US NATIONAL DATA MANAGEMENT REPORT

21st ADMT

September 1, 2020 – November 21, 2020

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 2,613 floats and sent more than 99,00 profiles to the GDACs (timeliness statistics can be seen in Figure 1). 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 94,000 Argo profiles to GTS in the BUFR format (excluded from this are NAVOCEANO floats, which are sent to GTS by NAVOCEANO), where 95 % of them reached the system within the 24 hours. If floats with large delays are excluded (e.g. new deployments and floats under ice), then 94% of the profiles are available in 12 hours and 99% of the profiles are available in 24 hours. These numbers are almost identical to their counterpart for GDAC transmission.

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 1,258,118 R-files, 1,079,421 D-files, 68,816 BR-files, and 72,365 BD-files. The corresponding numbers for non-profile files are 8,172 meta, 7,476 tech, 7,505 Rtraj and 2,120 Dtraj files.

The US Argo DAC added 444 new floats to the processing system, 59 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

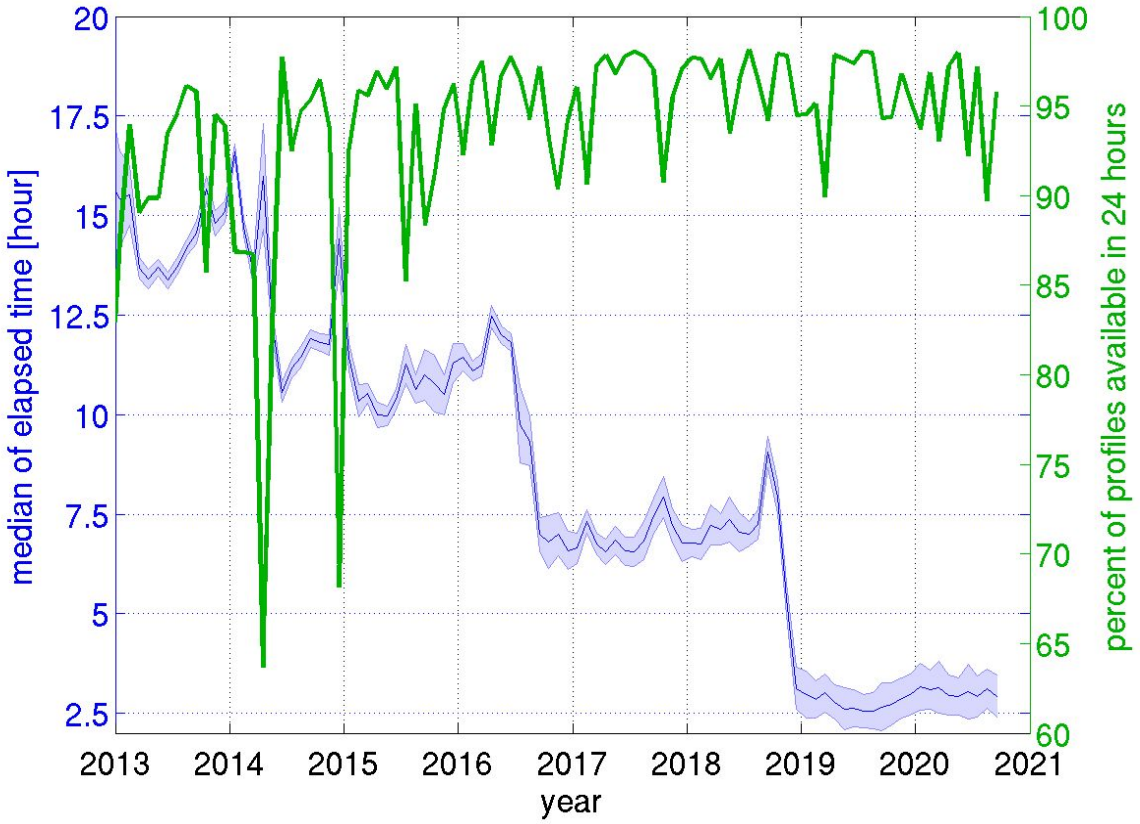

Figure 1: median of the difference between the time of the profile and the time of submission to the Argo GDACs (elapsed time, blue). The shading represents the standard error. The percentage of profiles available to users in 24 hours is also shown (green).

