DESIGN OF A COMPARATIVE STUDY ON DMQC METHODS USED FOR DEEP ARGO

Ref.: D3.2_v4

Date: 22/06/2020

Euro-Argo Research Infrastructure Sustainability and Enhancement Project (EA RISE Project) – 824131

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 824131. Call INFRADEV-03-2018-2019: Individual support to ESFRI and other world-class research infrastructures

Document Reference

Project	Euro-Argo RISE - 824131
Deliverable number	3.2
Deliverable title	Design of a comparative study on DMQC methods used for deep Argo.
Description	This report will review the existing DMQC methods, its initial appropriateness to be used in the deep Argo extension, what method each partner will use, and the floats to the evaluated.
Work Package number	WP3
Work Package title	Extension to deep ocean
Lead Institute	NOC
Lead authors	Kamila Walicka, Brian King
Contributors	Cecile Cabanes, Virginie Thierry, Damien Desbruyeres, John Gilson, Nathalie Zilberman, Sarah Purkey, Taiyo Kobayashi, Kanako Sato, Shigeki Hosoda, Annie Wong, Esmee Van Wijk, Annie Foppert, John Lyman, Gregory Johnson, Kjell Arne Mork, Matthew Donnelly, Pedro Vélez Belchí, Giulio Notarstefano
Submission date	22 June 2020
Due date	[M16] - 30 April 2020
Comments	Click here to enter text.
Accepted by	Click here to enter text.

Document History

Version	Issue Date	Author	Comments
v1	05/06/2020	Kamila Walicka	Preparation of report draft. Review of methods and tools used in real-time and delayed mode deep Argo floats data used by the global deep Argo community. Assessment of the current state requires agreement on establishing the most appropriate real-time and delayed mode procedures, validation of the CTD reference data and operator skill development.
v2	09/06/2020	Brian King	Review of the report draft
v3	12/06/2020	Kamila Walicka	Initial version
v4	22/06/2020	Kamila Walicka & Matthew Donnelly	Integration of partner's comments - Final version

EXECUTIVE SUMMARY

To assess the current status of Deep Argo delayed-mode quality control (DMQC) we invited comment from the full international Deep Argo community. This report incorporates responses from seven groups who between them have been examining data quality in around 100 floats. The first major issue that needs to be resolved to enable distribution of real-time data is to agree on a fleet average value for the response of conductivity cells to pressure – a coefficient referred to as CpCor.

After applying CpCor adjustment, delayed-mode adjustment of salinity can begin. Two international groups have experimented using the same 'OWC' software used for determining salinity adjustment for core DMQC. Those groups have also used the standard core distribution of reference data. Other groups have experimented with in-house methods for estimating float salinity errors, or have not yet begun delayed-mode determination of salinity error. Some groups have made use of ship-based CTD profiles collected at float deployments, but such profiles are not available for all floats, and will not generally be available for Deep Argo due to the limited capacity for undertaking high quality hydrographic castes on deployment. Where reference data are sparse in the deep ocean, it is likely that mapping scales in the OWC runs will need to be modified to reflect reference data distribution. Consensus on how to select appropriate scales for OWC will emerge as groups gain further experience. Of the groups that are estimating DMQC salinity adjustment, only one is uploading D-mode data to the Argo GDACs. Two other groups have undertaken experimental quality control for research purposes, with data files available on request. There has been little, if any, discussion between the groups about the appropriate way to assign the PSAL_ADJUSTED_ERROR parameter, the minimum adjustment that should be made, and whether the PSAL_ADJUSTED_ERROR is uniform in the vertical. Each active group is making these decisions for their particular group of floats.

To achieve the objections of Euro-Argo RISE T3.2, partners will need to contribute to 4 primary areas: determine CpCor values for their floats and contribute to assessing the fleet average CpCor value; collaborate in the development of the reference databases; contribute to agreeing deep Argo quality control procedures through a virtual workshop; and contribute to developing documentation on the agreed procedures. This should be achieved by working as part of the international Deep Argo Mission Team, and by bringing regional expertise to bear on the areas of focus mentioned above.

