

UK Argo National Data Management Report 2017

Report to the Argo Data Management Team – ADMT-18

Author list

The UK Argo data team (British Oceanographic Data Centre, National Oceanography Centre)

Contributing authors: Matt Donnelly, Katie Gowers, Robin McCandliss, Justin Buck

Other team members: Violetta Paba, Clare Bellingham, Paul McGarrigle, Elizabeth Bradshaw

With contributions from the wider UK Argo team by:

Jon Turton and Fiona Carse (UK Met Office)

Brian King (National Oceanography Centre)

Giorgio Dall’Olmo (Plymouth Marine Laboratory)

Status

The British Oceanographic Data Centre (BODC) is the data assembly centre for UK Argo funded primarily by the UK Natural Environment Research Council (NERC) and responsible for data management of UK, Irish and Mauritian floats. In addition, UK Argo is a member of Euro-Argo and has begun managing European Union floats as part of the MOCCA project. BODC has also made minor contributions to the developments of the GDACs and is also the lead for the Southern Ocean Argo Regional Centre (SOARC).

General Status

BODC Argo Team

The organisation of the Argo team at BODC has changed significantly in the past few months due to a range of staffing challenges, but has managed to make some progress throughout the changes. Violetta Paba joined the team in April 2017 from within BODC on a part-time basis, contributing to DAC operations. Robin McCandliss has joined the team from within BODC as the acting BODC Argo programme manager, pending a review of roles and responsibilities within the team. Katie Gowers remains involved with the team, primarily in a software development and advisory capacity, but is now working part-time. Matt Donnelly is still heavily involved in data science tasks and is taking the lead with NRT QC, delayed-mode QC and SOARC coordination. Justin Buck remains involved on a part-time basis, supporting the regeneration of delayed-mode QC and as a source of advice. Existing BODC team members Paul McGarrigle and Elizabeth Bradshaw have joined the Argo team on a part-time basis to provide support for DAC operations. We have recently recruited a new member of the team to BODC, Clare Bellingham, who brings a broad range of science and software skills to the Argo team. We have also benefitted from website development support provided by Craig Corbett at the National Oceanography Centre.

Funding outlook

National Capability funding from NERC is currently maintained for BODC at the same rate as previous years. In addition, it is expected that NERC-funded research projects deploying Argo floats will be providing additional sources of data management funding, such as from the ORCHESTRA, ACSIS, BoBBLE and RoSES projects. BODC also has funding from the Euro-Argo ERIC MOCCA project for the European Union floats now being managed by BODC, as well as some funding from the EU H2020 AtlantOS project to support delayed mode QC of Argo extensions. BODC continues to seek additional sources of funding to support SOARC functions: it has already identified a primary target and aims to make an application in 2018.

DAC Functions

Data acquired from floats

BODC retrieves data for all UK, Irish, Mauritius and EU MOCCA floats from a number of sources and archives these for further processing. Where possible, processing of arriving data is normally setup within one week of float deployment, and during the past year this has typically been achieved much more quickly. Please refer to table 1 for the types of communications used for different floats.

Progress in the past year:

During the past year, BODC has transitioned to retrieving all Iridium Rudics float data from the new CLS SFTP server, providing greater security and file integrity, and allowing us to terminate the FTP-push of data to our FTP by CLS. BODC has been working with CLS to address issues that arose during this migration and these have been resolved. BODC has also established a robust method for handling SBD Iridium messages received via email.

Data issued to GTS

BODC delivers core data in both TESAC and BUFR formats simultaneously to the UK Met Office four times a day, where it is subsequently issued to the GTS. Over 95% of these messages are delivered within 24 hours of the data being available to BODC. Coriolis is kindly providing the processing for a dozen PROVOR BGC floats and delivering the core data to the GTS on BODC's behalf until BODC can take on the management of these floats. The results of improvements to system performance, a proactive approach to loading floats to the BODC Argo System, and increased frequency of data processing and delivery can be seen in figure 1.

