A report on the adaptation of existing DMQC methods to marginal seas

Ref.: D2.7_v2.1

Date: 27/06/2022

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	D2.7
Deliverable title	A report on the adaptation of existing DMQC methods to marginal seas
Description	A report on the adaptation of existing DMQC methods to marginal seas (Arctic, Baltic and Mediterranean Seas).
Work Package number	2
Work Package title	Improvement of the core Argo mission
Lead Institute	BSH
Lead authors	Birgit Klein, Ingrid M. Angel-Benavides, Simo-Matti Siiriä, Malgorzata Merchel, Giulio Notarstefano, Antonella Gallo, John Allen, Matteo Marasco, Lara Díaz-Barroso.
Contributors	BSH, FMI, OGS, SOCIB Nilsen, Jan Even Øie (IMR, Norway)
Submission date	27th June, 2022
Due date	30th June, 2022
Comments	
Accepted by	Ingrid M. Angel-Benavides

Document History

Version	Issue Date	Author	Comments
0.1	10.12.20	B. Klein	Document creation, first draft and inclusion of existing texts
0.2	28.02.21	All groups	Addition of text from writing group
1.0	11. 03.22	All groups	Finalised table of content and identified remaining gaps in document
1.1	13.04.22	All groups	Addition of text and figures and other changes after several group meetings.
1.2	20.04.22	I.M. Angel-Benavides	Re-organisation of the contents
1.3	09.06.22	all groups	Final draft
2.0	20.06.22	I.M. Angel-Benavides and B. Klein	Version ready for review. After re-organisation of the contents, consolidation, edition, and writing of the executive summary.
2.1	27.06.22	I.M. Angel-Benavides	Final version after suggestions from the reviewers at the Euro-Argo Office

EXECUTIVE SUMMARY

The objective of Task 2.4 in the Euro-Argo RISE project is to improve the delayed-mode quality control (DMQC) process at the European level, with the ultimate goal to increase the Argo dataset overall quality. A specific focus is necessary for DMQC methods for the European marginal seas (Baltic, Mediterranean and Arctic) since they deviate from the open deep ocean characteristics (from which the global rules and methods for quality control of Argo floats were developed) by being shallower, having more variable hydrographic conditions and complicated bathymetry, and/or being seasonally covered by sea ice. The main focus of the DMQC procedure is to check for evidence of problems in the conductivity sensor (bias/drift) manifested in the salinity values reported by the float using the OWC software, and correct for it if possible. In this report we propose a set of procedures and rules for performing the DMQC in these marginal seas, using OWC and/or alternative methods as well as auxiliary tests.

For the Baltic Sea, where the Euro-Argo partners need to built a procedure from scratch since its characteristics are so unique, we propose two DMQC approaches detailed in Section 2: i) Comparing reference and float data, including a method to construct a time series of differences to evaluate potential salinity drift; and ii) Using lab calibrations, which is possible since the Baltic landlock nature allows float recovery. Additionally, recommendations regarding real-time quality control test, as well as the basis for the implementation of Coriolis min/max near real-time quality control in the region are provided.

For the Mediterranean Sea (Section 3), this report describes the modifications made to the OWC code to be used in this region and an approach for reference data selection that is necessary for its use to take into account the different hydrographic sub-basins. For the cases where OWC does not provide reliable results, additional qualitative quality checks are also presented. Also, the use of the Whitespace Maximisation tool for data intercalibration with Argo floats was explored. The tool, initially developed for inter-calibration of glider data, could be useful for assigning priorities for the DMQC activity in the region by allowing quick check of the float data quality.

For the Arctic (Section 4), the OWC software was used to perform salinity DMQC in floats operating in deep basins without modifications. However, despite efforts to improve the reference database in terms of geographical distribution and availability of recent profiles, there are few profiles available for comparison of the Argo profiles and even regions void of data. This makes the outputs of the procedure potentially unreliable. We propose to check the quality of the float data by comparing it with that of other Argo floats in the region, slowly generating a reliable Argo reference database that would complement the CTD reference database and improve the reliability of the OWC results. Also, we present our efforts to include the estimation of under-ice profile positions, which is crucial in the region both for the DMQC efforts as well as for the data quality, as an easy consistent step for all European partners.

For all regions the work necessary to establish and update the CTD reference databases is presented and the links to the Euroargodev repositories used to host all the code generated for the procedures and analysis described in the deliverable are provided.

TABLE OF CONTENTS

1 Context

The objective of Task 2.4 in the Euro-Argo RISE project is to improve the delayed-mode quality control (DMQC) process at the European level, with the ultimate goal to increase the Argo dataset overall quality. To do so, all partners have continued to develop DMQC procedures and materials for their area of expertise in the innovative framework for collaborative development. This framework (e.g. codes sharing repository, shared reference datasets, guidelines for specific marginal seas) was agreed upon at the beginning of the project and was implemented progressively afterwards.

A specific focus is necessary for DMQC methods for European marginal seas (Baltic Sea, Mediterranean and Black Seas, and the Arctic Ocean). Hydrographic properties in marginal seas differ from the open ocean in terms of variability, water depth, and environmental conditions, thus requiring adaptations of the global DMQC rules as well as the preparation of appropriate reference data sets.

Since Argo floats operate autonomously after their deployment at sea and are not usually recovered, the stability and quality of the float mounted CTD data providing core data on temperature and salinity has to be established by in-situ calibrations. While temperature and pressure sensors are believed to have long-term stability for the entire lifetime of the float operation, the conductivity cell is more prone to drift and has to undergo regular checks. The main procedure executed in the DMQC is thus to check for evidence of problems in the conductivity sensor (bias/drift) manifested in the salinity values reported by the float. The OWC software named after its authors, was originally proposed by Owens and Wong (2009) and augmented by Cabanes et al. (2016), is used for this purpose by those participating in the Argo programme. OWC is based on the principle of comparing high-quality reference data from deep ocean profiles (ship based CTD data and selected ARGO data) to the salinity profiles reported by the floats. The comparisons are done for individual float profiles selecting those parts of the water column which show stable water masses with well defined water mass characteristics and establish a time series of differences between the float data and the climatologies. The mixed layer and shallow profiles are therefore not used for this purpose due to their high natural variability and seasonal signals in the water properties. Due to the presence of eddies in the ocean, no 1:1 agreement can be expected with the climatological data. Therefore, the purpose of this comparison is to look for a degradation of the conductivity cell manifesting itself in continuously increasing disagreements. This software then proposes a correction to the data by piece-wise linear fits of the salinity differences between float and climatology. For the deep ocean it was established that the accuracy of the OWC method due to the variability in the reference data is ±0.01 PSU and only deviations larger than this threshold should be corrected. It was also decided to limit the magnitude of the corrections to ±0.05 and for corrections larger than this the sensor is considered too damaged to be corrected. The OWC software is semi-automated and needs manual intervention to select the proper search radii for the comparisons and to analyse the proposed corrections which are the main output of the procedure. This is done by a delayed mode operator, who decides if and how to correct the float data and subsequently creates updates to the netcdf-files provided by the data centres, producing the so-called d-files, where d stands for delayed in contrast to the files containing data that underwent only automated quality control in real-time (r-files). The rules and tasks to be performed by the delayed mode operators are described in the Argo Quality Control Manual For CTD and Trajectory Data, (Wong et al. 2022) which is continuously revised at the annual Argo data management meetings. Additionally, the DMQC cookbook for core Argo parameters (Cabanes et al., 2021) also published in the context of the Task 2.4 of Euro-Argo RISE, documents the end-to-end DMQC processing chain, provides guidelines on the existing manuals, and exemplifies best practices using case studies.

The rules for DMQC originally considered the deep ocean within ± 60° latitude exclusively. As the programme expands more and more the application of these rules to different environments needs to be checked and revised. In the context of European activities, this concerns mostly the Arctic and Baltic Seas, since the operation in the Mediterranean Sea has a longer tradition and even has already established its own Argo Regional Center (ARC) MedArgo. In the Arctic, the DMQC procedures have not been tested before and more and timely reference data need to be gathered, while in the Baltic no DMQC rules were established yet and, therefore, there is the necessity to define them and to select appropriate reference data**.**

Moreover, for the floats operating in the high latitude marginal seas (Arctic and Northern Baltic), the satellite positioning is missing in the profiles that the float collected under the seasonal ice. The under-ice profile locations are usually assigned automatically via linear interpolation between known positions. This method may render unrealistic positions, such as profiles positions located over land. Thus, the DMQC procedure for floats that went under the ice also involves checking the under-ice positions for correctness and their estimation with more sophisticated methods such as that proposed by Yamazaki et al. (2020).

The report is structured in three sections, one corresponding to each marginal sea. In them, we specifically address some of the remaining work necessary to fully implement DMQC in those areas seas, such as:

- Check existing real time quality control rules
- Establishing high quality reference datasets and performing regional updates
- Development of new methods and tools for DMQC, for example semi-automatic DMQC tool, and rules to apply them
- Adaptation/optimisation of the standard OWC code for the special characteristics of the marginal Seas

Each section is concluded with a summary together with plans and suggestions to guide the future work on this topic.

This work builds up on the recent Euro-Argo project Monitoring the Ocean Climate Change with Argo (MOCCA, 2015-2020), during which the reference dataset for DMQC in the Mediterranean and Black Seas was improved (Deliverable 4.4.1 Report on the reference dataset for the delayed-mode quality control activity in the Mediterranean and Black Sea - Notarstefano, 2019) and the foundations for the data management of Argo floats in the Baltic were laid (Deliverable 4.4.7 Data management for floats in the Baltic - Klein et al., 2020).

2 Baltic Sea

2.1 Generalities

The Baltic Sea is an intra-continental, coastal sea with relatively shallow water depth (mean depth of 55 m) compared to the global ocean. Water exchange to the open ocean is through the North Sea and is restricted by narrow straits and shallow sills (Darss and Drogden) separating deep basins (Arkona Basin, Bornholm Basin, Gotland Basin, Bothnian Sea and Bothnian Bay). Since the baroclinic Rossby deformation radius is small (<10 km) the hydrographic properties are varying on rather small scales. Despite the relatively shallow depth, permanent stratification exists in parts of the Baltic. During winter, a permanent halocline separates the less saline cold winter water from the more saline and warmer deep water in the deep basins. The standard use of OWC is inappropriate due to the shallow water depth, strong vertical salinity gradients, pronounced seasonality and small Rossby radius. But because of the small size and its landlocked structure, floats deployed into the deep basins are often recovered after moderate periods at sea. This allows for recurrent lab calibrations to check sensor drift. Currently, it is common practice in the Baltic to try to recover the floats after 1 to 2 years at sea. In contrast with the Argo floats operating in the open oceans, which are normally not recaptured and stay in the water for periods of 4 years or longer, recovering a float after a year limits the time period over which sensor issues can develop and salinity drift is expected to be small anyway. Since recovering floats is not a standard practice considered in the Argo Programme, it is also necessary to establish rules for DMQC operators on how to apply corrections considering lab calibrations.

For floats in the Baltic that are not recovered, establishing min/max climatologies based on high quality ship-based CTD data from the deep basins offers the potential to perform quality checks on profiles, and thus rules on how to apply specific quality flags should be provided. Potential methods for establishing the long-term stability of the non-recovered floats and to determine salinity corrections also have to be explored. These activities require: i) compilation of an appropriate reference database and a methodology to calculate the climatological statistics, and ii) the definition of selection criteria for the temporal and spatial search scales of reference profiles, as well as depth horizons to be evaluated. The existing international monitoring programmes (HELCOM) is a good source of CTD data for this purpose, from which many profiles are available (Figure 2.1) and which is regularly updated.

During mild and normal winters, ice cover occupies 15–50% of the sea area in the northeastern parts of the Baltic, but may extend to the whole sea during the infrequently occurring severe winters. So far, four floats that operated in the Bothnian Sea (WMO numbers 6902023, 6902025, 6902026 and 6902028[1]) have encountered ice and have profiled under it for periods ranging 1 to 6 months and thus the positions assigned were calculated via linear interpolation. Since the distances covered under-ice are short this estimation method seems acceptable for the region. In the future, as more profiles under the ice are available, other interpolation methods should be tested to assess if the position quality can be improved.

---

[1] https://fleetmonitoring.euro-argo.eu/dashboard?Status=Inactive&WMO=6902023,6902025,6902026,6902028

2.2 Requirements

2.2.1 Establishment of a CTD reference dataset

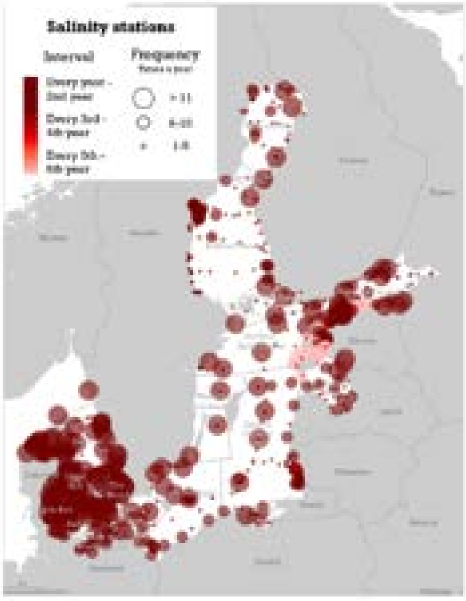


Figure 2.1: Coverage of salinity measurements in the Baltic within the HELCOM monitoring plan. The size of the circles indicates the measurement frequency within a year and the colour the repletion interval (source Helcom MORE report, 2013).

The environmental conditions of the Baltic Sea have been monitored in international cooperation under the Baltic Sea Environment Protection Commissions Helsinki Commission (HELCOM) monitoring programs since 1979 (HELCOM, 2017). All the coastal countries of the Baltic Sea are members of HELCOM and provide data to the HELCOM database, which is freely available at the International Council for the Exploration of the Sea - ICES (www.ices.dk). A total of almost 800 stations are regularly monitored and provide reference data for salinity. Observation station BY15 in the Gotland Deep (57.32N 20.05 E) is one of the most visited monitoring stations in the Baltic Sea Proper, and it is a benchmark of the state of the Baltic Sea.

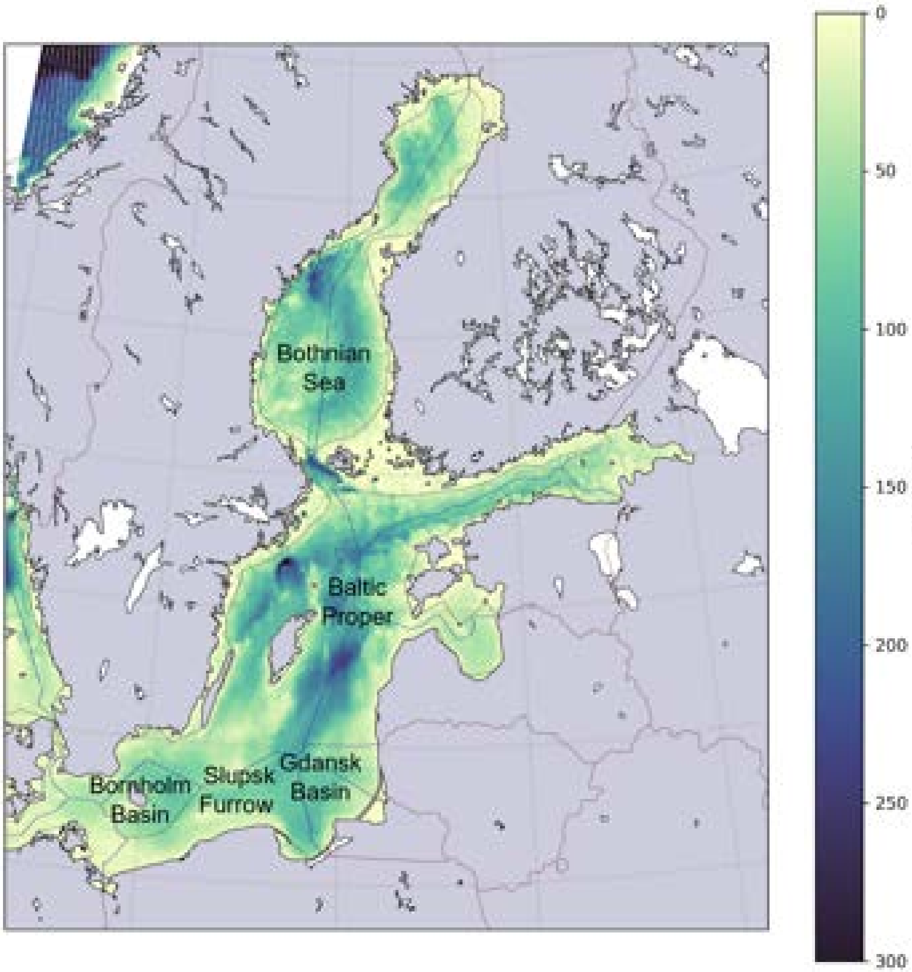


Figure 2.2. Map of the basins with min/max climatologies. Blue lines indicate the Exclusive Economic Zones (EEZ) of each country, and grey lines their internal waters.

In Euro-Argo RISE, min/max climatologies have been created for the deep basins shown in Figure 2.2 to be used for quality checks as the floats drift through. FMI was in charge of providing the min/max statistics for the northern basins (Baltic Proper and Bothnian Sea) using ICES/HELCOM data. IOPAN has traditionally created such statistics for the southern basins (Bornholm basin, Slupsk Furrow and Gdansk Deep) and was responsible for the calculation of these statistics there using data from their own cruises. These statistics could be used in the future to implement the min/max near real-time quality control (Gourrion et al., 2020a, 2020b), conducted by Coriolis operationally on the open ocean floats, also in the Baltic. Communication is ongoing with the team responsible to provide the min/max statistics for this test at POKAPOK (http://www.marineinsitu.eu/partners/pokapok/) to discuss and collaborate on this topic. In the future, some estimations of how many samples/profiles would be flagged using the statistics need to be generated to assess the applicability of this test.

Northern Basins

The salinity in the Baltic Proper is higher and has more variance compared with northern basins such as the Bothnian Sea. In the Baltic Proper, it ranges from under 6 to near 8 PSU on the surface, to around 10-14 PSU in lower layers, while in the Bothnian Sea it ranges from 3.5 to near 6 PSU on surface, to 6-7 PSU in lower layers. The Baltic Proper is more affected by salinity pulses through Danish straits, while the Bothnian Sea is more isolated, as well as shallower. The surface temperature varies considerably seasonally, between near 0 °C to occasional 20 °C in Baltic Proper, and from below zero (as the parts of the area frequently covered with ice) to 19 °C. The shape of the profiles varies seasonally considerably due to mixed layer forming and dissolving.

The climatological data for the Baltic Proper and Bothnian Sea are illustrated in Figure 2.3. The Baltic Proper dataset is cropped after 230 metres, and the Bothnian Sea dataset after 120 metres. Both areas have deeper sections, but there is considerably less data available, as only few areas go deeper than these cropping limits. In Baltic Proper, Landsort Deep is the only area deeper than the threshold. In the Bothnian Sea the deep area is more spread, but still small enough that the dataset gathered from the depths over the threshold is considerably smaller than for the depths above it. In both cases, including the deeper values directly would cause unphysical discontinuity on the dataset, as the amount and distribution of the deeper data is not comparable to the rest of the basin. For further improving both datasets, especially if/when Argo operations will reach the deepest Baltic Proper area, the dataset could be splitted so that the deep areas would have their own dataset.

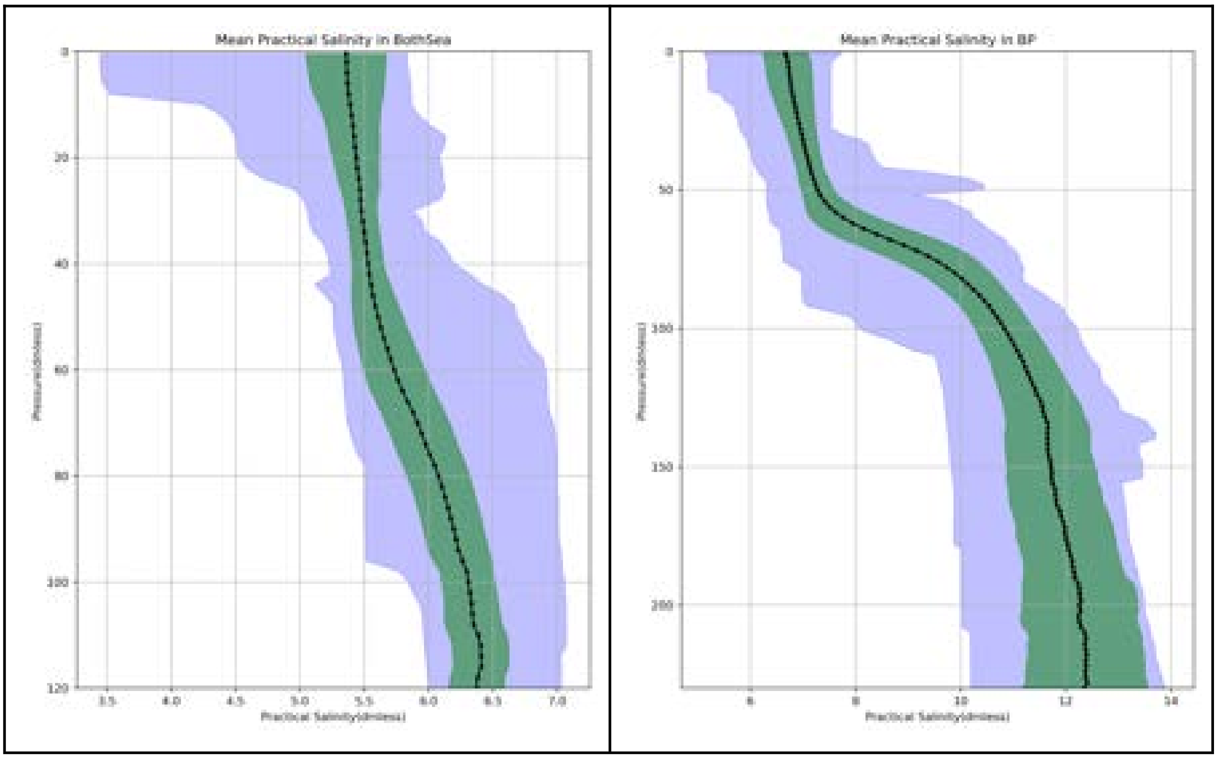


Figure 2.3. Min/Max climatologies for the Bothnian Sea and Baltic Proper. Mean value for each depth is indicated in black, one standard deviation in green. The lilac area shows the range of minimum and maximum values.

Southern Basins

IOPAN has created the min/max climatologies from their own monitoring activities and has been collecting CTD data during regular cruises of r/v Oceania in the southern Baltic since 1998. The frequency of the measurements is presented in Figure 2.4. There is some lack of data in the warmest months because r/v Oceania is in the Arctic at that time. The research area concerned selected basins in the Baltic Proper, such as Gdansk Basin, Slupsk Furrow and Bornholm Basin (Figure 2.5).

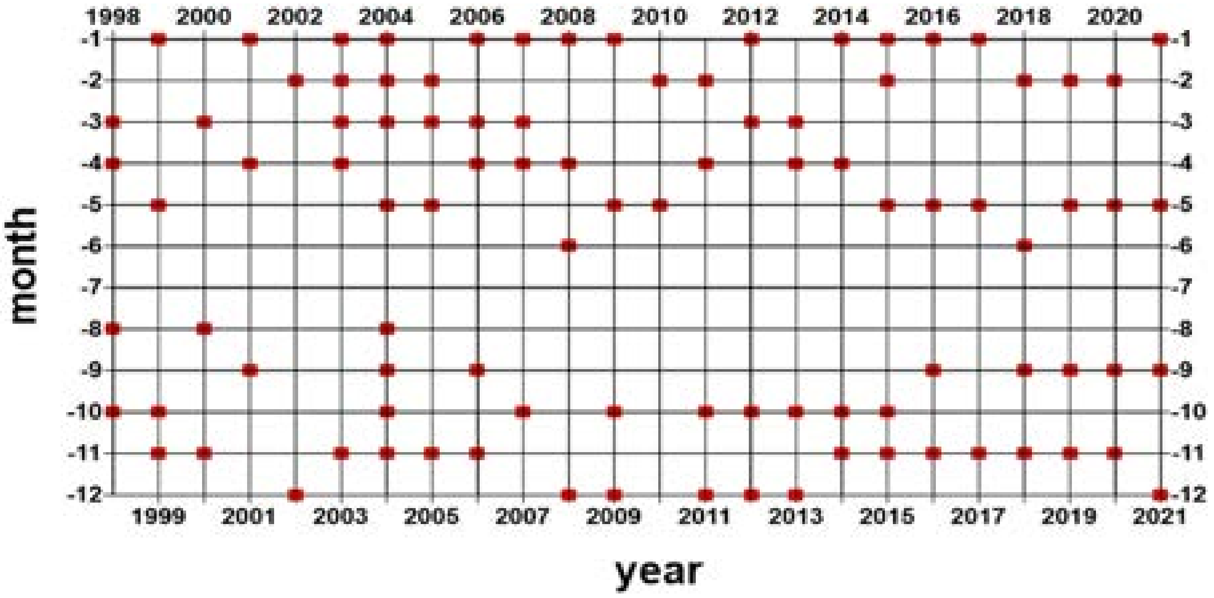


Figure 2.4. Monthly cruises schedule r/v Oceania 1998-2021.

