Argo National Data Management Report for ADMT-25

Peter Oke[1], Lyudmila Koziy[1], Lisa Krummel[2], Jenny Lovell[1], Annie Foppert[3], Pat McMahon[1], Gabriela Pilo[1], Steve Rintoul[1], Tatiana Rykova[1], Christina Schallenberg[1], Roger Scott[1], Dirk Slawinski[1], Peter Strutton[3], Esmee Van Wijk[1]

[1]CSIRO Environment; [2]Bureau of Meteorology; [3]University of Tasmania

1. Real Time Status

We run our real-time data processing system 8 times per day – four times at CSIRO and four times at the BoM – with scheduling 3 hours out of sync. Our real-time system processes data from all float types: Core, BGC, and Deep. We produce data in NetCDF and BUFR. NetCDF data is immediately transferred to the GDACs, and BUFR data are immediately transferred to the GTS.

How many floats are you currently processing & what type are they?

Float family	**Number of versions**	**Number of floats*** (*approximate)
APEX	9	98
ARVOR	6	44
PROVOR		
Navis	1	158
BGC Navis		
SOLO/S2A		
Deep SOLO	2	11
Deep Arvor		
BGC APEX	3	4
BGC PROVOR	5	13
ALTO	3	10
Other (customize additional rows as needed)		

- How many different sensors are you currently processing?

Parameters	**Type(s) of sensor for that parameter**
Temperature/Salinity	SBE41CP, SBE61, RBR3
oxygen	Aanderaa4330
NO3	ISUS, SUNA
pH	DURA, SEAFET
Chla	FLBB2, FLBB_AP2, FLBBCD
bbp	FLBB2, FLBB_AP2, FLBBCD
irradiance	OCR504_ICSW
New Sensors you have begun processing (either deployed in past 12 months or expected in the next few months)	**Have all the Argo vocabularies been implemented to accommodate the sensor? (Yes, No, In progress)**
We have purchased a float with a UVP and a RAMSES hyper spectral sensor. We plan to deploy this float in March 2025 on the Denman Voyage.	

- What is the status of BGC processing and RTQC test implementation?

parameter	**Processing cookbook version you are using (ie, current or version 2.0 Oct 2018)**	**QC manual version you are using (ie, current or version 2.0 Oct 2018)**	**Notes on when changes will be made to update to latest version**
oxygen	current	Current (adjusted error in progress)	
NO3	current	current	
pH	current	current	
Chla	current	current	
bbp	current	current	

irradiance	current	current	

- What is the status of RBR data processing (if applicable)? Are you adjusting salinity in real time?

RBRargo3 2K model	**Are you filling Adjusted data (A mode) following User Manual 3.8 instructions?**	**Notes or additional information**
pre-April 2021	Yes	
post-April 2021	Yes	

- Are you regularly applying real time adjustments for the following items:
    - Salinity adjustments
    - Cpcor for deep floats
    - BGC parameters (if so, which ones)

	Yes/No for current R files	**Are you going back to make adjustments on all available R files when new adjustment comes in?**	**Notes or additional information**
Salinity adjustment	Yes	Yes	
Cpcor adjustment for Deep floats	Yes	Yes	Individual Cpcor values (calculated using shipboard CTD data) used for Rfiles when available (and for Dmode files). Otherwise, the most up-to-date default value in manual is used.
oxygen	Yes	Yes	Using SAGE and WOA2018
NO3	Yes	Yes	Using SAGE

pH	Yes	Yes	Using SAGE
Chla	Yes	N/A	Following latest RT recommendations
bbp	Yes	N/A	Following latest RTQC recommendations
irradiance	No	No	

- What data are you sending onto the GTS?
  - Currently core only

- What data is going to the aux directory?
  - UVP and FL2BB

- Are you automatically greylisting questionable floats detected by min/max test?
  - No

- What is the status of the transition to v3.2 trajectory files?
  - This has been done, but currently only enabled for BGC floats

- When do you think you will be ready to stop acceptance of v3.1 Btraj files?
  - We are producing v3.2 traj files for BGC floats but not all the parameter data is complete. We have the ability to produce v3.2 for core floats but are not routinely doing this yet, it can be done quickly (e.g. days rather than months) if v3.1 is made obsolete.