TABLE OF CONTENT

1. Introduction

The array of core Argo floats sample pressure, temperature and salinity to a depth of 2000 dbar using a CTD sensor. The Deep Argo mission will extend this array to sample the deep ocean to depths of 6000 m, with Euro-Argo members contributing to pilot arrays in the Atlantic Ocean, the Southern Ocean, the Nordic Seas and the Mediterranean Sea. This report aims to investigate the accessibility and usage of methods and tools within the international Argo community from a range of organisations involved in the Deep Argo mission, enabling Euro-Argo RISE T3.2 partners to contribute to this international effort as part of the Deep Argo Mission Team. This report will contribute to further work to establish a uniform and appropriate method to adopt quality control data from Deep Argo floats. Moreover, this report will contribute to creating and agreeing on a framework for collaborative deep DMQC development. The report goes beyond the initial deliverable of design of a comparative study, by embarking on an initial survey of the groups responsible for the largest number of Deep Argo floats.

The survey covers matters such as the currently available deep Argo fleet, QC software used, approaches taken to adjust flags on data before the DMQC analysis, the configurations of the initial adjustments, thresholds and type of the reference database used to verify the quality of the Deep Argo salinity data. Moreover, this document highlights the issues reported by deep DMQC operators.

All groups represented in the international Deep Argo Mission Team were contacted and invited to comment on Delayed-Mode Quality Control of their data. This report summarises the interaction with those groups who responded. The interaction was through video conference meetings and summary reports delivered by operators. The review covers methods used in Ifremer (France), Scripps Institution of Oceanography (SIO, USA), Japan Marine Science Technology Centre (JAMSTEC, Japan), the University of Washington (UW, USA), Pacific Marine Environmental Laboratory (PMEL, USA), Commonwealth Scientific and Industrial Research Organisation (CSIRO, Australia), National Oceanography Centre (NOC, UK) and Institute of Marine Research (IMR, Norway). In addition, the Instituto Español de Oceanografía (IEO, Spain) and Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (OGS, Italy) have both deployed floats and will contribute to development of quality control methods following the collation of information captured in this report.

2. Deep Argo fleet

The measurements of pressure, temperature and conductivity data of deep Argo floats below 2000 dbar are being undertaken using five different types of platforms: Deep APEX, Deep ARVOR, Deep NINJA, Deep SOLO (built by Scripps) and Deep SOLO-MRV. The locations of deep Argo floats deployed by all deploying organisations are presented in Figure 1.

The floats carry a mix of two versions of the SeaBird CTD sensor: the SBE41CP rated to 4000 dbar and the SBE61 rated to 6000 dbar. Each measures pressure (synonymous with depth), temperature, and conductivity of the seawater, and together these three variables are used to calculate seawater salinity onboard the float. Understanding the effects of high pressures on sensor behaviour is key to developing accurate and reliable quality control procedures, including making adjustments to the data based on well characterised sensor response. At this stage, the evidence suggests that the approach to DMQC is the same for both versions of the sensor.

