US NATIONAL DATA MANAGEMENT REPORT

December 1st 2017 – November 16th 2018

19th ADMT Meeting

at Scripps Institution of Oceanography

STATUS

US Argo Data Assembly Center at AOML, summary

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 data originated from 2,600 floats and processed more than 92,00 profiles in real time.

The US Argo DAC added 326 floats to the processing system, for 65 of them the deployment was done in a collaboration between AOML and WHOI . Recent maps showing their positions with link to graphics of the data collected by the floats can be found at:

ww.aoml.noaa.gov/phod/argo/opr/php_forms/deployment_maps.php.

The distribution of data to GTS in TESAC format was discontinued on July 2018. During that period of time were submitted approximately 46,400 profiles and 95% of those profiles were transferred within less than 24 hours of transmission. The US Argo DAC also has distributed over 79,000 Argo profiles to GTS in the BUFR format (excluded from this are NAVO floats), where 95 % of them reached the system within the 24 hours .

The distribution of real-time profiles to both GDACs was improved with about 96% of them available within 24 hours. 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 US Argo DAC is also receiving the delayed-mode data from US floats and passes them on to the GDACs. The US Argo DAC maintains an ftp server for file exchanges between the DAC and Delay Mode operators (both for providing reprocessed R-mode files and for receiving D-mode files) as well as for real-time submission of data from Iridium floats and the submission of deployment information.

The US Argo DAC has continued its involvement in deployment planning by finding ships of opportunity and providing ship riders for selected cruises.

The US Argo DAC is maintaining a website that provides documentation and information about the operations as well as some information on Argo-based scientific research:

http://www.aoml.noaa.gov/phod/argo/index.php

Developments at the US Argo DAC

The software for processing data from a modified format originally only used for RUDICS data (msg format) was adapted to process a new instrument type NAVISIR_TSOPJ2 deployed by University Maine. In addition to this, the software was improved to handle files with corrupted drift measurements, missing sampling schemes in meta files and a mix of data from two floats within a submitted msg file.

Tracking Bufr files being sent to GTS was implemented by looking for the submitted data in downloads from GTS. To accomplish this robustly, the program to read the content of bufr files (argobufrdump) was improved (to avoid segmentation fault caused by problems related to memory allocation).

The process to receive and distribute netcdf files (D, BR) was modified to add proper handling of BD and AUX files. The DAC also adapted the code used to check phy files decoded by data providers for inconsistencies and formatting problems (.e.g, problems with measurement codes, or the direction of a profile).This software was also implemented on our remote mirror system.

Because AOML has to have the capability to create BR-files for some floats while not replacing BR-files provided by the float owner, a new system was set up that checks if the US DAC generated file should be sent to the GDAC or not. A similar process was also necessary for meta.nc files, because the float owner may provide a D-moded meta.nc file that should not be replaced in cases where US DAC creates a new real-time meta.nc file.

The US DAC is downloading the error ellipse data from CLS in real-time and stores them in position files. A system to add these to the traj files is being developed.

The US DAC is providing access to its real-time netcdf files to the US Argo partners, the software that does this has been improved to increase its speed and reliability.

The US DAC migrated the whole operational system from a Unix computer to a much more powerful Linux computer and was thus able to increase the frequency of processing the Iridium data from 3 times a day to 14 times a day (implemented 11/15/2018) . This new computer is mirrored daily to a second computer at AOML. In addition to that, the mirror server at a remote location (far from Miami) was replaced with a computer that previously ran the processing at AOML to replace an older computer.

Our qc and netcdf file generating software was adapted to process data from ALTO Iridium floats. In addition to that, the software for the traj.nc file generation was updated to improve the quality of these files. The software for writing meta.nc file was updated as well to fix issues causing GDAC rejections and to include information from new sensors in these files. With respect to profile files, to handle cycles without a primary profile and create correct nc files in

these cases (NPROF=1 empty). Updates were implemented to accommodate new parameters in the meta.nc and tech.nc files as well as measurement codes in the traj.nc files.

An updated version of the Bufr generation program was implemented to encode BR files along with data from R files. This was done in close collaboration with Anh Tran from MEDS who developed the code. Our feedback lead to improvements of the encoder.

Processing system was updated to process data files from deep Argo floats deployed by PMEL.

The program to update flags based on the objective analysis test performed at Ifremer was redesigned and implemented to include additional checks and thus improve its performance and reliability. AOML also made suggestions on how to improve the provided reports.

A quite significant update was the adaptation of our profile nc generator to make BR and R files in format version 3.1 files to replace the version 3.0 files of the majority of those oxygen-only bio-Argo floats the US DAC is responsible for. This required ensuring that all required parameters are included in the BR v3.1 file, along with the quality control flags. BR files of Bio-Argo floats deployed by UW with bio-sensors in addition to oxygen are handled by MBARI. The floats impacted by this upgrade were reprocessed to replace the version 3.0 files at the GDACs. The only remaining 3.0 files from AOML are now from a few specialty floats: (1) floats for which each even cycle contains on the order of 7 short bounce profiles (103 floats; their standard cycles are in format 3.1). The other specialty floats are prototypes with two or three oxygen sensors (8 floats).