- Do you have any code to share with other DACs? If so, where is that available?

  - We have completed the development and testing of a Python version of our Real-time data processing system (*PyRT*). We transitioned all of our floats across to the *PyRT* system and (anxiously) turned off our Matlab RT system. The transition was smooth and the *PyRT* system works beautifully (phew), producing real-time alerts, automatically generating easy-to-read webpages, and provides a valuable reference for our DMQC team to assess float status, performance, and other relevant factors.

  - The *PyRT* system was developed at CSIRO and implemented on CSIRO servers. Once completed, we ported this system to the Bureau of Meteorology, on different servers, in a different environment. The transfer was relatively smooth. If other DACs were interested, we could discuss sharing the *PyRT* System.

2. Delayed Mode QC status

This section of the report is for reporting on the status of DMQC in your country and is the place to share your progress, your challenges, your concerns and any links to shareable tools or code. The following questions to help guide you:

- What is the status of delayed mode trajectory files? Have you created any d-mode trajectory files? If not, what are the reasons? If you have, would you be interested in sharing your experiences with others?
    - Some of our DM effort has been allocated into this task and we have progressed with tools for visual assessment of some trajectory variables. However, we have not produced any submittable d-mode trajectory files as of this time. The effort so far has shown that the r-mode trajectory files produced by our *PyRT* system are of high quality and pass most of the standard tests for CYCLE_NUMBER and JULD. We have identified that a broad knowledge of float operations and the RT data systems is <u>required</u> to do this task well and efficiently. A plan is in place to progress this task and we anticipate producing d-mode trajectory files in the next year.
    - We have engaged with WHOI and Scripps for assistance and code sharing. As our RT system is Python based and our collaborators are Matlab based this task will focus on methodology and process sharing. Matlab and Python code sharing will be an adaptive process as we integrate the methods into our respective systems. Our codebase will be managed in a Git repository and could be shared at a later date.

- How are you implementing BGC d-mode - by parameter or one expert does all parameters?
    - We have one expert doing QC for all the BGC parameters, as such BGC DMQC is not as far along. The exception being DOXY, which has been DMQC'ed on >90% of our dead floats. We have prioritized getting the RTQC and RT adjustments correct for BGC parameters, including the use of SAGE to estimate the adjustments. The final step to follow will to DMQC the data. The code is ready to do DMQC of nitrate, pH and oxygen. We expect to start rolling this out for dead floats in the next 6 months, and then step it up from there.

- What challenges have you encountered and how have you dealt with them?
    - The CSIRO DM processing system is an integrated set of Matlab scripts, running on a Linux virtual machine. It relies heavily on graphics and GUI-based interfaces. The GUI elements of the code are incompatible with more recent versions of Matlab. Current Linux operating systems have been progressively removing support for the old Matlab version required to run our codebase. A significant effort this year has gone into uplifting the code to run using contemporary Matlab versions (2020 and beyond). This has involved re-writing much of the interface and plotting to adhere to modern Matlab standards and to

improve the speed of graphics rendering. The modified code is in the testing phase and will be brought into production mode over the next 2-3 months.

- Do you have any code or tools you'd like to share with other DM operators? If so, where is that available?
    - Our DM code is used by China-Argo and has been used by India-Argo (INCOIS). We have supported China-Argo to perform DMQC on floats with RBR CTDs. Once we have fully tested the new version of our codebase, we will make this available to China-Argo if they wish to upgrade.