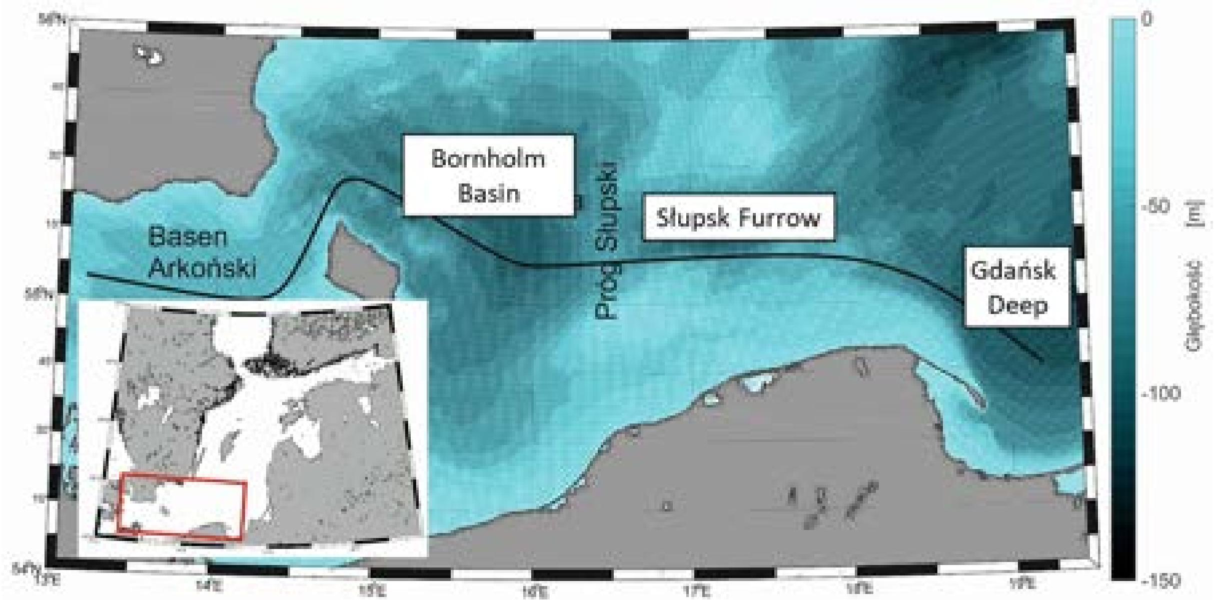


Figure 2.5. The main IOPAN hydrographic section through the deep basins of the Proper Baltic.

There are significant differences in the properties of the waters between the individual basins. This is due to the different distances from the Danish Straits, which are the only source of saltwater of North Sea origin. The Bornholm Basin is the closest to the Straits, Slupsk Furrow is a zone where transport of salty bottom water to Gdansk Deep and to the rest of the Baltic Sea takes place.

The annual heating-cooling cycle leads to the greatest temperature variability in the upper layer. The smallest temperature variation occurs in the bottom layer. The variability of salinity is opposite to the variability of temperature. The smallest salinity variation occurs in the upper layer, where this property shows a relatively homogeneous character. The highest variability in salinity is in the bottom layer due to the inflow of dense, highly saline water from the North Sea. The greatest changes in salinity can be expected in the areas closest to the Danish Straits and with the distance, these changes decrease (Figure 2.6).

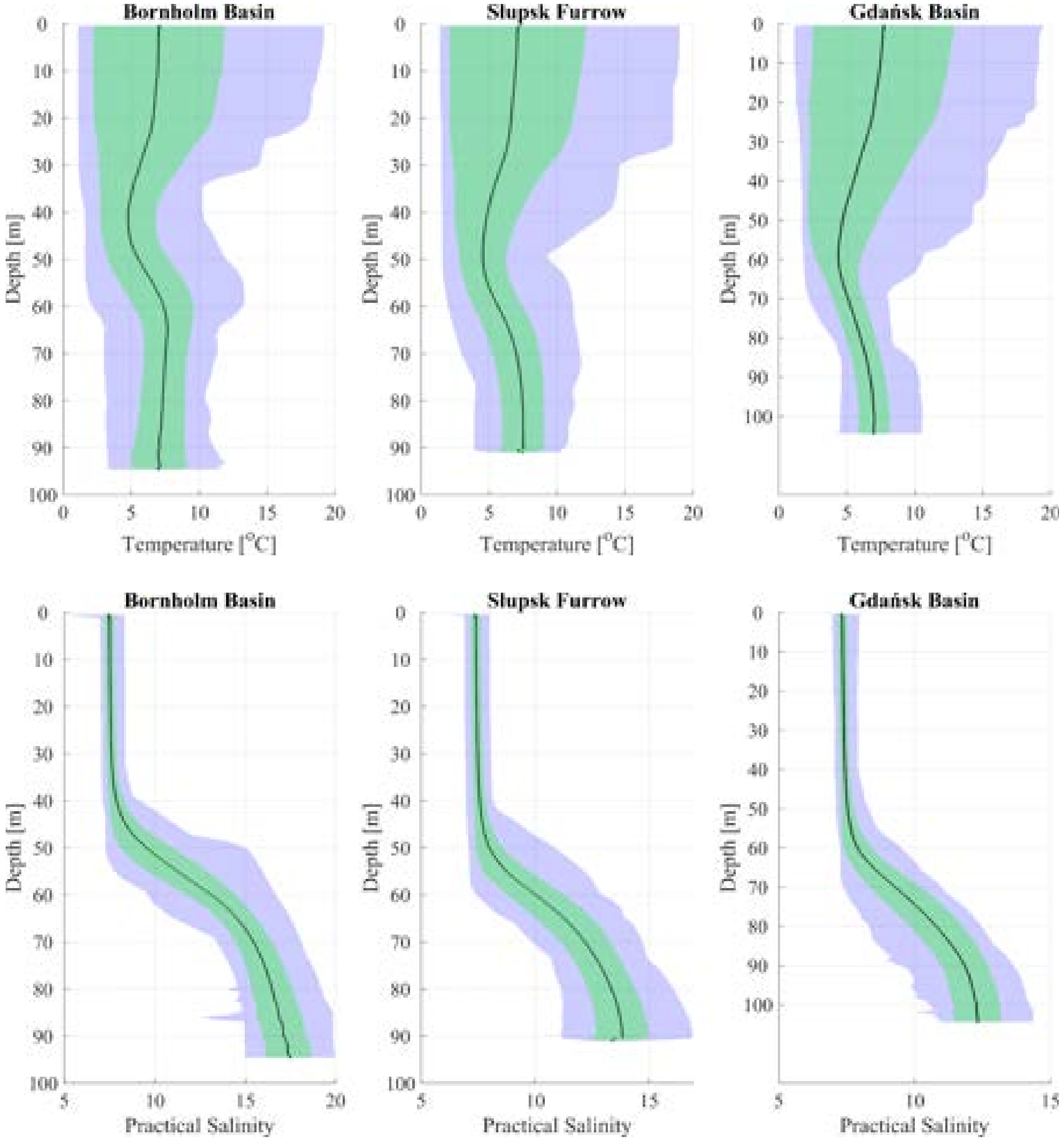


Figure 2.6. Min/Max climatologies for the Bornholm Basin, Słupsk Furrow and Gdańsk Basin from IOPAN. Mean value for each depth is indicated in black, one standard deviation in green. The Lilac area shows the range of minimum and maximum values on the reference dataset.

2.2.2 Updates to the CTD reference dataset

FMI has set up a directory (ear_dmqc) on their FTP server ftp.fmi.fi for the exchange of reference data with subdirectories for each of the contributing groups (FMI, IOPAN and BSH) to which the authorised personnel can access with a common password protected login. The statistics for temperature and salinity per depth are stored individually for each basin and parameter as spreadsheets (.csv or .xls) in the corresponding subdirectories. The stored statistics are: mean, standard deviation, minimum, maximum, 25%, 50% and 75% percentiles, and number of profiles used. Updates to the min/max climatologies should be provided every 5(10) years and FMI and IO-PAN will be in charge of their respective areas.

Additionally, all the profiles extracted from ICES are stored in their original format (ODV text) in the FMI subdirectory, while the r/v Oceania profiles from 2016 to 2021 are stored in the IO-PAN in a Matlab file (.mat) with each cruise in an individual structure variable. The ICES/HELCOM dataset will be updated in annual intervals by FMI. IOPAN data will also be updated yearly with the new cruises. Both FMI and IOPAN are responsible for the quality control of the profiles before delivering. BSH will store all currently available profiles in the usual format of the CTD reference database (as described Section 4.5 of Wong et al., 2022) and will deliver them to the responsibles for the global database maintenance (Coriolis/IFREMER) to make them available for all interested Argo Operators for the next release. Since the origin of the profiles is traced in the variable *qclevel,* the codes ICE (ICES) and SPI (scientists/principal investigators) will be assigned respectively.

2.2.3 Recommendations about the real-time quality control tests

An initial analysis of the real-time quality test has already been done during the MOCCA project and is published in the report D4.4.7 Data management for floats in the Baltic (Klein et al., 2020). This report was based on the findings of 32 floats in the Baltic which had delivered 6689 profiles until 22.04.2020 with 501.780 measurements of temperature (T), salinity (S) and pressure (P). More than 99.3% of the data at this time had been flagged as good in real-time applying the set of tests to the primary profile as described in the Argo QC manual (Wong et al., 2022). Since the real-time tests were designed for the global ocean and have been tested for the hydrographic conditions in 2000 m deep profiles, the results of the real-time flagging in the Baltic needed to be checked for false positives and negatives. Basically four tests raised concerns and were then investigated further: Density inversion, stuck value, digit rollover and gradient test. Moreover, a new test is suggested to flag data with wrong (too high) salinities near the surface.

2.2.3.1 Density inversion test

The real-time density inversion test uses a threshold of 0.03 kg/m^{-3} and catches some small hooks at the base of the mixed layer of the Baltic (see Fig. 2.7). The data points shown in red in Figure 2.7 are flagged as bad. Failed data points from the density inversion test most often occur at the base of the mixed layer where salinity shows a small fresh hook and density in that depth layer is therefore decreasing. The occurrence of density inversions is slightly increased in the descending profiles where the float measures downward. Based on the experience of FMI and IO-PAN the threshold used to flag the data as bad appears reasonable and it is suggested that DMQC operators use thermal lag corrections to check if this reduces the salinity hook.

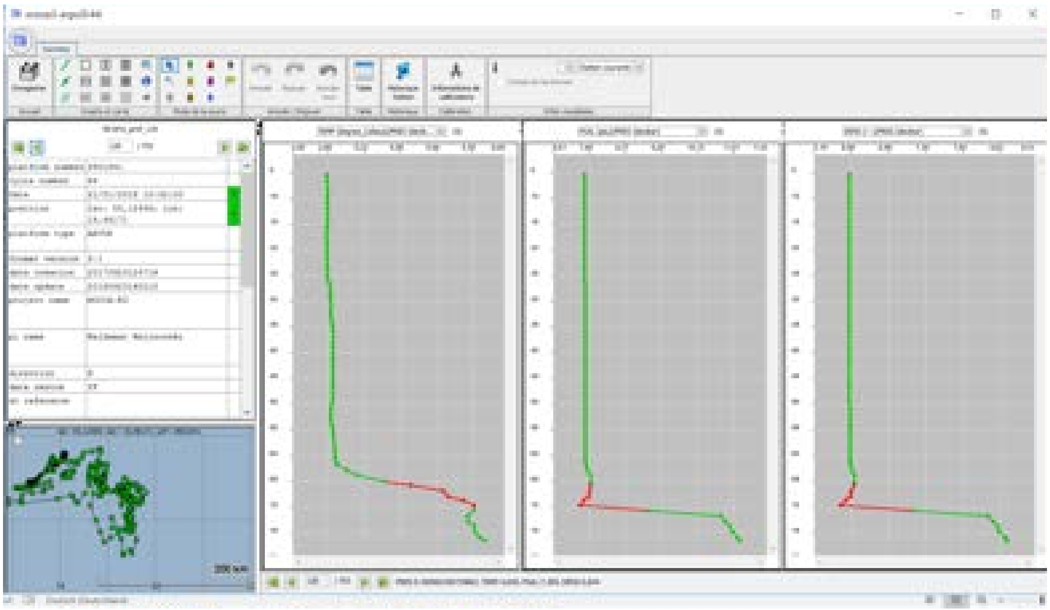


Figure 2.7. Visualisation with Scoop (Detoc et al., 2021) of a profile which failed the density inversion test. Shown is cycle 64D (descending mode) from float 3901941. From Klein et al., (2020).

2.2.3.2 Digit rollover test

The digit rollover test is a remnant from early Argo days when only a limited amount of bits were available for transmission in the Argos satellite system. The range of encountered temperature and salinity data however was not always large enough to accommodate them and when the range was exceeded stored values rolled over to the lower end of the range. To detect the rollover the test considers temperature differences between adjacent pressures > 10°C as a sign of rollover and salinity differences of >5 psu. This test is now longer necessary in modern floats with increased bandwidth for transmission. It was also never designed for strong, shallow thermo- and haloclines as encountered in the Baltic. In the example shown in Figure 2.8, the test fails and wrongly flags cycle 173A as bad because of the strong thermocline gradient of > 10 °C . This demonstrates that this test should definitely be disabled for the Baltic.

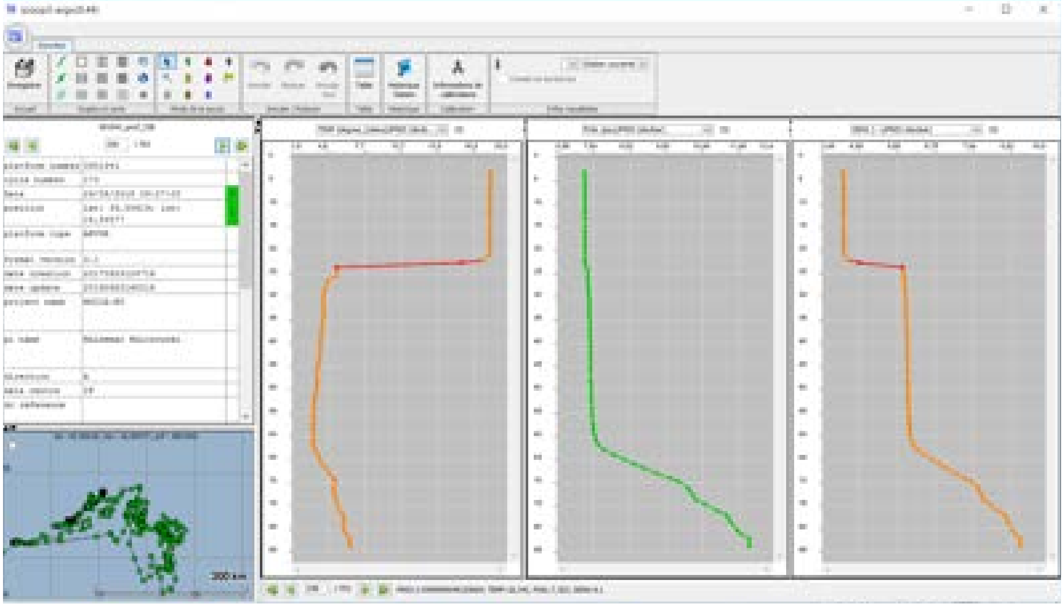


Figure 2.8. Visualisation with Scoop (Detoc et al., 2021) of a profile which failed the digit rollover test. Shown is cycle 173A (ascending mode) from float 3901941. From Klein et al., (2020).

2.2.3.3 Stuck value test

The stuck value test looks for measurements of temperature and salinity in a profile being identical. In the figure below (Figure 2.9) the profiles of temperature and salinity are nearly constant. All salinities are exactly the same and thus are flagged as bad, while temperatures at least show a 0.6 mK standard deviation and thus escaped a degradation in flagging. This test was never intended to work on short profiles and has the potential to catch homogenous winter profiles. It should be disabled for the Baltic.

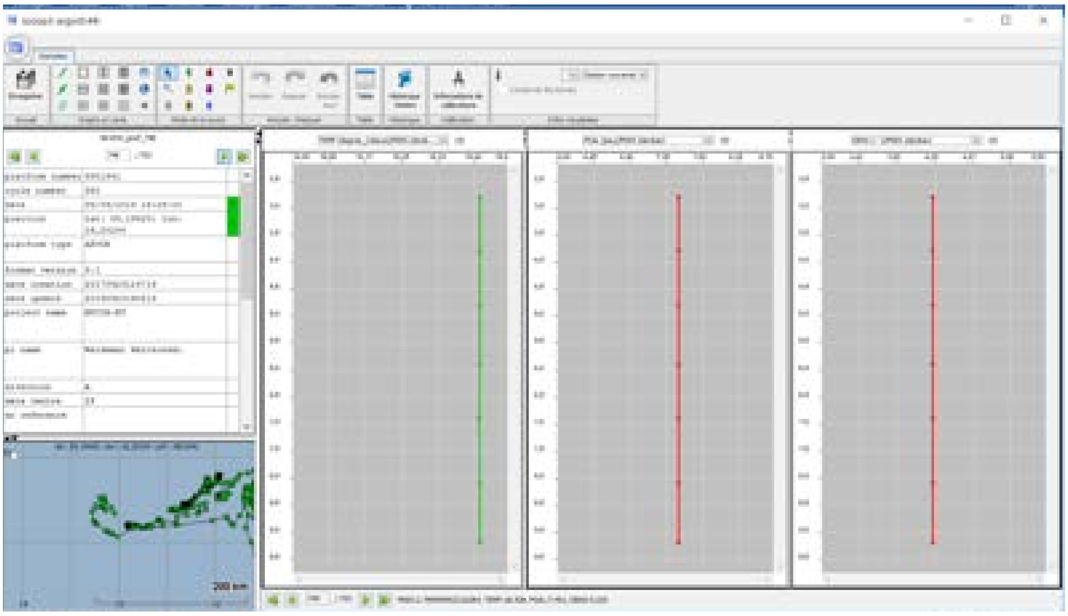


Figure 2.9. Visualisation with Scoop (Detoc et al., 2021) of a profile which failed the stuck value test. Shown is cycle 380A from float 3901941. From Klein et al., (2020).

2.2.3.4 Gradient test

This test had been designed to assign a 'bad data' quality flag when the difference between vertically adjacent measurements is too steep and has the potential to catch strong gradients in deep layers of the Baltic caused by the inflowing North Sea waters. It was declared obsolete in October 2019 at ADMT20 but it seems that the data from the Baltic floats need to be reprocessed to accommodate this decision, since the currently available netCDF files still have samples flagged because of this test. The DAC/GDAC will be asked to reprocess all data in the Baltic to remove those flags.

The example shown in Figure 2.10 shows that the gradient test was triggered in the bottom layer of profile 17A and marked all salinities below 35 m as bad. But in the Baltic strong gradients can be present at the bottom, because of the inflow of salty waters from the North Sea, and these should not be removed. In the way that the test is designed, it only considered temperature and salinity data between adjacent measurements and never considered the associated depth layer over which the change occurred. Having fine resolution data, as shown here, maximises the risk of false positives.

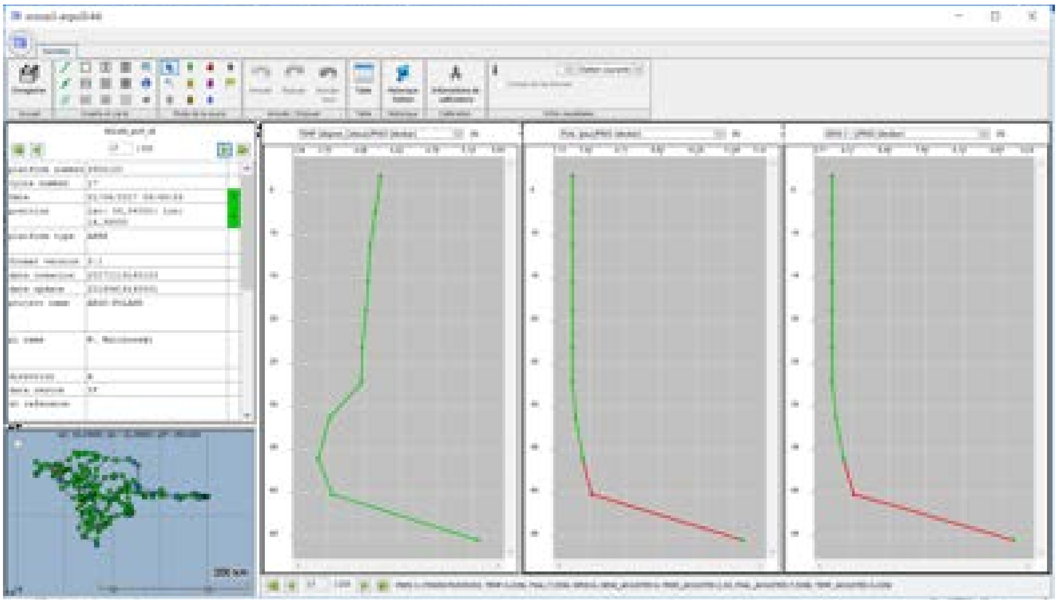


Figure 2.10. Visualisation with Scoop (Detoc et al., 2021) Example of a profile which failed the gradient test. Shown is cycle 17 A from float 3902100. From Klein et al., (2020).

2.2.3.5 New test required: Incorrect near-surface salinity

As noted already in the MOCCA report (Klein et al., 2020), the current real-time tests in the Baltic failed to detect unreasonable high surface salinities. Surface salinities in the Baltic are quite low because of the large fresh water supply to the Baltic. Sometimes, however, recorded float profiles exhibit surface salinities in excess of 8 salinity units. The issue was initially named clogging and was associated with a unsufficient flushing of the conductivity cell with salty water from the deep layer (Figure 2.11). Often these profiles, with salinity that is too high in the mixed layer, pass all real time tests and only individual data points flagged as 4 (bad data).

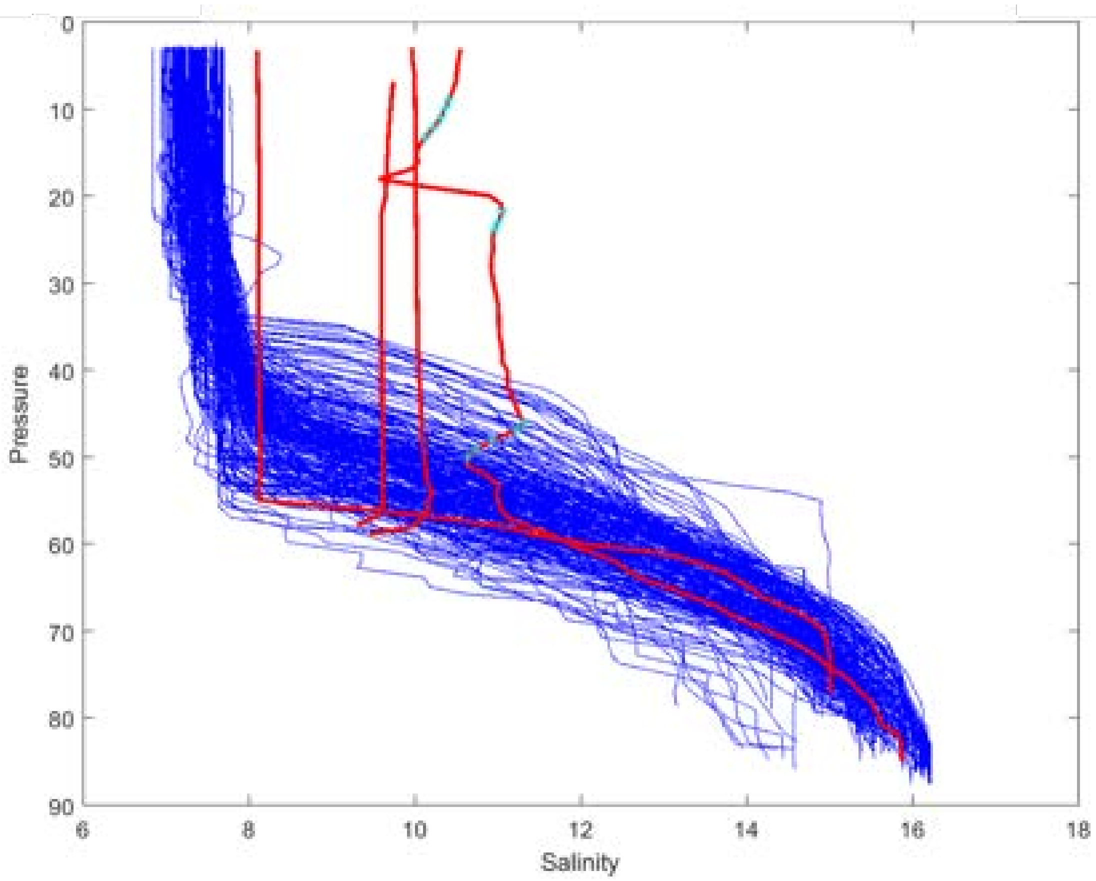


Figure 2.11. Salinity profiles from float 3902115. Profiles in red are showing incorrect near-surface salinities (too high) and correspond to cycles 58, 232, 246 and 247. Measurements marked with a cyan asterisk have received a PSAL_QC of '4' from real time tests indicating bad data.

These profiles are rare and the dataset downloaded in March 2022 contained only 69 occurrences in a total of 11746 profiles (less than 0.6%). The spatial distribution of these 'clogged profiles' is shown in Figure 2.12. They occur in the central parts of the Gotland Basin, as well as in the south along the deep spreading pathway of North Sea Water flowing from Arkona Basin to Bornholm Basin and finally into the Gulf of Gdansk. The particularly strong vertical gradients at these locations, between the salty North Sea Water at depth and the brackish surface waters, suggests a connection between the presence of these erroneous measurements with the strength of the halocline.