Developments at the US Argo DAC

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

During the current reporting period, two Argo team members retired and we hired replacements for them, who are making good progress. In addition to that we hired two team members with experience working with biogeochemical data that are taking on the development of the BGC data processing system in collaboration with the full US Argo team.

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. An upcoming change will be the addition of BGC data to the trajectory NetCDF file that currently only contains core Argo data (the redesign is currently being discussed and the specifications will be finalized after approval during the upcoming ADMT meeting). During this transition the DAC will also continue its work related to adding error ellipse information to the trajectory

NetCDF files.This became possible due to the transition to the web services approach for downloading the data from CLS (described below).

Implemented improvements to the software creating trajectory and meta files. For the former, the challenge was when certain measurement codes were missing in some data files (due to data transmission issues). For the latter the issue to tackle was discrepancies between the launch configuration provided by the float owner and the configuration transmitted by the float. The improved tracking improves the ability to have the best possible configuration information in the NetCDF meta files. for better data quality, error-tracking and messaging. Improvements were also implemented to trap cases with positions on land and/or times that are before deployment. This helps determine if the cause is bad data transmission or a mistake in the provided launch time or a combination of both.

The quality control reports based on an objective analysis continues evolving, partly due to our feedback. This evolution requires upgrades to our software handling the suggested revisions of the quality control flags.

US Argo deployed the first APEX Iridium floats with the RBR CTD. This float measures three pressure, temperature and salinity profiles. What remains to be added for such floats is the secondary temperature measured inside the conductivity cell. The DAC adapted the decoding and quality control software to accommodate these floats. The DAC also helped the hurricane floats group at AOML by developing a decoder for ALAMO floats purchased from MRV. The Argo team used its expertise running a parallel processing system for data from hurricane floats deployed by WHOI for this purpose. This decoder will come in useful as the SOLO Iridium float with BGC sensors are being developed.

With respect to BGC Argo, the US Argo DAC updated the oxygen processing in the decoders for APEX and NAVIS floats using Iridium for data transmission to comply with the specifications in the cookbook written by the BGC data management team. The impact of changes related to the computation of oxygen are shown in Figure AOML 8 (more information on this is given below). We also wrote the decoding components for pH, chlorophyll, backscatter, CDOM, and nitrate data from these floats (Figure AOML 7 shows which types of sensors have been installed on US floats) and came up with a design of the output files to be used during the quality control process. AOML is collaborating closely with MBARI, UW (the main groups with experience with APEX and NAVIS type BGC floats) and SIO (the developer of the SOLO type BGC floats) on standardized formats for this.

The US Argo DAC developed software to transition from downloading Argos data from CLS via telnet to a web-based system to comply with new security requirements implemented by NOAA. The Argo team helped the drifter group at AOML to implement this software, including producing output files that are compatible with the processing system used for their drifter data. Because there is a random issue with the web-based system, we are working with CLS to help them identify the problem by providing examples and sending them alerts when the problem occurs. For Iridium floats, the US Argo DAC transitioned the software to download emails with sbd data to use two-step verification for security reasons.

Information in the meta files controlling the software have been implemented to conform with ADMT requirements related to BGC sensors as well as now CTD sensors. This required changes to the software that creates meta NetCDF files.

Due to the 1024 week rollover problem of the GPS clock software changes were implemented to detect and correct the corrupted dates. The GPS rollover problem ended up resetting the float clock for some float types about 6 months after the rollover date (year was stored as two digits; therefore in October 2019 the year 2099 became 1999 which allowed a date reset by the firmware on the float).