- Do you have any concerns you'd like to bring to the ADMT?
    - We have switched off the sampling on ascent for our deep floats, because we learnt that the current firmware does not possible allow decoupling of sampling on descent and ascent sampling, (i.e., we can't just sample the top 500 m on ascent). This is a significant power drain, reducing the operational life of deep floats. We understand that IDG has some test floats in the water where ascent sampling can be programmed independently. When that firmware becomes available on the MRV floats we intend to implement it to sample the upper layer on ascent. We acknowledge (and regret) that this means that our Deep floats are not contributing to the OneArgo mission as intended.
    - Not a concern, but we can report that the CSIRO-led Argo DMQC Discussion series has continued, with a virtual meeting every two months. These meetings continue to be well-attended. A record of meetings and discussion topics is at: https://www.marine.csiro.au/argo/dmqc/html/ArgoDM-Disc.html.

3. Value Added items

- List of current national Argo web pages, especially data specific ones
  - IMOS-OceanCurrent plots Argo data in RT around Australia for comparison with other ocean data (satellite and in-situ). The images are used for decisions in marine operations (science, fisheries, aquaculture, recreational activities), training (universities, data management courses, national navy), and media/outreach. Example of Argo plot: https://oceancurrent.aodn.org.au/profiles/cycle.php?wmoid=5905513&cycle=85&depth=0

- Known National Argo data usage
  - Please list known operational centers using Argo data in your country in this table:

Operational center	**Contact (name, email), if known**	**What data do they use?** (for example, core, BGC, all profile data, trajectory data)
Bureau of Meteorology (OceanMAPS, Australia's operational short-range ocean forecast system, www.bom.gov.au/oceanography/forecasts/)	Lisa Krummel (lisa.krummel@bom.gov.au) Gary Brassington (gary.brassington@bom.gov.au)	Core Argo data (including T/S from Core, BGC, and Deep)
Bureau of Meteorology (POAMA, Australia's operational seasonal prediction system, www.bom.gov.au/climate/ocean/outlooks/)	Lisa Krummel (lisa.krummel@bom.gov.au) Debbie Hudson (debbie.hudson@bom.gov.au)	Core Argo data (including T/S from Core, BGC, and Deep)
Royal Australian Navy, CSIRO (ROAM, Australia's Relocatable Ocean Atmosphere Mode – a ocean-atmosphere-wave forecast system run by the Australian Navy)	Emlyn Jones (emlyn.jones@csiro.au)	Core Argo data (including T/S from Core, BGC, and Deep)

- Products generated from Argo data that can be shared

  Global Ocean Reanalysis
  CSIRO has perfoprmed a new global ocean reanalysis, called BRAN2023 (version 2023 of the Bluelink ReANalysis). The latest product is complete and is undergoing final quality checks before public release. Once released, BRAN2023 data will be freely available (alongside all previous versions of BRAN) at: https://dapds00.nci.org.au/thredds/catalog/gb6/BRAN/catalog.html.

  BRAN2023 includes several incremental improvements to the data assimilation system, and also includes biogeochemistry (BGC). This is the first version of BRAN that includes BGC. As a first step, the BGC fields are _not_ constrained by observations. We constrain the physical fields (temperature, salinity, velocity, and sea-level) and allow the BGC fields to evolve freely. We plan to use BGC Argo data, along with satellite data, to evaluate the BGC fields in BRAN2023. Initial qualitative assessments show that the BGC fields realistically respond to the reanalysed circulation (Figure 1).

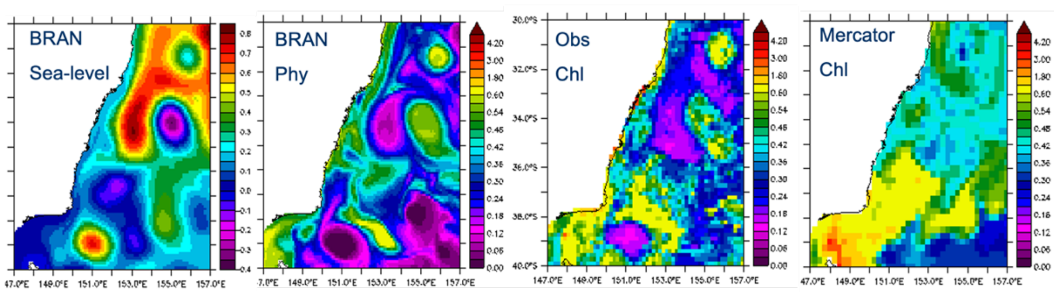


  Figure 1: A qualitative comparison of phytoplankton from BRAN2023, observed Chlorophyll-a from satellites, and reanalysed Chlorophyll-a from Mercator Ocean's reanalysis product (see labels). Sea-level from BRAN is included below to provide the oceanographic context.