Progress in the past year – general processing:

Processing of floats for the EU MOCCA project began at BODC in December 2016, increasing the total of core floats being processed by 46 at the time of writing. This included the deployment of decoder software provided by Coriolis to augment the BODC Argo System. During 2017, BODC updated automated data processing such that data are now submitted to the GTS four times a day, rather than twice a day as was previously the case. An issue identified with the application of the BODC greylist to a small subset of floats has now been resolved.

Progress in the past year – BUFR converter:

In support of a new BUFR converter, BODC has begun making all netCDFs available directly to the Met Office via SFTP and this is now occurring on a routine basis four times a day. The Met Office is currently testing a new Python netCDF to BUFR converter in a pre-operational mode. Once this testing is complete, the Met Office will assume responsibility for the BUFR generation prior to distribution on the GTS through this mechanism. Once operational, the Met Office will make the BUFR converter freely available for use. The code has been designed to be extensible, where capability for secondary temperature/temperature and salinity, and oxygen profiles will be added plus other BGC variables when required.

Current activity and future plans:

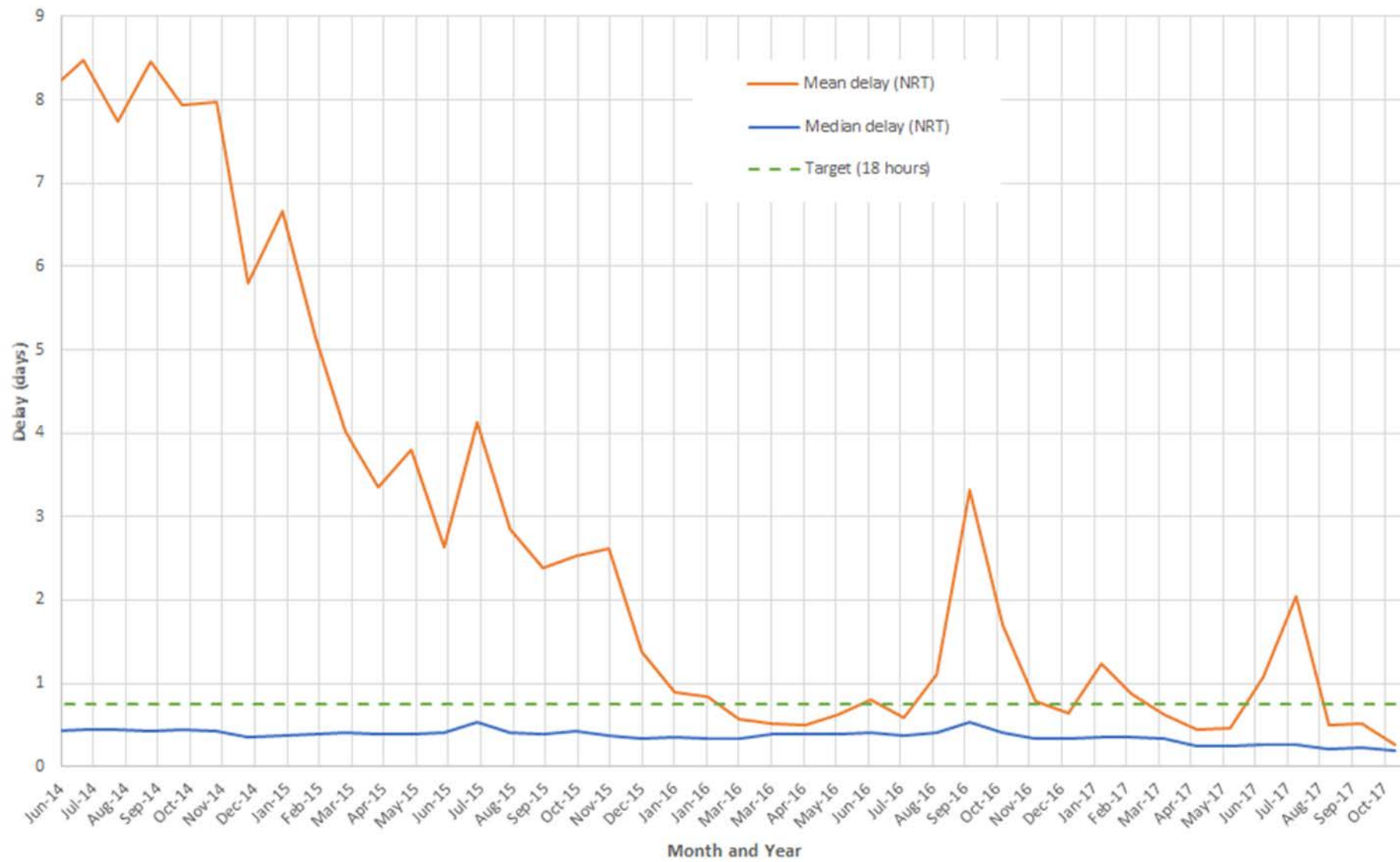
Distribution of all core data to the GTS from all BODC managed floats is a priority, including core data from floats with any type of Argo extension (deep, BGC or auxiliary data). BODC's current focus is to ensure all floats with an APF9I controller board are effectively managed within the BODC Argo System, and attention will then turn to floats with other controller boards such as the APF11. BODC is seeking to collaborate with other DACs in the development of APF11 decoders.

