

A beginners' guide to accessing Argo data

John Gould – Argo Director

Argo collects salinity/temperature profiles from a sparse (average 3° x 3° spacing) array of robotic floats that populate the ice-free oceans that are deeper than about 2000m. They also give information on the surface and subsurface currents. Each profile is made up of about 200 data points. The first Argo floats were deployed in 2000 and the array will be complete in 2007. Argo data are made available to users quickly and free of restriction.

The following document is a first draft for a brief user guide to the Argo data system. This is intended to describe to Argo users:

- (i) what they should expect from Argo data (real-time and delayed mode streams)
- (ii) how to find Argo data

Complete documentation of the Argo data system is contained in the "Argo Data Management Handbook" and "Argo Real-time Quality Control Tests Procedures", both of which are available at http://www.coriolis.eu.org/cdc/argo_rfc.htm.

Feedback from Argo users will be very valuable as we implement and test the remaining parts of Argo's data management system. Specifically, comments in the following areas are especially appreciated and should be sent to argo@ucsd.edu :

1. For users requiring real-time data (within about ~24 hours of measurement):
 - Are there inconsistencies between Argo's specified real-time QC tests and what is found in the data files (GTS or Global Data Centers)?
 - Are there formatting problems in the GDAC NetCDF files, with respect to the specifications in the Argo Data Management Handbook?
 - Are there additional or better (fully automated) real-time QC tests that should be applied? (For example it has been noted that a range check on pressure is needed).
 - Is an automated estimate of salinity sensor drift (presently only in delayed mode data) needed?
2. For users requiring near real-time data, but on a time scale of several days or longer:
 - Some DACS (e.g. US Argo DAC AOML) are presently implementing visual profile inspection in near real-time. It is unlikely that this will be done by all national DACS unless there is a strong requirement expressed for an intermediate quality dataset (better quality than real-time, faster than delayed-mode). Is there such a requirement?
3. For users requiring scientific quality data (delayed-mode):

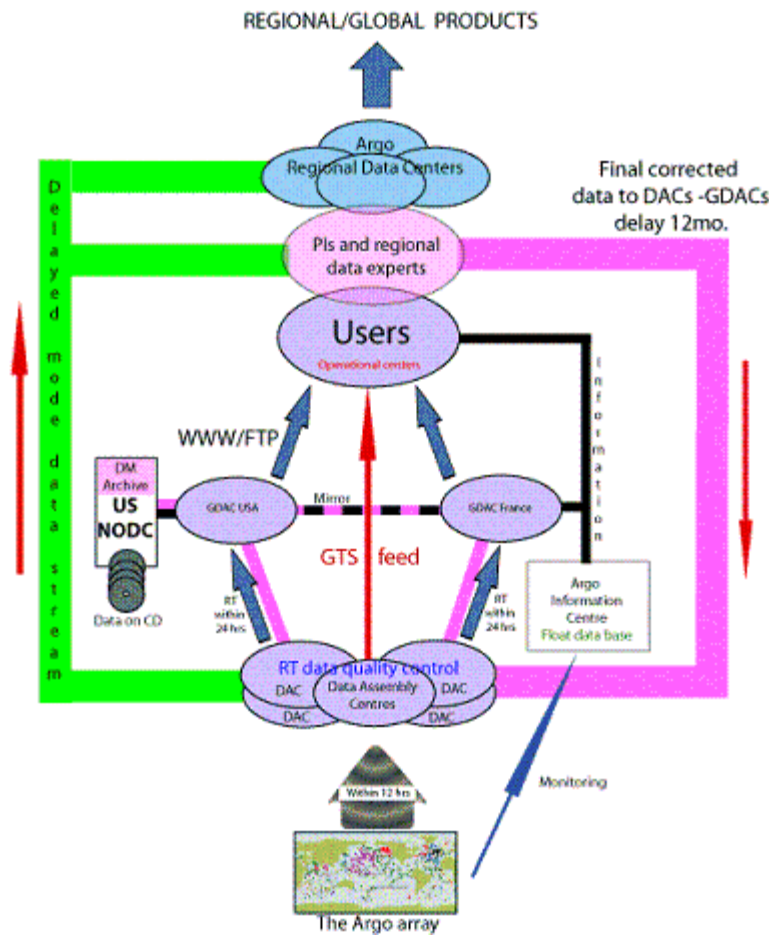
- At present, only about 30% of Argo profiles are available as delayed mode data (i.e. NetCDF filename begins with D), and the delayed-mode system is still under development. However, if problems are seen in delayed mode data, please let us know by contacting support@argo.net.