With respect to core Argo profile files, the software was modified to allow the processing of floats that record additional profiles (for example an extra profile for bio-data and two extra profiles for near-surface temperature and salinity). This required the capability to make profile files with N_PROF set to 3 or 4. Many other changes in 2018 involved adapting the quality control and file production software to run on a Linux computer, which required changes due to differences in compilers and libraries (e.g., format statements,, variable and array declarations and initialization of variables). Improvements were implemented to facilitate operational monitoring and more quickly and reliably identify problems as well as recognize floats that need to be grey listed due to failures of the frozen profile test. Another requirement that was implemented is the application for quality control flags from PSAL to CNDC. As required by ADMT, in cases where TEMP and/or PSAL is interpolated in secondary profile, these interpolated values are no longer being retained in the R files. Another thing that was improved was the determination of the resolution of data measured and reported by floats.

A float can have profiles for which and interpolated positions has to be derived. Typically, this is needed for floats under ice – but this is not always the reason. The processing system now attempts to distinguish under ice profiles from other cases, and to notify the operator about interpolated cases that are not identified as under ice for further analysis.

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 4 November 2018, PMEL had 161,550 D-files at the GDAC that were more than one year old, comprising 90% of the total of 178,865 PMEL profiles that were older than one year at that time. Last year, on 13 November 2017, PMEL had 113,795 D-files at the GDAC that were more than one year old, comprising 72% of the total of 158,072 PMEL profiles that were older than one year at that time. So, John Lyman's and Kristy McTaggart's DMQC efforts resulted in a net increase of 44,775 DMQC profiles for profiles older than one year, well over twice the 20,793 profiles that became older than one year during that time. Over the past two years, they have made excellent progress towards clearing the PMEL DMQC backlog.

That DMQC backlog arose mostly from delays owing to difficulties encountered during major maintenance and upgrading efforts on PMEL DMQC software in response to Argo format changes and internal IT requirements, as explained in previous reports. It took considerable time and effort to make these changes, and debug them.

John Lyman and Kristene McTaggart are continuing their DMQC work. John Lyman is also continuing work on streamlining our DMQC GUIs and processing. 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 OW Version1.1, currently with CTD (2014V01) and Argo (2014V04) reference databases, and adjust run parameters to get appropriate recommended salinity adjustments. Errors in OW are computed directly from the least squares fit. We accept or reject the OW recommendations on the basis of comparison with nearly historical profiles using a new PMEL GUI recently written for this step.

Scripps Institution of Oceanography

During the past year, Scripps Institution of Oceanography (SIO) has evaluated, as part of delayed-mode quality control (DMQC), a total of 231,512 Argo stations (profiles). This is an increase of 21,407 stations (586 nominal float years) since the previous Argo Data Management Team (ADMT) Report (November 06, 2017). At present, 98.8% of the SIO DMQC-eligible

stations have had their quality assessed. Here we define a station as being elibible for DMQC if the transmitted data is older than 12 months. The above numbers include SIO Core and Deep Argo floats, all Argo New Zealand floats, 7 NAVOCEANO floats deployed from the Peruvian vessel Zimic, and 2 floats donated to Argo Mexico.

SIO expects to be able to continue 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 2017.

During the year, the trajectory data from 52 end-of-life SIO Argos SOLO floats were delayed-mode quality controlled after which a Dtraj netCDF was created and passed to the GDAC. This process most notably included the estimation of float cycle timing, including float arrival and departure from the surface, and the full quality control of all Argos position data. There are now 956 DMQC trajectory netCDF ('Dtraj') data files available at the GDAC from SIO Argos floats (96% complete). DMQC on the few remaining Argos SOLO trajectory data will be ongoing as the floats cease transmitting data. The DMQC of trajectory files from SOLOII/Deep floats with Iridium data transmission is completed as part of the standard 7-9 month DMQC revisit pattern. The 'Dtraj' netCDF files from SIO Iridium floats include those cycles which have been DMQC'd as well as all subsequent transmitted realtime data, resulting in only a single necessary trajectory netCDF.

Although not often considered a DM file, the V3.1 meta file contains information shared between both the profile and trajectory netCDF, thus consistency across all three are required. Because of this fact, SIO has transmited DMQC meta files to the GDAC at the same rate as the trajectory files (97.8% total, 100% Iridium, 96.0% Argos).

