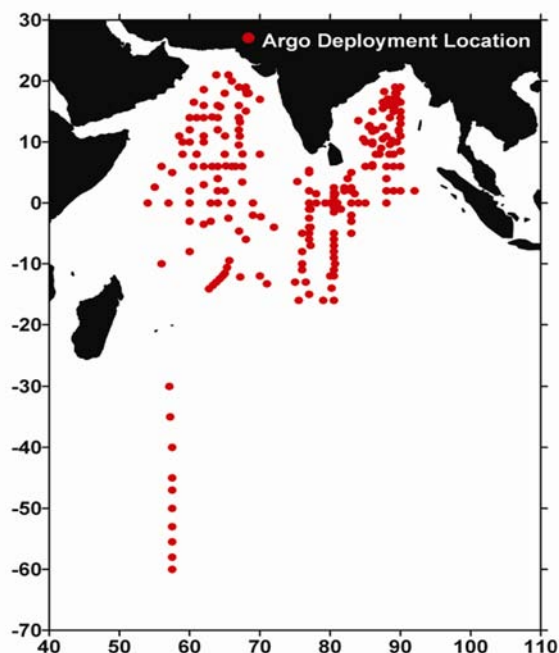


Fig. 1—Distribution of 243 Indian Argo floats by December 2011.

Table 1—Deployment of Indian Argo array given by year and meridional band.

Year	30°–20° N	20°–10° N	10°–0°	0°–10° S	10°–20° S	20°–30° S	30°–40° S	40°–50° S	50°–60° S
2002	0	0	2	7	1	0	0	0	0
2003	0	12	9	0	0	0	0	0	0
2004	0	8	17	3	2	0	0	0	0
2005	1	11	14	17	3	0	0	0	0
2006	0	7	3	5	0	0	0	0	0
2007	0	20	11	0	0	0	0	0	0
2008	0	0	5	10	0	0	0	0	0
2009	0	3	2	0	2	0	0	0	0
2010	0	7	5	4	7	0	0	0	0
2011	1	14	1	9	10	0	2	4	5

Primary objective of the real-time data processing is to provide salinity and temperature profiles that have undergone some quality control to the operational and research communities within 24 hours of collection. Delayed mode data experience more stringent review using statistical procedures and scientific knowledge.

The goal of this work is to show how Argo real-time procedure overlooks certain bad/good profiles and how the visual quality control system can be used to correct these profiles in near real time. In section 2 we will summarize the Real-Time quality control procedure. Section 3 discusses typical cases which are overlooked by RTQC procedures. Section 4 discusses the visual quality control system in detail.

Materials and Methods

Real Time Quality Control procedures

Data from Argo floats need to be checked for its quality, automatically and independent of any operator involvement so as to provide data to operational agencies within 24 hrs. For this a series of 19 automatic tests were proposed by the Argo Data Management Group (ADMT). Description of Real-Time Quality-Control tests are given in Appendix-1. Primary objective of the real-time quality control of Argo profiles is to identify erroneous data prior to insertion to GTS. Erroneous data are excluded from the profiles submitted to GTS, but they are flagged and forwarded to the GDACs. This is done so as not to burden the operational users with erroneous data.

Anomalies in RTQC procedures

Even though majority of the profiles pass the real time quality control test, there can be cases where in the profiles are incorrectly assigned flags. For example, incorrect metadata for the conversion from counts to physical values could be used. Profiles from such a float could pass all the automatic Argo tests although the data are not correct. Also as most of the

