Australian Argo National Data Management Report prepared for ADMT22, November 2021

Peter Oke[1], Joel Cabrie[2], Annie Foppert[3], Craig Hanstein[1], Catriona Johnson[1], Lisa Krummel[2], Jenny Lovell[1], Pat McMahon[1], Gabriela Pilo[3], Steve Rintoul[1], Tatiana Rykova[1], Christina Schallenberg[2], Roger Scott[1], Dirk Slawinski[1], Pete Strutton[3], Tom Trull[1], Esmee Van Wijk[1]
[1]CSIRO, [2]BoM, [3]UTAS

1. Status

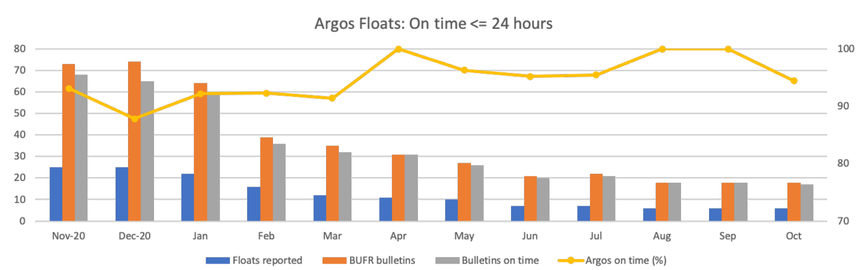
Data acquired from floats

We maintain an operational fleet of about 350 Core, 10 BGC, and 3 Deep floats for which we receive, process, and disseminate data 4 times per day. Our real-time (RT) system is run asynchronously at CSIRO and BoM (for redundancy), and CSIRO maintains primary real-time server in Hobart, and a secondary server in Perth.

Argo Australia celebrated the deployment of our 1000th Argo float on 10 May 2021 (Peter's 28th birthday). We raised a virtual glass to the team to celebrate, with a special acknowledgment to Susan Wijffels – who will forever be considered the queen of Argo Australia.

Data issued to GTS

All data is issued to GTS in near-real-time (NRT). Figure 1 shows the monthly timeliness of data delivery for floats with Argos and Iridium telecommunications processed by the legacy RT code (our Matlab-based system). The RT system continues to perform well, with an average of 91.4% of floats delivered on time – within 6 hours of measurement for Iridium floats.



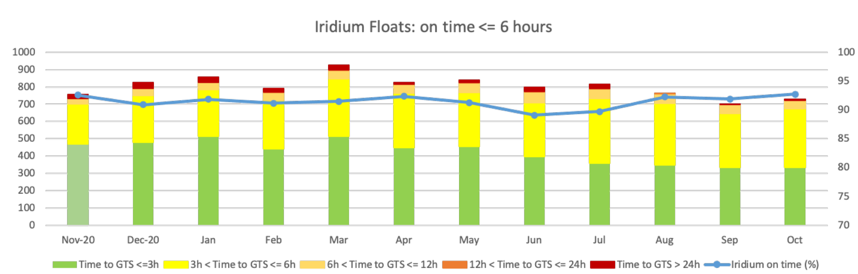


Figure 1: Timeliness of BUFR files delivery to the GTS for floats with Argos telecommunications (top), and floats with Iridium telecommunications (bottom). Data shown are monthly averages.

Figure 2 shows the daily timeliness of data delivery considering both the legacy RT code and our new RT code (our Python-based system). This figure includes floats with SBD telecommunications, in addition to the floats with Argos and Iridium telecommunications. Since June 2021, an average of 96% of the BUFR files were delivered to the GTS within 12 hours after float upload. In both February and July, CSIRO had a power outage, and no files were processed for a few days. This power outage, however, did not impact the monthly means substantially (Figure 1). The small dip in July was a result from internal upgrade that was performed on the servers.