  Argo trajectories under ice
  CSIRO produced estimates of Argo trajectories for floats that sampled under ice in the southern hemisphere. Data are available at: https://zenodo.org/records/6571146.CSIRO plan to provide an updated dataset in early 2025. The paper describing the method used is at:
  Oke, P.R., Rykova, T., Pilo, G.S. and Lovell, J.L., 2022. Estimating Argo float trajectories under ice. *Earth and Space Science*, *9*(7), p.e2022EA002312.

- For the information of the ADMT, we wish to report that a Special Issue of Frontiers in Marine Science on "Demonstrating Observation Impacts for the Ocean and Coupled Prediction", has been published (most articles being considered have either been accepted, or will soon be finalised). This special issue is an initiative under SynObs (led by Yosuke Fujii and Elizabeth Remy). Articles of relevance to the Argo community include:

- Balan-Sarojini, B., Balmaseda, M.A., Vitart, F., Roberts, C.D., Zuo, H., Tietsche, S. and Mayer, M., 2024. Impact of ocean in-situ observations on ECMWF sub-seasonal forecasts. *Frontiers in Marine Science*, *11*, p.1396491.
  - This article demonstrates that with-holding Argo data from coupled sub-seasonal predictions has an impact on the lower and upper atmospheric circulation. This is one of the first demonstrations of the impact of Argo on this type of system. The authors acknowledge that the skill of the coupled predictions is significantly degraded by problems with “initialisation”. This problem is common – and arises when the ocean and atmospheric states are not in balance. After the coupled system is initialise, both the ocean and atmospheric states adjust, and as a result, some information from observations (including Argo) is lost. But the impacts demonstrate provide clear evidence of the potential of Argo data for improving sub-seasonal coupled predictions.
- Balmaseda, M.A., Balan Sarojini, B., Mayer, M., Tietsche, S., Zuo, H., Vitart, F. and Stockdale, T.N., 2024. Impact of the ocean in-situ observations on the ECMWF seasonal forecasting system. *Frontiers in Marine Science*, *11*, p.1456013.
  - This article demonstrates that in situ ocean observations, including Argo, have a “profound and significant impact” on oceanic and atmospheric variables. The authors attribute changes of atmospheric variables to impacts on (i) air-sea interactions, (ii) dynamical imbalance, (iii) impacts of large-scale SST gradients, and (iv) impacts of warm pools (on deep atmospheric convection). The study found that Argo data in the Atlantic Ocean and extratropics had the most impact, and that many impacts on atmospheric variables were “non-local”.

- Publicly available software tools to access

  NA

4. GDAC Functions

NA

5. Regional Centre Functions

NA

6. Other Issues

For the information of the ADMT, we wish to report some metrics from our program.

Here we record the locations of floats deployed under the Australian Argo program over the last year (Figure 2), along with locations of planned deployments over the next year (Figure 3). We also record some details of the performance of our new real-time data system (Figure 4-6) to demonstrate the reliable performance we have achieved with this system.

Re: Deployments over the past year (Figure 2)

