US NATIONAL DATA MANAGEMENT REPORT

21st ADMT

December 2, 2021 - December 1, 2022

1. Real Time Status

US Argo Data Assembly Center at AOML statistics

The US Argo Data Assembly Center (DAC) at AOML is responsible for processing of Argo data obtained from all US floats. During the reporting period the DAC has received real-time data from 2,092 floats and sent more than 84,200 profiles to the GDACs. In addition to this, the US Argo DAC distributed meta, technical and trajectory files in the Argo NetCDF files to the GDACs as part of the real-time processing.

The DAC distributed over 83,400 Argo profiles to GTS in the BUFR format. Both for GDACs and GTS 86% of the profiles reached the system within 24 hours. If floats with large delays are excluded (e.g. new deployments and floats under ice), then 96% of the profiles are available in 12 hours and 97% of the profiles are available in 24 hours.

The DAC also passes the files on to the GDACs that come from delayed-mode processing, BGC float processing and auxiliary files. For this purpose, the DAC maintains an ftp server for file exchanges, both for providing reprocessed R-mode and meta files as well as for receiving D-mode files, real-time submission of data from Iridium floats and the submission of deployment information.

Overall, the US Argo DAC has 1,433,744 R-files, 1,219,128 D-files, 94,192 BR-files, and 87,389 BD-files. The corresponding numbers for non-profile files are 8,320 meta, 8,114 tech, 8,105 Rtraj and 2,129 Dtraj files.

The US Argo DAC added 323 new floats to the processing system, 44 of them were deployed in collaboration between AOML and WHOI . As part of this collaboration, the US Argo DAC is finding ships of opportunity and provides ship riders for selected cruises. Recent maps showing their positions with link to graphics of the data collected by the floats can be found at:

https://www.aoml.noaa.gov/phod/argo/opr/php_forms/deployment_maps.php

The US Argo DAC is maintaining a website that provides documentation and information about the operations: http://www.aoml.noaa.gov/phod/argo/index.php

Developments at the US Argo DAC

During the current reporting period, two Argo team members moved on to a new career and we added two team members.

As in the past, changes in float technology or core Argo floats, sensor configuration on BGC floats as well as decisions by the IADMT, of which AOML is a major contributing partner, will be the main reasons for changes to existing software and the development of new software. The trajectory NetCDF file format version 3.2 has been defined in user manual 3.41 (July 2021). Adaptations for writing trajectory files in format 3.2 (in this format core and BGC data are in the same file) were completed in September 2022 and will be activated. This upgrade will go operational once the GDACs accept that format. In January 2022 we finished a decoder for radiometer data from APEX Iridium floats. In March 2022 we implemented revision of quality control procedures because the pilot phase or RBR CTD ended. In April 2022 we completed adaptations for the processing of TEMP_CNDC. Quality control and netcdf software capabilities were expended in May 2022 to process radiometer data from two APEX Iridium floats and made it operational for two floats in test mode (without distribution of our BR netcdf files to the GDACs), because MBARI is producing the BR files for these floats. This will allow comparison of our files with the files produced by MBARI. Finished core data decoding for a NAVIS Iridium BGC float type in September 2022 and started distributing the BR netcdf files to the GDACs. We also finalized implementation of real-time quality control procedures for BGC data in November 2022 (tests are applied for chlorophyll A, CDOM, BBP, nitrate, pH, irradiance) .

AOML continued to collaborate closely with the US Argo partners on the expansion of our BGC capabilities and provide feedback related to the new SOLO BGC float data processing. Revisions o the quality control and netcdf software to process these floats was made operational in February 2022. Four such floats were deployed in March-April 2022. Their core data are quality controlled and the netcdf files are created at AOML. MBARI creates the BR files. Because of the number of profiles collected in each cycle, and how they relate to the BGC profiles, changes were made to the quality control and the netcdf file writing software.

For all floats, AOML continues to create and distribute the bufr files, including oxygen when available.

To date 120+ BGC profiles have been collected by three APEX floats in the AOML-led Gulf of Mexico pilot array. Two 5-sensor APEX operating in the basin (one with a failed pH and one with a failed FLBB sensor). One APEX float was recovered upstream of the Florida Straits region in June of 2022 by small boat. Testing of this float at AOML in October 2022 suggested a vacuum failure and it has been sent back to the University of Washington for testing and refurbishment. One Navis BGC was received in June 2022 and sent to the University of

Washington for hull testing at pressure in light of recent float implosions before and after production of this float. Two additional APEX floats have been ordered (FY22) and are expected to arrive in summer of 2023.