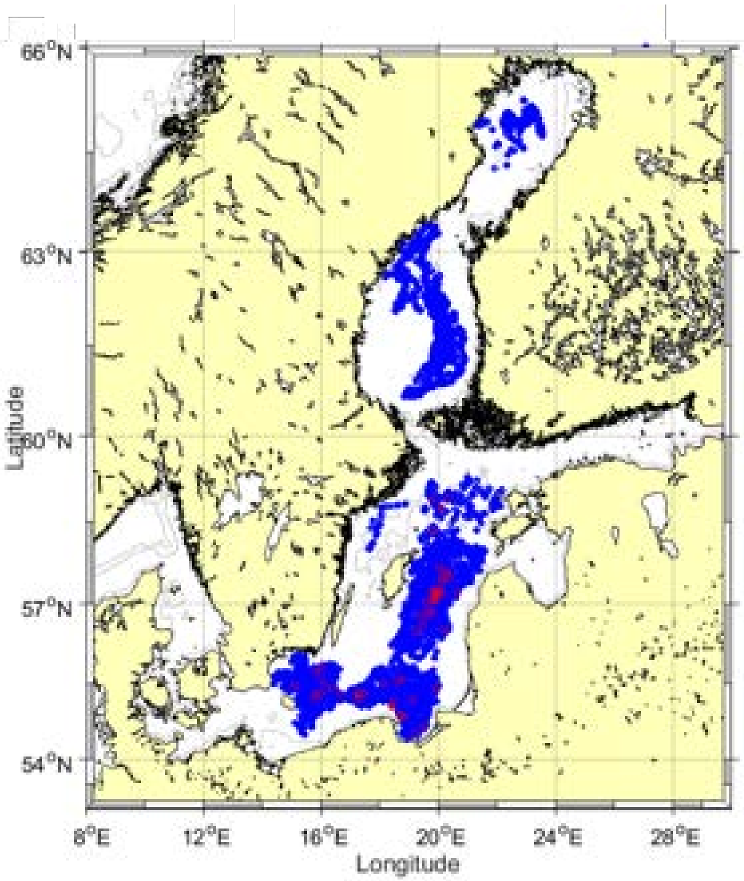


Figure 2.12. Location of Argo profiles in the Baltic as downloaded in March 2022. Profiles in red show excessively high surface salinities.

The causes of the wrong measurements are still not established properly. Although this issue was initially attributed to an imperfect flushing of the conductivity cell (clogging), it is also possible that they originate in an excessive ascending speed of the float, compared with the usual ~10 cm/s, when the float struggles to break the halocline by gaining sufficient buoyancy or air trapped in the cell. This needs to be investigated in more detail from the technical data reported back by the float. It was also noted that the issue also occurs in descending profiles which would not fit the explanations given above.

A simple real-time test is able to catch these profiles and flag them as probably bad in real-time, so the profiles are examined later by the DMQC-operators. The incorrect near-surface salinity test should examine PSAL at depth of 10-25 m and assign the entire profile a PSAL_QC of 3 (probably bad data) if these are greater than 8 salinity units or a regional varying threshold. The analysis showed that a common criterion such as 8 psu creates too many false positives and negatives. On the other hand, a regionally varying test examining salinities between 10-25 m using the thresholds shown in Figure 2.13 retrieved better results. The geographical limits for the test are as follows: The latitude limit between the Bothnian Bay and the Bothnian Sea was set to 63.5 °N. The Northern Baltic Proper was separated from the Bothnian Sea at 59.8 °N, and from the Baltic Proper at 58.4 °N. The Bornholm Basin and the Gdansk Basin are separated from the Baltic Proper at 56.2 °N and from each other at 17.5 °E. The westernmost limit for the Bornholm basin was set at 14.4 °E. The test identifies 67 out of the 71 manually labelled profiles correctly, but misses 4 profiles (2 from Bornholm Basin and 2 from Gdansk Basin). It also catches 2 more profiles not related to this issue with unflagged spikes in the profile. Since determining the thermocline depth to flag only the data above it as probably bad, we propose to flag the entire profile as bad, even below the mixed layer, and change the flags appropriately in delayed-mode. The code used to evaluate the performance of this test is available in a private repository in the Euroargodev (https://github.com/euroargodev/baltic-dmqc-tools).

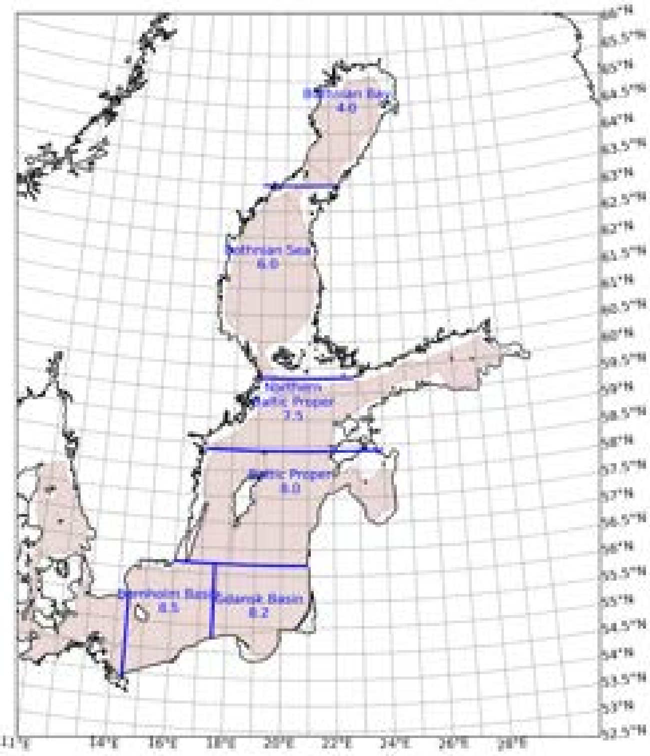


Figure 2.13. Salinity thresholds used for the '*Incorrect near surface salinities' test*

Recommendations for RTQC (real-time quality control tests)

- Disable the digit rollover and stuck values tests for the Baltic.
- Continue to use a threshold of 0.03 kg/m3 for density inversions in the Baltic and apply thermal lag corrections in delayed-mode.
- Add a regional real time test for Baltic to catch excessively high surface salinities. The proposed test would look at salinities at depth of 10-25 m and assign a QC of 3 to the entire profiles if these are larger than a regionally varying threshold.
- Communicate and discuss the proposed new set of rules at the next ADMT. If these are endorsed, then communicate with the GDAC/DAC about the reprocessing of all Baltic floats, for implementation.
- Explore if the implementation of the min/max near real-time quality control is appropriate for the Baltic.

2.2.4 Salinity calibrations based on lab calibrations

Because of the landlocked nature of the Baltic it is possible to recover floats at the end of their lifetime or earlier for technical inspection. FMI has successfully recovered most of its floats in the northern Baltic since the start of their operations. Table 2.1 lists all floats that have operated in the Baltic and the colour-code in the CTD serial number column is used to highlight the floats/CTD sensors that have been re-deployed after calibration, since a new WMO number is reassigned for each (re)deployment.

WMO number	Float serial. No	CTD serial no.	Float type	Country/ Programme	Deployment date	Date of last profile	Calib. of CTD after recovery
6901901	5397	3511	APEX	Argo Finland	17.05.2012	05.12.2012	y
6902013	5396	3503	APEX	Argo Finland	13.06.2013	02.10.2013	y
6902014	6711	4793	APEX	Argo Finland	14.08.2013	21.08.2014	y
6902017	5397	3511	APEX	Argo Finland	31.05.2014	24.10.2015	y
6902018	6710	5051	APEX	Argo Finland	31.05.2014	13.11.2014	y
6902019	7191	5699	APEX	Argo Finland	21.08.2014	05.08.2015	y
6902020	6711	4793	APEX	Argo Finland	05.08.2015	03.08.2016	y
6902021	6710	5051	APEX	Argo Finland	22.09.2015	13.05.2016	y
6902022	5396	3503	APEX	Argo Finland	13.05.2016	11.10.2016	y
6902023	5397	3511	APEX	Argo Finland	13.07.2016	25.01.2018	n
6902024	7191	5699	APEX	Argo Finland	03.08.2016	15.06.2017	y
6902036	7507	7248	APEX	Argo Poland	29.11.2016	01.02.2017	n
6902025	7958	8893	APEX	Argo Finland	09.05.2017	02.10.2018	n
6902026	7959	8894	APEX	Argo Finland	06.06.2017	02.06.2019	n
6902027	6711	4793	APEX	Argo Finland	15.06.2017	15.10.2018	n
6902028	6710	5051	APEX	Argo Finland	06.08.2017	04.09.2018	n
6902029	5396	3503	APEX	Argo Finland	06.08.2017	27.10.2017	y
3901940	AI2600-16FR083	8519	Arvor	MOCCA-EU	20.09.2017	04.10.2017	n
3901941	AI2600-16RF084	8498	Arvor	MOCCA-EU	21.09.2017	10.09.2019	y
3902100	7507	7248	APEX	Argo Poland	15.03.2017	07.01.2018	n
3902133	AI2600-16FR083	8519	Arvor	MOCCA-EU	03.11.2017	09.09.2019	n
6902030	5396	3503	APEX	Argo Finland	10.07.2018	04.03.2019	n
3902101	AI2632-17EU025	10114	Arvor	Argo Poland	06.02.2018	07.02.2020	n
3902104	AI2632-18EU005	4991	Arvor	Argo Poland	31.05.2018	10.09.2018	n
3902106	AI2632-18EU005	4991	Arvor	Argo Poland	11.09.2018	01.12.2020	y
3902134	AI2600-17EU010	10666	Arvor	MOCCA-EU	04.10.2018	11.01.2020	n
3902137	AI2600-17EU013	10073	Arvor	MOCCA-EU	09.11.2018	alive	-------------
6903696	AC0300-17FI001	10012	Arvorc	Argo Finland	30.09.2018	5.12.2018	n
6903697	7191	5699	APEX	Argo Finland	15.10.2018	18.08.2019	n
6903698	AI2600-18FI001	10946	Arvor	Argo Finland	30.05.2019	alive	-------------
6903699	8541	10306	APEX	Argo Finland	30.05.2019	alive	-------------
6903700	8543	10379	APEX	Argo Finland	01.06.2019	alive	-------------
6903701	8540	10305	APEX	Argo Finland	17.08.2019	alive	-------------
6903703	AI2600-19FI001	11996	Arvor	Euro-Argo RISE	10.06.2020	5.02.2022	-------------

WMO number	Float serial. No	CTD serial no.	Float type	Country/ Programme	Deployment date	Date of last profile	Calib. of CTD after recovery
6903704	8907	11793	APEX	Argo Finland	10.06.2020	alive	-------------
6903706	lovuse005b	5190	Provor	Euro-Argo RISE	19.05.2021	7.4.2022	-------------
6903707	8542	10309	APEX	Argo Finland	06.06.2021	09.09.2021	n, lost
6903708	6711	4793	APEX	Argo Finland	19.05.2021	alive	-------------
6903709	AI3500-20FI001	205366	Arvor	Euro-Argo RISE	19.05.2021	alive	-------------
6903710	AI3500-20FI002	205364	Arvor	Euro-Argo RISE	21.05.2021	alive	-------------
6903711	6710	5051	APEX	Argo Finland	21.05.2021	alive	-------------
3902110	AI2632-19EU033	12172	Arvor	Argo Poland	25.05.2020	alive	-------------
3902109	AI2600-19PL001	12170	Arvor	Euro-Argo RISE	03.06.2020	05.08.2021	n
3902115	AI2632-21EU012	13838	Arvor	Argo Poland	15.06.2021	alive	-------------

Table 2.1. List of floats in the Baltic. The colour code in column CTD serial number is a visual help to identify reuse of specific CTDs after recalibration.

In case that a float is recovered and returned to the CTD manufacturer SeaBird for lab calibration, the drift in the conductivity sensor can be estimated directly. Instructions on how to calculate the drift rate are given by SBE in their Application note 31: *"Correcting for Conductivity Drift Based on Pre- and Post-Cruise Laboratory Calibrations"*. The note recommends the use of the SeaSoft Software to determine the drift. However, not all drift rates require a correction of the salinity data recorded by the float. It is necessary to define a threshold to separate between the drift rates that are considered noise and those that demand salinity corrections. With an expected accuracy level of 0.05 to 0.1 PSU in salinity, a drift rate over a year (12 month) of 0.05/12, or 0.1/12 (PSU/month), should be acceptable and lab based drift rates lower than those would not require a correction.

Figures 2.14 and 2.15 show the calibration reports from SBE for the CTD sensors with serial numbers 4793 and 3503, respectively. From each sensor there are three calibration sets between missions.
In the calibration sheets can be seen that the calibration parameters (g-j) have obtained adjustments between each mission. Using the original and adjusted parameters to calculate the conductivity, one can estimate how much conductivity the sensor would measure in the absence of the subsequent calibrations. In the case of sensor serial number 4793 used in 2013, 2015 and 2017, the average difference between first and second calibration would have been in scale of 2-4 10^{-5} S/m (2-5 10^{-4} salinity) which is well within acceptable range. In the case of sensor 3503, the adjustments in the calibration parameters were larger. Recalculating the conductivity with the original values would have resulted in differences on the order of 2 10^{-2} S/M, which is equivalent to 0.1 in salinity. This value is in the limit of acceptable drift, but there is a caveat regarding the calibration method. The SBE's lab calibration is made for values around 3 S/m, while in the Baltic Sea the conductivities are below 1 S/m. This means that it is likely that the error is roughly one third smaller on the Baltic Sea conditions, making a 0.1 drift in salinity an acceptable level that would not require correction. However the present experience with the recovery of floats and recalibration does indicate that, for ensured accuracy of the measurements, it would be preferable to recalibrate the sensors between missions when possible.

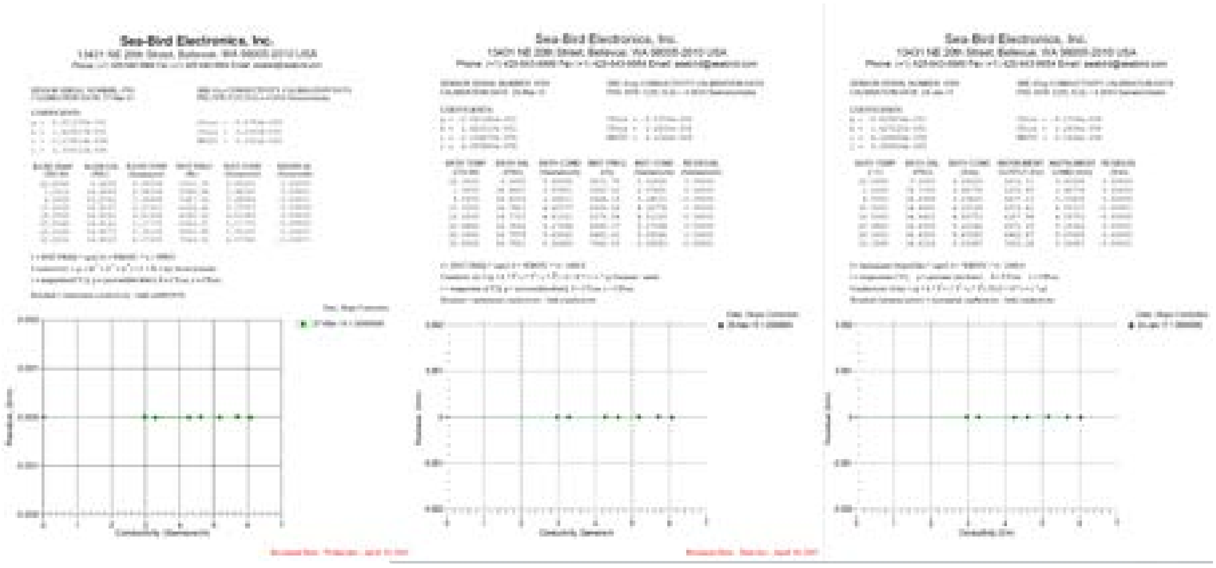


Figure 2.14. Calibration sheets of SBE SN 4793, from years 2013,2015 and 2017. First is calibration before WMO 69014, and the middle one after it and before WMO 6902020. Last one is after recovering it from that mission.

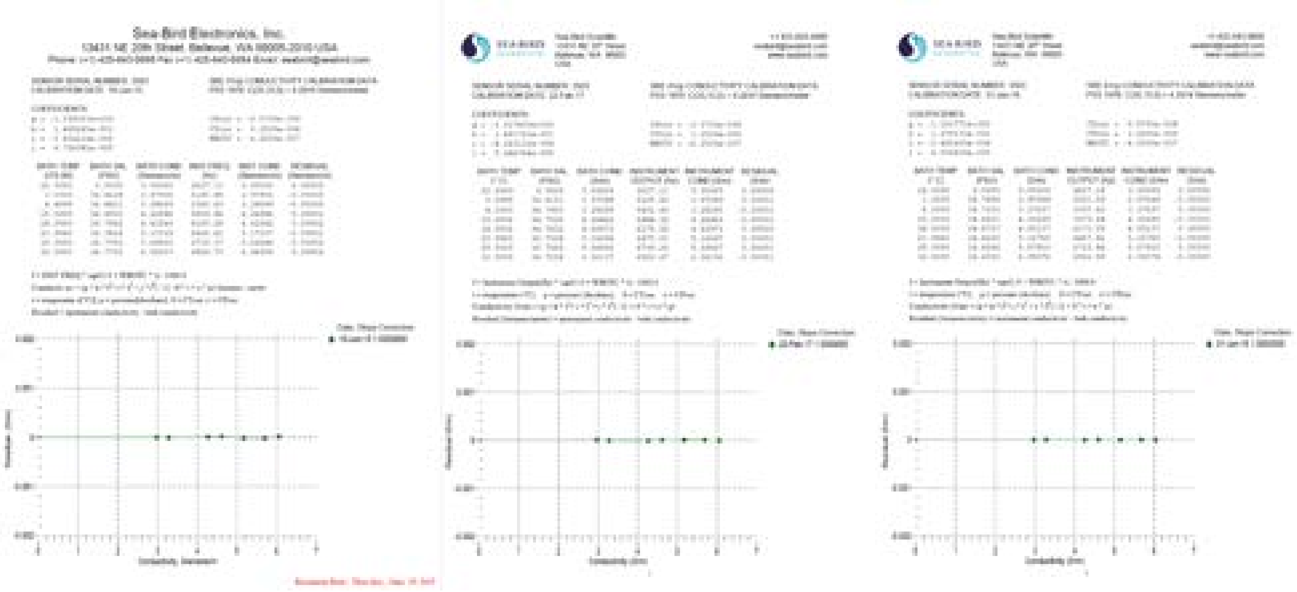


Figure 2.15. Calibration sheets from SBE SN 3503, from years 2015, 2017 and 2018. First one is after WMO 6902013 and before WMO 6902022. The middle one is before the next mission, WMO 6902029, and the last one after recovery from that mission.

The DMQC operators that carry out the quality assessments for the Baltic floats, and that are going to deliver data files with the calibrated salinity to the data centres, will need access to the calibration sheets and thus they should be stored at a central accessible directory within the Argo data management stream. The DMQC operator would thus be able to perform checks of coefficients entered in the metafiles and to calculate drift rates. We propose that the FMI FTP server, mentioned in subsection 2.2.2, will be used to share this information. The development of a python tool that performs the same drift rate calculations as the SBE software will be necessary in case the DMQC operator has no access to the SBE proprietary software.

As a final step of the quality assessment, the DMQC operator has to write netCDF files with the calibrated parameter fields for pressure, temperature and salinity. This step also involves filling the respective SCIENTIFIC_CALIB fields in the files which store information on the coefficients, equations and comments about the procedure (Wong et al., 2022). Appropriate standardised text should be recommended to dmqc-operator to fill the SCIENTIC_CALIB fields.

A suggestion for a standard use of the SCIENTIFIC_CALIB_COMMENT for recovered floats with no correction is: “*No significant salinity drift detected from laboratory calibrations of the recovered float. The quoted error is based on the laboratory calibration from xx/xx/20xx and xx/xx/20xx.*”

Recommendations for DMQC based on lab calibrations

- For deployments in the Baltic the best practise to be aimed for is annual or biannual recovery
- Corrections of the float salinity based on lab calibrations should only be performed if the drift per month is larger than 0.008 S/m/month (conductivity difference of 0.1 over 12 month)
- Further develop rules on how to correct salinity (if calibration differences are larger 0.1) when such a case appears. Decisions are needed if we should apply a linear drift correction or it can be determined from the data when the drift began. This has to be done float by float by the DMQC operator.
- The information necessary for the DQMC procedure based on lab calibrations needs to be accessible for the DMQC operators. We recommend that the lists of floats (Table 2.1) and the calibration sheets are stored at a central accessible directory. The directory should also host a list of people responsible for keeping the information up-to-date.

2.2.5 Alternative DMQC procedure for shallow waters (floats which are not recovered)

Detecting drift of the salinity sensor *in situ* is not an easy task in the Baltic because of the large natural variability, the strong seasonality and the shallow water depth. Therefore, drift estimates will result in large uncertainties. The ±0.01 accuracy of the OWC method was determined for the deep ocean and it is much too high for the Baltic. However, the much larger range of salinity values in this marginal sea does not require such levels of accuracy. Based on experience of CTD field work in the Baltic, values are only reported to second digit level and aspirations for accuracy are between 0.05 to 0.1. In Siiriä et al (2018), where the authors compared float and ship CTD data as explained below, the differences found were in those levels.

In case that a float is not recovered and thus can not be validated by the lab calibrations, nearby reference data need to be examined to calculate differences of the float data to the reference data. This requires appropriate search radii in space and time and also selection of appropriate depth levels. Early assessments of the data quality of floats in the northern Baltic were performed by Siiria et al. (2018)., using ICES database CTD profiles as reference data (Figure 2.16). The best match for each Argo profile was determined by comparing the spatial (km) and temporal (days) distance of the Ship CTD and Argo CTD. For the Gotland Deep area, it was determined that the closest match based on these would be the minimum of a relative distance defined as: *spatial distance + 3 (km/days)*temporal distance*. For the comparison, pairs with these relative distances over 35 km were skipped, as too far away for a meaningful comparison. In the examples shown in Figure 2.17, the best matches had temporal differences of under two days, and spatial distance of under 20 km, often under 10 km. In the final comparison, the upper (0 - 30 m) and lower (80-150 m) layers were compared separately, but the exact limits should vary based on the basin. It is reasonable to omit the mixed layer border as it can vary considerably in relatively short distances. In this example, the determined salinity biases were all under 0.1 PSU.

Automating a method such as described above could be used to notice a drift of salinity values. However, it is good to notice that such drift determination is based on statistics. A single pair, even if relatively close match, could easily have a true difference caused by different location or time. A systematic difference in the same direction between several pairs however is likely caused by actual sensor bias. The scripts developed by FMI for this analysis can be found in the baltic-dmqc-tools repository) in the Euroargodev Github.

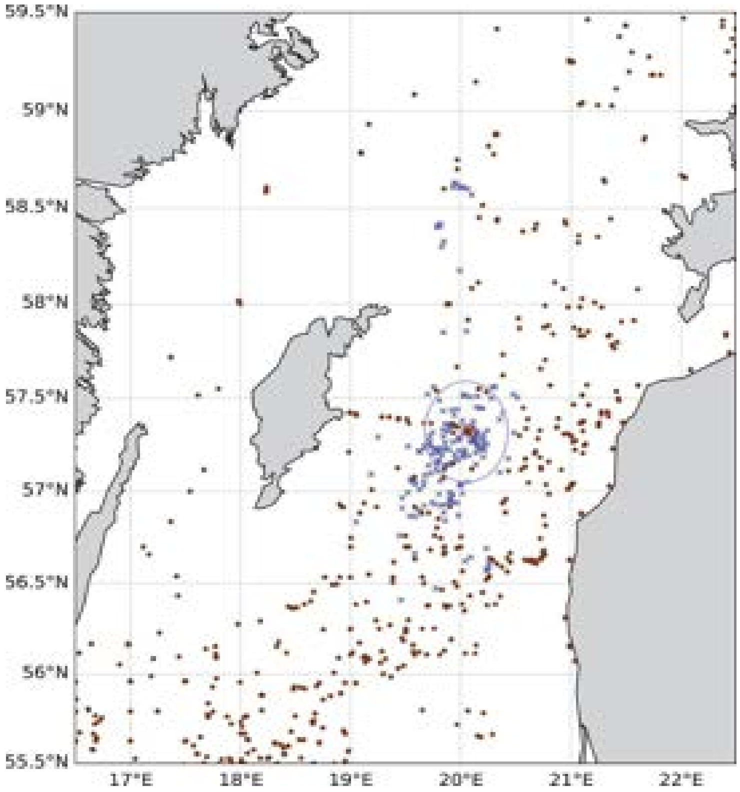


Figure 2.16. Figure 6 From Siiria et al, 2018. Red dots show locations of the ICES data CTD measurements, and blue crosses the Argo float profiles used in the study. Most frequent measuring point of the area, BY15 is encircled with a 30 km radius.