Table 1: Summary of all BODC managed Argo floats, with a focus on those that are currently active

Float type/controller	Comms	Total no. of deployed floats	Total no. of active floats	No. of active floats with ice detection	Mission of active floats					Total no. of active floats being fully processed
					Core only	Core + NST	Core with RBR CTD	Core + oxygen only	Core + other BGC	
Martec Provor	Argos	26	-	-	-	-	-	-	-	-
MetOcean NOVA	Iridium	1	-	-	-	-	-	-	-	-
NKE Arvor	Argos	5	1	-	1	-	-	-	-	1
NKE Provor *	Iridium SBD	13	11	-	-	-	-	-	11	11
NKE Arvor	Iridium SBD	47	46	-	-	46	-	-	-	46
NKE Deep Arvor	Iridium SBD	2	-	-	-	-	-	-	-	-
SBE Navis N1	Iridium Rudics	12	8	8	8	-	-	-	-	7
SBE Navis N1 with BGC	Iridium Rudics	4	1	-	-	-	-	-	1	0
SBE Navis N1 with oxygen	Iridium Rudics	6	6	-	-	-	-	6	-	0
SBE Navis N1 with radiometer	Iridium Rudics	3	3	-	-	-	-	-	3	0
TWR Apex APF7	Argos	7	-	-	-	-	-	-	-	-
TWR Apex APF8	Argos	252	2	1	2	-	-	-	-	2
TWR Apex APF9A	Argos	249	152	27	35	117	-	-	-	152
TWR Apex APF9I	Iridium Rudics	20	11	8	11	-	-	-	-	11
TWR Apex APF9I with BGC	Iridium Rudics	4	-	-	-	-	-	-	-	-
TWR Apex APF9I with STS	Iridium Rudics	4	4	-	-	4	-	-	-	0
TWR Apex APF11	Iridium Rudics	11	8	n/k	-	-	2	-	6	0
TWR Deep Apex APF11	Iridium Rudics	-	-	n/k	-	-	-	-	-	0
TOTAL		666	253	44	57	167	2	6	21	230

* = processing courtesy of Coriolis

Figure 1: Summary of all BODC processing times from time of profile to processing completion, indicative of GTS delays



Data issued to GDACs after real-time QC

All core data received for currently processed floats are distributed to the GDACs within one hour of the data arriving at BODC, with the real-time quality control tests applied. Any file that fails to be transferred is queued for the next transfer attempt. BODC has not yet completed the conversion to v3.1 for all file types. Please refer to table 1 for the types of float and whether they are being fully processed.