- Implemented sftp based gateway for PMEL to directly submit meta files to AOML using automated script.
- Updated AOML's ftp download programs to be able to receive "dura" and "isus" files from University of Washington

Decoding software for APEX and NAVIS iridium floats has been revised to capture flaws in the received data files (for example, binary characters). This ensures that profile data can be decoded even if there are some bad data inside the msg files. Warnings from the quality control software allow for fixes to the files if required configuration data are corrupted and can be corrected based on the corresponding information from other cycles or the pre-deployment meta files.

The automatized process for new meta files needed to process all floats has been updated to ensure that the vertical sampling scheme provided is consistent with the number of profiles sent by a float and follows ADMT formatting requirements.

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 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. Last year, on 25 September 2019, PMEL had 186,159 D-files at the GDAC that were more than one year old, comprising 94% of the total of 197,302

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 8,522 DMQC profiles for profiles older than one year, slightly more than one-third of the 22,936 profiles that became older than one year during that time. This reduction in the DMQC rate was partly owing to the challenges of COVID-19 and teleworking since March. McTaggart's DMQC contributions were also reduced this year owing to a GO-SHIP field season that was made quite challenging by the onset of the COVID-19 pandemic.

Lyman and McTaggart are continuing their DMQC work. Lyman is also continuing work on streamlining our DMQC GUIs and processing. As an alternative to the SIO GUI1 routine, he has developed an alternative flagging routine that displays more windows with more plotting options, including plotting just the profiles that have bad flags versus all profiles that have been autocorrected. There is an option to QC all or some profiles, an option to save data in order to come back to it later without losing your work, and an option to view previously QC'd profiles. A very useful option displays buoyancy frequency to identify density inversions. This GUI allows us to evaluate profiles faster and with greater accuracy.

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. We overwrite the raw Param_QC flags during this step as required. We use COWG Version1.1, currently with CTD (2018V02) and Argo (2018V01) reference databases, and adjust run parameters to get appropriate recommended salinity adjustments. Errors in COWG are computed directly from the least squares fit. We accept or reject the COWG 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 275,219 Argo stations (profiles). This is an increase of 26,221 stations (718 nominal float years) since the previous Argo Data Management Team (ADMT) Report (September 17, 2019). This count represents 97.7% 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, 5 NAVOCEANO floats deployed from the Peruvian vessel Zimic, and 1 floats donated to Argo Mexico.

SIO expects to maintain a high DMQC completion percentage during the coming year and will continue to revisit the profile data of floats every 7-9 months. The consensus standard DMQC procedures for SOLO/SOLOII/Deep profile data were continued in 2020.

To increase the timeliness of arrival of data to the GDAC, SIO floats now are

telemetered via directIP as well as the traditional SBD emails. The directIP route has proven more reliable, although both channels continue to be used for complete data return. The data from the two are merged prior to parsing the data.

SIO has actively participated in moving forward the priorities of the Argo Program during the year, most notably by Megan Scanderbeg's continued work with the BGC trajectory file. SIO continues to update the Argo Climatological Dataset for OW salinity calibration. John Gilson has worked with Susan Wijffels (WHOI), updating the change in behavior over time (serial number ranges) of the SBE41 and SBE41CP CTD sensor stability. Nathalie Zilbermann, Sarah Purkey, and Dean Roemmich have worked with Seabird to improve the calibration of the SBE61 CTD (0-6000dbar capability).

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. A float utilizing the RBR CTD has been developed and deployed (5906383). It's performance will be assessed over the next few months.

University of Washington

Delayed mode processing at UW has produced 43,318 D-files and 10,216 BD-files for the year 2020. 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.

Lastly, work has begun to investigate dynamic corrections of salinity data from RBR 2000dbar CTD.