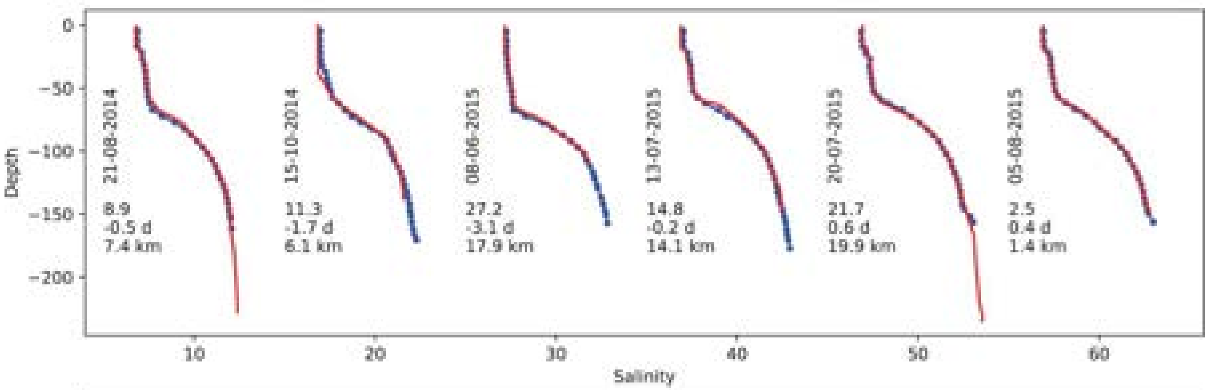


Figure 2.17. Panel from Figure 7 from Siiria et al., (2018) showing some of the salinity comparisons from the article. Red profiles are from ICES data, blue ones are from Argo float 6902019. Numbers listed on each profile show the distance of the pair. Lowermost is the distance of locations, middle one temporal difference in days, and the uppermost combination of these (relative distance), which is used to sort pairs from best match (lowest number) to worst (highest number).

Similar tests have also been performed by IOPAN for the southern Baltic. For this end, a long time series of CTD data from r/v Oceania has been used. As mentioned in section 2.2.2., all available data from 2016 to 2022 were collected and stored in Matlab *.mat files. Also, R-mode data (real-time quality controlled) from floats operating in the region were stored in separate *.mat files. In addition, the profiles whose average salinity in the 0-10 m layer was above 8.1 were removed (Figure 2.18) since they are affected by the issue discussed in subsection 2.2.3.5, and the near-surface samples should be flagged as 4 (bad data).

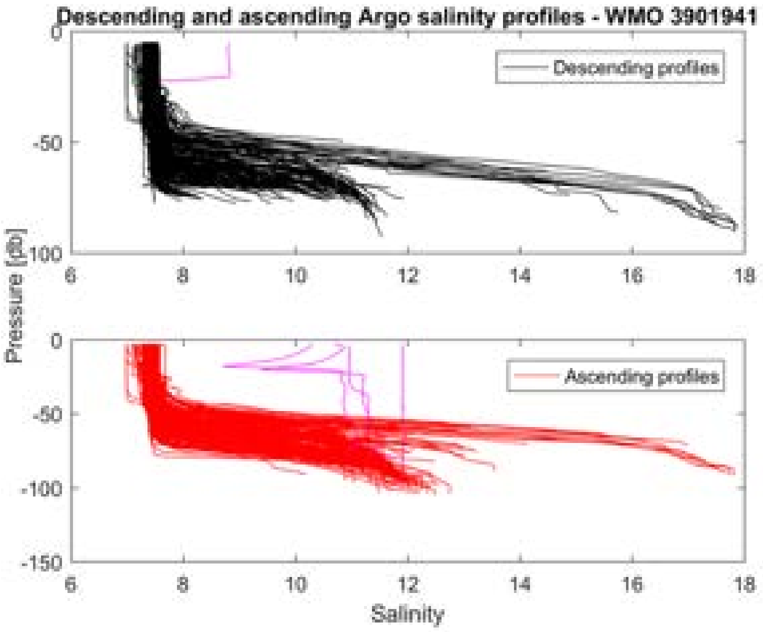


Figure 2.18. Descending and ascending Argo salinity profiles (WMO 3901941). The profiles in magenta are those with the average salinity in the layer 0-10 m greater than or equal to 8.1.

A dedicated Matlab code to do the data comparison and estimation of temporal trend was initiated at IOPAN can be found in the baltic-dmqc-tools repository in the Euroargodev Github. In the first step, the application calculates distances and time differences between each Argo profile and referenced CTD profiles. The closer the Argo profiles are to the reference profiles, the correlation between them is greater. The best results are given by a distance in space and time of no more than 30 km and 30 days. Then figures of Argo profiles and the closest CTD profiles are created (Figure 2.19). The next step was selecting the data from the layer between 10-30 m, which is characterised by its stable salinity. Below, there is a variable pycnocline as well as a deep layer, which is affected by the inflow of dense, highly saline water from the North Sea. In the 0-10 m layer, there is sometimes reduced salinity caused by the pump drawing air in as the instrument surfaced. For each Argo profile and the closest CTD profiles, the mean salinity in the selected layer and difference between them were calculated. Table 2.2. shows the mean difference for each of the analysed floats. Profiles that have salinity difference greater than 0.2 should be flagged as 4 (bad data) (Figure 2.20).

Float WMO	**Salinity (10-30 m)**	**Temperature (70-90 m)**
3902101	0.0508	0.6892
3902104	0.0316	0.8689
3902106	0.0706	0.4608

Table 2.2 Mean salinity and temperature difference between CTD and Argo measurements for three Argo Poland floats.

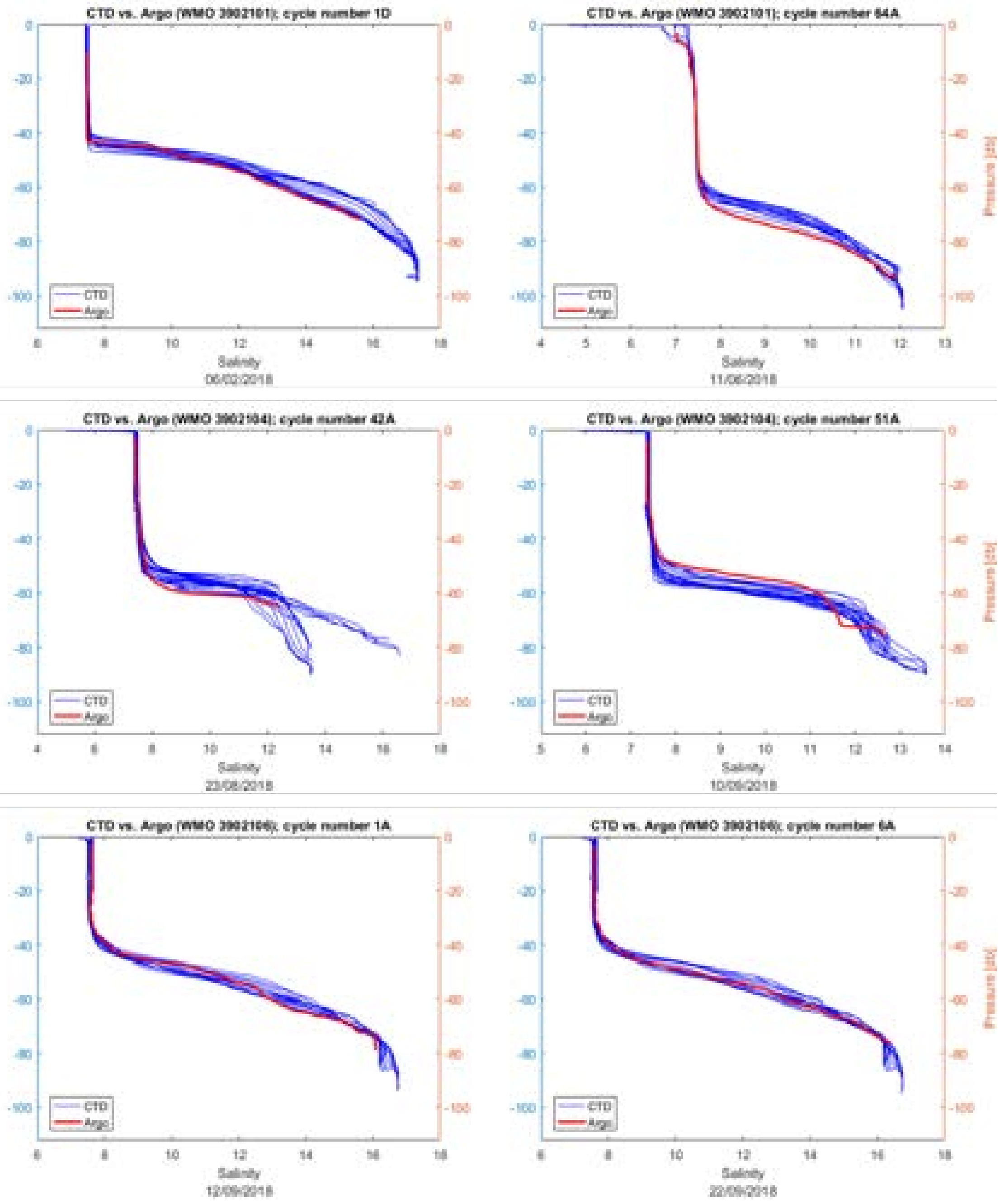

CTD vs. Argo (WMO 3902101); cycle number 1D
06/02/2018
CTD vs. Argo (WMO 3902101); cycle number 64A
11/06/2018
CTD vs. Argo (WMO 3902104); cycle number 42A
23/08/2018
CTD vs. Argo (WMO 3902104); cycle number 51A
10/09/2018
CTD vs. Argo (WMO 3902106); cycle number 1A
12/09/2018
CTD vs. Argo (WMO 3902106); cycle number 6A
22/09/2018

Figure 2.19. Sample comparisons of Argo profiles (red line) and reference profiles (blue lines) from the southern Baltic Sea.

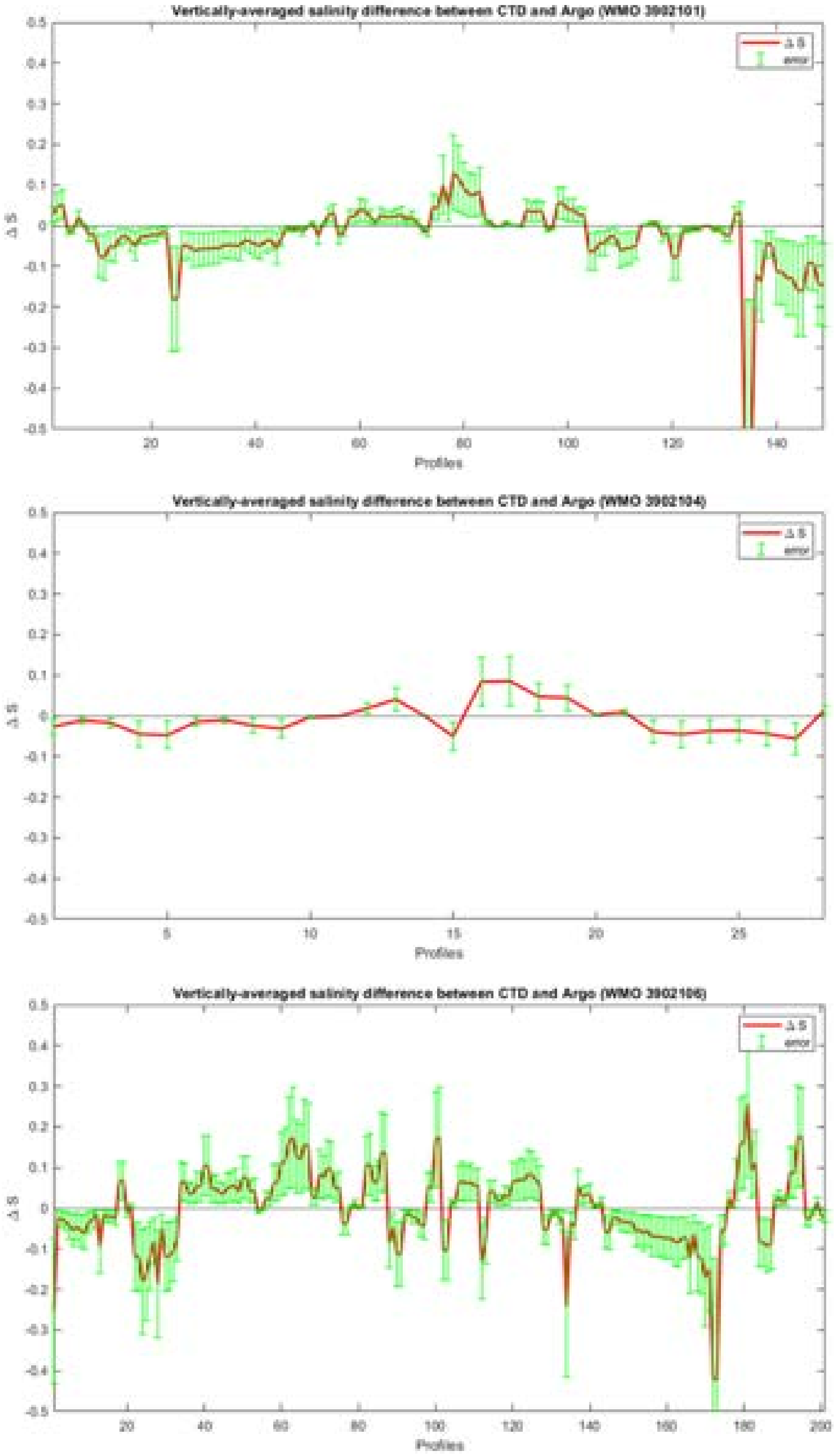


Figure 2.20. Vertically-averaged (10-30 m) salinity difference between CTD and Argo.

Recommendations for DQMC of non recovered floats

- Find the best matches of reference profiles for float profiles using an appropriate search radius (30 km and 30 days). Decide which layer to compare (shallower or deeper) depending on the region. In the southern Baltic only the layer 10-30 m should be taken for comparison.
- Calculate the absolute differences between the float and the reference data
- Check if there is a trend/systematic differences calculated from different profiles/moments in time.
- Assign corresponding flags and errors.

2.3 Summary and future steps

The Baltic Sea characteristics make it a unique place for Argo float operations. The large salinity variability, the unique hydrographic conditions and its shallow depth make alternative methods for float salinity DMQC imperative. But also its landlocked structure allows the operators to recover most of their floats for lab calibrations. In this report the basis for two DMQC approaches are proposed: i) comparing reference and float data, including ways to select the most appropriate reference data for each float profile, that result in the construction of a time series of differences to evaluate potential salinity drift; and ii) the use lab calibrations. Moreover, recommendations regarding real-time quality control tests are provided. This encompasses not only current tests, some which need to be disabled and some that are endorsed as they are, but also the proposal of a Baltic specific test to flag bad salinity data at the surface. Furthermore, the statistics necessary for the implementation of the min/max near real-time quality control were calculated, and collaborations have begun to test if this test is appropriate and should be used in the future. For all these efforts, the compilation of appropriate reference data was done and the dataset will be distributed through the central CTD reference database for Argo DMQC in a future release. Moreover, an exclusive FTP server was established in the FMI FTP server, where BSH, IOPAN and FMI can exchange and store data for Baltic DMQC activities. Finally all code developed and used for the work presented here is stored in a private repository in the Euroargodev (https://github.com/euroargodev/baltic-dmqc-tools). In the future, the partners will work collaboratively to further develop these codes to comment on them and make them user-friendly, in order to share them with the larger Argo community.
All of the suggestions and proposals presented here will be further discussed in the Euro-Argo RISE General Assembly and the upcoming Argo Data Management Team Meeting (ADMT-23) to get feedback and endorsement for their adaptation. Afterwards, the GDAC/DACs need to be contacted to ensure the reprocessing of the floats.

3 Mediterranean Sea and Black Sea

3.1 Generalities

The OWC method has been adapted for a semi-enclosed basin such as the Mediterranean Sea. Due to the different nature of the existing water masses and to the geography of the Mediterranean (Figure 3.1), the DMQC method is applied in the various sub-basins separately in order to avoid selecting historical data for calibration coming from completely different oceanographic regions. Therefore, theMediterranean Sea is divided in 10 areas (Figure 3.1), based on those as defined by the EU/MEDAR-MEDATLAS II project (http://nettuno.ogs.trieste.it/medar /climatologies/medz.html).

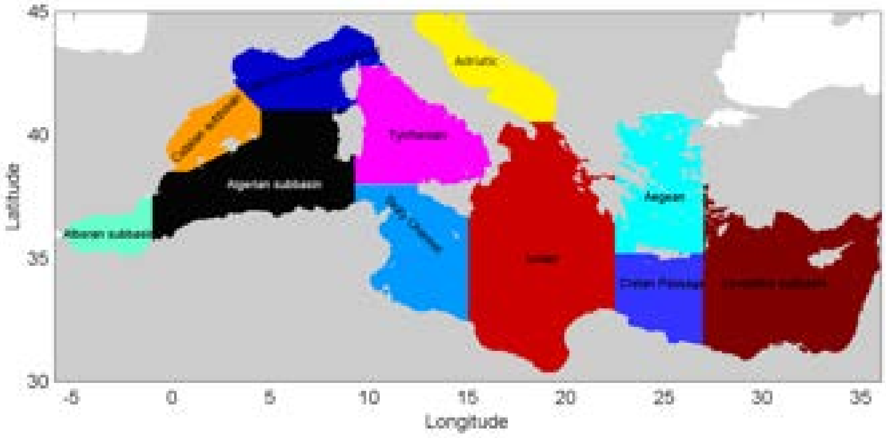


Figure 3.1. Geographical and climatological areas defined in the Mediterranean Sea.

Several water masses flow through the Mediterranean Sea and water mass properties can change dramatically over the years. So, it is crucial to have a co-located (spatial and temporal) historical dataset in order to separate differences between the float data and the historical data due to the sensor drift from those due to the change of the water mass properties.

The Black Sea (Figure 3.2) is almost an enclosed basin since a limited water exchange with the Mediterranean Sea occurs only through the Turkish Straits System. The Black Sea has a positive fresh water balance and hence it is a dilution basin characterised by low salinity values. A vertical stratification is maintained with colder and fresher waters overlying warmer and saltier waters. The halocline coincides with the pycnocline at the depth between 100 and 200 m and this level separates the oxic and the anoxic waters (no life, hydrogen sulphide). The main water masses of the Black Sea are: the surface water, the cold intermediate water, the intermediate water, the deep water. The Black Sea Deep Water (BSDW) is a water mass with stable TS characteristics, with vertical homogeneity of T and S from 1700 m to the bottom, where TS values collapse to a single point (8.90 °C; 22.32). Due to the Black Sea's particular hydrographic characteristics, some modifications of the OWC code and specific settings are used in order to obtain reliable results from the statistical method.

Figure 3.2. The Black Sea (source: https://fleetmonitoring.euro-argo.eu/).

3.2 Regional requirements for OWC

3.2.1 Regional updates to the CTD reference data

OGS, as responsible for the DMQC activities in the Mediterranean and Black Seas, reviews and improves on a regular basis the availability of high-quality ship-based CTD reference data for the quality control of core and deep Argo physical data. It collects CTD data in complement of the official CTD reference dataset, provided by the Coriolis Global Data Assembly Center (GDAC), using mainly two approaches: personal contacts and regional data services. This collection is particularly important for the Mediterranean Sea which is characterised by a complex bathymetry and geography and where water masses can go through dramatic changes over the years. For this reason, it is crucial to have an up-to-date reference dataset with the best co-location (in space and time) between the CTD reference profiles and Argo profiles. A demanding work has been done since the MOCCA project (Notarstefano, 2019) to collect data from different sources and make them compatible for use in DMQC analysis. Data was collected from several research institutes at regional level and the main European Marine Services and then it was converted in mat format to be used in OWC procedure. A quality control was applied such as an additional visual check to avoid spike or duplication. Data was merged and divided in subsets of WMO boxes according to the climatological areas of the Mediterranean Sea. The reference dataset in the Mediterranean and Black Sea for the DMQC activity has been updated in December 2021 (Figure 3.3) and consists of 67483 CTD profiles, more than about 10000 CTD compared to the previous one. A substantial improvement has been achieved and gaps in temporal and spatial coverage have been filled in some areas such as the Levantine and the Aegean. Data before 1995 were discarded because they were considered too old for quality control purposes. In 2020 CTD profiles are scarce due to the pandemic. The vertical resolution is about 1 dbar which is adequate for DMQC because the comparison between CTD reference data and floats data is done on deep and stable water columns. In the future, the dataset will be more thoroughly checked for duplicates using the code developed at BSH and available in the Euroargodev Github (**CTD-RDB_for2021v01 release**).

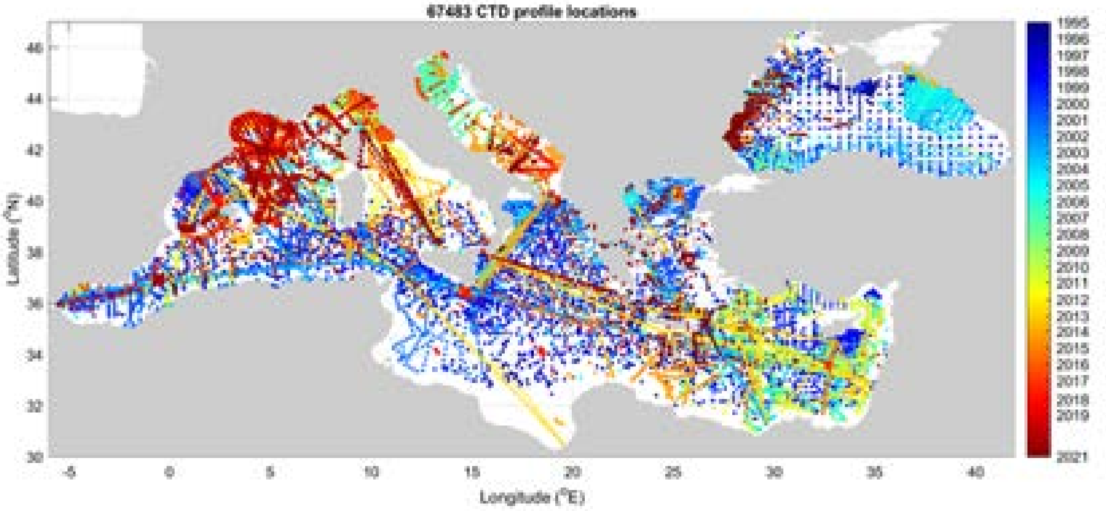
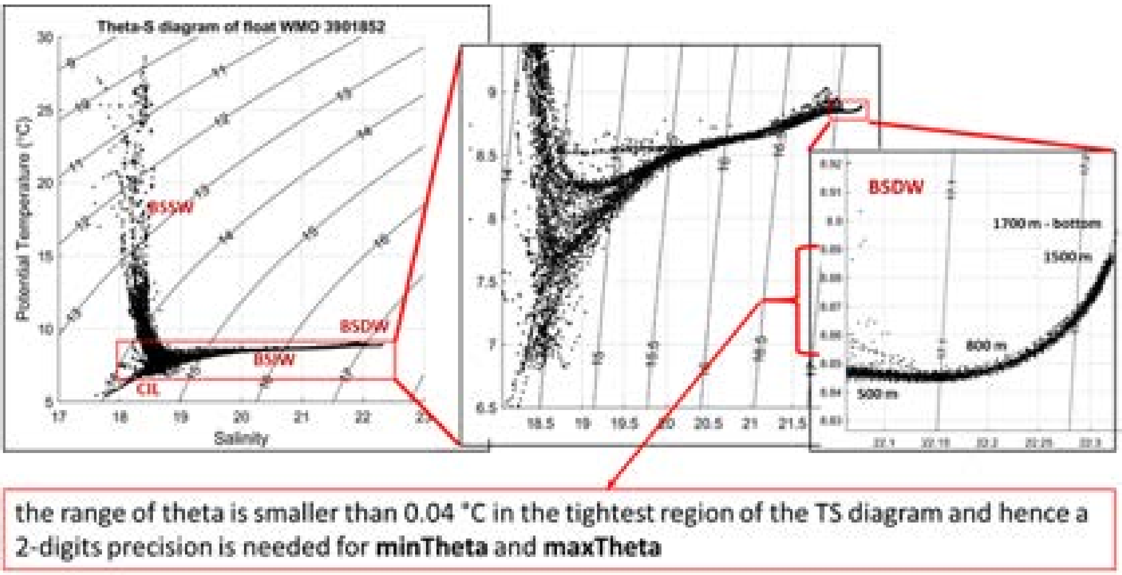


Figure 3.3. Spatial distribution, colour-coded for time, of the CTD profiles in the final version of the CTD reference dataset of the Mediterranean and Black Seas.

3.2.2 Modification of the standard OWC procedure

3.2.2.1 Script modifications

A manipulation of two OWC scripts was necessary in order that OWC produces not only some results but also results that are statistically reliable. The salinity thresholds used for quality control purposes in the *"update_salinity_mapping"* code had to be modified, taking into account the typical salinity of the Black Sea. The new threshold values are < 15 and > 25. Another code that needed a modification was *"find_10thetas"*. The problem was that the range of theta values where we look for the 10 levels is on the order of centidegrees (mostly in the Black Sea) and not decidegrees as in the open ocean. Therefore, the *if* statement "mintheta<maxtheta", as it was built, was never true because the theta values were truncated to the first decimal: the single-digit precision for minTheta and maxTheta caused minTheta to be always greater than maxTheta. Hence, a 2-digits precision was needed. (Figure 3.4). These modifications are now included in the last official software version of OWC (https://github.com/ArgoDMQC/matlab_owc).



the range of theta is smaller than 0.04 °C in the tightest region of the TS diagram and hence a 2-digits precision is needed for **minTheta** and **maxTheta**

Figure 3.4. Example of the θ-S diagram of a float in the Black Sea to show that the range of theta where we look for the 10 levels is on the order of centidegrees.

3.2.2.2 Selection of reference data

The reference data used in the standard OWC set-up are 10x10 ° degree squares and are too large for application in marginal seas. A MatLab function has been built to re-shaped WMO boxes according to the dimension of sub-basins defined (Fig. 3.3). Hence, the WMO boxes reshaping is useful to avoid selecting historical data for calibration coming from completely different oceanographic regions. The “WMO-boxes-reshaping” function is available on the Euroargodev GitHub at https://github.com/euroargodev/WMO-boxes-reshaping.