Accessing Argo data

The data are available to users through 3 routes:

- To operational centers via TESAC messages on the Global Telecommunication System (GTS)
- By ftp, http, LAS downloads from two Global Data Assembly Centers
- From an archived data set at the US NODC.

The pathways of data flow from the floats to the data centers and users are shown in the following diagram.



Schematic of Argo data flow

Preparation, availability and quality of Argo data

GTS data stream

These data are subjected to a number of quality checks in national data assembly centers and only those measurements within a profile that pass all tests (see below) are inserted onto the GTS. The checks are also used to set quality flags in the data that go on to enter other data streams detailed below.

In the TESAC format, temperature and salinity are truncated to two decimal places. The vertical co-ordinate is depth not pressure (as measured). No corrections are made to salinities in the GTS data stream.

Real-time data quality checks

These checks will be reordered and all DACs will apply tests in same order to the profile data.

1. Platform ID *
- 2 Impossible date *
- 3 Impossible location *
- 4 Position on land *
- 5 Impossible speed *
- 6 Global range test *
- 7 Regional parameter range*
- 8 Pressure increasing
- 9 Spike test
- 10 Top – bottom spike - obsolete
- 11 Gradient test
- 12 Digit rollover
- 13 Stuck value
- 14 Density inversion
- 15 Grey list
- 16 Gross salinity or temperature drift
- 17 Visual QC – not mandatory in real time
- 18 Frozen profile
- 19 Pressure not greater than Deepest_Pressure +10%

Test 17 is not mandatory in real time. Tests marked * are also applied to trajectory data (see below).

An Argo “Grey list” of floats on which some sensors may have problems is available at the GDACS. Data from floats on the “Grey list” is not sent out on the GTS and should be treated with caution.

Profile data, data flags (see below) and probably trajectory data will at some stage become available using the BUFR format that is less restricted than TESAC.

Data from GDACS

For many users, GDACS should be the route to access Argo data. National data centers make Argo data available at the same time as on the GTS and to the GDACS.

The data on the GDACs are held in NetCDF format that contains profile and trajectory data and associated metadata and quality control flags.

These flags are:

- 0 No QC tests have been performed
- 1 Observation good
- 2 Observation probably good (implies some uncertainty)
- 3 Observation thought to be bad but may be recoverable
- 4 Observation thought to be bad and irrecoverable

Data are available at the GDACs in two versions:

1) **Fast mode data** has been subjected to initial quality control at national Data Assembly Centers (DACs). These data should be free from gross errors in position, temperature and pressure. The uncalibrated salinity data values are available and where a salinity offset is known it may appear as an “Adjusted salinity” variable in the same file on GDAC FTP servers. These data are identified with R in the “Data stream” variable if no adjustments were made and are identified with A if an adjustment was made. These files are available on the GDAC FTP sites. In general these data should be consistent with ocean climatologies even though no climatology tests have been performed.

2) **Delayed mode.** These data profiles have been subjected to detailed scrutiny by oceanographic experts and the adjusted salinity has been estimated by comparison with high quality ship-based CTD data and climatologies using the process described by WJO, Böhme and Send or OW¹. This process is carried out on a 1 year long “data window” and so no “Delayed Mode” observations can be less than 1 year old.

Accessing data

The data are held at the GDACs in NetCDF format and the GDACS provide a number of means of selecting data:

- From single floats
- From floats within a range of Latitude/Longitude and time
- Data type (R – Fast mode Data and or D – Delayed mode)

Which GDAC should I use?

This depends on a number of factors but the first issue is location. North American users will likely use the Monterey site and European users the Coriolis site. Beyond this the choice is a matter of which GDAC has a mode of access that best suits the user’s needs.

These access modes are described below:

USGODAE Monterey (See Annex 1)

Access to the complete Argo data collection, including float metadata, detailed trajectory data, and geographic and float specific Multi-Profile collection is provided via a DODs LAS, HTTP and FTP.

¹ Wong, A.P.S., G.C. Johnson and W.B. Owens, 2003: Delayed-mode calibration of Autonomous CTD profiling float salinity data by Theta-S climatology. *Journal of Atmospheric and Oceanic Technology*, 20(2), 308-318.

Böhme, L. and U. Send, 2006: Objective analyses of hydrographic data for referencing profiling float salinities in highly variable environments. *Deep Sea Research Part II* 53(1-2), 246.