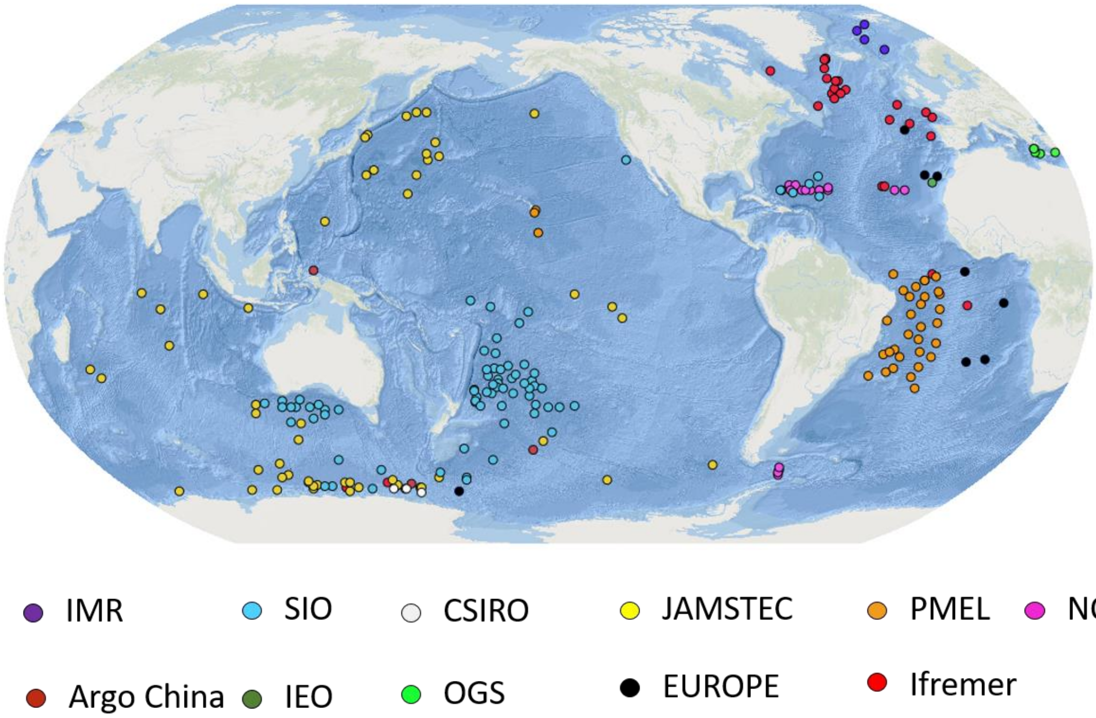

Figure 1: Locations of the deep Argo floats deployed by 11 organizations involved in the Deep Argo Mission Team.

The largest European fleet of deep Argo floats is managed by Ifremer. They have 21 Deep ARVOR floats, with nine operational and 12 inactive floats. 17 of these floats are located in the subpolar North Atlantic, two are in the equatorial region of the South Atlantic and another two are in the Southern Ocean, respectively (www.jcommops.org, 05 May 2020). In addition, Ifremer is responsible for the quality control of 5 Deep Arvor floats funded by the European Union AtlantOS project which were deployed into the South Atlantic and Southern Oceans.

The world's largest deep Argo fleet belongs to SIO, with 69 floats. SIO is using exclusively Deep SOLO floats. Most of these floats are deployed in the southwest Pacific Ocean. Moreover, a few of their floats are located in the southeast Indian Ocean, and the Southern Ocean (to the south of Australia and New Zealand), and some in the western North Atlantic.

JAMSTEC has deployed 56 floats. This organisation has deployed two types of floats Deep NINJA (29 floats) and APEX Deep (27 floats). Deep Ninja floats were deployed mainly in the Southern Ocean, while the APEX deep floats are collecting data in the Pacific, the Indian and the Southern Oceans.

CSIRO has three Deep SOLO-MRV floats, deployed in the Southern Ocean, in the Australian-Antarctic Basin.

PMEL has 31 deployed floats located in the Brazil Basin and three in the North Pacific Ocean. All of these floats are Deep SOLO-MRV types.

NOC has 17 Deep APEX deployments; seven were deployed in the Southern Ocean (Drake Passage), but three of those were recovered. NOC has deployed 10 floats in the western North Atlantic, along the 26˚ N RAPID array. This includes three floats recovered from Drake Passage and re-deployed.

UW has deployed three Deep APEX, none still active, and is an active member in advising and making decisions in the DMQC process.

IEO is responsible for the quality control of 4 European Union Deep Arvor floats deployed into the North Atlantic Ocean near the Canary Islands as part of the E-AIMS and AtlantOS projects.

OGS has deployed 5 Deep Arvor floats into the eastern Mediterranean Sea.

IMR has deployed 4 Deep Arvor floats into the Nordic Seas, with each deployment having been accompanied by a ship's CTD cast on deployment of the float.