Sub-basins are defined by polygons created by the user (i.e. the Adriatic Sea polygon is highlighted in red in Figure 3.4). In the example in figure 3.4, it is shown that if we want to work with an Argo float in the Adriatic Sea we have to use the WMO box 1401. OWC procedure considers the entire box and hence CTD reference data coming from the northern Ionian, Thyrrhenian and small part of the eastern Ligurian Seas are used in the DMQC analysis. The result is that the OWC estimate will be misleading since completely different water masses are taken into account. However, if a re-shape of the WMO box is done according to the Adriatic Sea dimension (using the “WMO-boxes-reshaping” function) prior to the run of OWC, the statistical method will use only the selected subset of data. This will allow a more reliable comparison between the Argo and the reference data.

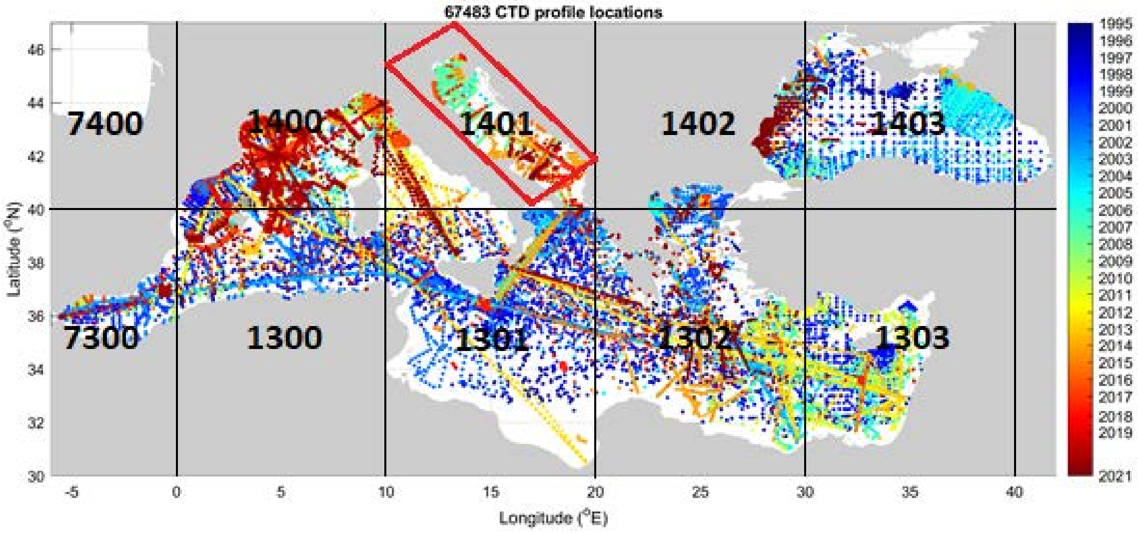


Figure 3.3. WMO boxes in the Mediterranean and Black Sea area (WMO numbers are in black and bold). The Adriatic Sea polygon is shown in red.

3.3 Quality control decision: additional procedures

The OWC method, which is based on tight θ-S relation for calibration, sometimes turned out to be problematic for the Mediterranean Sea. The main reason is that the Mediterranean Sea is characterised by high variability and various water masses as previously described. Thus, if the CTD reference dataset is scarce, the OWC method can give an estimate, but these estimates will have quite large uncertainties, so large that they could be useless for quality control. Moreover, the portion of the water column in the Mediterranean Sea that has a uniform θ-S relationship is mostly at pressure larger than 700 dbar. Unfortunately, in some cases, profiles are not deep enough to be useful. The result is that the ten “best” θ levels used for the calibration are spread only through the water column of the near-surface and intermediate layer. Hence, we often use additional qualitative analysis in support of the OWC method. The following examples explain better the purpose of these procedures.

The OWC procedure was applied to the float WMO 6901816 deployed in the Mediterranean Sea. During its life the float crosses at least two main water masses. Despite testing several OWC configurations (e.g. multiple breaking points, splitting time series), a satisfactory fit was not found, which suggests that the issue is the reference database. The OWC suggests a correction larger than the Argo requested accuracy (0.01)(Figure 3.4). However, the salinity is quite constant on the selected θ levels suggesting a good behaviour of the conductivity sensor. The jump observed between the two time series (Figure 3.4 right panel) is related to the sampling of two distinct water masses.

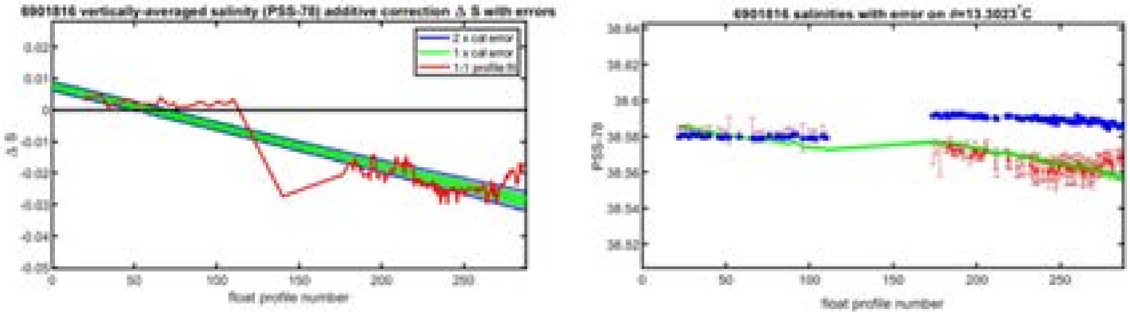


Figure 3.4 Evolution of the suggested adjustment with time (left panel) and evolution of the salinity with time in a selected theta level (θ = 13.3023°C) with minimum salinity variance (right panel).

Hence, we applied the following additional analyses in support of the OWC outputs:

1. *Comparison with the closest (in time and space) historical profiles* (Figures 3.5, 3.6). This step shows a good agreement and confirms that the float worked well.

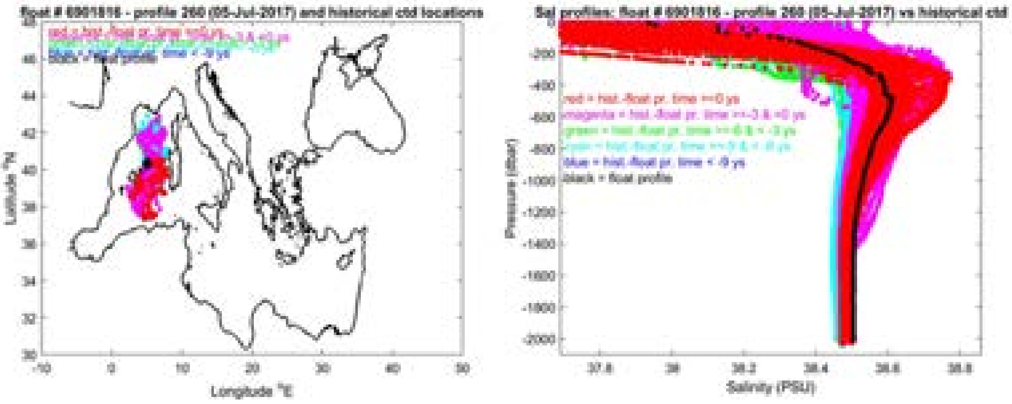


Figure 3.5 Locations of the salinity float profile number 260 and historical CTD data (right panel) and the respective salinity profiles (left panel).

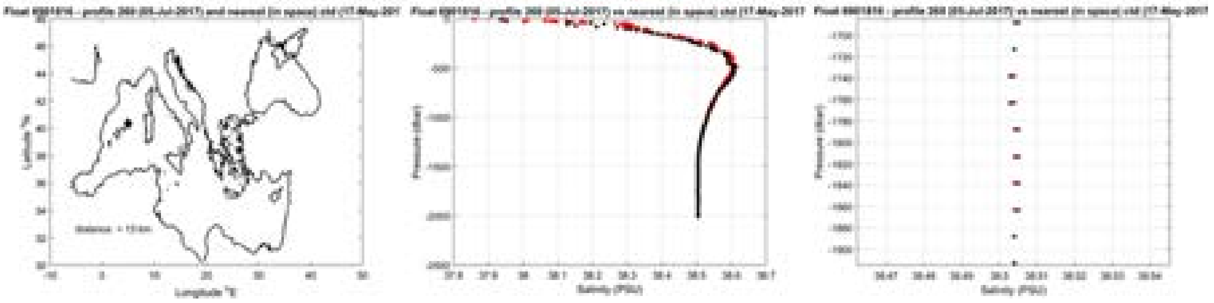


Figure 3.6 The salinity float profile number 260 (black dots) are compared to the nearest in space reference profile (red dots). On the right the zoom in comparison in the deeper layers. The locations of the two profiles and their distance is given in the left panel.

2. *Analysis of the* θ-S *relationship in the deeper (> 700 dbar) float profiles*. The water column exhibits uniform θ-S relations and gives no evidence of any potential drift. The θ-S relations of these deeper profiles are very tight and the limited spread of salinity on theta level gives an indication that any potential conductivity sensor drift (that is quite easily detectable by a systematic full shift in the θ-S measurements) is well below that the accuracy requested by Argo (Figure 3.7).

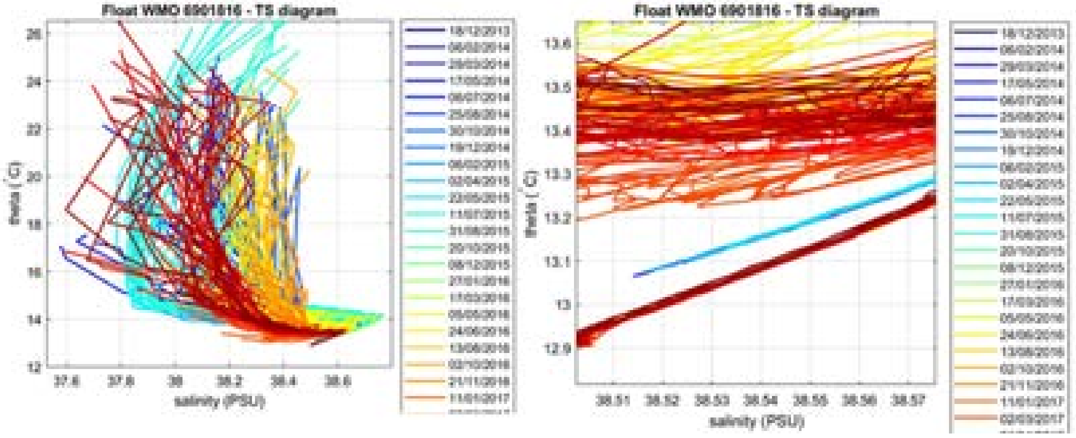


Figure 3.7. θ-S diagram colour-coded per cycle number (left panel) and magnified view of the area where the θ-S relationship is more uniform (on the right). Two distinct water masses are identified: cluster in blue and cluster in dark red below it.

3. *Comparison of the* θ-S *curves of the float and the reference data as mapped by OWC.* This comparison gives information about the initial potential sensor calibration offset highlighted by OWC: the salinity difference between the two curves (red and black just below in figure 3.8) is larger than the variability of the reference dataset and hence, this difference could represent the calibration offset. However, this reference data is not adequate to describe the hydrographic characteristics of the Ligurian and Algerian sub-basins sampled by the float (the black θ-S profiles below 13.05 θ in figure 3.8) and therefore this comparison can be misleading. Indeed the difference in time between Argo and reference profiles is more than six years and therefore the reference data is not capturing the natural variability of the area. Therefore the presence of this difference between datasets cannot confirm the presence of an offset, because it could be caused by natural variability of the water masses.

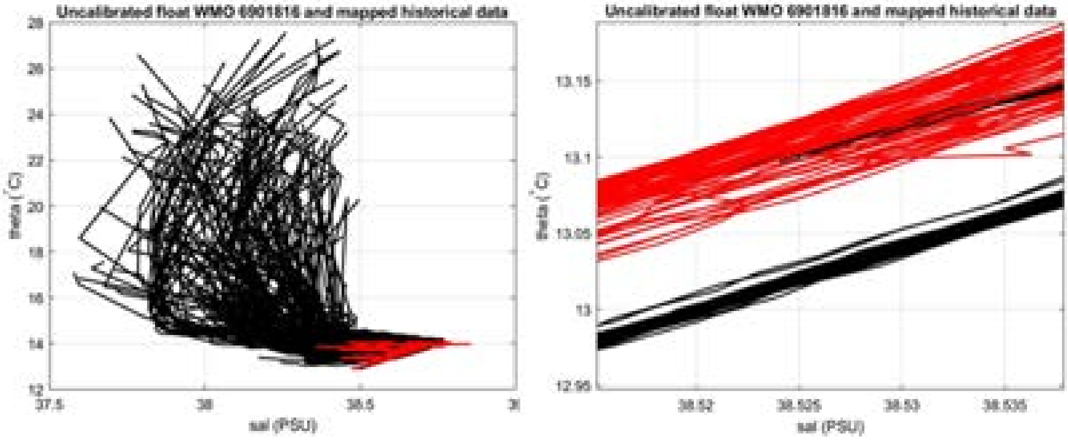


Figure 3.8 θ-S diagram showing uncalibrated float salinity profiles (black lines) and mapped historical data (red lines): whole θ-S ranges (left panel) and magnified view of the area where the θ-S relationship is more uniform (right panel).

4. *Comparison of the* θ-S *curve of the analysed float with the curve of other floats with a similar trajectory.* The floats used for comparison in Figure 3.9, have been already checked in delayed-mode and have been deployed quite close in time and position. The salinity difference between the two θ-S curves gives indication about the conductivity sensor behaviour. As shown in figure 3.9, this comparison is quite good and indicates that the float worked well.

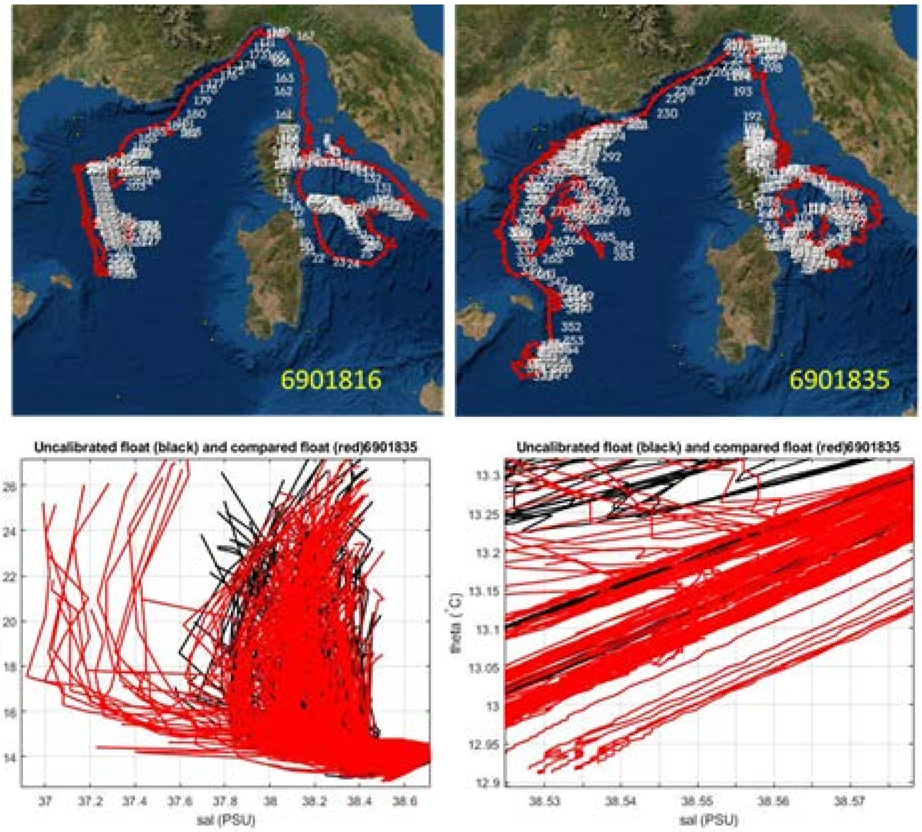


Figure 3.9 Trajectory of the floats used for the comparison (upper panel). Comparison between float 6901816 (black line) and 6901835 (red line) in the θ-S space (lower panel): in the whole θ-S range (left panel) and in the most uniform part of the θ-S curve (right panel)

These additional analyses suggest that salinity data of Float WMO 6901816 are accurate and don't need a delayed mode correction.

3.4 Whitespace Maximisation: machine Delayed Mode intercalibration of datasets

At the Balearic Islands Coastal Observing and Forecasting System (SOCIB) a semi-automatic routine has been developed for the field calibration/correction for salinity data using whitespace maximisation image analysis of θ/S data (adiabatic potential temperature - theta / Salinity - θ/S). This technique enables machines to carry out the same analyses that learned oceanographers have performed for many decades back into the 20th century: that is by overlapping θ/S diagrams against the light from a window pane, or a light box if one was available, a correction factor could be derived for the gathered data versus the reference data.

The application was originally created for gliders in the framework of the JERICO-Next EU project and is available in the Ocean Best Practices System Repository (http://dx.doi.org/10.25607/OBP-430). A peer-reviewed paper describing the method has been published in Frontiers in Marine Science Best Practices in the Ocean Observing Research Topic (Allen et al., 2020). As part of Euro-Argo RISE, SOCIB has started the adaptation of this application for its use on Argo float data and the first tests have been conducted. Although the application itself was originally designed for gliders and not to provide corrections that evolve on a long timescale (many months to years), as OWC does, it is potentially useful as an auxiliary tool for DMQC of Argo floats operating in the Mediterranean and Black Seas. Due to the user-friendly front end (operational at SOCIB), it could be used in the future to do a semi-automatic preliminary float analysis and such results would be useful in assigning priorities for the DMQC activity.

3.4.1 Whitespace maximisation principle

As recommended in SeaBird, Application Note 31, for their CTD instruments the Argo or glider salinity values should be corrected as a conductivity slope (www.seabird.com/application-notes, AN-31), equation (1), where A is the conductivity correction coefficient. It is worth noting that in ocean regions where salinity and temperature ranges are small an offset correction can provide equivalent results.

$$\text{Conductivity\_corrected} = A * \text{Conductivity\_measured}, \qquad (1)$$

Generally the slope correction is iteratively determined as required to get a good overlap between the glider/Argo and reference CTD datasets on the θ/S diagram. A semi-automated procedure of maximising whitespace in overplotted corrected CTD and glider/Argo θ/S data has been developed. This involves the creation of a θ/S diagram of "background" ship data, and "test" glider/Argo data as is shown for float WMO number 6901245 in Figure 3.10. The whitespace maximisation is then carried out, whereby the test data is iteratively moved along the x (salinity) axis by changing the value of coefficient A in equation (1), and a calculation of the whitespace area, i.e the number of empty or white pixels, in the θ/S diagram figure is carried out. The iterative procedure continues adjusting the conductivity correction coefficient (i.e. the coefficient that ultimately leads the shift in the test data to the left or right) until the whitespace area of the figure has reached a maximum. This occurs when the test data is most closely overlying the background data. Confidence in the glider/Argo salinity values is then typically better than 0.010 psu assuming the normal confidence in reference CTD calibrated datasets of better than 0.005 psu, as will be demonstrated in the examples below.

By default the automatic whitespace iteration procedure is carried out three times with different initial guesses for the correction coefficient, A, in equation (1). One of these guesses will normally be 1 (i.e. no initial movement of the glider/Argo data) or a number close to 1 derived from initial inspection. For the other two initial guesses, the user is advised to use guesses that move the glider/Argo data to the left and to the right of the background data to begin with. SOCIB's software uses by default, A = 0.9999, 1.0000 and 1.0001 as the three initial guesses (Figure 3.10), this is equivalent to offsets of ±0.004 in salinity at Mediterranean water salinities of 38.500; in more characteristic Atlantic waters these initial guesses would be equivalent to offsets of ±0.003. The iteration procedure has three-step levels, major steps of 0.0001, minor steps of 0.00001 and miniscule steps of 0.000001, with equivalent offsets of 0.004, 0.0004 and 0.00004 respectively at 38.500 salinity. Each of the three starting scenarios (Figure 3.11) begins by moving in major steps to the left (lower salinity) or right (higher salinity), testing the whitespace pixel count each time, the steps then continue in the most whitespace increasing direction until the whitespace begins to decrease. At this point the programme changes to the minor step level and reverses direction until the whitespace stops increasing and this same procedure is then repeated with miniscule steps.

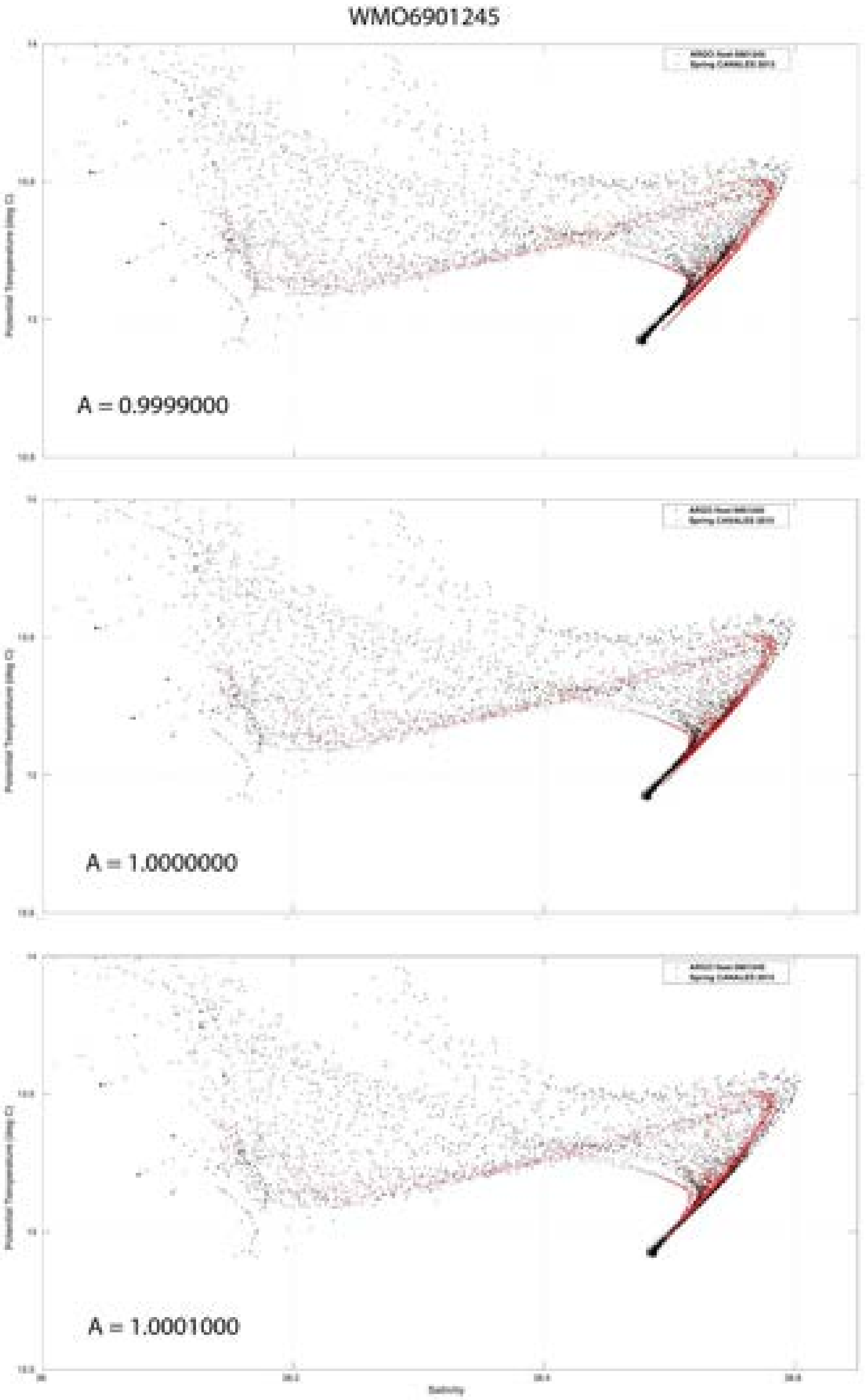


Figure 3.10: Three iteration examples from the selection of gradient A, for the intercalibration of Argo float WMO6901245 against the Canales April 2015 salinity bottle calibrated CTD data.

