JAMSTEC position Quality control
-- Suggestion of a standard procedure for Argo position QC --

Tomoaki Nakamura, Naoko Ogita, and Taiyo Kobayahsi
Argo Group, IORGC, JAMSTEC

1. Introduction

We are suggesting a standard procedure for the position QC of Argo trajectory data based on checking float speed between positions fixed by Argos system.

This procedure is a draft of it, which is modified from the procedure used at JAMSTEC and has been introduced by TK at the 2nd trajectory workshop in Seoul, S. Korea. After then we have checked all schemes and criteria of the QC procedure and improved them considering the comments at the workshop. This document describes the details of the QC procedure and some results obtained by this QC scheme.

This program reads netCDF files of Argo trajectory in direct and then outputs QC flags of float positions by text file. Also this program creates a KML file of the results for “Google Earth”, which helps us to check results easily by visual inspection. Some parameters of this procedure are changeable by its configuration file (details of usage are described in section 5). The source file of this procedure is written by “Ruby”, which is not so familiar for most of you, unfortunately. In later days we will convert it into more familiar language like Matlab (see section 6. Future plan).

2. Basic concepts of position QC

We suppose that this position QC procedure is the first check of Argo trajectory data to discard (or flag) very “strange” positions automatically (or instead of “visual check” by experts), and some of more sophisticated interpolating schemes will succeed it in order to estimate actual float trajectories and locations where floats arrive at sea surface or dive into subsurface (e.g., Davis et al., 1992, Park et al., 2005), or “delayed-mode QC” for trajectory data. Since “strange” positions distort these estimations largely if they are used in any interpolating scheme in the same manner as normal data. On the other hand, it is well known that the more data we can use for the interpolations the better results we can have. Thus, it is desirable that this position QC scheme discards only very “unusual” position data for most of experts.

Another purpose of this suggestion is to standardize data quality of Argo trajectory data by a single QC scheme. Trajectory files are generally used to estimate velocities on the sea surface and/or at the parking depth of floats directly. Huge amount of data are needed regardless of PIs (DACs) of floats when basin (global) scale velocity fields are calculated (e.g., YoMaHa dataset, Konstantin et al., 2007). Quality of such atlases may depend on the lowest level of original data. Thus, standardized quality of Argo trajectory data brings lots of benefits to all data users.

To standardize the quality of trajectory data, we think a standard program is needed definitely considering the present circumstance of Real-time QC of float profiles. We know that its results are very different by each DAC even though they follow the standard QC criteria defined by Argo Program completely. It means that only definitions of criteria (like the table of Real-time QC flags for float profile) might not be enough to dissolve differences between DACs in the case of complicated data like float profiles and float trajectories.

3. Position QC procedure

This position QC scheme identifies “bad” positions considering the float speeds which are required by the float positions fixed by Argos system.

If we find a segment which satisfies the following 2 criteria, at least one position which composes the segment will be flagged 3 ("bad"). See also Figure 1.

- Float speed along a segment exceeds 3 m/sec.
- The length of the segment is longer than the critical error length which is determined by the nominal errors of Argos system as follows:
  the critical length = 1.0 * sqrt (Error_A ^2 + Error_B ^2).
  Here, Error_A and Error_B are the radii of position error of Argos system (150m, 350m, and 1000m for Argos class 3, 2, and 1, respectively) at the float positions A and B, respectively.

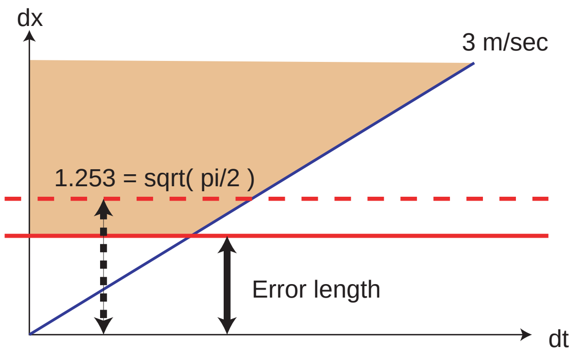


Figure 1: Definition of erroneous segment which includes at least one float position to be flagged 3 (red hatched region). Dx and dt are horizontal distance and time interval between 2 float positions fixed by Argos system. Error length is determined by 1.0 * sqrt(Error_A ^2 + Error_B ^2).