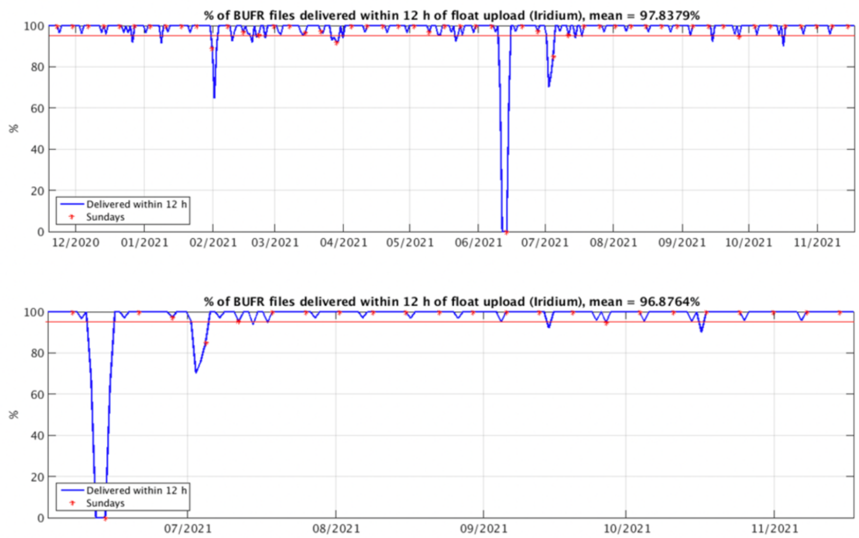


Figure 2: Daily percentage of BUFR files delivered within 12 hours of float upload between December 2020 and November 2021 (top) and June 2021 and November 2021 (bottom). The red vertical line indicates the 95% submission rate.

Data issued to GDACs after real-time QC

All data is issued to GDACs in NRT, with 96% of the data reaching the GTS within 12 hours of float upload. In the last year, a number of improvements have been made to our RT system including:

- Implementation of the global range test for temperature and salinity (in both our Matlab and Python system);
- Implementation of a RT PSAL adjustment, with values provided by the DMQC operators;
- Replacement of gradient test with MEDD test (in our Matlab system only, for now);
- RTQC and RT adjustments are now operational for NITRATE and DOXY variables on BGC floats;
- Correction of some format issues that were reported by the GDACs;
- Cross-validation of data produced in parallel by our Matlab and Python systems and addressed all inconsistencies;
- Implementation of CPcor salinity corrections for data from Deep floats in RT;
- Application of the greylist flags to trajectory files;

- Release of data from 10 Argo-equivalent floats that were deployed on the continental shelf off Antarctica to the GDACs;
- Removal of fully D-moded floats from the greylist; and
- We've addressed most of the actions assigned to us (and other DACs) from ADMT-21. Details have been included in the actions log.

Challenges in real-time processing and operations

We currently maintain two RT processing systems (legacy and new), adding some redundancy in the work related to RT processing. The transition between systems within CSIRO has started – with all floats deployed after May 2021 being only processed by the new system. The next big step for us is to implement the new RT system in the BoM, to ensure redundancy for all the new floats. There has been a delay in implementing the MEDD test in the new RT Python-based system, because the code available in the Argo community is in Matlab, and we still need to translate it to Python.

Over the past year, we've had several issues with float operations. We've had issues with battery passivation in floats, resulting in floats not sampling during ascent. We're referring to this issue as a "CTD switch fault". We have changed these float's mission, as per manufacturer advice, and some of the batteries have recovered. We currently have 16 floats under watch, hoping to recover their battery from passivation. We've also had two floats with problems in the high pressure valve that stops the float from descending to the full profile depth from park depth. We also have 17 floats with bladders that are not fully deflating, and are losing air during parking. This issue impacts the ability of floats to communicate when they reach the surface.

We've also encountered a large number of our APF11 floats are not getting a GPS fix when they surface after their initial profile. As a result, we haven't submitted position data for those profiles until the float transmits data after its second cycle. This is impacting our delivery of data to the GTS.

Data issued for delayed QC

A detailed response was provided for all floats identified in the core DMQC Audit (August 2021) and action has been taken to correct data as required.

Alerts from the Coriolis Objective Analysis are processed regularly and feedback is sent. All floats listed in the summary table of the monthly Anomalies Monitoring report are inspected promptly for addition to the grey-list or other appropriate action and feedback is sent.

There are 31 Australian floats currently listed on the FSD spreadsheet for which details were updated during October 2021.