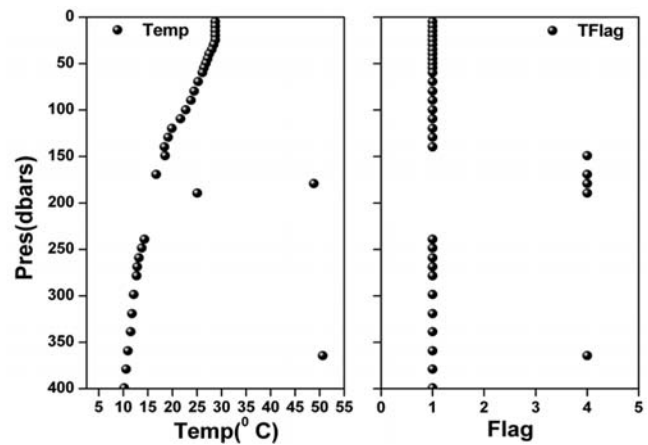


Fig. 2—Temperature profile showing flags after RTQC.

quality control tests rely on the records above and below for assigning a flag, a bad flag in one record might cause record above and below also to be assigned as bad. Typical cases of profiles where good/bad data are assigned bad/good flags are discussed below:

Bad flag assigned due to spike

Consider a typical case of a temperature profile shown in Fig. 2 obtained from float WMOID 5901171 (Observation date: 12/02/2012, Longitude: 99.335, Latitude: -13.574). This profile has a bad temperature value at pressure 364.3 dbars (V_2), the temperature value being 50.633° C. This bad pressure value can be attributed to a corrupt message packet. Since pressure values not used in the evaluation of spikes, this bad temperature value, caused the temperatures above and below to be assigned a bad flag. When the profile is passed through the automatic test a flag of 4 (indicating it as bad) is assigned to the temperature at depth 149.1 (V_1) and 169.2 (V_3) by the spike test. To make matters simple let us work out for the depth 364.3. The spike test is defined as follows:

$$\text{Test Value} = |V_2 - (V_3 + V_1) / 2| - |(V_3 - V_1) / 2|$$

where V_2 is the measurement being tested as spike and V_1 and V_3 are the values above and below V_2 .

Temperature: The V_2 value is flagged when

- The test value exceeds 6°C for pressure less than 500 dbars or
- The test value exceeds 2°C for pressure greater than or equal to 500 dbars.

Salinity: The V_2 value is flagged when

- The test value exceeds 0.9 psu for pressure less than 500 dbars or
- The test value exceeds 0.3 psu for pressure greater than or equal to 500 dbars.

Here in this case $V_1 = 18.478$, $V_2 = 50.633$ and $V_3 = 16.665$, which when used in the above expression gives a values of 16.976°C which exceeded the prescribed limit for depth less than 500 dbar. Hence the temperature at 364.3 dbar is assigned a bad flag 4. As this bad temperature value is used while evaluating the flag in spike test for temperature at 149.1 dbars it will also be assigned a flag of 4 which is also the same case with the temperature at 169.2 dbars. Similarly cases of salinity can also cause assignment of bad flags to good data.

Setting good flags for profile having off set in salinity:

This is rare case but can happen some times. There can be a float giving profiles which are exactly offset (biased) by a value from the mean (Fig. 3). One such

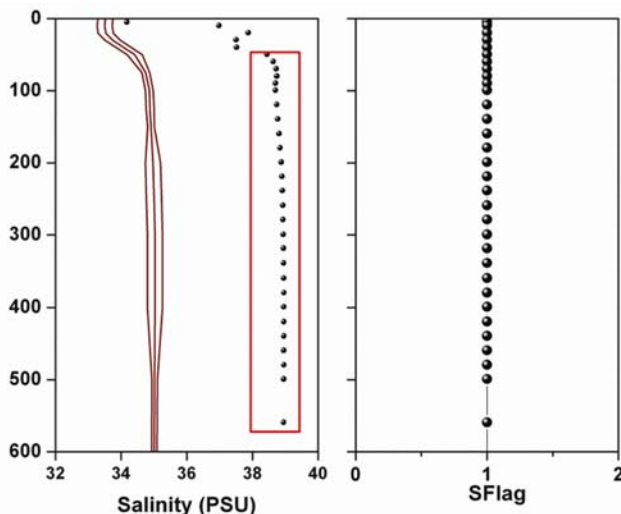


Fig. 3—Profile with salinity offset and flags assigned by the RTQC. Black dots represent the Argo CTD records. Thick blue lines indicate climatology. Broken line indicate 2 times standard deviations from the mean.

typical case is the float with WMOID 2900783 deployed in Bay of Bengal by India. This float is observed to having some problem with the salinity sensor which caused the salinity profile to be having an offset value of 4 psu. Unfortunately, these profiles can never be caught wrong by any of the automatic test. The only solution to such profiles is to check them visually.