The AOML DAC and AOML BGC D-mode operators are working together to develop a system to apply Sage and Sage-O2 determined adjustments in real-time to produce A-mode data.

2. Delayed Mode QC status

The US Argo DAC receives the Delay mode Argo profiles from US delayed-mode operators and verifies their contents to ensure soundness of the files if requested.

Each US Argo institution has provided information on their delayed-mode processing which was added to this report.

NOAA/AOML

AOML has conducted DMQC for the BGC parameters using Sage-O2 and Sage but continue to refine DMQC for the Gulf of Mexico array, specifically by testing the application of a Gulf of Mexico pH and nitrate reference layers and correction of pH profile offsets that occur around 1000 dbars when the CTD pump turns on in continuous mode. DMQC for core parameters is also underway, including an AOML effort to build out available high quality CTD data in reference databases for the Gulf of Mexico.

NOAA/PMEL

As of 28 November 2022, PMEL had 214,113 D-files at the GDAC that were more than one year old, comprising 82% of the total of 260,138 PMEL profiles that were older than one year at that time. Last year, on 28 November 2021, PMEL had 206,593 D-files at the GDAC that were more than one year old, comprising 86% of the total of 241,040 PMEL profiles that were older than one year at that time. So, Kristene McTaggart's DMQC efforts over the past year resulted in a net increase of 7,975 DMQC profiles for profiles older than one year, about 42% the 19,098 profiles that became older than one year during that time. This reduction in the DMQC rate was largely owing to the continued challenges of COVID-19, teleworking, and John Lyman working on another project last year that took much of his time. A focus on difficult cases identified by automated checking has also slowed progress. Next year McTaggart will continue her DMQC work, rejoined by Lyman.

The PMEL float DMQC procedure currently consists of the following steps: We perform an automated correction, with visual check, of reported pressure drifts and correction for the effect of these pressure drifts on salinity, as well as an automated correction of conductivity cell

thermal lag errors following Johnson et al. (2007). We do visual inspection and modification of quality control flags for adjusted pressure, temperature, and salinity using the SIO GUI and the Lyman GUI. We overwrite the raw Param_QC flags during this step as required. We use OWC Version1.1, currently with CTD (CTD_2021v1) and Argo (2020v03) reference databases, and adjust run parameters to get appropriate recommended salinity adjustments. Errors in OWC are computed directly from the least squares fit. We accept or reject the OWC recommendations on the basis of comparison with nearly historical profiles using a PMEL GUI written for this step.

Scripps Institution of Oceanography

Scripps Institution of Oceanography (SIO) has evaluated, as part of delayed-mode quality control (DMQC), a total of 335,000 Argo stations (profiles). This is an increase of 39,372 stations (1079 nominal float years) since the previous Argo Data Management Team (ADMT) Report (December 2021). This count represents 98.3% of the SIO DMQC-eligible stations (older than 12 months). The above numbers include SIO Core and Deep Argo floats and all Argo New Zealand floats for which SIO does DMQC.

SIO expects to maintain a high DMQC completion percentage during the coming year and will continue with a 7-9 month revisit schedule. In the past year, resources were allotted to allow the group to reduce the DMQC backlog that built over the past few COVID-impacted years. The DMQC backlog of SIO's Deep SOLO floats, following the Argo best practice CPcorr estimation, was reduced by about half. The SW Pacific and Indian Ocean Pilot-arrays were the initial Deep floats to be DMQC'd during 2022. The Southern Ocean and Atlantic arrays will follow in 2023. The consensus standard DMQC procedures for SOLO/SOLOII/Deep profile data were continued in 2021.

The timeliness of SIO real time data arrival at the GDAC this year has been uneven. SIO profile data collected between 1 Jan 2022 and 28 Nov 2022 reached the GDAC within 24 hours 95.1% of the time, and 86.9% of the time within 6 hours. Several hardware failures at SIO were the primary cause of the lower percentages as compared to previous years. During June 2022 the percent arrival was only 82.3%/70.5% within 24hrs/6hrs, due to the loss of SIO's primary Argo processing computer. The above timeliness calculation used the float surfacing time, so the temporal span includes the time of transmission, SIO SBD/directIP processing, and AOML DAC netCDF creation.

In 2022 the first IDG developed prototype 6-sensor BGC SOLO was deployed. The software development necessary to parse the transmitted data and distribute the result to AOML/MBARI DACs for netCDF creation was a significant endeavor.