Delayed data sent to GDACs

Between 1 January 2021 and 1 December 2021, we have performed DMQC on Core variables for 366 floats, producing almost 67,000 profiles that have been sent through to the GDACs. As of 1 December 2021, we estimate that of our fleet of 1021 floats (of which 964 are greater than 12 months post-deployment and are eligible for DMQC) that have collected almost 192,000 profiles, DMQC is currently required on 61 floats (with about 10,000 profiles). This indicates that data from about 95% of eligible floats have been DMQC-ed.

Web pages

We maintain several technical web pages that we use to monitor the status of our fleet, and the performance of each component of our operation. Details can be provided if anyone from the Argo community wishes to examine these, but they are intended for internal use.

Statistics of Argo data usage

Australian operational systems that use Argo data include:

- OceanMAPS, for ocean forecastwww.bom.gov.au/oceanography/forecasts/;
- POAMA/ACCESS-S, for seasonal prediction (poama.bom.gov.au);
- ReefTemp, for estimating the risk of coral bleaching for the Great Barrier Reef region (www.bom.gov.au/environment/activities/reeftemp/reeftemp.shtml);
- OceanCurrent, for estimating current ocean properties around Australia (oceancurrent.imos.org.au); and
- BoM's SST Analysis (www.bom.gov.au/marine/sst.shtml).

Scientific applications include:

- BRAN2020 (https://research.csiro.au/bluelink/global/reanalysis/)
- Tasman Sea reanalysis that is produced by Colette Kerry, from UNSW (Kerry et al. 2020).
- CAFÉ (https://research.csiro.au/dfp/cafe-csiro-decadal-prediction-system/) that is a contribution to a suite of WMO-endorsed annual-to-decadal climate predictions (https://hadleyserver.metoffice.gov.uk/wmolc/).

Products generated from Argo data

We've developed a system for estimating Argo float trajectories under ice at high southern latitudes. Details of this system are being prepared for publication. The system has been applied to the entire Argo array, including 521 floats and 2370 "gaps", where no observed position is recorded. An example of an estimated trajectory for a float that spent most of its operating life under ice is presented in Figure 3. The system is simple – identifying continuous trajectories that approximately follow contours of properties that we expect may be approximately preserved along a floats trajectory.

We compare results from our method to estimates that assume a linear trajectory for periods when trajectories are measured using GPS, and find a modest reduction of position error. But we find a significant reduction in the number of inconsistencies (e.g., when the bottom depth is less than the depth of the deepest measurement – see Figure 3a between year 2 and 3). The novel aspect of the new system is the use of multiple different constraints to estimate trajectories. We use potential vorticity, mean sea-level, and density at 1000 m depth. The system can currently produce three estimates of a floats trajectory under ice. We wonder if these estimates could be provided to the community routinely – perhaps at the end of each winter season – and could be considered for inclusion into D-files after assessment by a DMQC Operator. We're also considering using estimated trajectories from this system for R-files instead of linear estimates.