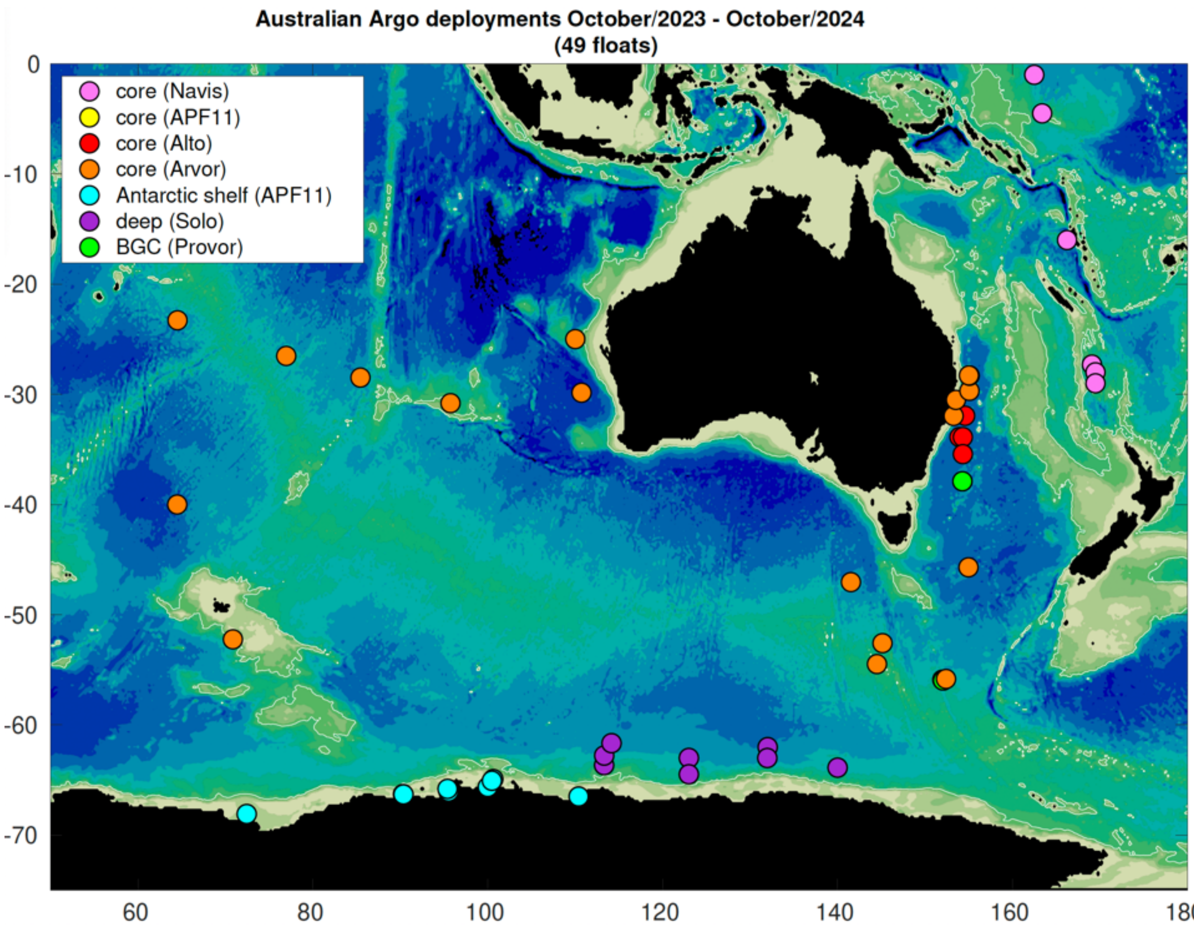


Figure 2: A map of the locations of Argo float deployments between October 2023 and October 2024, indicating float types

Re: Planned deployments over the next year (Figure 3)