MBARI (Monterey Bay Aquarium Research Institute)

Since ADMT20 (Oct 10, 2019), MBARI has incorporated 48 new deployments into its BGC processing system. These included 45 APEX (APf11) and 3 Navis floats. 34 of these (including all 3 Navis) were deployed in the Southern Ocean as part of the SOCCOM project with oxygen, nitrate, pH, chlorophyll fluorescence and particle backscatter sensors, except for 10 that did not include bio-optics. Data from 3 similar APEX floats deployed recently in the California Current by NOAA/PMEL were also part of this year's cohort of floats incorporated into processing. Data from 11 APEX floats equipped with oxygen, pH, chlorophyll and backscatter, as well as PAL acoustic sensors that have been deployed over the past year in the equatorial Pacific for the NOAA TPOS program were incorporated into the processing system as well.

pH and nitrate sensors deployed on APEX floats are a mix of MBARI or Sea-Bird sensors, and all are calibrated at MBARI. MBARI data activities over the last year have been focused primarily on: (1) maintaining production for sensor build and real-time operational BGC float processing requirements, (2) working toward improvements to sensor design & performance, and (3) enhancing real time and delayed mode processing methods for select BGC sensor data. More details on these topics can be found in related documents and reports from ADMT21 BGC sessions. Scaling up the current system at MBARI has also been a focus. NSF funding for the SOCCOM project has been extended to 2024, and the NSF

Mid-Scale Research Infrastructure proposal to fund the deployment of 500 additional BGC profiling floats throughout the global ocean, a multi-institutional project termed the Global Ocean Biogeochemistry (GO-BGC) Array, has been granted. In anticipation of GO-BGC, the UW/MBARI team is starting to prepare for the processing of radiometry data which will be included on GO-BGC floats (in addition to the standard suite of BGC sensor data as is returned from SOCCOM floats).

Additionally, MBARI continues to generate Argo B-files in real time and delayed mode for submission to the GDAC through AOML and has been supporting AOML in the development of their own institutional BGC-processing capabilities. MBARI continues to conduct the oxygen data audit and has recently QCd and adjusted 30 historic oxygen floats deployed by WHOI and submitted the B-files to the AOML DAC. MBARI-developed BGC data quality assessment tools and reports, such as the oxygen audit and MATLAB SAGE and SAGE-O2 GUIs for DMQC of chemical data, are being utilized by other DACs as well to increase the level of adjusted BGC data on the GDAC available to users.

Wood Hole Oceanographic Institution

During the period October 1, 2019 to September 30, 2020, WHOI deployed 40 Argo floats and reported 15866 profiles to the GDAC from 416 unique platforms. The total number of WHOI profiles at the GDAC is now 221987 profiles: 173546 D-files (an increase of 30,000 since prior year), and 48441 R-files. Of the profiles eligible for DMQC, 83.8% have been completed.

The last of the SOLO WHOI ARGOS floats ceased reporting this year so the entirety of the WHOI fleet now communicates via IRIDIUM. The majority of floats are MRV S2A platforms (343 active) however WHOI continues with testing and limited deployments of the MRV ALTO platform with currently 22 ALTOs operating.

Deb West-Mack has made great progress against the backlog of overdue R-files and continues development of software to DM R-trajectory files for the suite of WHOI platform types. Sachiko Yoshida at WHOI continues to work on DMQC of core WHOI floats and NAS-funded floats in the Gulf of Mexico. Pelle Robbins has improved tools for tracking flow of real-time data and highlighting problems as they arise. Frank Bahr performs profile DMQC on WHOI floats in the Atlantic and has been developing tools for filling missing position information.

WHOI has helped coordinate the Argo RBR Data Task Team. Several Task Team virtual meetings have been held, and progress is being made on assessing drift, conductivity/pressure calibrations, testing dynamic error (both fast and slow)

correction schemes. Work is currently focussed on new pilot float data from CSIRO, which have just recently been deployed.

WHOI has also put in place the decoders and processes to receive, manage and forward deep and BGC data, in anticipation of deployments in early 2021. One issue that should be considered by the ADMT is the management and access of ship-based deployment profiles for deep, BGC and also RBR pilot floats. Should Argo work towards a more uniform approach to the storage and access to these data?