Hence profile of such nature and many more can be accurately checked for the correctness by cross comparing them with the best available climatology. Further the assigned quality flags can be double checked for their correctness which will improve the confidence of the users of Argo data.

Results and Discussion

Visual Quality Control System-modules

The Visual Quality Control System (VQCS) discussed here is developed based on the experience in processing the Argo T/S from the tropical Indian Ocean. Pankajakshan *et al.*⁴ developed some what similar tool for quality control of XBT data archived at Indian Oceanographic Data Centre. However the VQCS discussed here is capable of handling both temperature alone and T/S profiles. Argo floats being different from conventional ship borne CTD measurements (unmanned measurement, sensor degradation due to bio fouling) Argo T/S are prone to errors. The general errors which Argo T/S encounter are:

- (a) Sensor degradation due to bio fouling.
- (b) Pressure sensor error due to micro leakage.
- (c) Salinity hooks (bottom most salinity value appearing as hook).
- (d) Erroneous values due to corrupt message packets.
- (e) Missing CTD triplet due to loss of message packets.

Most of these errors, could be identified and corrected (or flagged for the type of errors) while passing through the 19 automatic quality control checks. But in few cases some of the good measurements are flagged as bad, while some bad measurements pass the quality checks (as described in the previous section). These types of errors can be corrected by using the VQCS. Therefore, a user interactive and computer based quality control system, which can be used to set right the quality flag, would be quite useful tool for quality control of Argo

data. The quality control system discussed in this paper is developed with the objective to help the operator at the data centers to identify such errors and to apply the possible corrections in near real time as well as delayed mode data processing. Though this visual quality control tool is designed for quality control of Argo CTD data it can be used for quality control of other oceanographic data like ship borne CTDs, XBTs, XCTDs etc. Various controls of VQC and their use in quality control are discussed in section 4.1.

Visual Checks

Visual quality checks incorporated in this system are T/S spikes, density inversions, comparison with climatology and checking with previous profiles for checking sensor consistency (water fall plot) among others. A provision is given for graphical presentation of T/S profiles obtained from Argo float (WMOID chosen) with the quality flags assigned by the RTQC. Individual profiles are chosen randomly by choosing the WMOID from the dropdown box (Fig. 4). The cycle number, date, time, longitude, latitude, float type etc are also displayed. A provision is given for editing the flag associated with the record for possible correction, where it is found to be erroneous.

Spikes

Spikes in the T/S profiles are caused by instrumental, communication problems. Big enough

spikes are definitely caught by RTQC test and accordingly flagged. However there can be small spike which pass the RTQC and still appear as spikes visually. These small spikes can be visually identified and flagged.

Inversions

Inversions in the temperature some time can be real. In certain regions, especially the region of cold low saline water over warm high saline waters (south eastern Arabian Sea in winter) inversions are usual features. The profile is visually inspected to distinguish the real inversion from inversion-like feature.

Comparison with climatology

For comparing the T/S profiles with climatology the World Ocean Atlas 2001⁵ is used. The monthly mean climatology of $1^\circ \times 1^\circ$ grid is used. The longitude, latitude, month of the CTD profile under consideration is used to compare with the climatology. The climatological mean profile, 2 time standard deviation from mean obtained from nearest grid and month to that of the Argo observation are used as background to qualify the observed profile. All the records appearing as spike or deviating grossly from the 2 standard deviation envelope are flagged as bad.