Scripps has actively participated in moving forward Argo Program priorities during the year, most notably by Megan Scanderbeg's continued work with the Version 3.1 trajectory file. SIO continues to semi-annually update the Argo Climatological Dataset for OW salinity calibration. John Gilson has worked with Susan Wijffels (WHOI) and Annie Wong (UW/CSIRO) to assess and document the change in behavior of the SBE41 and SBE41CP CTD sensor stability within Serial Number ranges. Nathalie Zilbermann and Dean Roemmich have worked with Seabird to improve the calibration of the SBE61 used within the Deep Argo Program.

Scripps continues to work with float developers (IDG[1], MRV) to add capabilities to the SOLOII/S2A/Deep SOLO float types. The battery passivation evident in earlier SOLOII/S2A Iridium floats has been overcome with the transition to a new battery manufacturer: Tadiran hybrid lithium batteries. The first Tadiran battery pack float has been active for 3 years (completing 157 cycles), with no evidence of battery passivation nor other float reliability issues related to the introduction of the new battery type. All deployments since 25 Aug 2017 have been equipped with Tadiran batteries (215 to date). Several firmware upgrades have been added to the

SOLOII and Deep SOLO, including the diagnostic measurement of CPU temperature and relative humidity and the ability to profile on both ascent and descent (Deep SOLO only).

At present, Scripps has 47 active Deep SOLO floats spread over 4 Deep Argo pilot arrays. During the year, Scripps deployed 11 Deep SOLO floats to renew and expand the Southwest Pacific Deep Argo array. In addition, during the latest Deep SOLO deployment cruise on the R/V Kaharoa to the SW Pacfic array, two Deep SOLO were recovered in order to replace their failed SBE61 CTDs. After successful replacement, the floats were redeployed. During the year, a new Southern Ocean pilot array was initiated south of Australia with the deployment of 5 SIO Deep SOLO and a mix of other Deep Argo float models. All Deep SOLO data is reaching the GDAC/GTS within 24 hours of being received.

University of Washington

D-mode profile files for 304 floats were received.

MBARI (Monterey Bay Aquarium Research Institute) by Tanya Maurer.

Biogeochemical data from SOCCOM and pre-SOCCOM-equivalent floats are currently being processed and subjected to real-time and delayed mode quality control by MBARI (a total of 181 BGC floats). BR- files are being generated and transferred to the Argo GDACs at a frequency of twice per day. "Delayed mode" assessment of oxygen, pH and nitrate data is performed on a bi-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 now publically available through the SOCCOM github at https://github.com/SOCCOM-BGCArgo/ARGO_PROCESSING and is starting to be utilized by the international community. On July 15-19, 2018 a training workshop focused on real-time and delayed mode BGC processing and quality control procedures was held at the Second Institute of Oceanography in Hangzhou, China, led in part by MBARI personnel.

During the period of October 1, 2017 – October 30, 2018, 32 BGC floats were deployed as part of the SOCCOM array. During this same period, 9,712 BR* files and 8,779 BD* files were submitted to the GDAC. In addition, documentation outlining pH processing at the DAC level (including real-time and delayed mode procedures) was produced and published on the Argo Data Management Team web site.

Wood Hole Oceanographic Institute

During the period Oct 1st 2017 to Oct 30 2018, WHOI deployed 69 Argo floats and reported 14084 profiles to the GDAC. The total number of WHOI profiles at the GDAC is now 188980 profiles (138476 D-files, 50504 R-files). Of the profiles eligible for DMQC, 80.4% have been completed (138094 D-files, 33717 R-files).

In 2018, WHOI began limited deployments of a new platform, the MRV ALTO, an instrument with similar size of a SOLO-2 but with a potentially more efficient hydraulic system. The majority of the WHOI fleet remains composed of MRV S2A instruments (334 floats) and there are still a few older SOLO-WHOI floats active (23 floats).

WHOI also added a new employee, Deb West-Mack, to assist in data management and operations. She has been learning the OW analysis for DMQC as well as developing a protocol for addressing the backlog of R-trajectory files from the early SOLO-WHOI floats. The telemetry from this hardware provided very few time stamps. Using the known mission files, we are filling in timing information and MC codes to the best of our ability. Other improvements to software include additions to the WHOI decoder so that trajectory MC codes are now reported correctly in the real-time data stream. Sachiko Yoshida at WHOI continues to work on DMQC of NAVO floats, so far completing DMQC on 71 floats in the Arabian Sea.

Wijffels worked with Gilson, Robbins and Wong to do a global analysis of salinity drift against CTD serial numbers. This analysis confirmed the DMQC results seen in the UW and SIO fleet of a particular SN cohort developing a salty bias faster than normal and with larger frequency. This analysis has been shared with SBE and will be updated for the ADMT meeting.

In collaboration with float CTD manufacturer SBE and RBR, WHOI is also testing both RBR and SBE CTDs against a ship board CTD system on the RV Armstrong in November. Both dynamic and static errors will be analyzed. Three RBRArgo’s will be tested, and two SBE41CPs, one with a Kistler and one with a Druck pressure sensor. If the comparison is successful the results might be available for the AST-20 meeting.