3. Real-Time analysis

3.1. Adjusting flags on data

Adjusting flags is one of the first steps before the DMQC process, where an operator undertakes a visual quality control of profiles and applies flags to suspicious or problematic profiles. All of the involved organisations are using their in-house codes to visualise the raw data and adjust flags on pressure, temperature and salinity data. The typical detected issues are various spikes e.g. due to strong gradients in the water column in the upper 2000 dbar as well as some fresh spikes in the bottom part of the profiles. SIO reported that in working with both core and deep float data, the real-time tests are not very effective to appropriate identifying large density inversion issues and additional checks done by operators are necessary. Improvements to the real-time tests have been approved for core Argo, but have not yet been validated for deep Argo.

3.2. Estimates of deep Argo CpCor value

Deep Argo salinity data are affected by a bias resulting from the conductivity cell compressibility with increased pressure. This bias leads to an offset in salinity data. Nearly all real-time salinity data are reported fresher than true values at deep ocean depths. To reduce the effect of this bias the pressure conductivity coefficient (CpCor) needs to be changed from the default value used onboard the float and in real-time data telemetered ashore. The key aspect that needs to be resolved is to agree on the most appropriate value of the CpCor.

Establishing a new fleet-average default value for CpCor is a current focus for the Deep Argo Mission Team. A value for CpCor has been recommended, however, this needs to be internationally assessed and approved for implementation for real-time adjustment.

Ifremer has multiple steps for estimating the CpCor value and for float salinity adjustment. Their CpCor value has been estimated using eight deep Argo floats, exclusively deployed in the North Atlantic:

- In the first step, the quality of salinity data is determined by comparing the first ascending deep Argo profile with the reference CTD data.
- In the second step, the salinity data are verified by running the DMQC analysis using OWC software.

Both analyses showed that the average CpCor value calculated from eight deep Argo floats was very similar to those recommended by SBE. At present, CpCore value of around -13.5 x 10^{-8} dbar $^{-1}$ is used to correct the deep Argo data. Overall, for the deep float data affected only by bias, the average CpCor value is applied. For those data affected by offset and drift, the CpCor value and OWC correction are used.

In SIO, despite a good coverage of CTD reference data in the south-west Pacific, for active floats, operators use the average CpCor parameter estimated from the global fleet by SBE. When floats reach the end of life or have long time series e.g. 5-6 years, the CpCor can be estimated separately for each float.

In JAMSTEC, the estimates of CpCor with climatology such as WOA are difficult to obtain. That is associated with insufficient historical reference data coverage in some regions where their floats operate and due to a lack of CTD cast at deployment for some floats.

IMR has undertaken the initial calculation of CpCor for a float deployed in the Greenland Sea based on comparison with ship CTD on deployment casts, obtaining a value of approximately -16 x 10^{-8} dbar^{-1} depending on the exact method used.

3.3. QC flags on real-time data

Another important aspect in real-time data below 2000 dbar that needs to be agreed is QC flags of raw data. Currently, these data have been greylisted and flagged with QC=3, (probably bad data) for Sea-Bird salinity data.

SIO has recommended changing the QC flag of salinity data from QC=3 to QC=1 (good data) forreal-time adjusted data that has had CpCor applied, after passing all other real-time QC tests. This suggestion is also agreed by UW and NOC, however, to justify the validity of this change of QC flag, the analysis confirming recommendation from SIO need to be provided. Moreover, the need for ‘A’ mode adjustment for pressure data for SBE61 is not clear and need to be further discussed.

After finalising the procedures associated with estimating the CpCor value, UW and NOC have recommended introducing the real-time ‘A’ mode adjustment for PSAL, which will include data with automatically applied CpCor correction.

The instructions explaining the motivation and procedures of applying QC flags to will be prepared by deep Argo team in the second half of 2020.

