US NATIONAL DATA MANAGEMENT REPORT

21st ADMT

November 21, 2020 - December 1, 2021

STATUS

US Argo Data Assembly Center at AOML

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 87,000 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 86,900 Argo profiles to GTS in the BUFR format. Both for GDACs and GTS 92% 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 98% of the profiles are available in 12 hours and 99% 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 1,346,053 R-files, 1,131,360 D-files, 86,038 BR-files, and 81,448 BD-files. The corresponding numbers for non-profile files are 7,941 meta, 7,777 tech, 7,775 Rtraj and 2,079 Dtraj files.

The US Argo DAC added 287 new floats to the processing system, 28 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

2.3: Software Development at the US Argo Data Assembly Center (DAC)

During the current reporting period, one Argo team member moved on to a new career.

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). The format checker will be updated to check such files (target date June 2022). We already started developments needed for adding BGC data to the traj 3.2 files. We started this after setting up the data processing system for APEX Iridium BGC floats deployed in September 2021 by AOML (to receive their data the DAC set up a Rudics server; a secondary Rudics server is hosted by UW). This required expanding the decoding and quality control capabilities to include the processing of BGC beyond oxygen. The processing system for these floats handles nitrate, pH, chlorophyll-a fluorescence, and suspended particles. The data are processed following the cookbooks and the quality control manuals for oxygen as well as these new parameters and the resulting data files are distributed via the GDACs. Beyond revising the decoder, this required (1) updating the software that checks and expands the information in the meta files containing all information needed to process the data and write the meta files in compliance with Argo NetCDF standards; (2) updating the quality control software package; and (3) updating the software generating the profile NetCDF files. At this time, we are working on the expansion of the BGC processing system to handle radiometer data. AOML continues 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. Thanks to the recent hiring of two new team members that will get our team back to the size we need in the near future we will be able to make faster progress on pending work (e.g. MEDD test, adding error ellipse information to the trajectory NetCDF files).

We implemented improvements to the software checking the content of the ASCII meta files that are needed for the processing of float data, and we set up a system to expand their content by adding sensor pre-deployment calibration information from various files provided by US Argo groups. We also interacted with US Argo partners expanding to BGC to help them create such ASCII meta files for their floats (APEX, NAVIS, SOLO Iridium floats). The expansion to processing radiometer data is underway.

The US DAC moved the mirror processing system from a Unix server at a fixed location to a virtual linux server in the cloud. This mirror system can handle real-time processing and distribution as well as pushing delayed-mode data to the GDACs. It is running all the time, and can take over distribution if necessary (i.e., if the main processing system is down for more than a few hours).

Adaptations throughout the processing system were made to accommodate data from floats with the new RBR CTD. This sensor has an additional temperature measurement taken inside the conductivity cell, and it requires special treatment during the quality control. Changes to accommodate these data were needed in the profile, meta and trajectory data processing modules.

For float profiles that are received without a position, additional steps were implemented to allow adding an interpolated position if a profile with a position comes in at a later time.

All software was revised to accommodate cycle numbers exceeding 999, which became necessary due to improvements of float life times. The global range tests for pressure, temperature and salinity were updated after revisions were approved during ADMT-20. Similarly, adaptations were implemented to expand capabilities for storing more extensive information on the batteries in the floats in the meta NetCDF files.

We gave recommendations to facilitate using the relatively new quality control reports based on an min/max test.

DELAYED MODE QC:

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/PMEL

As of 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. Last year, on 18 November 2020, PMEL had 194,681 D-files at the GDAC that were more than one year old, comprising 88% of the total of 220,238 PMEL profiles that were older than one year at that time. So, John Lyman's and Kristene McTaggart's DMQC efforts resulted in a net increase of 11,912 DMQC profiles for profiles older than one year, about 57% the 20,802 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 a hard drive failure in early August which left us without a computer for months.

Lyman and McTaggart are continuing their DMQC work. Lyman is also continuing work on streamlining our DMQC GUIs and processing. Lyman has updated mapping in OWC for use with deep Argo floats. This was done by limiting the reference data set to CTD greater than 3000 m that are in the same basin as a given deep Argo profile. The process is made possible by using deep basin definitions provided by Sarah Purkey.

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 new PMEL GUI recently 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 295,628 Argo stations (profiles). This is an increase of 20,409 stations (559 nominal float years) since the previous Argo Data Management Team (ADMT) Report (November 23, 2020). This count represents 96.7% of the SIO DMQC-eligible stations (older than 12 months). The above numbers include SIO Core and Deep Argo floats, all Argo New Zealand floats, presently active NAVOCEANO floats deployed from the Peruvian vessel Zimic and 1 float donated to Argo Mexico.