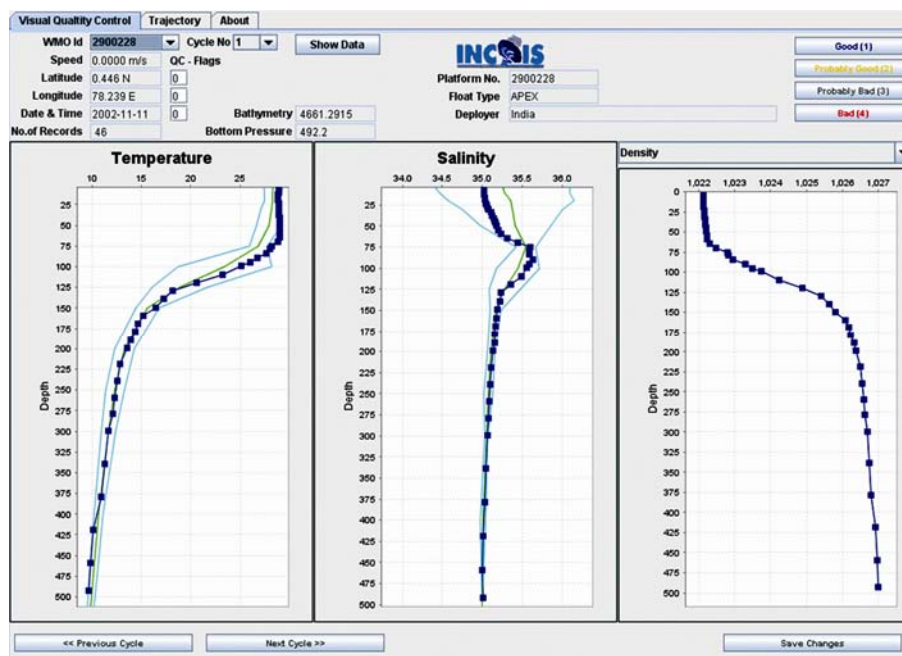


Fig. 4—VQCS window for checking the quality of Argo T/S profile. T/S profile is plotted in dark blue with the climatological mean (green color) and 2 time standard deviation (cyan color) envelope at the background. Profile is checked for correctness based on comparison with the climatological profile.

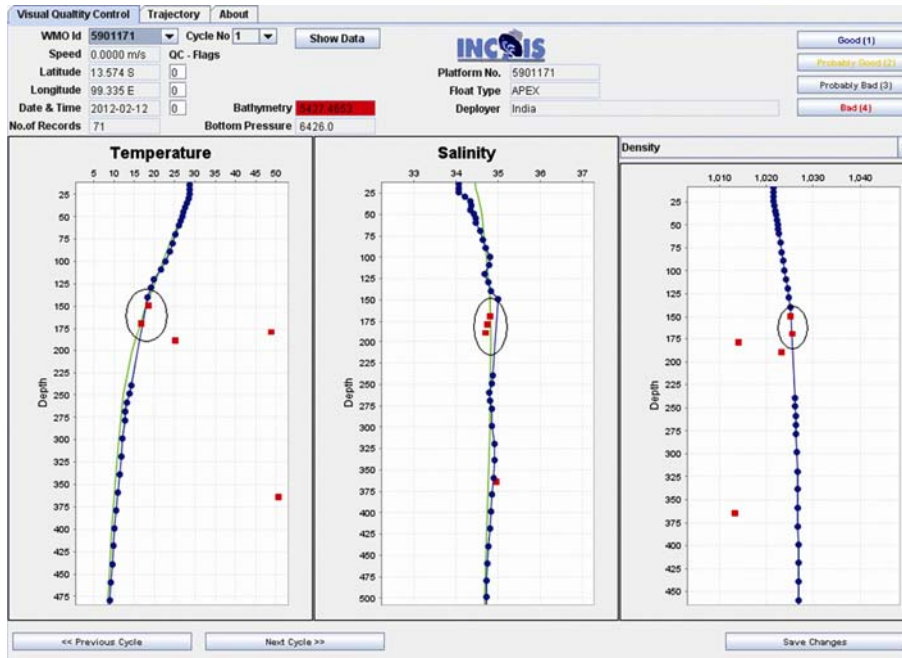


Fig. 5—View of profile with good values flagged as bad.