Progress in the past year:

BODC has effectively completed the transfer to v3.1 for metadata files, and has made further progress towards delivering core profile files in v3 netCDF files (now > 60 % compared to 30 % at the last ADMT). Progress on the conversion has been slowed by disruption to IT infrastructure and changes to the BODC Argo Team.

Current activity and future plans:

There remains further work to complete the conversion of the remaining core profile files, with the conversion to v3.1 for technical and trajectory files due after this work has been completed. We are not currently issuing any BGC-Argo files for UK floats due to the current focus on core profile data. The exception to this is the dozen PROVOR floats kindly hosted for BODC by Coriolis until such time as BODC can take over the real-time processing.

Data issued for delayed-mode QC

All delayed-mode QC on BODC hosted floats is performed within BODC, with the exception of some floats funded under the Euro-Argo MOCCA project. Currently BODC is only capable of providing data for delayed-mode QC for core data, with work required to finish the delivery of biogeochemical parameters in v3.1. Again, the exception to this are a dozen PROVOR floats that Coriolis is kindly hosting on BODC's behalf. See section 2 of this report for the status of delayed-mode QC.

For any given float, if the R-mode or A-mode file is available following real-time QC, then any profiles that have been through delayed-mode QC will be available as D-mode files. This applies to float profile files that are in either v2.2 and v3.1 format.

Progress in the past year:

Progress on v3.1 profile files has made more delayed-mode files available in v3.1. BODC has begun liaising with relevant Euro-Argo partners to manage the delayed-mode files for MOCCA floats as they become eligible for delayed-mode QC.

Delayed-mode data sent to GDACs

All delayed-mode QC on BODC hosted floats is submitted to the GDACS the same day that delayed mode QC is complete for a profile. See section 2 of this report for the status of delayed-mode QC.

Web pages

BODC continues to maintain the UK Argo website (www.ukargo.net) along with a Facebook page (www.facebook.com/UKArgofloats/) and a Twitter account (twitter.com/ukargo). Work has also progressed to split the UK Argo website into separate UK Argo and SOARC websites – providing SOARC with a distinct web presence – and the new SOARC website is now available (www.soarc.ag). For further details see the SOARC report.

Data use and data products

Statistics of Argo Data Usage

National Oceanography Centre

Argo data are used widely within NOC science with the following regional leads for float deployment and science:

- Alex Sanchez Franks (Indian Ocean)
- Yvonne Firing (Southern Ocean)
- Penny Holiday (Sub-polar N Atlantic)
- Brian King (everywhere else)

Elaine McDonagh is also engaged in using Argo data, bidding for float funds, planning strategies, leading analyses and mapped products

The applications of Argo data at NOC include:

- Measurement of evolution and drivers of mixed layer processes in the (Indian Ocean);
- Inventory and evolution of heat and freshwater establishing controls on budgets (both regional and global);
- Deep heat content (N Atlantic).

UK Met Office

Argo data (received over the WMO GTS) are routinely assimilated into the Met Office's FOAM (Forecasting Ocean Assimilation Model) which is run daily. The FOAM suite runs daily in an early morning slot and produces 2 analysis days (it has a 48 hour assimilation window) and a 7-day forecast. The 3-D temperature, salinity and current fields from the global model run are used as boundary conditions for the regional models run on the following day. There are 4 different configurations: ¼ degree global, 1/12 degree European North West Shelf, 1/12 degree North Atlantic, 1/12 degree Mediterranean and 1/12 degree Indian Ocean. More details are at: <http://www.ocean-sci.net/12/217/2016/os-12-217-2016.pdf> and <http://www.geosci-model-dev.net/7/2613/2014/gmd-7-2613-2014.html>. A coupled ocean/atmosphere prediction system is being developed in weather forecasting timescales, including assimilating Argo data in a coupled data assimilation framework (Lea et al., 2015). A demonstration coupled numerical weather prediction (NWP) system is being run operationally at the Met Office. The timeliness constraints on Argo for this application are more stringent (shorter, 24 hour, assimilation window). The impact of Argo on this system was assessed as part of an E-AIMS EU project (King et al., 2015).