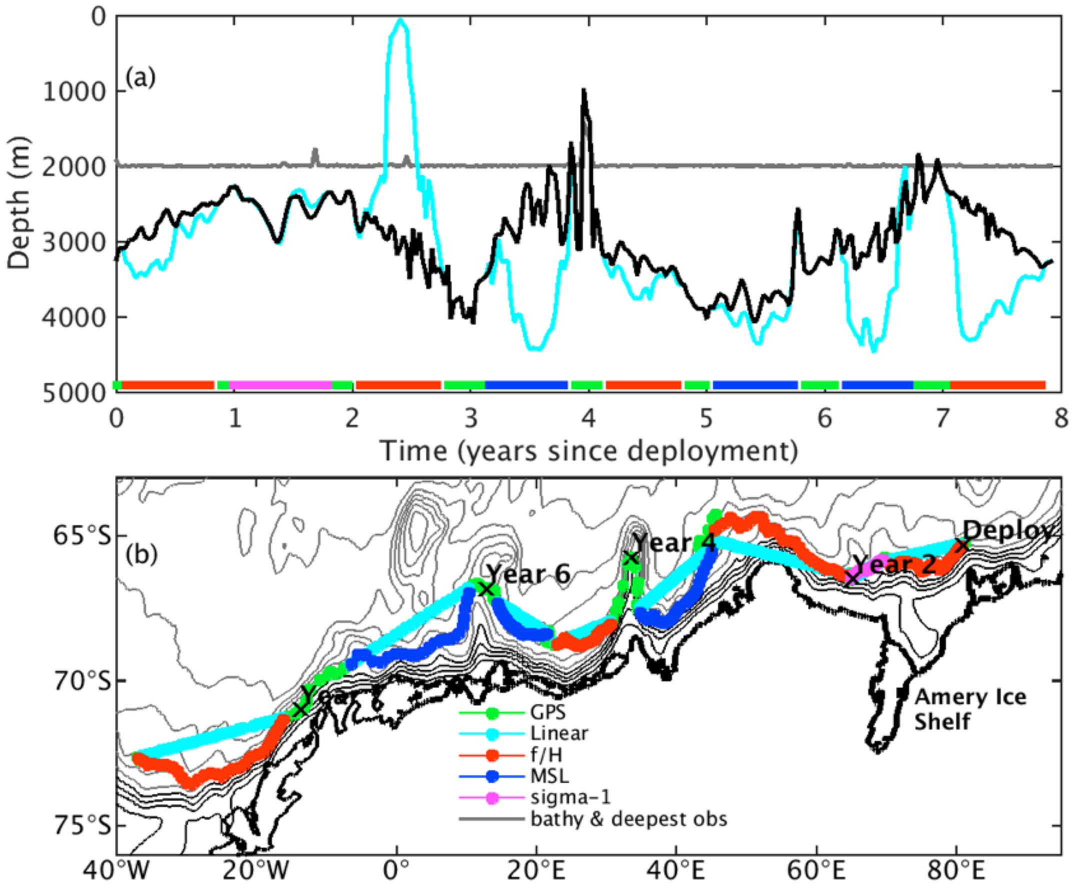


Figure 3: Example of the estimated trajectory for float 7900310. Panel (a) shows time-series of the depth of the deepest measurement (grey), and the bottom depth beneath the trajectory using a linear method to fill gaps (cyan) and using the multiple-constraints method (black) that is introduced here. Panel (b) shows the float trajectories measured by GPS (green), and estimated from the linear method (cyan), and the multiple-constraint method - showing when the constraint used is f/H (red), MSL (blue), and sigma-1 (magenta). The coloured lines along the bottom of panel (a) also show which method is used for estimating oat positions as a function of time. Time references along the trajectory are shown in panel (b) for several points along the float's path.

The Australian Argo team have developed a new product that is called Blue Maps. Blue Maps is a gridded product with 1/10 degree horizontal resolution that combines Argo data with satellite data to produce weekly maps. To date, Blue maps has been used to produce fields for the period spanning March 2006 and December 2020. Blue Maps is still an immature product that needs careful assessment, but it offers some features – including high-resolution, and synthesis of multiple observation types in a single step – that may prove worthwhile. Data from version 1p0 of Blue Maps is available at:
https://dapds00.nci.org.au/thredds/catalog/gb6/BRAN/Blue_Maps/MAPS-v1p0/catalog.html

The Argo Australia team were also involved in the production of BRAN2020 (Chamberlain et al. 2021). This is the latest high-resolution (1/10th degree) ocean reanalysis performed by the Bluelink team in Australia. The new aspect of BRAN2020, compared to previous versions, is the adoption of a multi-scale data assimilation. This approach has virtually eliminated the model bias that has plagued all previous versions of this reanalysis product. Data is available at:
https://geonetwork.nci.org.au/geonetwork/srv/eng/catalog.search#/metadata/f9372_7752_2015_3718

CSIRO has also updated the CSIRO Atlas for Regional Seas (CARS – internally referred to as ARGO_for_DMQC_2020V03) – producing an Argo-only product that grids all Argo data, using the methods described by Ridgway et al. (2002). Argo Australia use this climatology for DMQC.

Deployments

Argo Australia has deployed 68 floats between October 2020 and October 2021, including 9 floats with BGC sensors. Figure 4 shows the deployment location of these floats. Nine different vessels contributed to these deployments, and the floats were funded by IMOS, CSIRO, AAPP, IMAS, BOM, and the Royal Australian Navy.