The available selection and display tools are:-

- A profile location plot for all profiles returned by the query (may be plotted with or without float ID for queries returning many profiles)
- Download of selected profiles (in NetCDF Multi-Profile format) as a TAR file
- Plots of T-P and S-P for individual profiles
- Plots of float tracks for individual floats

Coriolis, Brest (See Annex 2)

Access to the complete Argo data collection, including float metadata, detailed trajectory data, and geographic and float-specific multi-profile collections is provided via a Distributed Oceanographic Data System Live Access Server (DODS LAS), HTTP and FTP. All float profiles can be visualized using a GIS tool.

A subsetting tool allows selection by

- profile type
- time and lat/long windows,
- measured parameter
- platform type
- real-time or delayed mode QC data

Access to additional non-Argo data (XBT, CTD, drifters, moorings, thermosalinograph, ADCP) is available from Coriolis through the same interface.

At present neither GDAC has the facility to select data from all floats that **have been** within a latitude/longitude box (including profiles from those floats when they are outside the box).

Argo data archive at US NODC

This is not yet fully operational but the NODC has plans to distribute Argo data on CD/DVD to groups without easy access to the internet.

Argo Regional Centers (ARCs)

Argo is establishing a number of regional centers. These will have multiple functions that will include:

- Performing regional analysis of all the Argo data in the region to assess its internal consistency as well as its consistency with recent shipboard CTD data.
- Providing feedback to PIs about the results of the regional analysis and possible outliers.
- Facilitating development of a reference database for delayed more quality control. This includes assembling the most recent CTD data in their region.
- Preparing and distributing Argo data products on a regular basis. The main data product will be a consistent Argo delayed mode dataset for their region, but other products might include weekly analyses of temperature, salinity and currents calculated from floats. Documentation of these products will also be provided.

The centers are identified as follows:

Pacific ARC: <http://apdrc.soest.hawaii.edu/argo/>

North Atlantic ARC: <http://www.coriolis.eu.org/cdc/Argo-NA-ARC.htm>

South Atlantic ARC: <http://www.aoml.noaa.gov/phod/sardac/index.php>
Indian ARC: http://www.incois.gov.in/Incois/argo/argo_dataregional.jsp
Southern ARC: http://www.bodc.ac.uk/projects/international/argo/southern_ocean/

Tools for assisting with Argo data handling.

Some people have difficulty working with NetCDF format files on the Argo GDAC servers. To encourage wider use of Argo data the Argo Information Centre has made an inventory of the tools that are available and links to the sources of this information. The links can be found from UCAR website (<http://www.ucar.edu/ucar>) or NetCDF documentation (<http://www.unidata.ucar.edu/packages/netcdf/index.html>). Users are encouraged to share the tools they develop with the rest of the Argo Community.

Annex 1

US GODAE Web site

Global Ocean Data Assimilation Experiment

Home | Data Sets | L.A.S. | USGODAE | Projects | Links | News | Contacts

Year: 2004, Month: 10, Day: 28, North: 90, West: -180, East: 180, South: -90

END: 2004, 10, 28

DAC: ALL, aoml, bodc, cstro, gts

FloatID: ALL

Output Type: Text List Only, Text List and Location Plot, Text/Loc Plot with ProfileIDs

Return Delayed Mode Profiles Only

Go Reset

USGODAE Argo GDAC Data Browser

NOTICE: 20020.414. Several changes have been made to the Argo GDAC data management. For details see the USGODAE Main Argo Page

This page provides access to the global collection of delayed-mode data from the Argo temperature-salinity profiling floats. The query results page provides:

- A profile location plot for all profiles returned by the query (may be plotted with or without float ID for queries returning many profiles)
- Download of selected profiles (in NetCDF Multi-Profile format) as a TAR file
- Plots of T-P and S-P for individual profiles
- Plots of float tracks for individual floats

For access to the complete Argo data collection, including float metadata, detailed trajectory data, and geographic and float specific

Annex 2 Coriolis Web page

The screenshot displays the Coriolis web interface for "global ocean profiles". The main feature is a world map showing the distribution of oceanographic stations, color-coded by type. A search panel on the right allows filtering by Zone (Lat/Long), Start date, End date, and Platform type. Below the map, there are controls for Platform type, Stations including, Processing level, Meta Data, Periodicity, Quality flags, and Output format. A data summary table provides counts for various platform types and their weekly distribution.

Platform type	Total number of stations
Argo floats	1036
XBT, CTD	105
Buoys	530
Moorings	1330
Others	246

Weekly distribution for selected platform types:

- Argo floats: 883 stations (Year/Week: ALL)
- XBT, CTD: 530 stations (Year/Week: ALL)
- Buoys: 1330 stations (Year/Week: ALL)
- Others: 246 stations (Year/Week: ALL)

Footer: Last Update 22/03/2005 © Coriolis