Overcoming RTQC anomalies with VQCS

By using the comparison with climatology functionality of VQCS, some of the above mentioned anomalies in the real time quality assignment can be overcome. For instance we consider the same float (WMOID: 5901171) for which some good values have been assigned a bad flag due to spike test. These values are clearly seen to lying very much near to the mean T/S profiles from WOA. Visually checking them give a fair idea about their quality and can be easily corrected using the functionalities provided in VQCS.

Figure 5 shows the profile under consideration being viewed in VQCS. Clearly one can observe that some of the good T/S records are flagged erroneously as bad by the real time procedures. Profiles like this can be easily corrected by the dedicated operator having a second look after the profiles are passed through RTQC.

Miscellaneous functionalities

The VQCS is provided with miscellaneous functionalities like, bathymetry, trajectory, erroneous position by the way of monitoring speed, facility to view the data in ASCII format, buttons to traverse to previous and next profiles and buttons to save/modify the corrections done to the quality flags on records (Fig. 6). The bathymetry field shows the bathymetry at the profile location which gives a idea whether the



- (1) General information panel.
- (2) Bathymetry at observed location and deepest pressure of profile.
- (3) Instrument specific information.
- (4) Buttons to assign/modify Quality flags of the records.
- (5) Button to view the data in ascii format.

Fig. 6—Different panels of VQC software. Each panel serves different purpose.

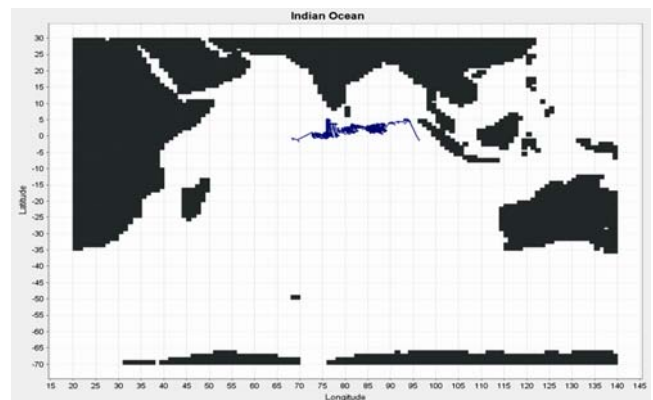


Fig. 7—Panel showing the trajectory of the chosen float during its life time.

float had grounded (profiling pressure \geq bathymetry) or some bad pressure values are recorded. The trajectory panel can be used to view (Fig. 7) the trajectory of the float under view. This gives an idea as to where all, the float have traveled during its life

time and also suspected positions if any (appearing as spike in the float trajectory).

Conclusions

Argo is internationally coordinated program aimed at seeding the ocean with autonomous profiling floats. These profiles are quickly passed through 19 real time quality control (RTQC) steps and flags are assigned accordingly. However there can be cases where the RTQC assigns bad (good) flags to good (bad) profiles. These can only be caught while checking the profiles visually. To meet this purpose software for visualization and quality control of temperature and salinity profiles obtained from Argo floats is developed. This software is capable of handling all those profiles which are erroneously assigned flag by the RTQC procedures.

Acknowledgments

Authors are grateful to Dr. SSC Shenoi, Director, INCOIS for encouragement and facilities to carry out the work and to Dr Christine Coatanoan and Thierry Carval of Coriolis/IFREMER for their useful discussion on the VQC. Argo data is made freely

available by the Argo community. This is INCOIS contribution number 155.