Near-surface Argo data are used to validate the output from the Met Office's OSTIA (Operational Sea Surface Temperature and Sea Ice Analysis) – the OSTIA fields are in turn used as a lower boundary condition in numerical weather prediction models run by both the Met Office and ECMWF.

Argo data are also used in the initialization of ocean and sea-ice conditions in models run to make monthly to decadal predictions, see: <http://www.metoffice.gov.uk/research/modelling-systems/unified-model/climate-models>.

Plymouth Marine Laboratory

Giorgio Dall'Olmo is the lead PI for BGC data in the UK. Bio-Argo data from 13 Provor floats are now available from the GDACS, thanks to processing courtesy of Coriolis.

Core-Argo data are used at PML for:

- providing a description of the physical environment in the framework of biological (e.g. mapping eel migration routes) and biogeochemical studies;
- developing techniques to generate 3D fields of biogeochemical variables by merging ocean-colour and in-situ data;
- investigating mesoscale structures by combining altimetry and in-situ profiles with a special focus on Agulhas rings.

BGC-Argo data focuses on investigating new methods to:

- efficiently monitor the ocean biological carbon pump;
- quantify particle flux attenuation;
- vertically-resolve seasonal remineralisation rates;
- and to better understand the nitrogen cycle in oxygen minim zones.

Data Products

National Oceanography Centre

Elaine McDonagh is engaged with 4-D global fields of mapped Argo T and S, but they are not currently publicly available, however Elaine can be contacted by any interested parties.

UK Met Office

The Hadley Centre maintains two data products that incorporate Argo observations:

- EN4 contain in-situ ocean temperature and salinity profiles and objective analyses. It is updated monthly using real-time Argo profiles, and annually using delayed-mode Argo profiles. EN4 is freely available for scientific research use (see <http://www.metoffice.gov.uk/hadobs/en4/>);
- HadIOD is an integrated database of surface and sub-surface temperature and salinity observations for the period 1850 to present. It includes quality flags, bias corrections and uncertainty information (Atkinson et al., 2014). At present, HadIOD obtains sub-surface profile data from EN4. Public release of the data are expected during 2017. HadIOD is expected to supersede the HadGOA data product, which has not been updated for approximately 6 years (<http://www.metoffice.gov.uk/hadobs/hadgoa/>).

The datasets are used for climate and global change studies, including ocean heat content analysis.

Delayed Mode QC

In recent months BODC has regenerated delayed mode QC capability through a software and procedural review, accompanied by knowledge transfer primarily from Justin Buck to Matt Donnelly. BODC use OW software for delayed mode quality control with the latest reference data available from Coriolis (CTD climatology and Argo profile climatology for guidance).

Following advice from the wider UK Argo team, and particularly from Brian King, we are currently working to the following prioritisation:

- To facilitate increasing knowledge/experience of regional oceanography we are addressing DMQC on an ocean basin-by-ocean basin basis. DMQC has started with the Indian Ocean, and will move onto the South Atlantic next, followed by the Southern Ocean and North Atlantic.

- To facilitate increasing knowledge/experience of the use of the OW software and to avoid applying excessive corrections, we are tackling floats with simple pathologies first, then moving onto floats with increasingly complex pathologies as DMQC skills improve within BODC.

During October/November 2017 we have performed our first significant batch of DMQC since 2013, totalling c. 1700 profiles. At the time of writing, 49% of BODC hosted floats profiles eligible for delayed mode QC have been processed and submitted to the GDACs in delayed mode. BODC expects to continue improving this situation in the coming year.