3.4. Sea surface pressure corrections

All types of floats measure pressure while on the surface. These values can be used to adjust the pressure to remove sensors offsets. However, different float types use this information differently. Some types of deep Argo floats (Deep ARVOR, Deep SOLO, Deep SOLO-MRV and Deep NINJA) apply the sea surface offset to the pressure data on-board the float. Deep APEX floats do not make any internal pressure correction and require the adjustments. The method of applying sea surface pressure correction can be found in Argo User's Manual version 3.3, 2020 (https://archimer.ifremer.fr/doc/00228/33951/32470.pdf). The sea surface pressure adjustments need to be applied to Deep Apex floats before running the DMQC analysis.

However, the analysis made by Ifremer found that Deep ARVOR floats sometimes show a relatively large offset of the sea surface pressure data, from around 2 dbar to 8 dbar. These values can lead to bigger offsets at depths below 2000 m. To obtain the most accurate measurements, Ifremer has suggested the use of a similar method of correction to the sea surface pressure data, as those used for APEX floats, where the sea surface pressure data can be adjusted by data from the next cycle (n+1). This issue needs some further investigation and agreement on procedures.

SIO for Deep SOLO floats also have found some large drift of sea surface pressure data of around 4-6 dbar. The drift is stronger for the first 30-40 cycles, where further drift it stabilised. Both float types use Kistler SBE61 pressure sensors. Additionally, the CTD on a Deep SOLO is mounted below the buoyancy sphere, so the instrument remains slightly submerged when the float measures 'sea surface' data. This means the 'surface pressure' is in fact measured 0.5 dbar below the surface, and SIO adjusts the 'surface pressure' by a constant value of 0.5 dbar. However, no analysis of the effect of the sea surface pressure drift on pressure data at depth has been undertaken.

4. Delayed-mode analysis

4.1. Quality control of temperature data

The temperature data are only analysed by visual inspection in the near real-time process. There are no available QC methods to determine temperature issues. These data are assumed to be not affected by external factors.

4.2. Quality control of salinity data

The DMQC process covers the assessment of the magnitude of offset or drift in Argo float salinity data and provides an estimate of the associated error compared to the reference data. In DMQC analysis Ifremer and SIO are using the OWC software, as they do for the core data quality control https://github.com/ArgoDMQC/matlab_owc/releases. At JAMSTEC, OWC software is also going to be implemented for deep floats analysis in the near future. Up to now, NOC and JAMSTEC have used in-house tools for examining deep float data. CSIRO and PMEL will start the DMQC analysis after the consensus on exact value defining the CpCor value and procedures in applying it will be reached. Both of these organisations are going to adapt their in-house developed software for core Argo analysis and modify it to be usable for the deep Argo floats.

4.2.1. CTD reference data

In Ifremer and SIO, the key source of reference data used to compare with the deep Argo float salinity data is CTD data (CTD_for DMQC_2019v01) distributed by Ifremer via the FTP site at /coriolis/data/DMQC-ARGO/.

This CTD reference dataset has relatively good coverage in the North Atlantic, where currently Ifremer has conducted their DMQC analysis. That can make the DMQC analysis of deep floats relatively accurate and efficient in this region. However, due to high natural variability in the North Atlantic and various sources of CTD data may lead to higher uncertainty in quality checks of the deep Argo floats. To assess the quality of the CTD profiles in the North Atlantic, more comparisons with other datasets e.g. from Go-Ship dataset lines is strongly recommended.

Adequate coverage of CTD profiles is also available in the south-west Pacific, where SIO deployed their floats. They have reported no major issues with the quality of the CTD reference data in this region that can be successfully used for deep floats analysis. SIO and PMEL do not have a sufficient amount of CTD casts from float deployment to allow comparison with all their deep Argo floats, which may limit their analysis.

In the Mediterranean Sea, OGS has explored the availability of deep reference data as part of EuroArgo RISE WP3, and this work is ongoing.

However, in the case of other regions in the world ocean with relatively sparse data coverage, the CTD reference database distributed by Ifremer may be insufficient. One of those regions is the Southern Ocean, where many of the deep Argo community have deployed their floats.