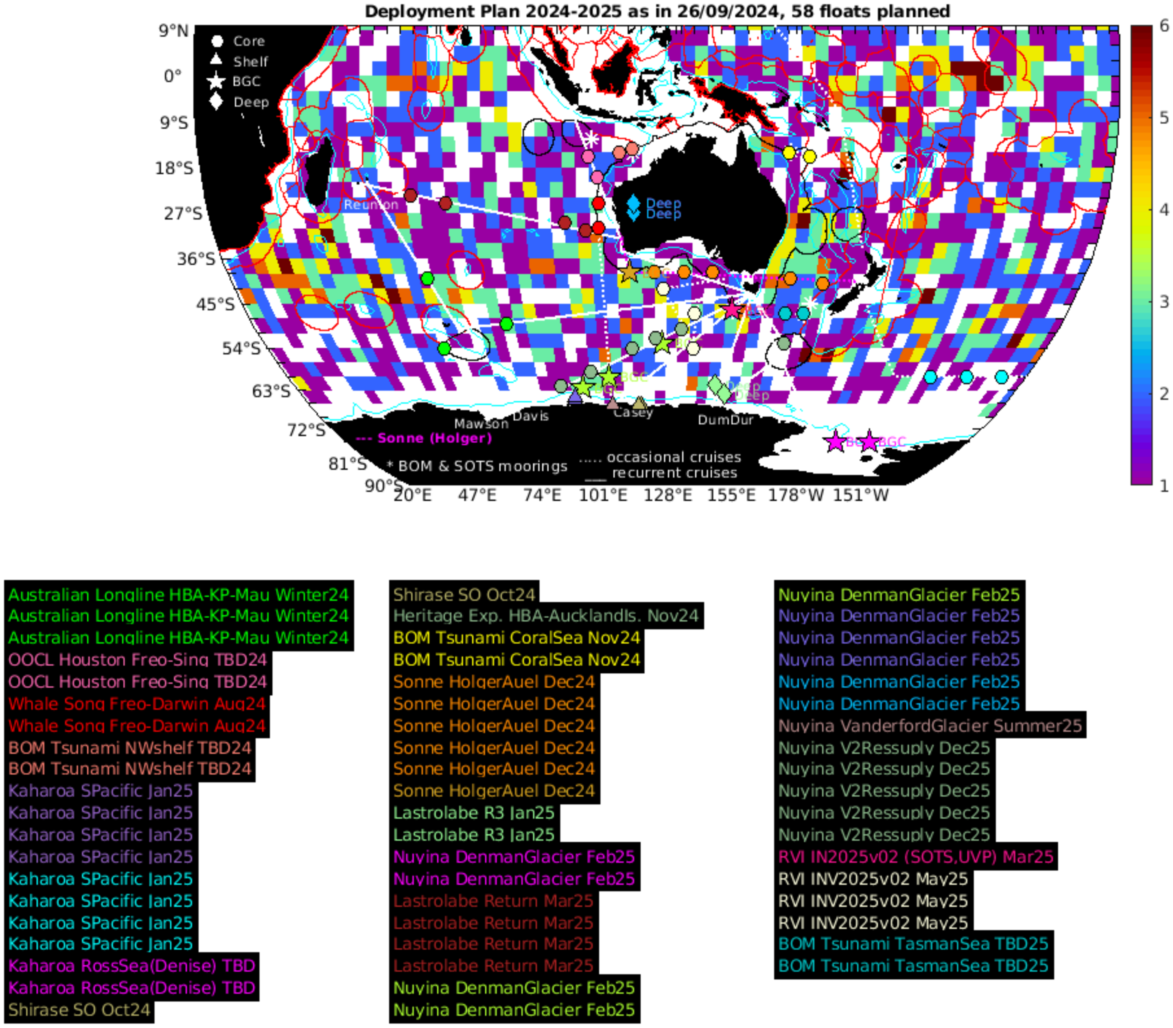


Figure 3: Deployment plans for 2024-2025. We plan to deploy 58 floats, including 8 BGC floats (stars, 6 from Australia and 2 from New Zealand), 4 deep floats (diamonds, 2 are still to be allocated), and 44 core floats (circles and triangles). Background colours denote the number of floats within each 3x3deg box in July-August 2024; the cyan lines denotes the 2000 m isobath; the red lines denote international EEZs; black lines indicate Australia's EEZ. Solid and dashed white lines indicate common ship routes.