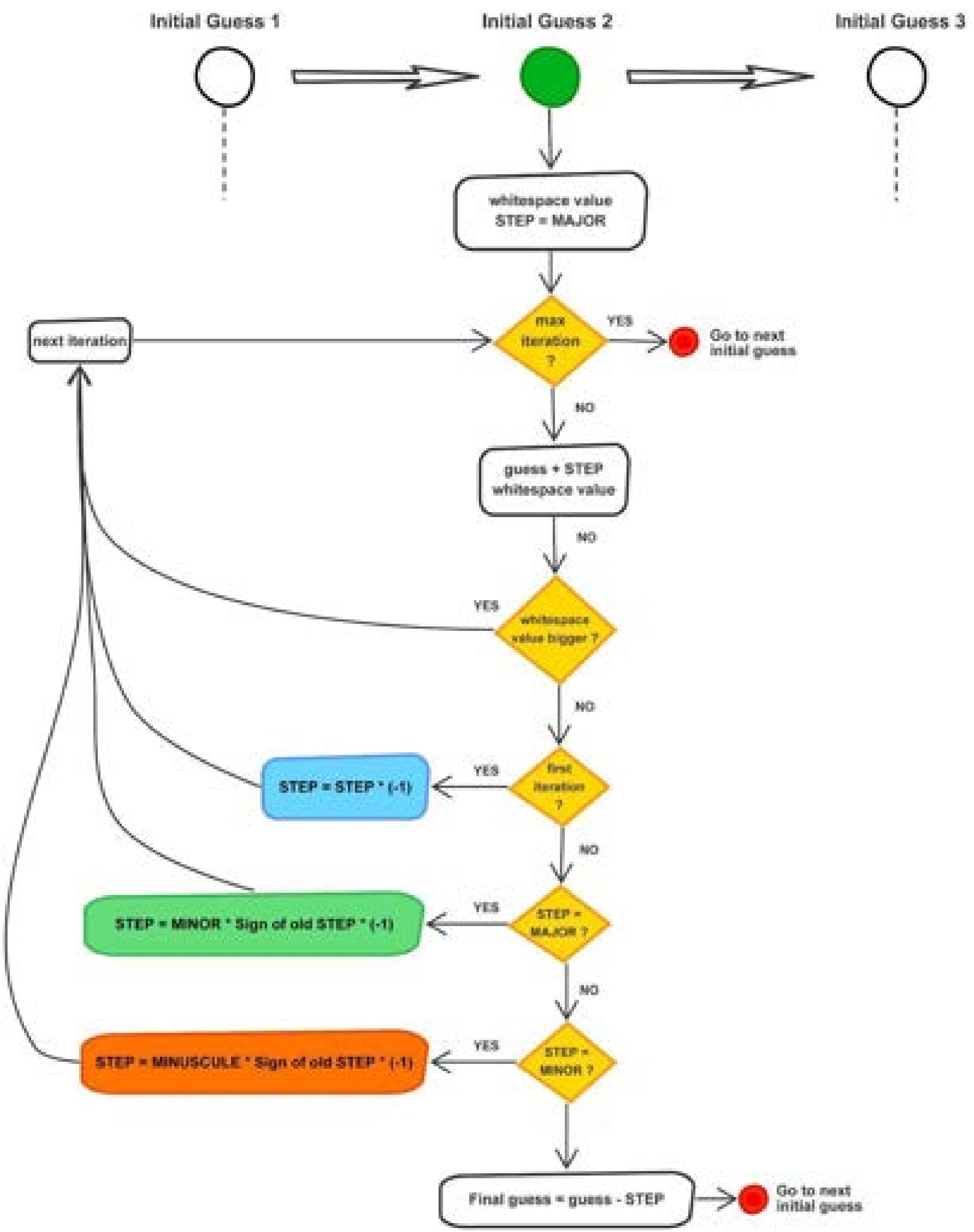


Figure 3.11. Flow diagram for the whitespace maximisation component of the SOCIB glider salinity delayed mode scientific correction toolbox. For the sake of brevity, only the second initial guess procedure is shown as all initial guess procedures are the same.

The inter-calibration correction procedure is described as semi-automatic because preparation for the whitespace maximisation may require several subjective expert decisions which can influence the outcome of the automated correction. Appropriate background reference data need to be used, where ideally both spatial and temporal separation of the datasets should be as small as possible. If the most contemporary delayed mode corrected reference CTD datasets and the glider/Argo CTD dataset are not quasi-synchronous, a difference in time of more than a few weeks for example, the glider/Argo θ/S diagrams are then analysed from a longer term point of view. Analysis of the seasonal evolution during that year, using previously corrected glider/Argo and ship missions before and after, and the inter-annual evolution, corrected ship and glider/Argo missions from a similar seasonal time period in previous years, can provide critical data for DM calibration. Any inter-annual evolution must of course bear in mind the possibilities of published long term water mass changes and decadal scale variability.

3.4.2 Use of the toolbox with Argo floats and example studies

Recently SOCIB has begun to adapt our Delayed Mode correction technique to accept and work with Argo float data. The difficult bit for Argo data sets is the quasi-'random' path of the profiler and the resultant difficulty in searching databases and reference datasets, for suitable DMQC CTD data profiles with which sensible comparisons and inter-calibrations can be made. These reference dataset selections need to consider and be aware of known spatial and temporal trends and changes (oceanographic knowledge) that may exist in water mass characteristics prior to any attempt to fit salinity corrections in line with OWC procedure. Inter-calibration needs to be considered to be an evolutionary process; during an Argo mission the float may be carried through regions for which suitable CTD data for inter-comparison exist in databases, and also through regions for which they do not. Therefore, inter-calibration should be considered a 'rolling' process throughout the lifetime of the float. If instrumental changes whilst in continuous operation can be assumed as being slow and linear, the inter-calibration results could be interpolated across regions where no inter-comparison CTD data exist in the databases.

Recognising that ARGO floats do not report conductivity, in the beta version of the Whitespace maximisation toolbox that handles Argo data, the conductivity needs to be re-calculated from the salinity reported by the float using the CSIRO standard SeaWater algorithms (Fofonoff and Millard, 1983). As of the writing of this report, the data available to test the beta version of the Whitespace maximisation toolbox is restricted. In its current version, the software only digests files disseminated through the SOCIB API since the automatic routines, which identify the different paths, names and cross-reference ship CTD vs float, cannot be run when using external data sources. This limits the selection of Argo data to those floats deployed by Argo Spain and the reference data to SOCIB cruises, including both a monitoring program conducted by SOCIB since September 2013 and targeted science cruise programmes. Therefore, the first tests were run on floats WMO6901242 and WMO6901245, which were the only two Spanish floats deployed after September 2013.

3.4.2.1 Float WMO 6901242

Apex float WMO 6901242 had a very short lifespan of just 10 cycles, approximately 44 days over April, May, and June 2015. However, as a result, it spent the entirety of its short mission in the Ibiza channel region between the island of Ibiza and the Spanish mainland (Figure 3.12). Therefore, SOCIB has two salinity bottle corrected shipboard lowered CTD campaigns with which independant calibrations were carried out; the spring Canales monitoring cruise on board R/V SOCIB in April 2015, and the SHEBEX (Balearic Shelf Exchanges) cruise on R/V SOCIB in May 2015.

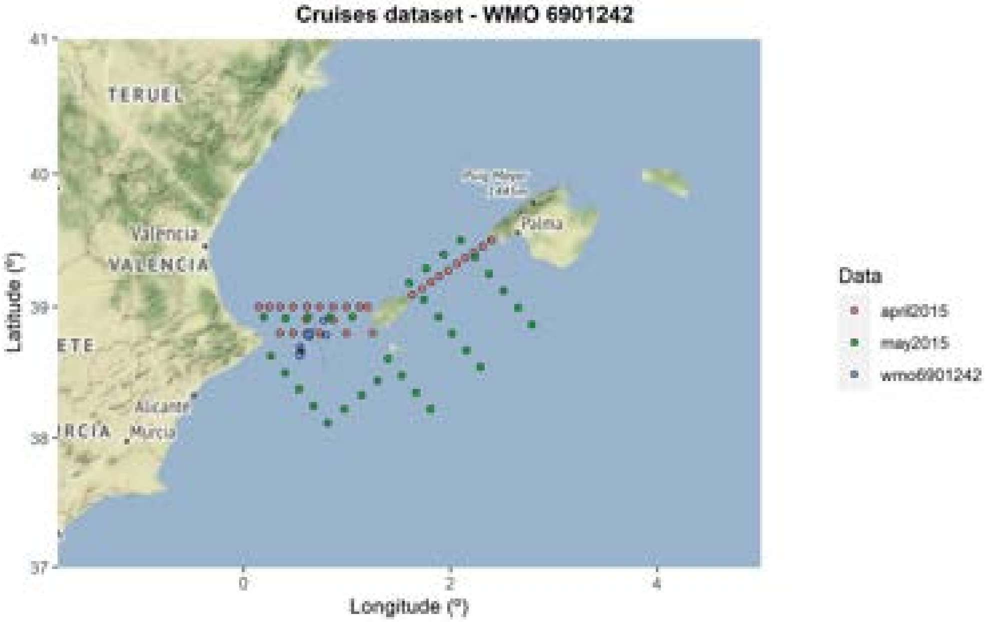


Figure 3.12. Float cycles analysed and spatial distribution of the CTD profiles from the dataset used in the intercalibration.

Intercalibration against the Canales April 2015 CTD dataset provided a conductivity correction factor of A = 1.0001000 (Figure 3.13), where Cond_corrected = A* Cond_measured as above.

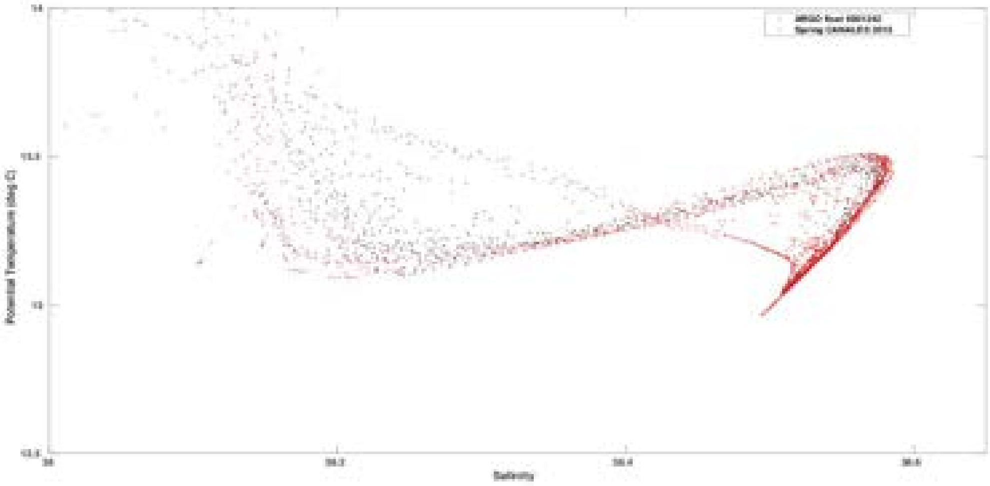


Figure 3.13. Intercalibration against the Canales April 2015 CTD dataset using a conductivity correction factor of A = 1.0001000 for the float (WMO 6901242) data.

Intercalibration against the SHEBEX May 2015 CTD dataset provided a conductivity correction factor of A = 1.0000700 (Figure 3.14), where Cond_corrected = A* Cond_measured as above.

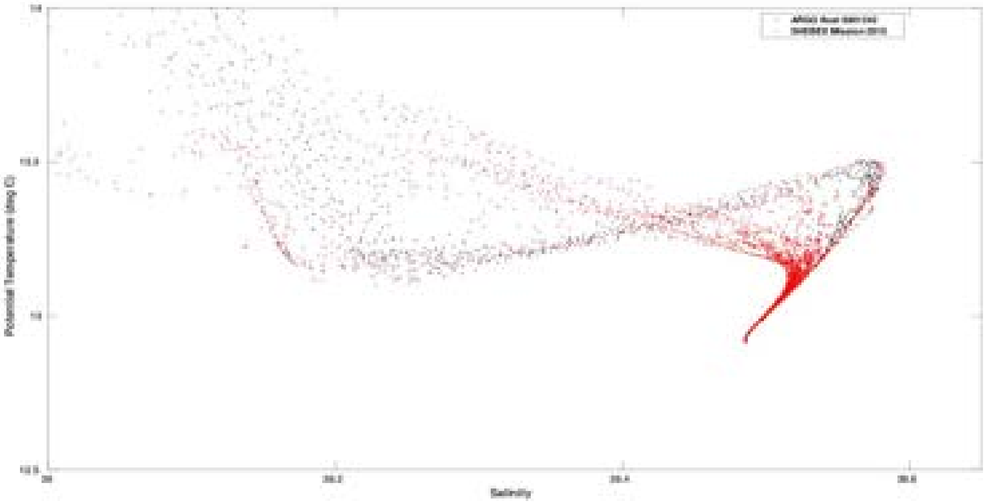


Figure 3.14: Intercalibration against the SHEBEX May 2015 CTD dataset using a conductivity correction factor of A = 1.0000700 for the float (WMO 6901242) data.

At a conductivity of 45.000 mS/cm, for example, we have an equivalent offset difference between these calibrations of 0.00135. So, taking the mean, we could consider the resulting relative salinity calibration to be better than ±0.001 (~ 0.0007), relative to the cruise salinity bottle calibration confidence levels of ±0.0046 (Canales) and ±0.0059 (SHEBEX). Therefore, we can consider a confidence level significantly better than ±0.01 absolute. The OWC results are in agreement with the Whitespace Maximisation method, as shown in figure 3.15. Indeed, the proposed salinity correction is on the order of a few thousandths.

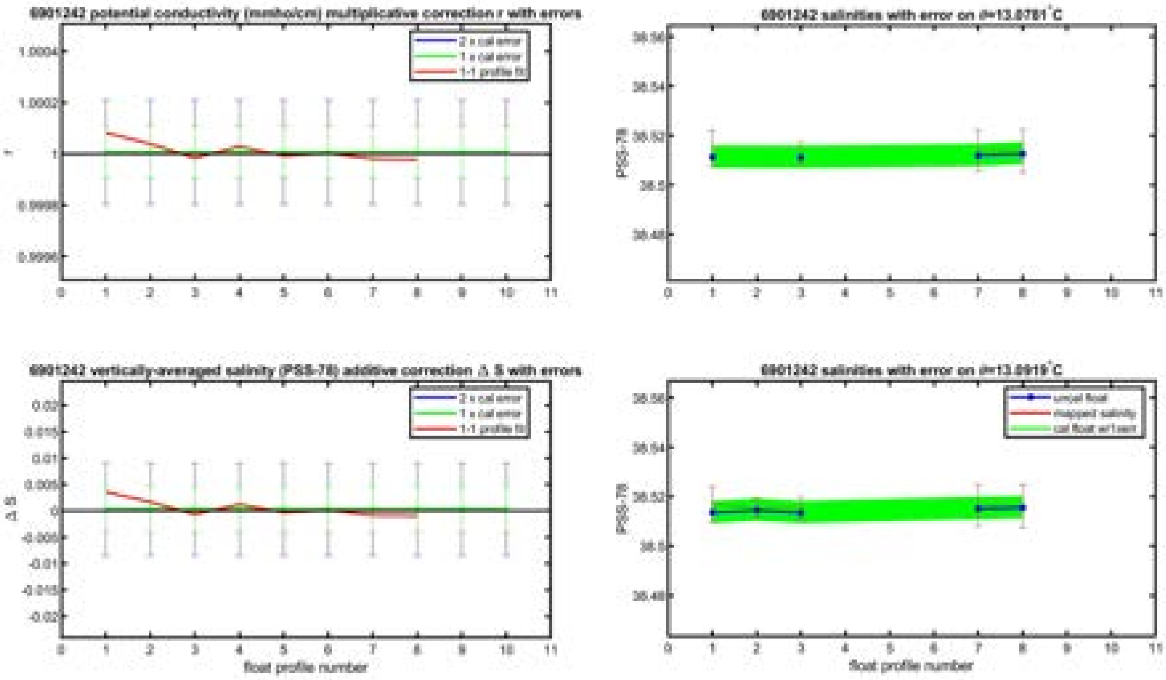


Figure 3.15: OWC results for the WMO 6901242 float. Evolution of the suggested adjustment with time (left panel) and evolution of salinity with time along selected theta levels with minimum salinity variance (right panel).

3.4.2.2 Float WMO 6901245

Arvor float WMO 6901245, also launched in the Ibiza Channel, had a long lifespan of approximately 1422 days beginning in November 2014, for which all but the first 400-500 days were spent in circulation round the Alboran Sea gyres of the Western Mediterranean (Figure 3.16). SOCIB has 4 salinity bottle corrected shipboard lowered CTD campaigns with which independant calibrations were carried out; the autumn Canales monitoring cruise on board R/V SOCIB in November 2014, the spring Canales monitoring cruise on board R/V SOCIB in April 2015, the SHEBEX (Balearic Shelf Exchanges) cruise on R/V SOCIB in May 2015, and the summer Canales monitoring cruise on board R/V SOCIB in July 2015. The first 67 cycles of the float between November 2014 and October 2015 were used for the intercalibration as these were reasonably related to the cruise datasets in terms of expected water masses, time and geography.

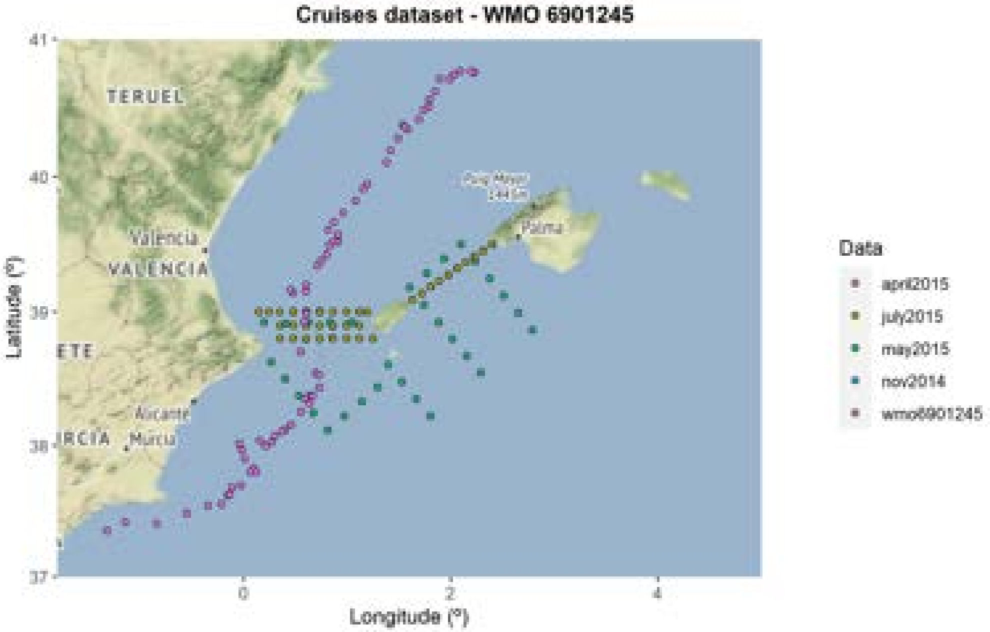


Figure 3.16. Float cycles analysed and spatial distribution of the CTD profiles (November, April and July dataset are overlapping) from the dataset used in the intercalibration.

Intercalibration against the Canales November 2014 CTD dataset provided a conductivity correction factor of A = 1.0001090 (Figure 3.17), where Cond_corrected = A* Cond_measured as above.

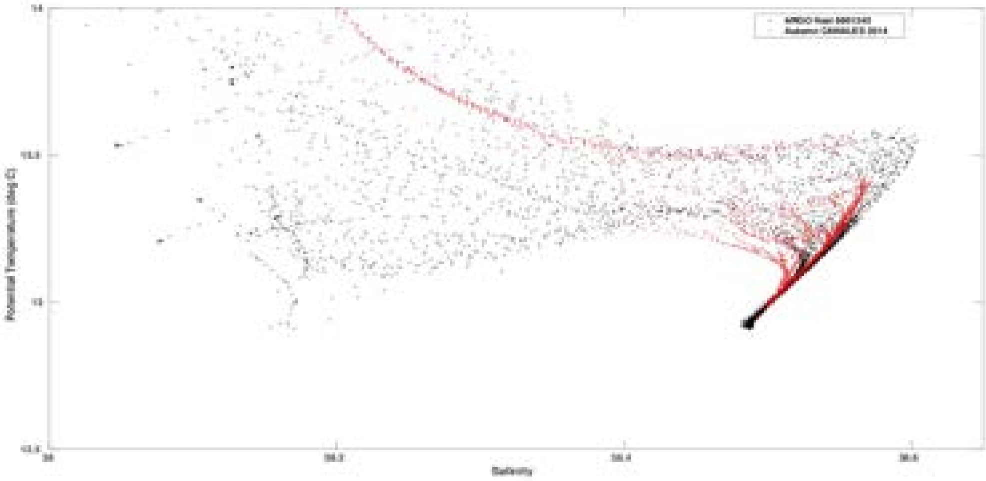


Figure 3.17. Intercalibration against the Canales November 2014 CTD dataset using a conductivity correction factor of A = 1.0001090 for the float (WMO 6901245) data.

Intercalibration against the Canales April 2015 CTD dataset provided a conductivity correction factor of A = 1.0000700 (Figure 3.18) , where Cond_corrected = A* Cond_measured as above.

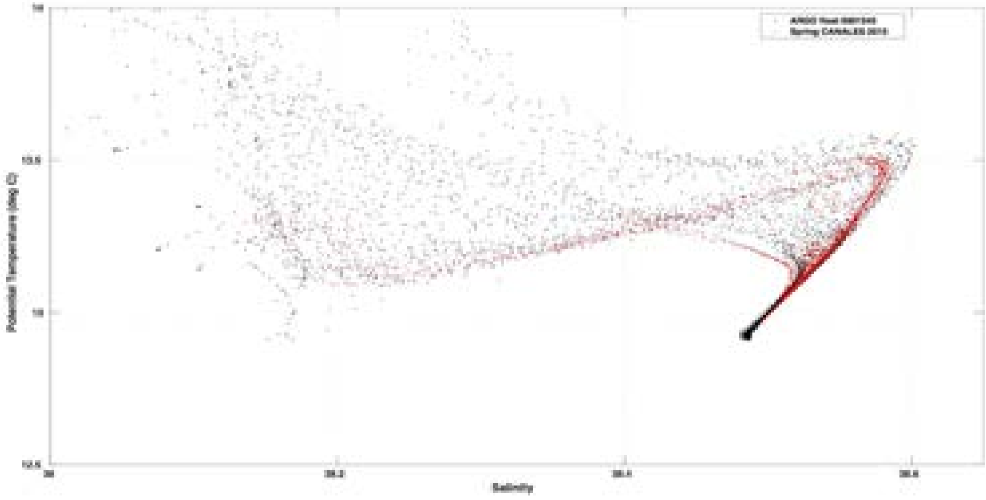


Figure 3.18. Intercalibration against the Canales April 2015 CTD dataset using a conductivity correction factor of A = 1.0000700 for the float (WMO 6901245) data.

Intercalibration against the SHEBEX May 2015 CTD dataset provided a conductivity correction factor of A = 1.0000200 (Figure 3.19), where Cond_corrected = A* Cond_measured as above.

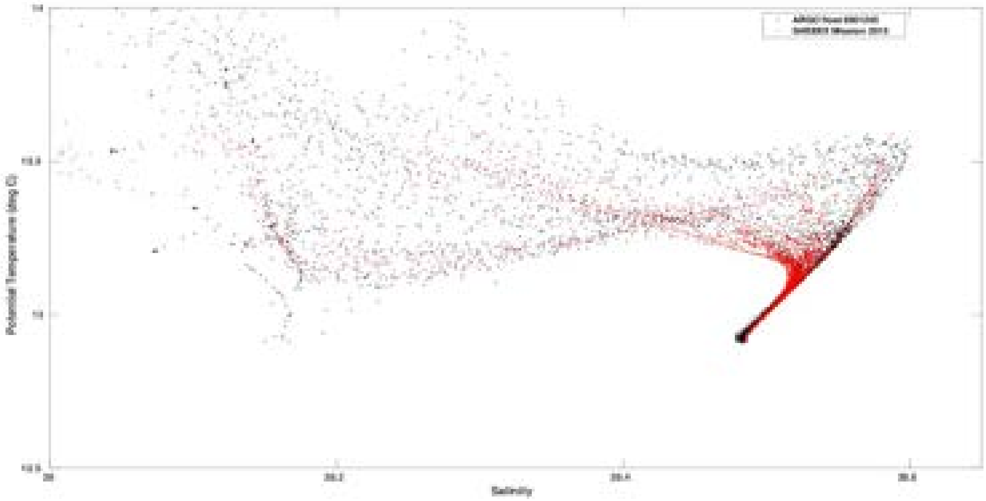


Figure 3.19. Intercalibration against the SHEBEX May 2015 CTD dataset using a conductivity correction factor of A = 1.0000200 for the float (WMO 6901245) data.

Intercalibration against the Canales July 2015 CTD dataset provided a conductivity correction factor of A = 1.0000910 (Figure 3.20), where Cond_corrected = A* Cond_measured as above.

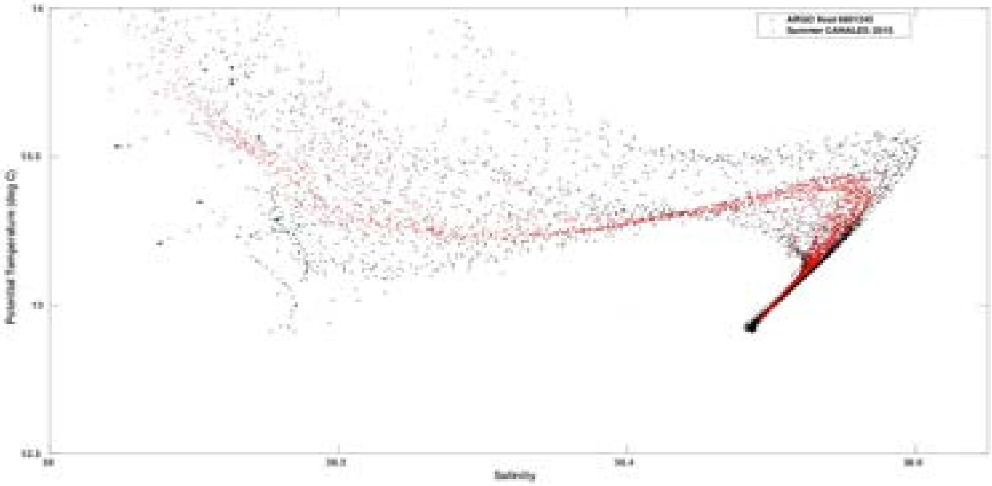


Figure 3.20: Intercalibration against the Canales July 2015 CTD dataset using a conductivity correction factor of A = 1.0000910 for the float (WMO 6901245) data.