"Bad (flagged)" position of the questionable segment is identified by the following scheme:

- If Argos classes at both positions are different, the less accurate position is flagged.
  Accuracy of Argos class: More accurate <= 3, 2, 1, 0, A, B => Less accurate
- If Argos classes at both positions are the same, the position which requires more quick movement along the continuous segments is flagged (see also Figure 2).
  - In the case of 4 continuous position data can be used for the check, we compare average speeds along the segments via A and via B (yellow and cyan trajectories in Figure 2a).
  - In the case of 3 continuous position data are available for the check (e.g., segment includes the first/last position), we compare average speeds on the segments terminated at A and at B (yellow and cyan trajectories in Figure 2b).
  - If we use only 2 positions for the check (e.g., trajectory of a cycle for the check is composed of 2 positions), both positions are flagged 3 because we can not determine which position is worse than the other (an exceptional case).

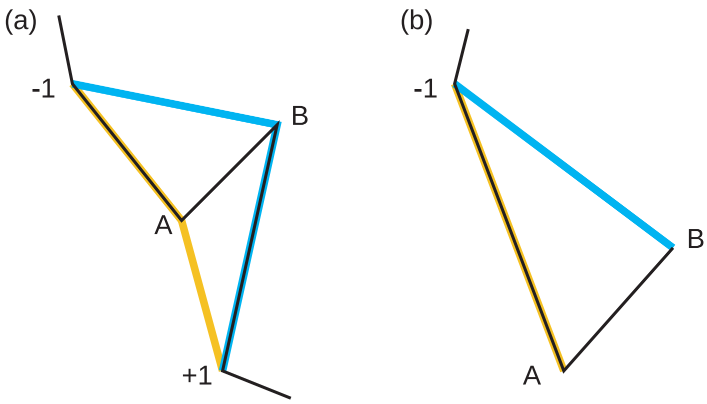


Figure 2: Schematic figures of identification of the position to be flagged in the cases of (a) 2 next positions (signed by -1 and +1) of the questionable segment (AB) can be used and (b) only one next position (-1, the previous or next) of the segment can be used.

Note:

- The minimum time interval is set 1 sec to avoid “dividing by 0”.
- The original JAMSTEC position QC procedure uses 5 knots for an absolute speed criterion, which is changed to 3 m/sec due to the latter criterion is used in “Argo delayed-mode user manual”. 3 m/sec is about 6 knots, so this criterion is more relaxed in this QC scheme.
- The original JAMSTEC procedure also uses another criterion of 2.5 times the average speed along whole the float trajectory. In the new scheme this criterion is removed because in most cases the average speed is less than 0.5 m/sec, this criterion does not work in distinction schemes generally. The removal of this criterion makes us possible to check the trajectory with only 2 positions.
- The length criterion works only to relax the speed criterion for the pair of positions which are fixed in very short time (less than 10 minutes in most cases).

Figure 3 shows scatter plots of horizontal distance and time interval between any 2 positions in the same cycle of a float (n = 4379 for 290 float*cycle). The criterion of 3 m/sec is not so severe except for the position pairs with shorter time interval (about less than 10 minutes). In this time range the position errors of Argos system can not be ignored to calculate actual speed along float trajectory. Thus, we concluded that Argos error of float positions be considered only in this range in this position QC procedure.