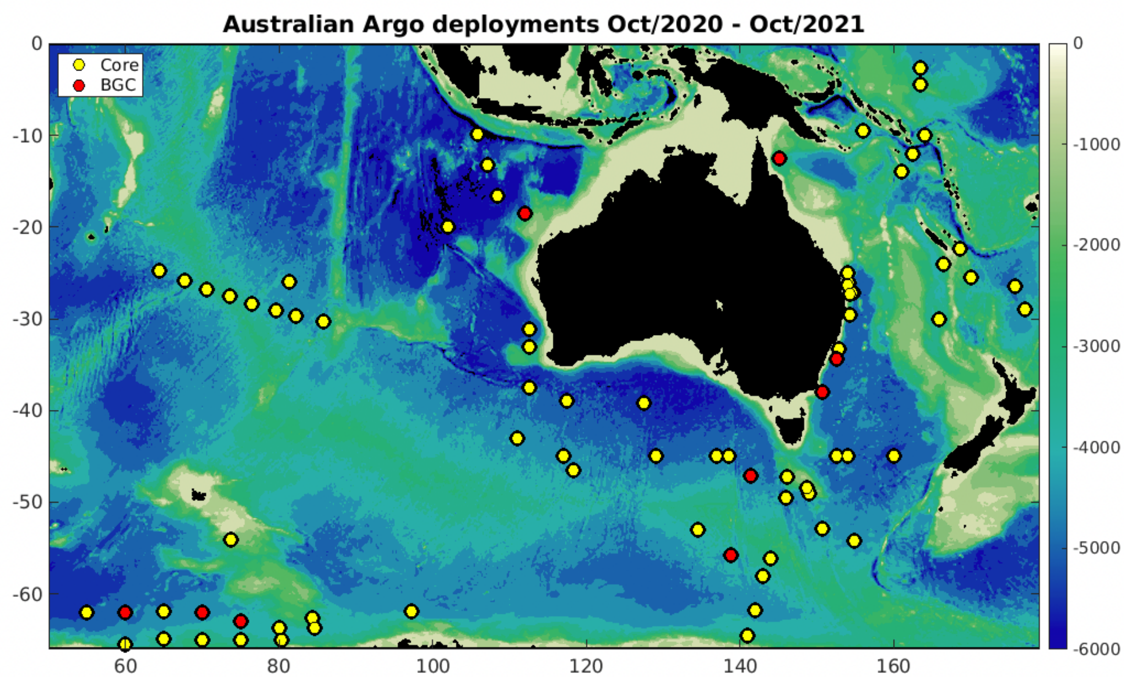


Figure 4: Map of Argo Australia deployments between 1 October 2020 and 1 October 2021, showing deployment locations for BGC floats (red) and Core floats (yellow). The background is bathymetry.

2. Delayed Mode QC

We have maintained a Matlab-based DMQC system that is employed by three DMQC Operators. Our DMQC system is maintained by Dirk Slawinski, and our active DMQC team includes Catriona Johnson, Jenny Lovell, Tatiana Rykova. The team is supported by Esmee Van Wijk for problematic floats. Our Matlab system is also used by Argo China (CSIO, Xiaofen Wu) and has previously been used by Argo India (INCOIS). Uday Bhaskar has recently obtained an updated version so we expect that he will resume using it again. The team provides occasional support to these external users to enable their activity, including code changes to accommodate different float specifics as well as trouble-shooting errors. These collaborations have resulted in a more robust and flexible code base.

Our DMQC system has long been used to perform DMQC on pressure, temperature, salinity, and dissolved oxygen (DOXY). Recently the workflow has been modified to allow the DOXY QC to be done by a separate operator after completion of the physical parameter QC, this is still in testing phase. We introduced this option to allow Christina Schallenberg to perform DMQC on our DOXY measurements without having to perform DMQC on core variables at the same time. We understand that decoupling the DMQC of DOXY and T/S/P isn't ideal, with a loss of information that may inform DOXY DMQC decisions. We may revert to always performing DMQC on T/S/P/DOXY together. But we're trying to play to the relative strengths of different team members, and we'll revisit our strategy regularly.