At a conductivity of 45.000 mS/cm, for example, these correction factors give an equivalent offset range of 0.00401. Therefore, selecting perhaps a mean of the four correction factors, we could consider the resulting relative salinity calibration good to approximately ±0.002, relative to the typical cruise salinity bottle calibration confidence levels of approximately ±0.005. So again we can have confidence to a level better than ±0.01 absolute. Figure 3.21 shows the OWC results. The correction proposed by OWC is quite small and below the Argo requested accuracy (0.01) from profile 1 to 199. Then a suspicious behaviour of the conductivity sensor is observed but it reflects the natural variability of the area, hence no correction is recommended.

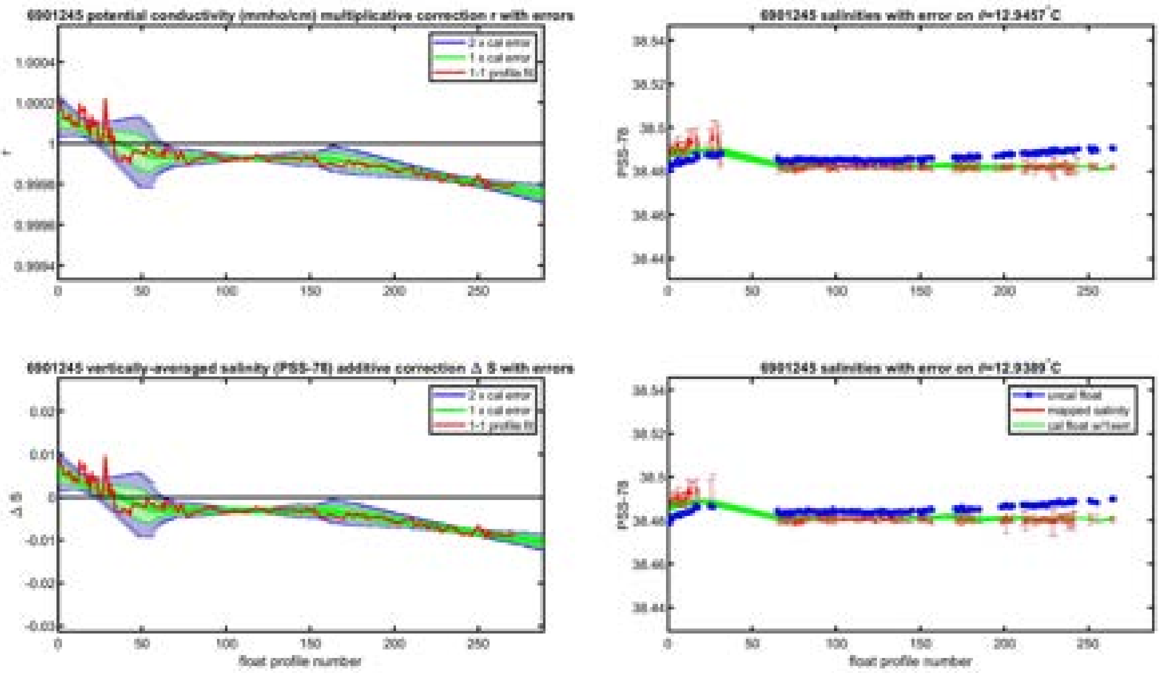


Figure 3.21. OWC results for the WMO 6901245 float. Evolution of the suggested adjustment with time, on the left. Evolution of salinity with time along with selected theta levels with minimum salinity variance, on the right

Interestingly, if we take all 1422 days of the float data, i.e. including all the profiles further west and indeed dominated by the Alboran Sea, the deep mixing line between Levantine Intermediate Waters and Mediterranean Deep Waters remains very similar in slope albeit with a little more scatter in absolute values. Intercalibrating this total dataset with the Canales November 2014 CTD dataset produces a conductivity correction factor of A = 1.0000100. This would suggest a drift high in conductivity (i.e. salty drift), of perhaps as much as 0.004 over the nearly 4 year life of the float; this is perhaps counterintuitive if one expects a slow biofilm accretion in the conductivity cell, which would cause a drift low. However, since Argo floats spend little time at the surface, salty drifts are attributed to ageing of the electronics and other technical issues with the data. In any case, we need to be careful and remember to consider climatic trends in salinity and temperature (Schroeder et al., 2016; Allen et al., 2020) and geographical distances between the datasets when interpreting the results.

3.5 Summary and future steps

After some modifications in the code thresholds and intermediate steps for reference data selection taking into account the different hydrographic sub-basins, the OWC method can be used in the Mediterranean. The reference dataset has been improved by OGS in the last years by using both freely available data and acquiring data directly from scientists, and also by allowing the selection of all profiles, disregarding the maximum recorded pressure minimum of 900db. Currently, the CTD reference database has more than 65000 profiles but it is very likely that it includes many duplicated profiles and therefore it will be checked thoroughly for duplicates in the future to further improve its quality.

Most of the time, the OWC method provides reliable results, however, there are cases that require additional quality checks to verify the OWC output. These cases could be when the relation between salinity and theta is not the tightest or when the reference data are not adequate. At OGS the following qualitative analysis are in place: comparison with the closest historical profiles, analysis of the θ-S relationship in the deeper (> 700 dbar) float profiles and comparison of the θ-S curves of the float and the reference data as mapped by OWC, as well as with other floats with similar trajectory.

The Whitespace Maximisation routine could also be useful in such situations and used to complement the OWC toolbox, also providing a quantitative analysis. The initial tests using datasets disseminated through the SOCIB API, showed results that are in line with OWC outputs but further tests using other data sources are needed, specifically the Argo GDAC/DAC and the CTD reference database for Argo DMQC. Until the end of the project the beta version of the code used for the tests presented here will be shared on the Euroargodev github in the repository ARGO_withespace_method. SOCIB will contribute in future collaborations with all interested DMQC operators and community in general, to define a roadmap of the steps needed to reuse or take advantage of the code to ingest data from the GDAC and from the CTD reference database for Argo DMQC. Future versions of the DM calibration routine could be used for preliminary float analysis, in order to faster detect floats with suspicious conductivity sensor behaviour. This task would be very useful in assigning priorities for the DMQC activity and would be further eased if the user-friendly front end application at SOCIB would be also adapted for Argo inputs. Long term drift, if a float has been measured over a long period of time, can currently only be dealt with by the SOCIB DM calibration analysis, through selecting the float dataset by time ranges and calibrating each time range with the closest CTD datasets in time and space. The future development of semi-automated reference data selection routines would support both the SOCIB Argo DM salinity correction routine to achieve better results. This activity is a priority development theme at SOCIB for DM glider data.

4 Arctic

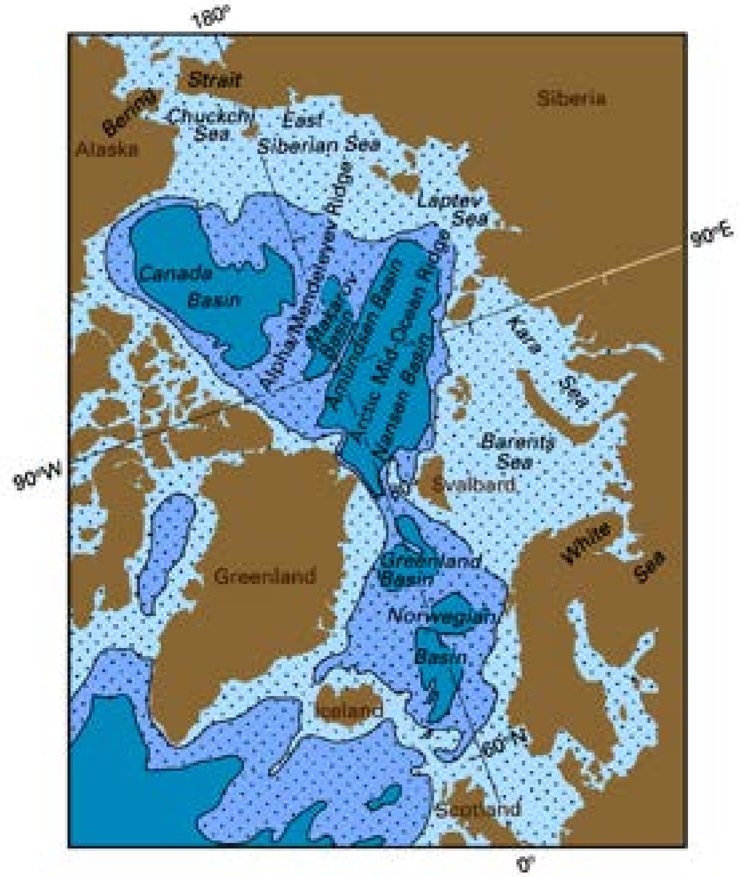


Figure 4.1. Bottom topography of the Arctic Sea, with its main deep basins and shallow seas. The 1000, 3000, and 5000 m isobaths are shown, and regions less than 3000 m deep are shaded. From Tomczak & Godfrey (1994)

4.1 Generalities

The Arctic consists of several deep basins (Nansen Basin, Amundsen Basin, Makarov Basin, and Canada Basin) separated by ridges, and large shallow shelves, especially along the Russian coast (Figure 4.1). Floats deployed in the Nordic Seas have reached the Nansin Basin in the past, but only recently floats have been deployed directly there, as a result of renewed efforts by the Euro-Argo partners. The large shelf seas of the Barents Sea, that borders Europe, has already been sampled by ice-capable Argo floats as part of pilot projects. One float has even reached the Kara Sea. The use of Argo floats is limited geographically by the presence of permanent sea ice in the central Arctic, since there they can't surface and therefore are not able to communicate with the satellites. However, the long-term steady decline of sea ice coverage in the region has increased the seasonal sea ice region, thus making the use of Argo floats possible in a larger area.

In principle, the OWC software for salinity DMQC should work without major adaptations in all the deep basins, provided that enough reference data can be found. Therefore, our first action regarding

DMQC in the region is to test if the OWC does work in the deep basin, using selected floats deployed in the Nansen Basin, which also implies the compilation of adequate reference data.

The Barents Sea and the Nansen Basin have been subject of several pilot deployments of European floats (Poland, Germany, Norway and France) in recent years. For details about these deployments the reader can check this dashboard in the Euro-Argo fleet monitoring website . The Barents Sea, and other shelf seas like the Kara Sea where floats deployed there can drift to, is mostly shallower than 300m. Therefore alternative salinity DMQC methods need to be applied. At the moment of writing, none of these floats have D- files, meaning that no DMQC has been performed yet. We recommend that the methods developed for salinity DMQC in the Baltic, reported here in Section 2, are tested in these floats in the future. For this purpose, the appropriate reference profiles need to be gathered and the reference min/max climatologies should be calculated. Seasonal resolution might be crucial, given the strong cooling and the sea ice formation. Outside of the EA-RISE project, colleagues from the Norwegian Argo national program (NorArgo) have started to use a similar test based on binned statistics from the World Ocean Atlas, but including visual control, for biogeochemical parameters. At the moment, this analysis is done for many regions, where the Arctic as a whole is one of them, but it would be possible to divide it to isolate the Barents Sea region. Moreover, a possible source of reference data is a regional atlas under development at the Norwegian Institute of Marine Research (IMR, personal communication). For now, quick checks of Norwegian floats that revisited the same region have shown consistent values which suggest well-behaved salinity sensors. In the following, we will focus on the floats that have operated in the Nansen Basin, which depth allows the use of OWC.

4.2 Assessment of DMQC requirements

4.2.1 Estimation of under-ice positions

Since the OWC method relies on the selection of reference profiles acquired in the vicinity of the float profiles, the quality of float profiles positions is crucial. Floats operating in the Arctic will eventually encounter seasonal sea ice and therefore some of their positions will be missing or have been automatically assigned using linear interpolation.

Several methods that effectively improve the estimated positions have been proposed, using different approaches: interpolation in potential vorticity coordinates (Chamberlain et al., 2018), bathymetry constrained interpolation aided by float groundings (Wallace et al., 2020) and terrain (bathymetry) following interpolation (Yamazaki et al., 2020). All avoid obvious errors as estimated positions over land, and render more accurate positions than linear interpolation. As part of this project, the author of terrain-following method, Kaihe Yamazaki, was approached and asked to share the corresponding code in the Euroargodev Github, to make it available to the European DMQC operators and the Argo float community general. The python scripts are now available in https://github.com/euroargodev/terrain-following. This method has the advantage of having an easy implementation and is therefore recommended to estimate the under-ice positions of the floats in preparation for the OWC method tests. Since the DMQC methods “migration” from Matlab to Python has just started with the Euro-Argo RISE activities in Task 2.4, an example Jupyter-Notebook to facilitate the use of the terrain-following method is under development in this repository: https://github.com/euroargodev/terrain-following-example. As of writing, the available version is meant to be used locally and the estimated positions are replaced in the locally saved netcdf file, as one would need them to feed the OWC software. In the future, another version based on argopy, instead of locally saved files, will be written. This will allow the online use of the Jupyter-Notebook, through Binder, which will further facilitate its use for DMQC operators that are still building their python expertise. This effort will provide the basis for the inclusion of this interpolation method into the argopy package. The use of this method to assign under-ice positions as a standard practice will be put forward for discussion at the next Argo Data Management Team Meeting (ADMT-23) at the end of 2022.

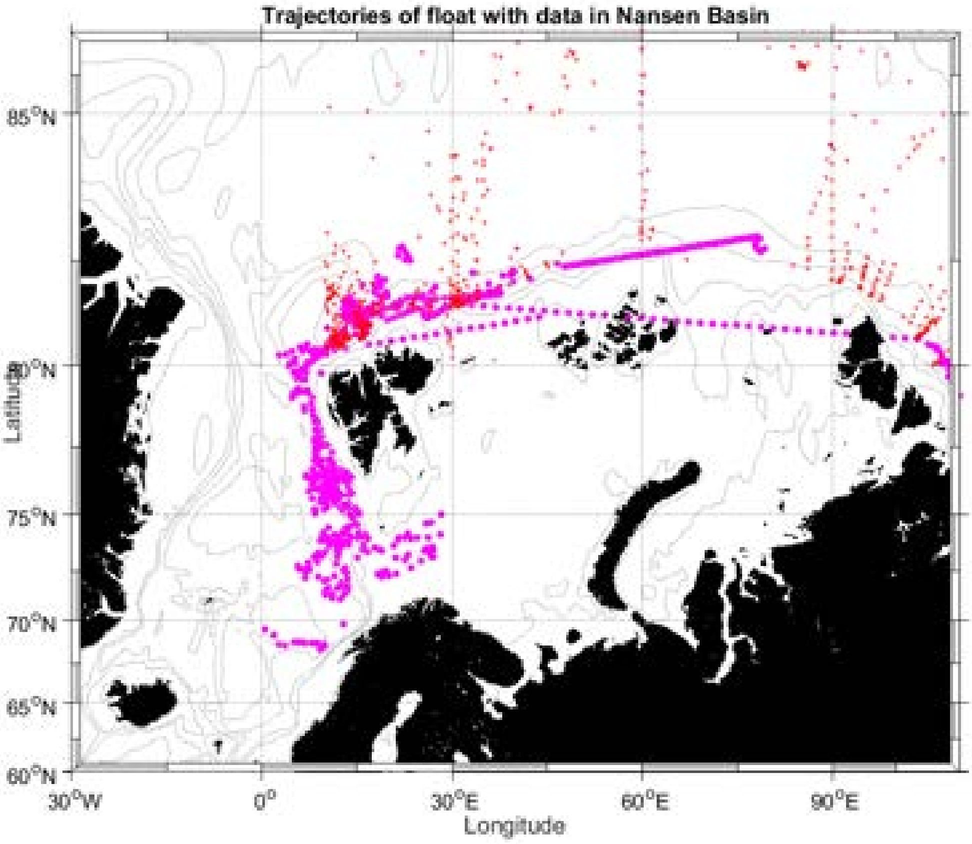


Figure 4.2. Positions of Argo profiles (pink) and reference data in the Nansen Basin (red). More detail about the floats included in this figure is found in Table 4.1.

Eleven floats have been identified which either drifted into the Nansen Basin through the Fram Strait or were directly deployed in the Sophia Basin. The float trajectories mostly follow topography and are confined to the continental slope. The float trajectories in Figure 4.2 that describe a straight line correspond to under-ice profiles, whose positions were linearly interpolated, but they most likely followed the continental slope while drifting with the boundary current. Although some floats already underwent delayed quality control procedures (Data-Mode D in Table 4.1), in some cases more realistic under-ice positions could be assigned in DMQC. A first attempt to obtain better estimates using the terrain-following software was made at BSH using its default configuration. This exercise retrieved 4 code errors, 6 good trajectories and 1 bad trajectories (Table 4.1).

WMO number	Programme	Number of cycles	Data-Mode	Salinity corrections	Quality flags of missing satellite positions	Terrain-following under-ice positions w. default parameters
3901910	Poland	89	D	N	QC 8	good
3902105	Poland	70	D	N	QC 8	good
3902107	Poland	113	D	N	QC 8	code error
3902108	Poland	71	D+A	N	QC 8,9	code error
3902112	Poland	72	D+A	Y	QC 8	bad
6901907	Germany	57	D	N	QC 8	code error
6902729	France	280	D	N	QC 8,9	good
6903548	Norway	45	D	Y	QC 8	good
6904087	Germany	26	R	N	QC 8	code error
7900549	Germany	70	R	N	QC 8,9	good
7900550	Germany	23	R	N	QC 8,9	good

Table 4.1. Floats with trajectories leading to the Arctic. Data mode refers to the Argo definitions and can be either R for real time data, A for real time data with on-the-fly adjustments and D for quality controlled files after DMQC was performed. For technical details about these floats the reader can access this dashboard in the fleet-monitoring website.

4.2.2 CTD reference database in the Arctic

The CTD reference database for quality control of Argo floats salinities is maintained by the Coriolis/Ifremer team for operational oceanography and it is updated by them at least once a year. The origin of the profiles is traced in the variable *qclevel* using a three letter code. Usually these updates include data delivered directly by scientists/principal investigators (SPI) and data obtained via downstream services: the updates of the Ocean Climate Laboratory/NODC to the World Ocean database WOD (OCL), and data from the International Council for the Exploration of the Sea ICES (ICE), the CLIVAR and Carbon Hydrographic Data Office CCHDO (CCH) and their product GO-SHIP (GSH). Also, visual quality control by period and platform are performed to ensure the good quality of the data. The updates and changes are reported at the ADMT every year.

Since 2019, BSH has worked to improve the CTD reference database, as reported in the Argo Data Management Team meetings (ADMT20, ADMT21 and ADMT22). As part of MOCCA, a local improvement focused on the Nordic Seas was done and is described in Angel-Benavides et al. (2020). The main action was to include profiles from the Unified Database for Arctic and Subarctic Hydrography UDASH (Behrendt et al. 2018) and ICES in the boxes surrounded by a dotted blue line in Figure 4.3., which include some boxes in the Arctic. The UDASH data, a very comprehensive collection of profiles from different databases, was stored with the code UDA and more detailed information was stored in the variable *source* (database of origin, cruise name, station number, etc). The profiles there were extensively quality controlled by the authors. ICES data is included in UDASH, but data was directly downloaded from their webpage for 2015 onwards, since they are not included in UDASH. At the moment, the ICES data was distributed without quality flags and needed extensive quality controls before being added. Therefore, it was stored with the code BSH. This update was included in the CTD_for_DMQC_2020v1 release of the database. Then, under EA-RISE, the European and Asian Arctic regions, shown in the figure in green and red respectively, were updated with UDASH profiles (2005 to 2015). This update was released in the CTD_for_DMQC_2021v1. During the merging of all datasets, there was a careful check for duplicates and near-duplicates, using Matlab scripts that are available on the repository **CTD-RDB_for2021v01 release** hosted in the Euroargodev Github.

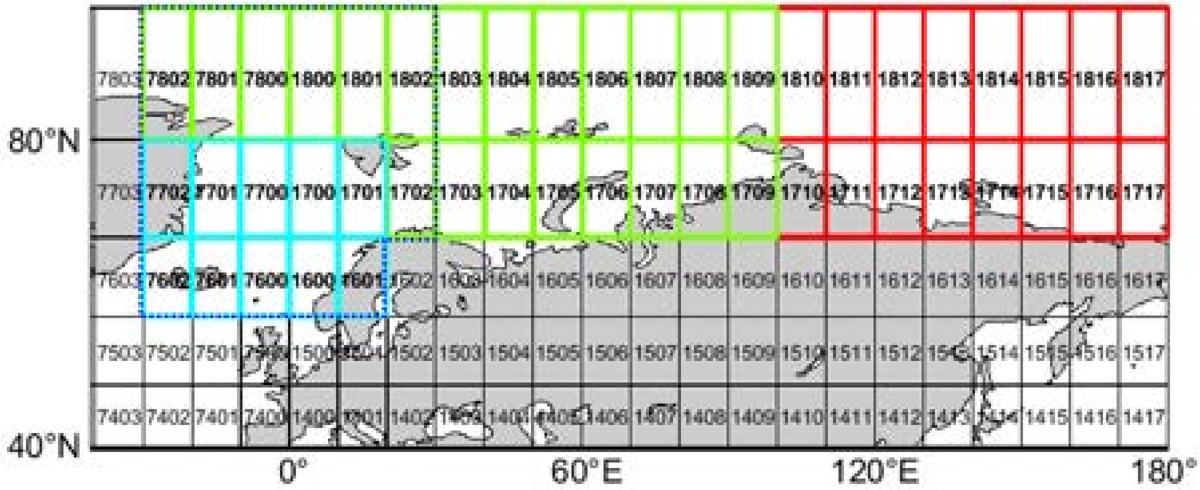


Figure 4.3. Map showing the regions and WMO boxes that were updated by BSH until 2021. The Nordic Seas (cyan), the European (green) and Asian (red) Arctic were the regions targeted during Euro-Argo RISE, and the area surrounded by a blue dotted line was the one targeted in MOCCA.

A diagnosis of CTD_for_DMQC_2021v2, the latest version of the reference database as of writing, was performed using the **check_CTD-RDB** toolbox, also developed in the context of MOCCA and EA-RISE's Task 2.4. The boxes selected (Figure 4.4) for this analysis are all those north of 70°N (codes starting with 17, 77, 18 and 78) except for those between -100°W and 20°E (Baffin Bay and Nordic Seas).

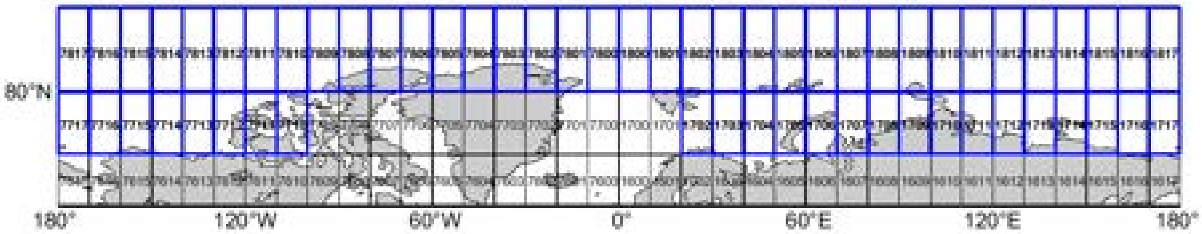


Figure 4.4. Map showing WMO boxes used in the diagnosis of CTD_for_DMQC_2021v2

There is no obvious gap in the spatial coverage of the data (Figure 4.5) and is densest in the Nansen Basin. However, the total number of profiles is only 1608 covering the time span of 1979-2018 (Figure 4.6) and therefore the actual number of reference profiles available for each float profile is very limited (FIgure 4.2).

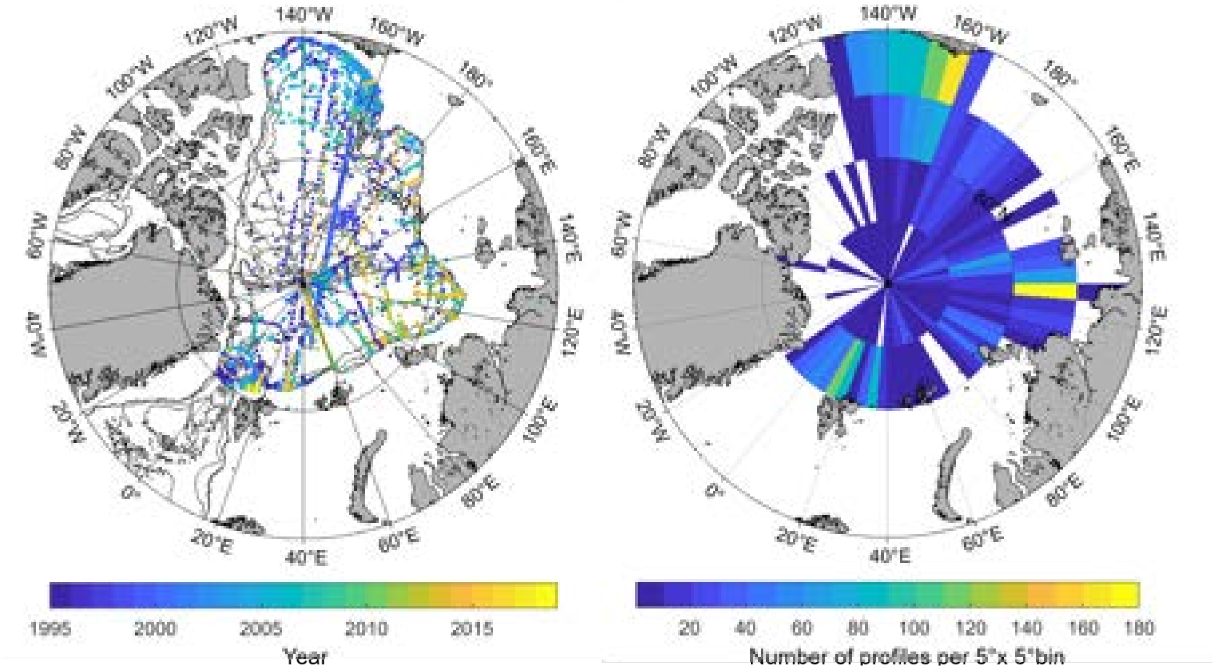


Figure 4.5. Maps showing the position of the profiles colour-coded by year (left panel) and the number of profiles per 5° x 5° bin (right panel) in the CTD_for_DMQC_2021v2.