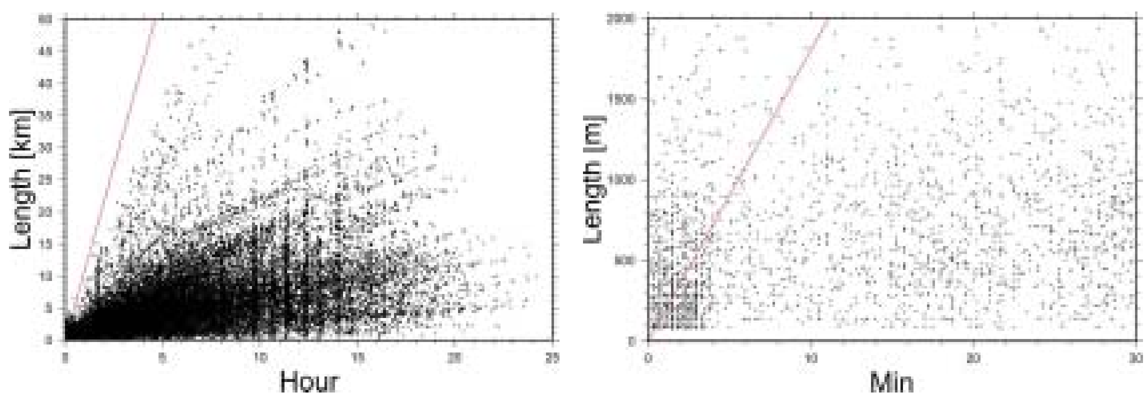
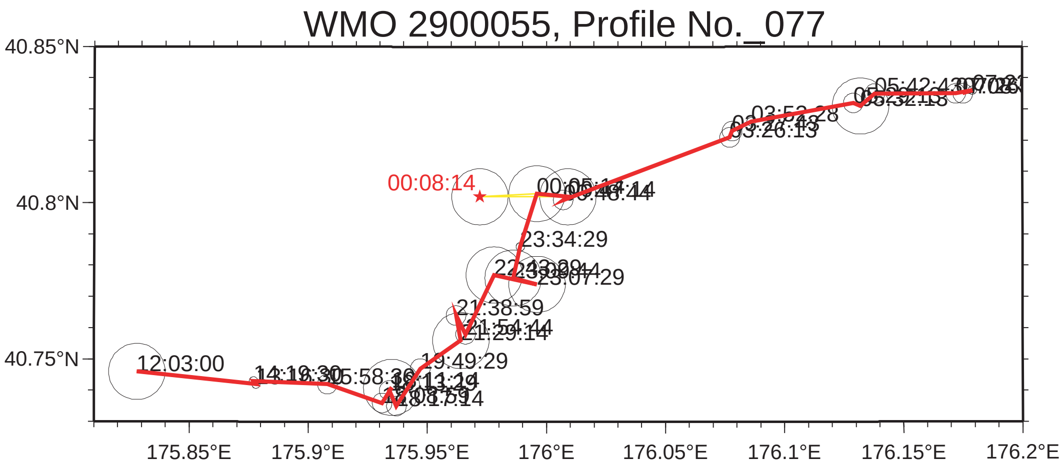


Figure 3: Scatter plots of horizontal distance and time interval between any 2 float positions in the same cycle (n = 4379 for 290 float*cycle). The red line represents the relation of float

movement by 3 m/sec.

Results

The Figure 4 shows some examples of the results of this position QC. This program identifies several % of position data to be flagged 3 in average. By our preliminary examinations, this ratio varies largely (about 1 % to more than 10%) by float and also PI (or DAC). The differences among DACs are mainly caused by the features of their trajectory files (e.g., positions with less accurate Argos class (0/A/B) are included or not).