At present, our DMQC system cannot handle floats with data from five or six BGC variables, but a work-around that will allow us to still perform DMQC on core variables (T, S, P) and DOXY for those BGC floats is in progress. We plan to trial an intermediate file reformat so that the data can be provided to the software system in the form that is expected and then once BD files containing D-mode DOXY_ADJUSTED are created, they can be re-merged with the other BGC parameters to create the final BD files for submission to the GDAC and as a basis for DMQC of the other parameters. This script-based reformatting and merging will be undertaken in coming months and is necessary to do all the DMQC of Core variables of our BGC floats. Currently only two floats are impacted.

Except for DOXY, we don't plan to perform DMQC for BGC variables with our DMQC system. We use the SAGE system for DMQC of NITRATE and PH (with NITRATE adjustments made in RT based on the latest DMQC; the same is in the process of being implemented for PH), while DMQC of CHLA and BBP will be done offline at the end of a float's life. For CHLA and BBP, the adjustments and QC that are envisioned in real time (not fully implemented at our DAC yet, but work is in progress) should allow these variables to be used with confidence for scientific applications, so it's less crucial for DMQC to be done periodically throughout a float's life.

DMQC for the 3 deep floats has been performed independently (outside of our current DMQC software) and science results and data quality is reported on in Foppert et al. 2021 (submitted to JGR). We intend to update our DM system to cater for Deep floats in 2022.

Reference data for OWC has been updated to the latest releases. We are using OWC-V3 with ARGO_for_DMQC_2020V03 and CTD_for_DMQC_2021V01. We have not started using pyOWC but have engaged with the developers periodically to make comments on priorities. We have installed pyOWC and intend to test it during the next year. Delayed-mode PSAL adjustments are now passed to the RT processing system so that PSAL_ADJUSTED can be produced in real-time. We maintain a high priority for DMQC of floats that may be susceptible to fast salty drift and have updated the details for identified fast drifters on the FSD spreadsheet (currently 31 Australian floats are on the spreadsheet).

DMQC issues we've encountered include:

- We have several floats with RBR CTD sensors that are eligible for DMQC. But we're waiting on final recommendations from the RBR working group before we implement these corrections in our DMQC system.
- Considerable effort has been taken to separate the physical parameter workflow from DOXY DMQC and gain calibration to enable a specialist BGC DMQC operator to work on this data, work is ongoing to accommodate more complex data formats as they unfold
- During the installation process of the pyOWC code it was discovered particular versions of Python and libraries were required. Installing this on an "open" system seems to be fine, however, in an organisations like CSIRO, were our systems are managed to a specific security model, our installation required some effort by our IT Staff to install compliant versions of the various sub-components. A potential solution could be a Docker or Singularity container so all the required components were in place without the need to build pyOWC.
- For the PH variable, some uncertainty remains about how to implement the SAGE output with our RT Python code. But we're in the process of figuring it out (hopefully!).

3. GDAC Functions

Nothing additional to report.

4. Regional Centre Functions

Nothing additional to report.

5. References

Chamberlain, M.A., Oke, P.R., Fiedler, R.A., Beggs, H.M., Brassington, G.B. and Divakaran, P., 2021. Next generation of Bluelink ocean reanalysis with multiscale data assimilation: BRAN2020. *Earth System Science Data Discussions*, pp.1-31.

Foppert, A.,Rintoul, S.R., Purkey, S.G., Kobayashi, T., Sallee, J.B., van Wijk, E.M and Wallace, L.O. (2021). Deep Argo reveals bottom water properties and pathways in the Australian-Antarctic Basin. submitted to JGR.

Kerry, C. G. and M. Roughan, Powell, Brian, Oke, Peter (2020). A high-resolution reanalysis of the East Australian Current System assimilating an unprecedented observational data set using 4D-Var data assimilation over a two-year period (2012-2013). Version 2017. UNSW.dataset. DOI: 10.26190/5ebe1f389dd87.

Oke, P.R., Chamberlain, M.A., Fiedler, R.A., Bastos de Oliveira, H., Beggs, H.M. and Brassington, G.B., 2021. Combining Argo and satellite data using model-derived covariances: Blue Maps. *Frontiers in Earth Science*, *9*, p.485.

Ridgway, K.R., Dunn, J.R. and Wilkin, J.L., 2002. Ocean interpolation by four-dimensional weighted least squares—Application to the waters around Australasia. *Journal of atmospheric and oceanic technology*, *19*(9), pp.1357-1375.