The supporting reference database in deep Argo DMQC analysis can be the most updated, high quality "GO-SHIP Easy Ocean" hydrographic database, which is currently developed by SIO and JAMSTEC. The GO-SHIP dataset will be used by PMEL in the Brazil Basin, which includes recent deep data in this region. The plan is to make their curated dataset available in the correct format to use in the OWC software. Apart from this, CSIRO is also planning to use other reference databases such as CARS 2006, CARS 2009 and Gouretski and Koltermann.

Another solution to verify the quality of deep Argo floats is the use of the CTD profiles collected during the float deployment. That procedure is undertaken by JAMSTEC and NOC. However, not all of their floats have associated CTD casts due to some failures or drift of floats out from the location of the nearby shipboard CTD casts. As the deep Argo array expands it will not be practical to undertake sufficiently high-quality ship hydrographic casts for all deep Argo deployments.

The key role for partners in Euro-Argo RISE is to ensure that there is a clear flow of reference data supplied to both the ship CTD reference database managed by Ifremer and whatever preferred mechanism emerges for making a further curated database available for use with Deep Argo. European partners are well placed to facilitate this as coordinating members of Argo Regional Centres in the Atlantic Ocean, the Southern Ocean and the Mediterranean Sea.

4.2.2. Objective mapping parameters and theta levels

In DMQC analysis of deep Argo floats, the density distribution of the CTD reference data below 2000 dbar plays a crucial role in the objective salinity mapping calculations. The length and time scales configurations selected in the OWC software configurations need to be compatible with the availability of reference data for the specific regions in the deep ocean. Where reference data are sparse in the deep ocean, it is likely that mapping scales in the OWC codes will need to be modified (e.g. extended) to correspond with the reference data distribution. The DMQC operators of deep Argo floats need to have extensive awareness about the reference data coverage in areas of deep float operation and pay special attention in selecting decorrelation length scales there.

Ifremer and SIO use the same spatial and time length scales as those they use in the Argo core data analysis. This approach is suitable for the region of the location of the float, where in both the subpolar North Atlantic and the southwest Pacific Ocean, respectively, CTD reference data are sufficient for the deep Argo float comparisons.

Both organisations are using the automatically selected theta levels by OWC software, to compare the deep Argo salinity data with reference data. That selection is mostly using the theta levels from the deepest part of the profiles. Additionally, in SIO the OWC software runs twice, firstly operators allowing the software to use the deepest water masses around 4000-5000 dbar, while in the second run water masses from shallower depths are forced by the operator e.g. 2000 dbar. That two-run allows an additional check of the salinity drift at different water masses and also allows the possibility to determine different offsets at different pressures, which would be characteristic of a badly determined CpCor. These analyses are exclusively done for the south-west Pacific Argo deep floats.

IEO has conducted DMQC on 2 floats near the Canary Islands, correcting initial bias on both floats using ship CTD on deployment casts. One of the floats also had a large drift in salinity, and the DMQC method used in core Argo (OWC) was applied to correct the data. It was found that the method was able to correct with an accuracy less than 0.005 at 4000 dbar. This was possible as 4 ship CTD casts were undertaken in close proximity to the float during its roughly 2 years of life. During the OWC analysis it was apparent that it was very sensitive to the choice of temporal mapping scales due to the scarcity of deep ocean reference data in the current DMQC database.

4.2.3. A threshold in deep Argo salinity data

Similar as for core data, the analysis of deep Argo data also requires a threshold of applying salinity corrections that is appropriate to the natural variability of the ocean and the quality of various ship data.

In Ifremer, the lower boundary where salinity data are not corrected is below 0.002. The salinity correction is applied when offset or drift exceeds 0.002. Currently, the smallest corrections applied by Ifremer was 0.0023 and 0.003. In SIO, the minimum error that is used to correct salinity data was 0.004.