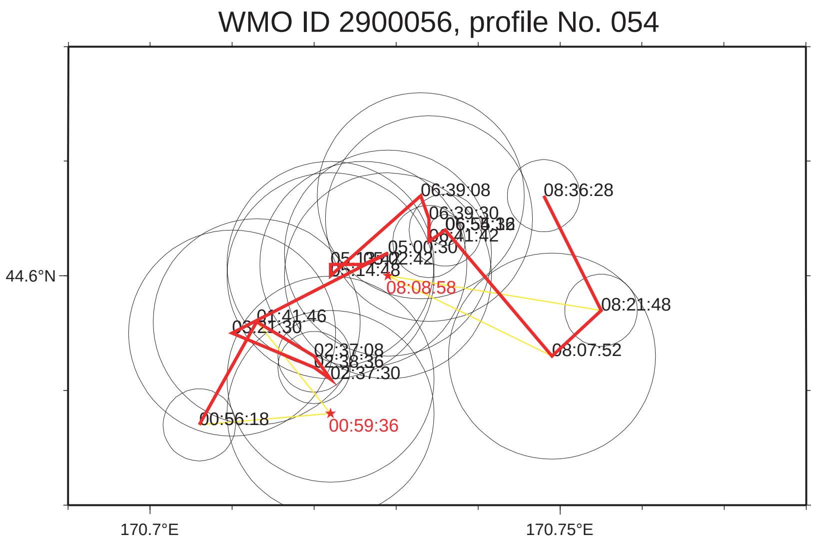


Figure 4: Examples of results of this position QC program. The flagged positions are shown by red stars, and positions though QC are connected by red line. The time when the position is fixed by Argos satellites is also shown. Circles mean error radii of Argos system for the positions (1000m, 350m, and 150m for Argos class 1, 2, and 3, respectively).

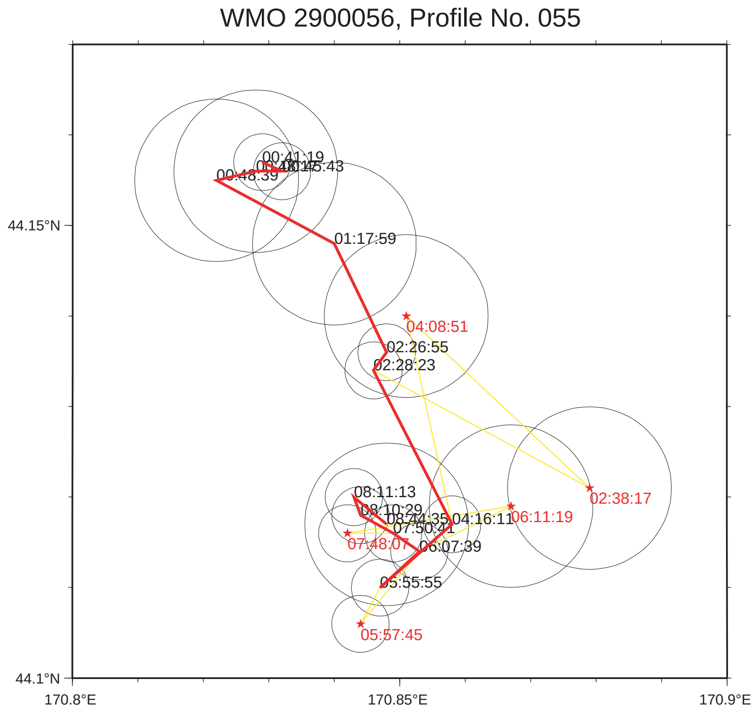


Figure 4 (continued)

Other bad data identified in this scheme:

- The following position data are flagged 4, and they are initially removed before the checking procedures mentioned above.
    - ➢ Duplication of position data (both location and time)
      => They are flagged 4 except for a single data. If different positions are fixed in the same time, all positions are checked.
    - ➢ Positions which are apart from the first position by more than 24 hours in time
      => Flagged 4. It is considered that surface-drifting data in another cycle be involved.
    - ➢ Float speed in subsurface (from the last “reasonable” position of the previous cycle to the first position of the cycle) exceeds 3 m/sec.
      => Flagged 4. Too fast movement of float drifting in subsurface.
- Dates of float positions fixed are 50 years fast in some trajectory data.
  => Adjusted to the “right” dates (this procedure will be changed to “range check” and it returns some warning if such future dates are found in files).
- Argos class is written by octal notation.
  => Adjusted to the decimal notation (this procedure will be changed to output some “warning” only if such the cases are found).

Note: Most of them will be removed from newer version because we consider that such the “strange” data should be removed before we deal with QC flags of position data (see also section 6. Future plan).

4. Concerns in this program

- Argos class and its error radius
    - ➢ This program uses all classes (3, 2, 1, 0, A, B) except for error (e.g., Z), and the error

radii of the positions with Argos class of 0, A, B are set as 1500m uniformly. We need to have some consensus about the Argos classes to be used for position QC (or included in netCDF file) and their error radii.