SIO has actively participated in moving forward the priorities of the Argo Program during the year. A non-exhaustive list follows. Megan Scanderbeg's continued work to improve data access descriptions for users and to communicate more often with operational users. SIO continues to update the Argo Climatological Dataset for OW salinity calibration. John Gilson has worked with Annie Wong to provide an Argo-wide audit on the profile netCDF salinity adjustments. Nathalie Zilbermann and Dean Roemmich have worked with Seabird to improve the calibration

of the SBE61 CTD (0-6000dbar capability). Sarah Purkey and Jeff Sherman (IDG lab) have led the development of the SIO BGC SOLO.

University of Washington

Over the course of the present reporting period (January 1 through November 15, 2022), it was estimated that 559 floats were “active” in the University of Washington fleet and reported a total of 17,236 profiles. Active floats are defined as having recorded and sent at least one profile during the reporting period. A total of 31,691 profiles were processed in delayed mode during the reporting period, including 28,256 (89 %) from active floats and 3,435 from floats that are no longer active and presumed dead. Delayed mode profiles written during the reporting period included both newly-recorded profiles, profiles that required re-processing (e.g., change in calibration), and profiles that were accidentally or intentionally skipped during previous reporting periods.

Floats associated with the SOCCOM program recorded 2,837 new profiles and a total of 2,701 profiles were processed in delayed mode (includes both new and reprocessed files).

Floats associated with the GO-BGC program recorded 1,120 new profiles and a total of 628 profiles were processed in delayed mode (includes both new and reprocessed files).

Three RBR floats were deployed during the present reporting period. A total of 632 profiles from these floats have been processed in delayed mode, including 257 that were determined to have good salinity that required no adjustment (40 %), 36 that were determined to have good salinity after an adjustment (6 %), and 339 that were determined to have bad and unadjustable salinity (54 %).

In addition, a total of 4,664 dissolved oxygen profiles were written during the reporting period; these profiles were recorded by floats equipped with Aanderaa 4330 optodes and are not processed by the Monterey Bay Aquarium Research Institute. These floats are separate from the SOCCOM and GO-BGC programs. Approximately 3,390 (73 %) of these profiles were newly recorded during the reported period and 1,274 of the profiles were re-processed from older files that required new calibrations.

MBARI (Monterey Bay Aquarium Research Institute)

Biogeochemical data from 237 operational five- or six-sensor BGC-Argo floats are currently being processed and subjected to real-time and delayed-mode quality control by MBARI. This includes 119 active SOCCOM floats in the Southern Ocean, 77 active floats deployed as part of the Global Ocean Biogeochemistry (GO-BGC) array, and 41 active “SOCCOM-equivalent” partner floats in various locations. All float data is managed by Tanya Maurer, Josh Plant and Emily Clark.

A total of 92 BGC-floats managed at MBARI have been deployed throughout 2022 across all programs. 53 of these were GO-BGC floats deployed across 12 different cruises, and 29 SOCCOM across six cruises. These deployments included 57 five-sensor APEX, 21 five-sensor Navis, and four six-sensor BGC-SOLOII. Additionally, MBARI continues to assist with data processing and management for floats outside of these programs, including four APEX oxygen-only BGC-Argo floats deployed in the low-oxygen zone of the Eastern Tropical Pacific and two Navis five-sensor floats deployed near the Bermuda Atlantic Time-series Study (BATS) site. MBARI has also been involved in processing and management of data from various test-floats within the past year, including two dual-optode APEX test floats in the Eastern Pacific and two triple-optode Navis test floats in the North Pacific (dual- and triple-optode data are not yet available at the GDAC).

BR- files are being generated and transferred to the Argo GDACs for all five- and six- sensor operational floats at a frequency of twice per day. Delayed-mode quality control assessment of oxygen, pH and nitrate data is performed multiple times per year. BD-designated files generated at MBARI signify that at least a preliminary DM assessment has been performed, although BD* files are subject to updates periodically throughout a float's life. MBARI-developed MATLAB software used to perform BGC DM assessment is publically available through the SOCCOM github at https://github.com/SOCCOM-BGCArgo/ARGO_PROCESSING and methods are described in Maurer et al (2021), https://doi.org/10.3389/fmars.2021.683207.