References

- 1 Argo Science Team, 2001. The Global Array of Profiling Floats, in observing Oceans in 21st century. C.Z. Koblinsky and N.R. Smith, eds., Godae Proj. Off., *Bur. Meteorol.*, Melbourne, Australia, (2001) 248 – 258.
- 2 Ravichandran M., P. N. Vinayachandran, S. Joseph and K. Radhakrishnan, 2004, Result from the first Argo float deployed by India, *Curr. Sci*, 86, 651 – 659.
- 3 Roemmich, D., 2001: Argo: The global array of profiling floats, observing the oceans in the 21st century, C.J. Koblinsky and N. R. Smith, Eds., Melbourne Bureau of Meteorology, 604 pp.
- 4 Pankajakshan Thadathil., Aravind K Gosh, J.S. Supriya and V.V. Gopalakrishna, An interactive graphical system for XBT data quality control and visualization, *Comp & Geos.*, 27: (2002) 867 – 876.
- 5 Conkright, M.E., R. A. Locarnini, H.E. Garcia, T.D. O'Brien, T.P. Boyer, C. Stephens, J.I. Antonov, 2002: *World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures, CD-ROM Documentation*. National Oceanographic Data Center, Silver Spring, MD, 17 pp

Appendix – 1

Automatic Real Time Quality Control procedures

1. Platform identification: Each float should have a unique valid identifier provided by World Meteorological Organization (WMO).

3. Impossible location test: The latitude (longitude) must be in the limits -90 to 90 (0 to 360).

5. Impossible speed test: Surface and subsurface drift speeds must not exceed 3 m s^{-1} .

7. Regional range test: Temperatures from floats in the Red Sea (Mediterranean Sea) must range from 21.7° to 40.0° C ($10.0^\circ - 40.0^\circ \text{ C}$) and salinity ranges must be from 2.0 to 41.0 (2.0 to 41.0 psu).

9. Spike test: $|V_2 - (V_3 + V_1)/2| - |(V_3 - V_1)/2|$ for a value V_2 , where V_1 and V_3 are the values above and below V_2 , which may not exceed prescribed limits. Above 500 dbar, the limit for temperature (salinity) is 6°C (0.9) and below 500 dbar the limits are 2°C (0.3).

2. Impossible date/time test: Year must be greater than 1996; month in the range of 1 to 12; date must in the expected range for the month; hours in range 0 – 23; minutes in the range 0 – 59.

4. Position on land test: The floats must be located in the ocean. ETOPO2 bottom topography is used for this test.

6. Global range test: Temperatures must be in the range of -2.5° to 40.0° C and salinity must be from 2 to 41 psu.

8. Pressure increasing test: The pressure must increase monotonically.

10. Top and bottom spike test: This test is obsolete now.

11. Gradient test: The test value $|V_2 - (V_3+V_1)/2|$ for a value V_2 may not exceed prescribed limits. Above 500 dbar, the limit for temperature (salinity) is 9.0°C (1.5) and below 500 dbar the limits are 6.0°C (0.5).

13. Stuck value test: This test checks for constant temperature or salinity values throughout the profile

15. Grey list: A list generated based on the history of a float. When a float sensor has systematic problems it is placed on this list.

17. Frozen profile test: If floats produce five consecutive profiles with very small differences throughout the entire water column (i.e., of the order of 0.001 for salinity and of the order of 0.01°C for temperature) they are candidates for the gray list.

19. Visual Quality Control: Subjective visual inspection of float values is done by an operator.

12. Digit rollover test: A specific number of bits are allocated for the storage of temperature and salinity values in a float. When the number is exceeded, stored values rollover to the lower end of the range. This rollover when detected is compensated for in the processing algorithm.

14. Density inversion : This test computes the density at all pressure levels from the observed temperature and salinity values and tests for hydrostatic stability.

16. Gross salinity or temperature sensor drift: If the average temperature (salinity) from the last 100 dbar of two adjacent profiles exceeds 1°C (0.5), then the profile is considered to be bad.

18. Deepest pressure test : This test requires that the profile has pressures that are not higher than DEEPEST_PRESSURE plus 10%. DEEPEST_PRESSURE value comes from the meta-data file of the float.