- Length criterion
  - The length criterion used in this program is much simplified to be factor 1, by which our concept is understood easily. The preliminary statistic examinations gave us another candidate of the factor, 1.253 (= sqrt (pi/2)) (see Figures 1 and 5). We would like to have any comment about this issue.

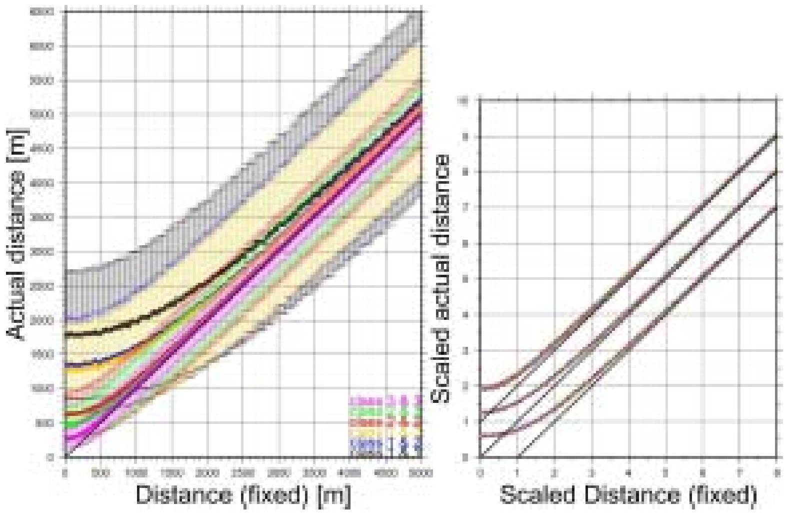


Figure 5: Estimation of "actual" distance between float positions fixed by Argos system with Monte Carlo method. (a) Average and standard deviation of the actual distance between 2 positions considering Argos error distance. Color means each combination of float positions with different Argos class. (b) Average +/- standard deviation of the distances normalized by the factor of sqrt(Error_A ^2 + Error_B ^2).

5. How to use this program

This program (position_qc.exe) works on Windows XP only (sorry). The batch file (posqc.bat) makes position_qc.exe run and deal with all netCDF trajectory files in the same folder automatically. The results of position QC are output by text file.

DOS > posqc.bat (press return)

Also, this program produces KML files of the QC results for Google Earth, which helps you to check the results by visual inspection (see Figure 6).

You can modify the criteria of this program (the maximum float speed, error radius of Argos system) by change of configuration file (position_qc.cfg). The program position_qc.exe uses netcdf.dll as library file, so this library file should be in the same folder of position_qc.exe.

If you would like to modify this program largely, its source programs are also available from the following web site.

PARC-JAMSTEC: http://www.jamstec.go.jp/ARGORC/

From page of "Tools and Links"

They are written by "Ruby", which is not so familiar with most of you, unfortunately. Some of additional documents are also prepared for the setup of "Ruby" circumstances on PC with Windows XP, but we are not responsible for any of your troubles.