Current activities and future plans:

Whilst BODC has restored DMQC capability using existing tools linked with the BODC Argo System, an assessment of current capabilities and workflows has highlighted a need to explore other available DMQC software, particularly with regard to biogeochemical variables. BODC would welcome an assessment of existing DMQC software tools within the Argo community to inform our decision-making.

GDAC Functions

The UK does not currently make a direct contribution to GDAC functions, although BODC has contributed to the development of the Argo DOI and has begun supplying some auxiliary data files to the new auxiliary directories at Ifremer for floats funded by the EU MOCCA project.

Argo DOIs

BODC-NOC and Ifremer won a small 15 k Euro grant to progress the Argo DOI from Research Data Alliance (RDA) Europe. Ifremer migrated the Argo snapshots to a single DOI (<http://doi.org/10.17882/42182>) in March 2016. A '#' key is used to identify the monthly snapshots within the archive on SEANOE, e.g. <http://www.seanoe.org/data/00311/42182/#45420>. The '#' means that the identification of the snapshot is evaluated on the browser rather than the DOI resolving server making it possible to use a single DOI without a change to the DOI system.

The approach was presented at International Data Week 2016. Although it meets the needs of the Argo community there is informatics community concern with the usage of the '#' key. It means there is a secondary resolving service creating an additional dependency. The RDA data citation working group is to become a formal interest group in the RDA and dynamic data has been identified as being in need of further research. This small project is now being brought to a close with the method being written up in a technical paper for publication. The first citation of the new DOI has been identified in the literature (Piron et al. 2017). Also, the approach has been used in other observing systems as part of the EC H2020 AtlantOS project, see:

<https://www.atlantos-h2020.eu/download/deliverables/7.1%20Data%20Harmonization%20Report.pdf>

Regional Centre Functions

BODC continues to provide the coordinating role between the SOARC partners.

The new SOARC website (www.soarc.ag) has recently been launched to provide a distinct SOARC web presence. This has been delivered after a series of delays caused by a combination of factors which include other pressures on the BODC Argo team, an election 'purdah' period, and ill-health. Thanks go to Jcommops for the new Southern Ocean status and planned deployments map. Feedback on the website is welcome and can be submitted either via the website contact form or direct to argo@bodc.ac.uk.

With the new website now in place, BODC is now reviewing with SOARC partners which priority areas to focus attention upon within its contribution to SOARC functions. Efforts continue to secure dedicated additional funding for BODC to carry out a greater range of regional centre functions. Please refer to the Southern Ocean Argo Regional Centre report for further details.

References

Atkinson, C. P., N. A. Rayner, J. J. Kennedy, and S. A. Good (2014), An integrated database of ocean temperature and salinity observations, *J. Geophys. Res. Oceans*, 119, 7139–7163, doi:[10.1002/2014JC010053](https://doi.org/10.1002/2014JC010053).

King, R.R., M. Martin, A. Stearl, 2015. Weather, seasonal and decadal forecasting: OSE/OSSE results and recommendations. E-AIMS deliverable report D3.323. http://www.euro-argo.eu/content/download/88659/1093576/file/E-AIMS_D3.323-v2.pdf?version=1

Lea, D. J., I. Mirouze, M. J. Martin, R. R. King, A. Hines, D. Walters, and M. Thurlow, 2015: Assessing a New Coupled Data Assimilation System Based on the Met Office Coupled Atmosphere–Land–Ocean–Sea Ice Model. *Monthly Weather Review*, 143, 4678-4694, doi: [10.1175/MWR-D-15-0174.1](https://doi.org/10.1175/MWR-D-15-0174.1).

Piron, A., V. Thierry, H. Mercier, and G. Caniaux (2017), Gyre scale deep convection in the subpolar North-Atlantic Ocean during winter 2014–2015, *Geophys. Res. Lett.*, 44, doi:[10.1002/2016GL071895](https://doi.org/10.1002/2016GL071895).