SIO expects to maintain a high DMQC completion percentage during the coming year and will hope to return to a 7-9 month revisit schedule. The past year has been a challenge. Remote DMQC is a slower process due to bandwidth issues. Even with the ability to DMQC through remote access, personnel availability has been squeezed. The DMQC of SIO's deep floats have been delayed while the CpCorr and the pressure sensor studies are ongoing. DMQC of all SIO Deep SOLO floats should commence early 2022. 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 has been excellent. A recent analysis of SIO data collected between 1 Oct 2020 and 30 Sep 2021 showed 97.6% of profile data reaching the GDAC within 24 hours and 89.9% reaching the GDAC within 6 hours. This calculation used the float surfacing time, so the span includes the time of transmission, SIO SBD/directIP processing, and AOML DAC netCDF creation.

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.

The SIO IDG built and designed SOLOII/Deep SOLO float firmware has been unchanged over

the course of the year, except for minor internal bug fixes.

University of Washington

Delayed mode processing at UW has produced 28,817 D-files and 26,244 BD-files for the year 2021. These total file numbers include newly processed files as well as those re-processed due to changes in calibration or identification of anomalies that were subsequently corrected. In terms of active floats (those that have transmitted new data within the last two years), the UW fleet is >90 % up to date with respect to core (P, T, S) Argo profiles. With respect to both active and legacy floats equipped with DOXY, UW floats that are not handled by MBARI are >95 % up to date. Work has begun to prepare for the production of V3.2 Dtraj files, which will include surface air O2 for SOCCOM and GO-BGC floats. For other pilot data, UW has been active with other Argo groups in developing data management procedures for Deep Argo and RBR CTD. One activity of note at UW in 2021 was a global Argo delayed-mode salinity audit, which was done in collaboration with SIO, and which resulted in a best practices chapter for the QC Manual.

MBARI (Monterey Bay Aquarium Research Institute)

Biogeochemical data from 153 operational five-sensor BGC-Argo floats are currently being processed and subjected to real-time and delayed mode quality control by MBARI. This includes 102 active SOCCOM floats in the Southern Ocean, 31 active "SOCCOM-equivalent" partner floats in various locations, and 20 active floats deployed as part of the recently funded Global Ocean Biogeochemistry (GO-BGC) array. All float data is managed by Tanya Maurer, Josh Plant and new MBARI team-member, Emily Clark.

Despite impacts from COVID19, float deployments continued throughout 2021, although numbers were reduced compared to previous years. Only 1 SOCCOM cruise went out, resulting in the successful deployment of 10 5-sensor floats at the start of 2021 (all are operating). The GO-BGC multi-institutional float program was successfully launched, starting with the initial cruise along the A20/A22 GOSHIP lines in March-April, 2021. One of 12 APEX floats was declared dead upon deployment; all other floats continue to operate, despite a few isolated sensor issues which have been appropriately handled in the data system (through RT and DM QC measures). GO-BGC deployments are ongoing. In addition to SOCCOM and GO-BGC, MBARI continues to assist with data processing and management for floats within programs outside of these programs, including 2 WHOI 5-sensor Navis deployed in May, 2021 in the northeast Atlantic, as well as 1 Navis and 1 APEX NOAA/PMEL float deployed in the North Pacific in October, 2021.

MBARI has also been involved in processing and management of data from various test-floats, including 2 APEX with SBS83 optodes, 2 APEX with OCR, and one 5-sensor APEX with a Gasket DuraFET (GDF) pH sensor (a new pH sensor design developed by the Johnson lab at MBARI). All test deployments have been successful thus far.

BR- files are being generated and transferred to the Argo GDACs for all 5-sensor operational floats (Navis and APEX) at a frequency of twice per day. Delayed-mode quality control assessment of oxygen, pH and nitrate data is performed on a tri-annual basis. 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).

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/). Documentation outlining quality control methods for NITRATE (including real-time and delayed mode procedures) was produced this year as well with contributions from MBARI and recently published on the Argo Data Management Team web site (https://doi.org/10.13155/84370). Additionally, 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

During 01 October 2020 – 30 September 2021, WHOI deployed 57 new Argo floats representing three different hardware platforms: 48 MRV S2A, 6 MRV Alto, and 3 SBE NAVIS-BGC. Over this time period the active fleet consisted of 416 unique platforms reporting a total of 15,852 new profiles. The WHOI contribution to the GDAC now comprises 23,5958 total profile files (201,477 D-mode files, 3,4481 R-mode files). Of those older than one year (219,475 in total),199295 files are in D-mode status representing a completion rate of 90.8%. WHOI currently has 116 floats on the greylist for sensor issues (85% of which are active floats). We have identified at least 24 CTDs with conductivity sensor drift consistent with the 'fast salty drifter' cohort.