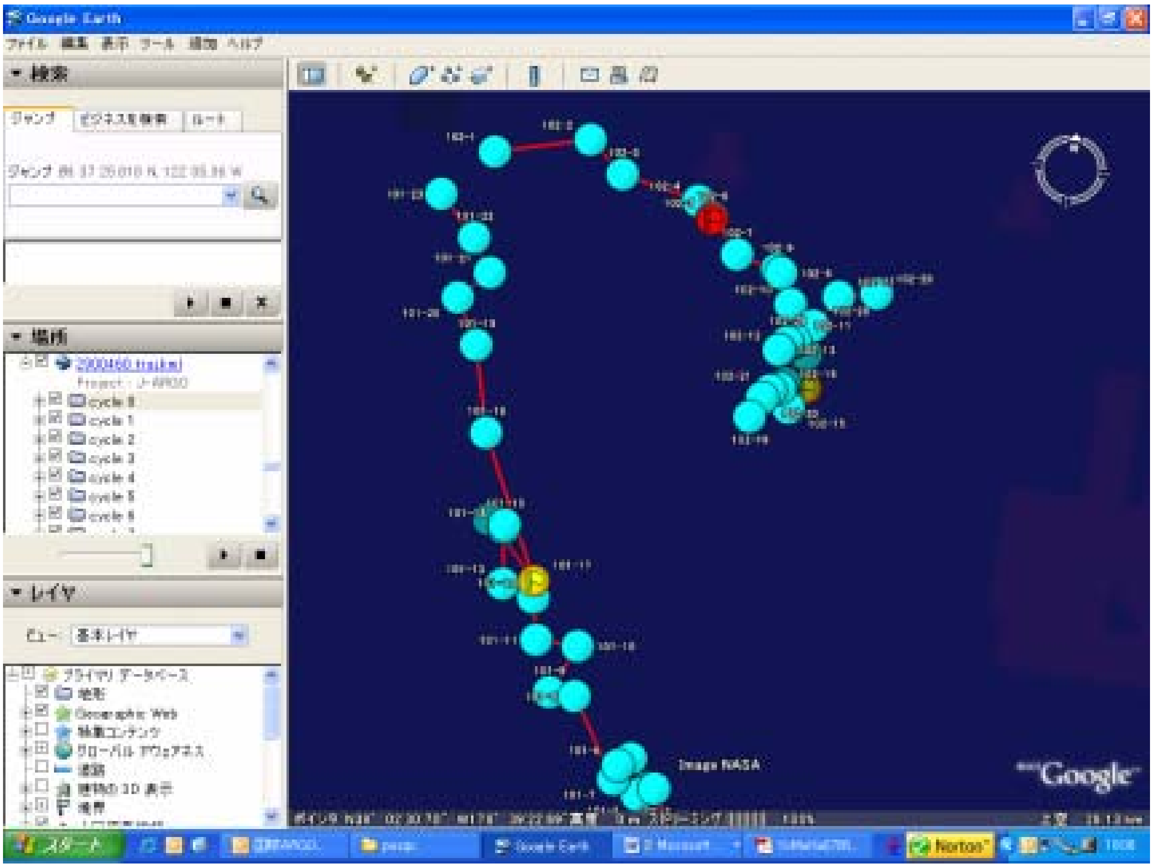


Figure 6: An example of output of position QC results for Google Earth. Yellow and red marks represent the positions with flag 3 and 4, respectively.

6. Future plan

We have a plan to convert this program into a more familiar language, Matlab. At that time, most of the schemes for dealing with “strange” data will be removed, because we believe such “strange” data be removed before position QC.

The data which satisfy the following criteria are ideal to be checked by position QC, we think. These give us much simpler code of the position QC procedure and also lots of benefits to all persons who use Argo trajectory data directly.

- All data are filled by the format determined officially.
- All the data (e.g., date, longitude, latitude …) fall in reasonable values.
- All the data are ascending ordered by time.
- Most of the data of positions and Argos class have values.
- No complete duplications (all the data of date, position, Argos class are the same).
- No large (temporal) gaps in the same cycle, because such the data can be considered as miss involving of another cycle data.

Acknowledgement

We are grateful to all the attendees at 2nd Argo Trajectory Workshop to have given us fruitful comments about our original position QC scheme. Especially, Dr. Hiroshi Yoshinari gave us lots of additional comments in other day, which are valuable very much to improve the program. Also, we thank to Dr. Brian A. King for his advices and encouragements, especially for TK.

References

Devis, R.E., D. C. Webb, L. A. Regier, and J. Dufour, 1992: The autonomous Lagrangian Circulation Explorer (ALACE). *J. Atmos. Oceanic Technol.*, **9**, 264-285.

Konstantin V. Lebedev, Hiroshi Yoshinari, Nikolai A. Maximenko, and Peter W. Hacker, 2007: YoMaHa'07: Velocity data assessed from trajectories of Argo floats at parking level and at the sea surface, *IPRC Technical Note,* **4**(2) 16pp.

Park, J. J., K. Kim, B. A. King, and S. C. Riser, 2005: An advanced method to estimate deep currents from profiling floats. *J. Atmos. Oceanic Technol.*, **22**, 1294-1304.