In addition to the processing and dissemination of float profile data, over the past year the data team at MBARI has begun processing biogeochemical data from park phase on floats for which it is available. This includes bio-optical data from 279 APEX floats as well as pH and oxygen data on select Navis floats. Both APEX and Navis floats sample temperature, salinity, and pressure data at park depth. These data are being subject to real-time quality control and are being stored internally at MBARI. Integration into combined v3.2 trajectory files will begin in 2023 in collaboration with AOML.

The data team at MBARI is also working closely with sensor developers to prepare for modified and new biogeochemical sensor models. Deployment of Sea-Bird FLBBFL sensors which excite chlorophyll fluorescence at two wavelengths (standard 470 nm, and new 435 nm) will begin shortly. Preparations are also underway to process data from floats carrying the Pyroscience Pico pH sensor. Data from these new sensors will initially go in the auxiliary directories at the Argo GDAC until approved.

MBARI continues to generate a semi-annual audit on DOXY profiles to assist DACs with furthering the amount of adjusted DOXY data at the GDAC. Work is ongoing and international response to this audit has been successful thus far (information on the audit can be found on the MBARI ftp: ftp://ftp.mbari.org/pub/BGC_argo_audits/DOXY/). In addition, the team has been involved in further testing the application of the Bittig et al (2014; https://doi.org/10.4319/lom.2014.12.617) oxygen response-time correction on Argo floats with oxygen sensors and plans to endorse the use of this method in delayed mode or as an external product, depending on community feedback.

An updated temperature correction for NITRATE was recently developed at MBARI and represents an improvement to nitrate accuracy from the Sakamoto et al (2009) correction currently referenced in Argo documentation (doi 10.13155/46121). A manuscript for this method is in preparation, to be submitted to L&O methods within the first quarter, 2023. Additionally, documentation outlining processing and quality control methods for PH_IN_SITU_TOTAL were enhanced this year by the MBARI team and should be available publicly within the first quarter 2023.

MBARI continually supports the ADMT; Tanya Maurer serves as co-chair of the BGC-ADMT task team and MBARI data team members remain active in ADMT working groups focused on various BGC parameter topics.

Wood Hole Oceanographic Institution

WHOI Argo data management report covering the time period Oct 1, 2021 thru Sep 30, 2022. During this time, the WHOI Argo group deployed 123 floats. Of these there were 104 MRV S2A, 2 MRV Alto, 2 MRV Deep Solo, and 15 Navis-BGC. The size of the standing fleet averaged about 425 platforms. There are currently 258510 profiles reported to the GDAC, of which 239611 are eligible for DMQC. Of the eligible profiles, 95.8% have completed DMQC.

WHOI maintains two instances of our real-time telemetry decoder. The first operates on a server in Woods Hole while the second backup server operates in the cloud. Both of these servers are configured to submit data to the primary AOML DAC as well as the DAC's backup server. This system provides redundancy which has been exercised several times in past year with good success as we have managed to maintain data flow despite numerous downtime events of the primary servers at WHOI and AOML.

WHOI continues to contribute to the RBR Data Task Team, and the paper documenting recommended corrections. We are just beginning to work on applying these to our RBR float data. We have acquired 8 RBR ALTO's with onboard dynamic correction capability, which will be deployed in the coming year.

Work continues to identify Fast Salty Drifters. Of the 69 WHOI floats carrying SBE CTDs in the suspect serial number range, we have identified 21 CTDs suffering fast salty drift. New and improved software has been developed to deal with increasing diversity of operating float types including Deep and BGC. Additionally Deb West-Mack has made significant progress in development of protocols and software for performing DMQC on trajectory files. Salinity calibration in the Gulf of Mexico has been aided by development of a carefully screened set of bottle reference data. Other contributions to Argo data management include Global audits of salinity drift of each individual CTD (https//argo.whoi.edu/argo/sbedrift_wmo/) and global maps of fleet coverage (https://argo.whoi.edu/argo/maps/sparse).

3. Value Added items

University of Washington

A manuscript is currently being prepared that summarizes physical and biogeochemical data recorded from floats deployed in the Argentine Basin. It is intended that this manuscript will be submitted to a peer-reviewed journal in early 2023. Another manuscript that described the bias in Argo salinity was submitted to Earth System Science Data in September 2022. UW also contributed to a manuscript about the RBRargo^3 CTD, which was published in July 2022 (DOI 10.1175/JTECH-D-21-0186).

In ADMT activities, UW and SIO collaborated on an Argo salinity audit. This activity is expected to continue in future years.

4. GDAC Functions

US GDAC is up and running. Details will presented during ADMT23

5. Regional Centre Functions

Not applicable

6. Other Issues

Nothing to report