Re: Real-time data processing and dissemination (Figures 4-6)

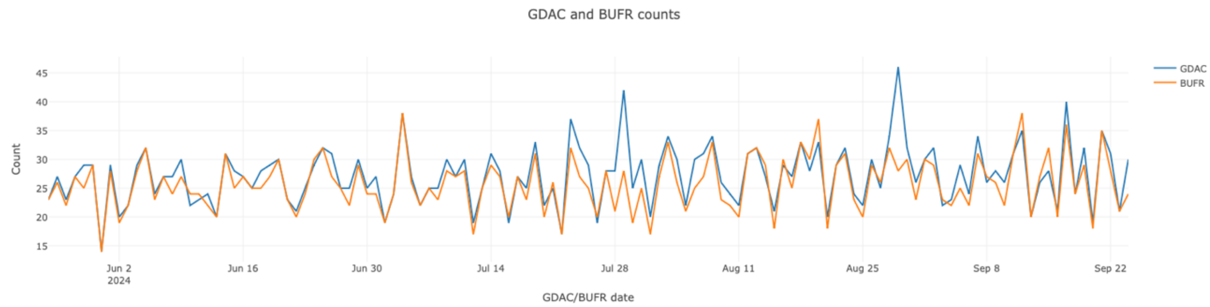
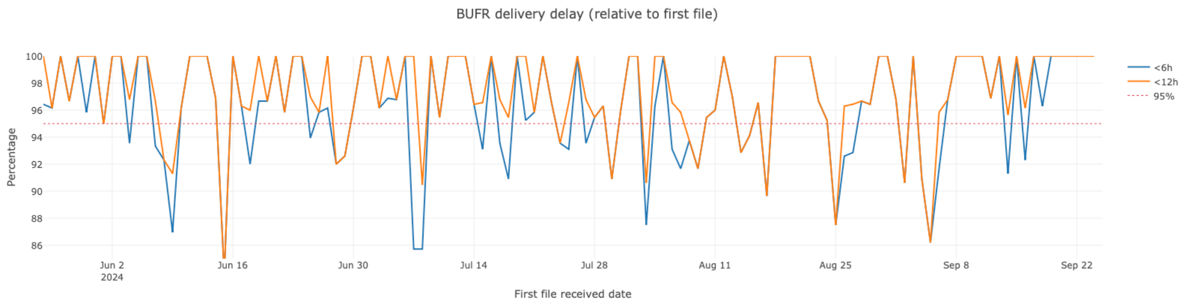


Figure 4: Time-series of the GDAC and BUFR file counts since the end of May 2024 (when we switched to the *PyRT* system for real-time data processing), showing the number of GDAC (blue) and BUFR (orange) files created daily.

Figure 5: Time-series of the percentage of data processed and disseminated within 6 hours (blue) and 12 hours (orange). The 95% level is denoted by the dashed line. The top panel shows the percentages delivered relative to the time the “first file” is sent; and the bottom panel shows percentages delivered relative to the “last file” sent. The difference between the results in the top and bottom panel indicate that delays are due to problems with data transmission, and not due to delays related to the data processing system.



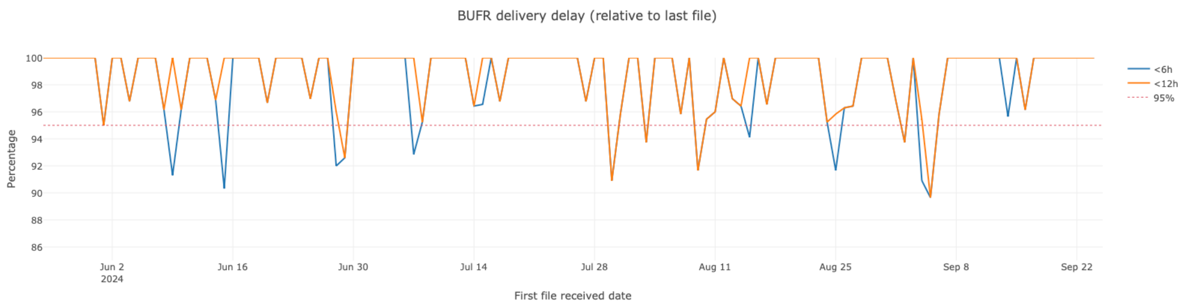


Figure 6: As for Figure 5, except for BUFR files.