The upper threshold above which salinity data should not be corrected for deep Argo floats is not clear. Further investigations of the upper threshold should be undertaken

4.2.4. Error-bars associated with depth dependence of salinity data

Some groups have reported small vertical variations in the salinity offset from the SBE61, even after CpCor is applied. These changes potentially could reveal due to ocean variability or could be part of the sensor drift characteristics, or a combination of both. The small vertical variation of offset in the salinity data could be also present in the SBE41/41CP used in upper 2000 dbar. However, due to a much larger variety of data in this part of the water column, these changes will be very difficult to distinguish.

IEO have undertaken analysis of floats in the stable central waters around the Canary Islands, making use of the stability of the isotherms and isopycnals, and did not find any dependence with pressure of the drift or the initial bias, but other groups have not yet replicated this result.

This issue reveals a need for appropriate adjustment method to estimate the vertical variation of the salinity data for deep Argo floats. Currently, the OWC software corrects only depth-independent offsets, therefore is not able to adjust the bias due to the vertical variations.

So far, there has been little discussion to address this issue. UW and NOC have suggested the use of the depth-dependent error bars in delayed-mode (PSAL_ADJUSTED_ERROR). Otherwise, assign the vertically uniform PSAL_ADJUSTED_ERROR parameter.

As part of the upcoming comparative study, Euro-Argo RISE partners will need to develop a clear understanding of pressure dependence and the impact on estimation of error associated with those measurements and any adjustments. This will be a key element to resolve to be able to make Deep Argo data available in delayed-mode.

5. Availability of D-mode deep Argo data

The DMQC analysis done by Ifremer is currently in an experimental testing stage, and so they have not submitted deep Argo floats to the GDAC. Their quality control analysis aimed to prepare some deep Argo floats for further research purpose. Similarly, in JAMSTEC the development of quality control for deep floats are at the preliminary stage, however, some Deep NINJA floats have been already quality-controlled for scientific researchers. These data are only available on JAMSTEC website and can be accessible with their permission. None of their data has been submitted to the GDAC. SIO has applied delayed-mode corrections to some of their floats and these data are available in the GDAC. It is important within EuroArgo RISE that data is only made available in delayed-mode once we are satisfied that the data is suitable for scientific use.

6. Conclusions

The report provides a review of the existing methods and tools used for the delayed-mode quality control of deep Argo floats. Moreover, this document expanded the initial scope of deliverable by including the review of the real-time mode procedures used by the deep Argo international community.

This report revealed several issues with sensors on deep Argo floats and a need for agreement on establishing the most appropriate methods used in real-time and delayed mode quality control.

The key areas of focus for EuroArgo RISE partners to contribute to are:

- Agreement on fleet average CpCor value needed to correct the real-time deep Argo data. This is the subject of an international working group established after the Virtual Argo Steering Team meeting in April 2020. When the group reaches an agreement, expected later in 2020, then a value will be published with instructions on how to implement it and instructions on how to assign QC flags for raw data and real-time adjusted data. EuroArgo RISE partners should contribute to this with respect to their floats.

- Verify the quality of the CTD reference data for all regions where deep Argo floats are or going to be deployed. An international group is preparing a subset of the CTD reference database that contains the highest-quality reference data from GO-SHIP research cruises. This will be ready for distribution later in 2020. Once available, EuroArgo RISE partners should work to enhance the amount of data included in this dataset;

- Improve the overall consistency of the Argo delayed-mode quality control procedures and operator skills development. We recommend that in the second half of 2020, after the CpCor working group has reached a conclusion and the higher-quality reference database is available, e.g. in October 2020, a virtual workshop should take place to promote consistency between groups on methods and tools, mapping scales, reference data, interpreting estimates of salinity drift, and assignment of QC flags and _ERROR fields. All EuroArgo RISE partners should contribute to this workshop, with the aim of contributing experience to-date, and implementing decisions in the European Argo data system.

- The comparative study should result in the development of comprehensive documentation regarding Deep Argo quality control, including advice regarding regional implementation, such as in European marginal seas. The aim should be to move from experimentally exploring quality control considerations to an operational capability for Deep Argo DMQC.