It is concerning that the most recent data in the database are from 2018 (Figure 4.5 and 4.6) which makes for a delay of 4 years at the release of the reference data. Thus, it is also necessary to find other data sources to obtain more recent profiles. This includes trying to obtain data directly from scientists, thus avoiding the delays on delivery to the open databases and their release to the public. Moreover, some data that is not meant for public access or is being withheld temporarily (moratorium). Therefore the principal investigators need to be approached in the future to persuade them to share their data for Argo DMQC purposes.

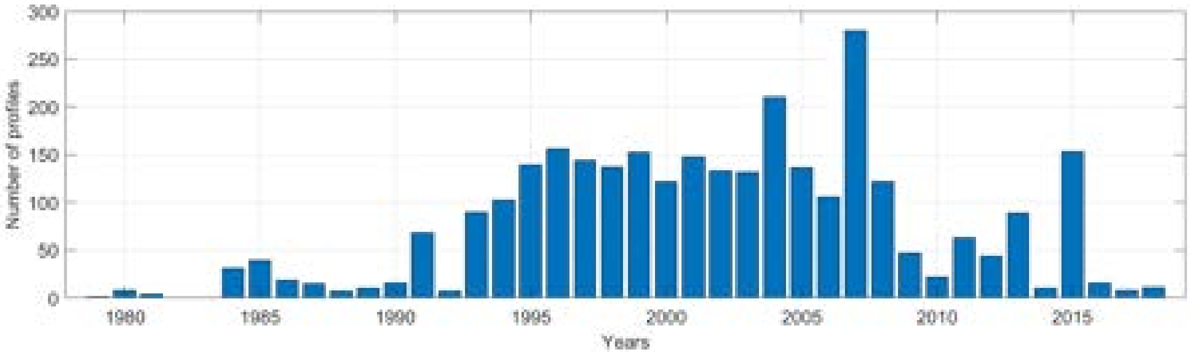


Figure 4.6. Number of profiles per year the CTD_for_DMQC_2021v2.

Taking into consideration that several floats in the Nansen Basin drift with the boundary current along the shelf edge or to adjacent shallower areas (Figure 4.2), the number of reference profiles available for DMQC is also strongly limited by the lack of profiles with maximum recorded pressure (MRP) shallower than 900 dbar in the CTD_for_DMQC_2021v2. This is because, according to the rules established for floats operating in the open ocean, described in Section 4.5 of the Argo Quality

Control Manual for CTD and Trajectory Data (currently in v3.5, Wong et al., 2022) prescribes the CTD-RDB selection, aggregation, and quality control procedures, only profiles with MRP > 900 dbar are included in the database. This selection criteria is likely excessively restrictive for a marginal sea like the Arctic. Similar to what has been done for the Mediterranean (Notarstefano 2019, MOCCA report D4.4.1 and this document), it is useful to define a shallower maximum recorded pressure criteria for the Arctic.

4.2.3 Updates to the CTD reference database

As a complement to the yearly checks for new profiles performed by Coriolis/Ifremer, BSH has acquired data from alternative sources to improve the status of the reference database in the boxes shown in Figure 4.4. These sources are: i) PANGAEA, for the data acquired by the German Research Vessel and Icebreaker Polarstern, since they do not deliver their data directly into the WOD (personal communication); and ii) NABOS, the Nansen and Amundsen Basins Observational System, is part of the Arctic Observing Network and is funded by the National Science foundation and the National Oceanic and Atmospheric Administration in the US. Profiles with MRP>700 db were selected. The usual sources UDASH and ICES, which recently has incorporated quality flags into their dataset, were also used. The data merging included thorough duplicate checks as in previous BSH updates. The resulting database will be referred to as CTD_for_DMQC_2021v2_plus.

UDASH and ICES

Profiles with MRP between 700 and 900 dbar and collected between 2005 and 2015 were selected from the UDASH database, since profiles with MRP>900 dbar were already present. All ICES profiles with MRP>700 dbar collected from 2015 onwards were selected, to ensure that possible updates in their database was included, but no new data with MRP>900 was found. In both cases the quality flags were applied and only good samples were kept. A total of 178 profiles were added, 101 from UDASH and 77 from ICES. Most of the UDASH data (60%) and all of the ICES data is from the region North of Svalbard (boxes 1800 to 1803).

POLARSTERN/PANGAEA

A campaign list of Polarstern including cruise maps can be found in the PANGAEA website (https://www.pangaea.de/expeditions/bybasis/Polarstern) and from this list links are provided to cruise tracks and data. Since 2015 the cruises PS93/1+2, PS94, PS99/1+2, PS100, PS101, PS106/1+2, PS107, PS108, PS109, PS114, PS11[BK1] 5/1+2, PS121, PS122/1-5, and PS126 have conducted. From these 14 cruises only 6 (PS93, PS94, PS106, PS107, PS109, and PS114) are available (as *oce.tab* files). The rest of them are either under moratorium or the data has not been yet submitted or made available. Communication with the Alfred Wegner Institute (AWI) is ongoing to ensure easy and more timely access to all PANGAEA data for Argo DMQC. Data from PS93 and PS107 were already in the database. The Table 4.2 summarises the new 61 profiles added from this source.

Pressure range\ Cruise	PS94 -2015	PS106-2017	PS109-2017	PS114 - 2018
700 to 900 db	1	1	1	5
900+ db	8	37	7	2

Table 4.2 Summary of the new profiles from Polarstern cruises in the CTD_for_DMQC_2021v2_plus.

NABOS

The goal of NABOS is to compile a cohesive picture of climatic changes in the Eurasian and Makarov basins of the Arctic Ocean. The multinational cruises started in 2002 and the most recent, but not yet freely available, started in September 2021. NABOS cruises started in and the most recent one was started in September 2021. The data for the NABOS cruises in 2013, 2015 and 2018 have been downloaded and were added. The duplicate checks discovered that 100 of the 126 stations of the 2013 and 2015 cruises were already in the database (from UDASH). Thus only 89 new profiles were added mostly from the 2018 cruise and only 5 shallower than 900 db. In this first exercise, the UDASH data was replaced with the NABOS because this is the original source, but before deciding which to keep it is probably worthwhile to compare the profiles, since UDASH profiles underwent another round of quality control. If the UDASH data is considered better, its metadata should be updated (*source*) since it is incomplete at the moment.

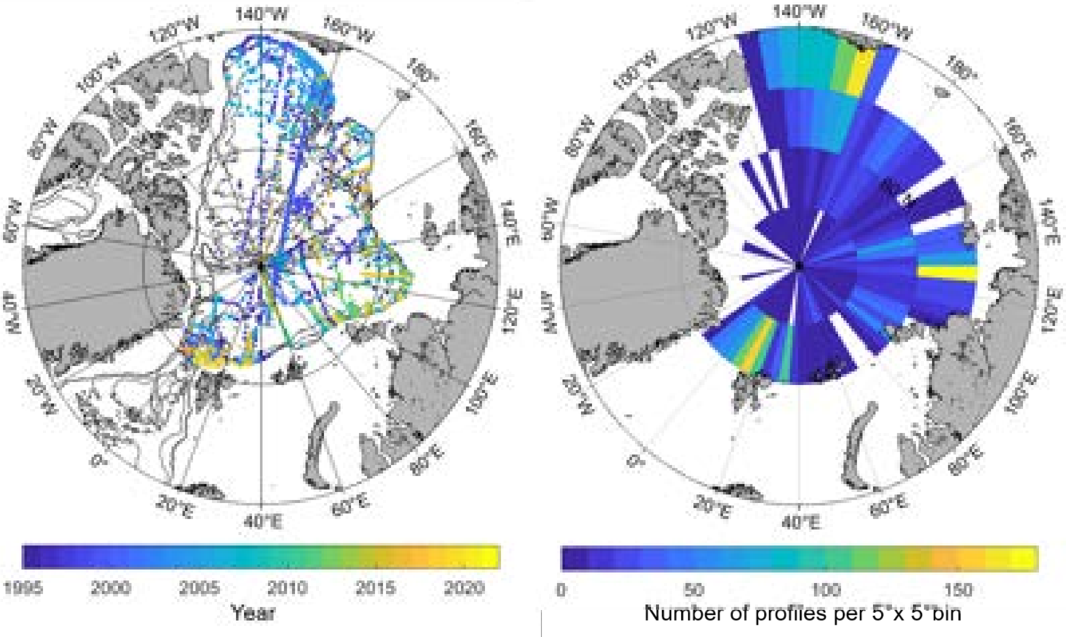


Figure 4.7. Maps showing the position of the profiles colour-coded by year (left panel) and the number of profiles per 5° x 5° bin (right panel) in the CTD_for_DMQC_2021v2_plus..

The profile positions and the number of profiles in the CTD_for_DMQC_2021v2_plus is shown in Figure 4.7. Figure 4.8 shows the year of the latest profiles in each 5° x 5° bin before and after the update, where it is evident that the update has provided more recent data (brighter yellow bins) specially in the Sophia Basin and the Yermak Plateau, north of Svalbard where data from 2021 were added. However the number of profiles added was low in general (300) from which half were profiles between 700 and 900 dbar. The total and new profiles per year (Figure 4.9) show that there are only little more than 200 profiles from the last five years and that most of the data added was for 2018, showing an important lag in the data availability that needs to be shortened by, as mentioned before, contact scientist directly to access their data before their release to the public. Nevertheless, due to the fact that there are few and spaced cruises in the Arctic, the lack of reference profiles will remain an issue for the DMQC in the region, even if we manage to have access to most of them.

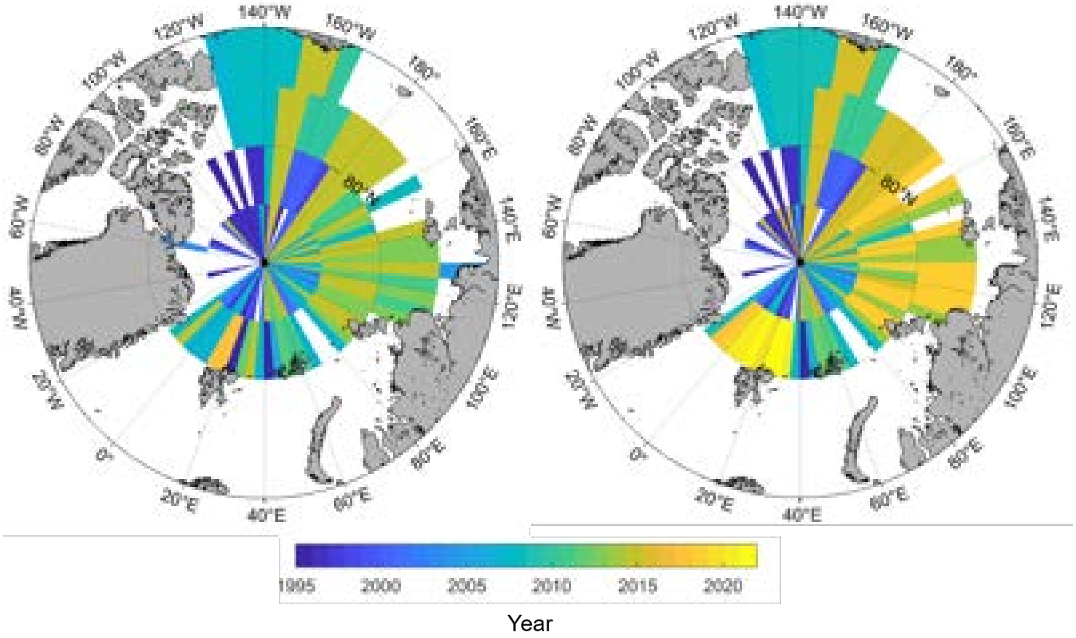


Figure 4.8. Maps showing the year of the latest profile per 5° x 5° bin in the CTD_for_DMQC_2021v2 (left panel) and CTD_for_DMQC_2021v2_plus (right panel).

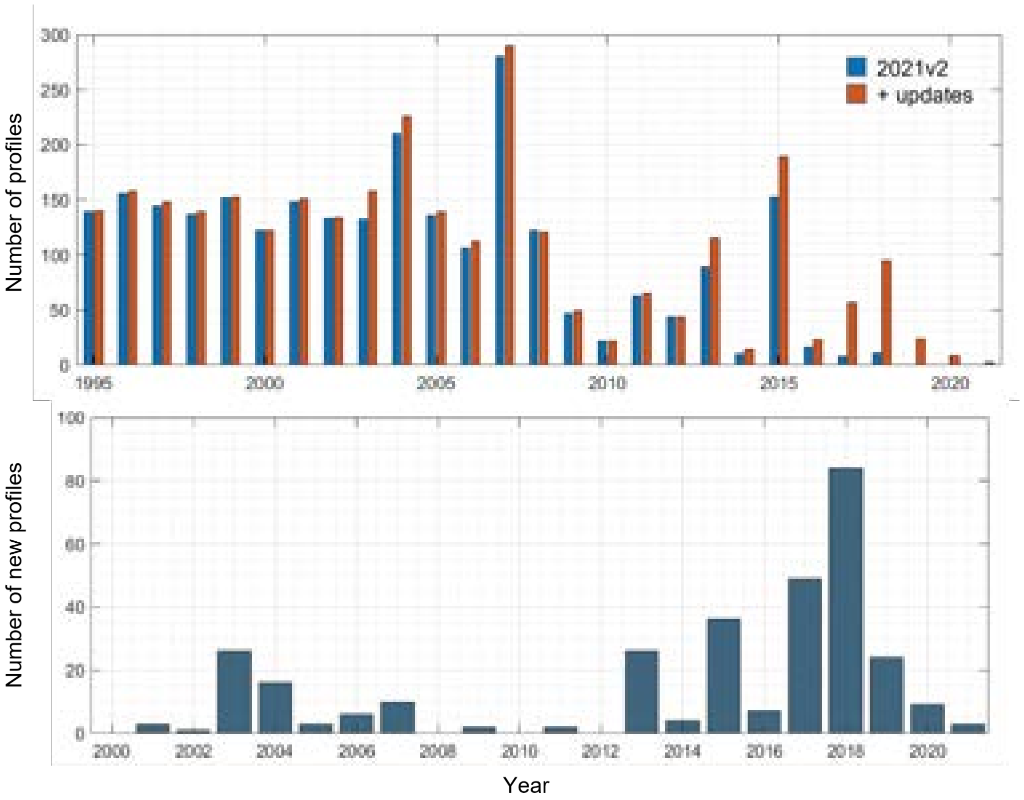


Figure 4.9. Number of profiles per year in the CTD_for_DMQC_2021v2 and CTD_for_DMQC_2021v2_plus (top panel) and the number of new profiles per year in the CTD_for_DMQC_2021v2_plus.

4.2.4 Quality of the CTD reference database

The thorough procedures to check for duplicates in the reference database, available here, are very important for the Arctic and its updates. During the preparation of the CTD_for_DMQC_2021V01, 468 and 266 duplicated profiles present in the European and Asian Arctic regions (see green and red boxes in Figure 4.3) were identified from the database. These duplicates were either already there in the previous version of the database, or were profiles that were in the ones selected for the update but were already in it. This procedure led to a lower number of profiles in CTD_for_DMQC_2021V01 than in CTD_for_DMQC_2020V01. That update and the one described in this document show that having a robust duplicate check procedure in place allows us to search and add data from all possible sources without worrying about data redundancies and artificial increase in the number of profiles.

The CTD data in the present version of the reference data set (CTD_for_DMQC_2021V02) has been optically tested using plots by period and platform at Coriolis, and per boxes at BSH. However, further more detailed quality control is required. During the DMQC procedures of floats in the Nansen Basin, described below and shown in sub section 4.2.6, the presence of strange profiles was noticed (profiles in land/fjords, too high surface salinities, strange profile shapes, pressure that is non monotonically increasing, etc). These profiles, mostly from ICES via Coriolis, were removed before obtaining the counts reported above.

Moreover, in most boxes in the Nansen Basin, many salinity profiles are either noisy or have strange, highly variable shapes. An example for box 1804 (east of Sophia Basin) is shown in Figure 4.10. In his box 22 of 143 profiles are noisy. They come from the OCL (14) and UDASH (8) and are mostly from 2001 or before. A quick check reveals that this issue affects boxes from 1803 to 1808, and from 1811 to 1817, but seems to affect mostly older profiles. In box 1710 some “wiggling” salinity profiles are found. BSH will seek the advice of Arctic oceanography experts to decide if such profiles are normal or if they are erroneous. Although further quality control is required, all profiles were kept for the time being, since the reference data is very scarce.

A question that remains open, due to the deep convection in the area, is how the presence of sea ice could affect the results of OWC. This could be explored by calculating the distance to the ice edge for each CTD reference profile and then select subsets better suited for comparison in dmqc or check the established corrections.

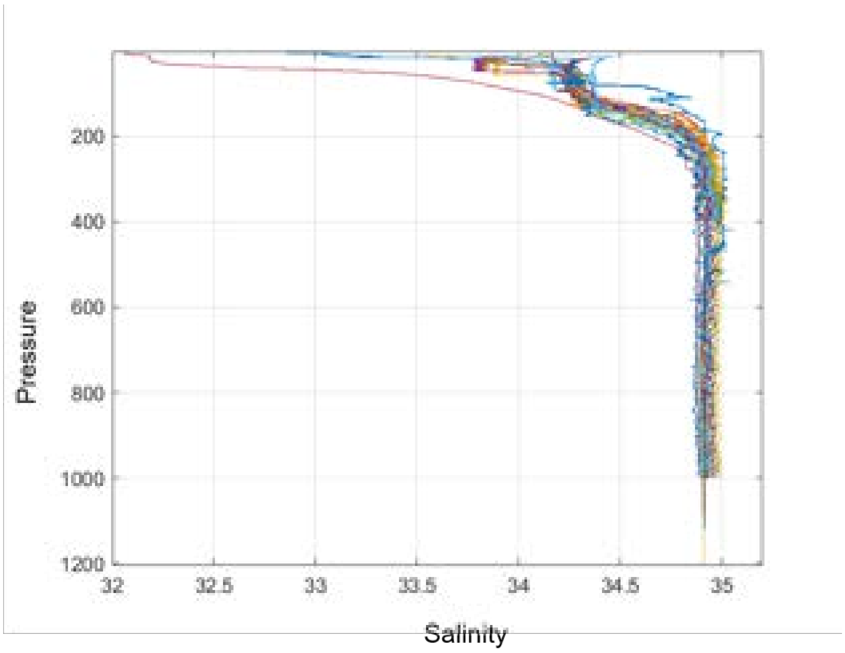


Figure 4.10. Selected noisy profiles from box 1804.

4.2.5 OWC test performance

Before testing the OWC method in floats operating in the Nansen Basin, the real-time quality control tests were checked. On a first assessment these tests perform well and no changes are required.

Tests of the applicability of the OWC method have been performed on floats 7900549 and 7905540 from the deployment in the Sophia basin which were not D-moded yet and seem to have issues with their salinity data. Both floats have received an update to their interpolated positions using the method by Yamazaki et al. (2020) prior to the runs of the dmqc to make the selection of the reference data as good as possible (Figure 4.11). Especially 7900549 which travelled along the continental slope needed a more sophisticated estimation of the under ice positions because the extended period under ice led to obviously wrong linear interpolated positions which crossed land.

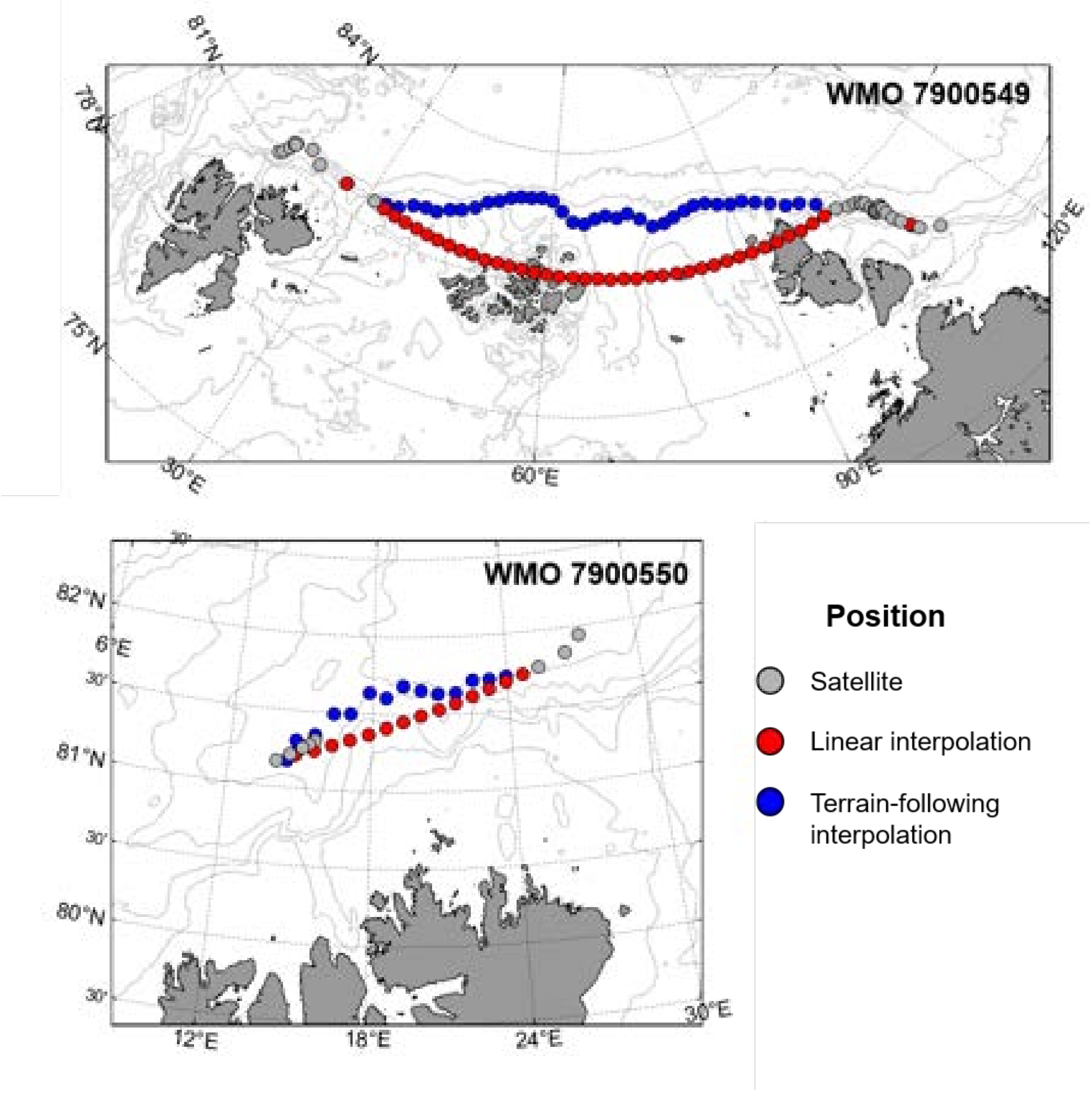


Figure 4.11. Profile positions from floats 7900549 (above) and 7900550 (below). The different colours show the GPS positions and the under-ice positions estimated using linear interpolation and the terrain-following interpolation method.

The runs with the OWC software were performed twice, first with the present version of the reference database (CTD_for_DMQC_2021V02) and were repeated with the updated version CTD_for_DMQC_2021V02_plus, including PANGAEA and NABOS data and profiles deeper than 700 db. The updates will be made available to the entire Argo community in the next release of the CTD reference database at the end of 2022. Due to the small amount of reference data added, there were only minor changes in the results and only the second run is shown and discussed below (Figures 4.11 and 4.12).

Float 7900549 travelled a considerable way from Svalbard to the island archipelago of Franz Josef Land and onward to the Laptev Sea beyond Servernaja Semlja. The position of available reference profiles is given by the blue dots in the Figure 4.12 below. The mapping parameters for the OWC follows settings for the Nordic Seas with small search radii (Table 4.3), but because of the sparsity of reference data the time constraint has been relaxed a bit. Reference data are generally sparse and after leaving the Sophia Basin 7900549 does not have any reference data for long stretches of its trajectory.

Figure 4.12 from the OWC diagnostics for float 7900549 would indicate that this float starts with a fresh bias in the order of 0.007, which was also apparent in comparisons with other float data in the same area (Figure 4.14). But given the sparsity of the reference data it remains unclear if the fresh bias increase to 0.02 for 7900549 around cycle 50 is reliable when the float resurfaced in the Laptev Sea after its long stay under ice. The data transmitted from the float do indicate changes in water mass characteristics, putting the reference data under suspicion. In this case it is advisable not to follow the output from the OWC analysis too strictly and either wait with dmqc until more reference data are established, increase the salinity error or change the salinity flag.

OWC Parameter	Selected value
p_delta	50
p_exlcuded	100
scale_age	1.3404
scale_age_large	4
scale_lat_large	1
scale_lat_small	0.25
scale_long_large	1.60
scale_long_small	0.4
scale_phi_large	0.5
scale_phi_small	0.1
use_pv	1
use_saf	0

Table 4.3. Parameter settings used in the OWC runs performed for float 7900549 and 7900550.

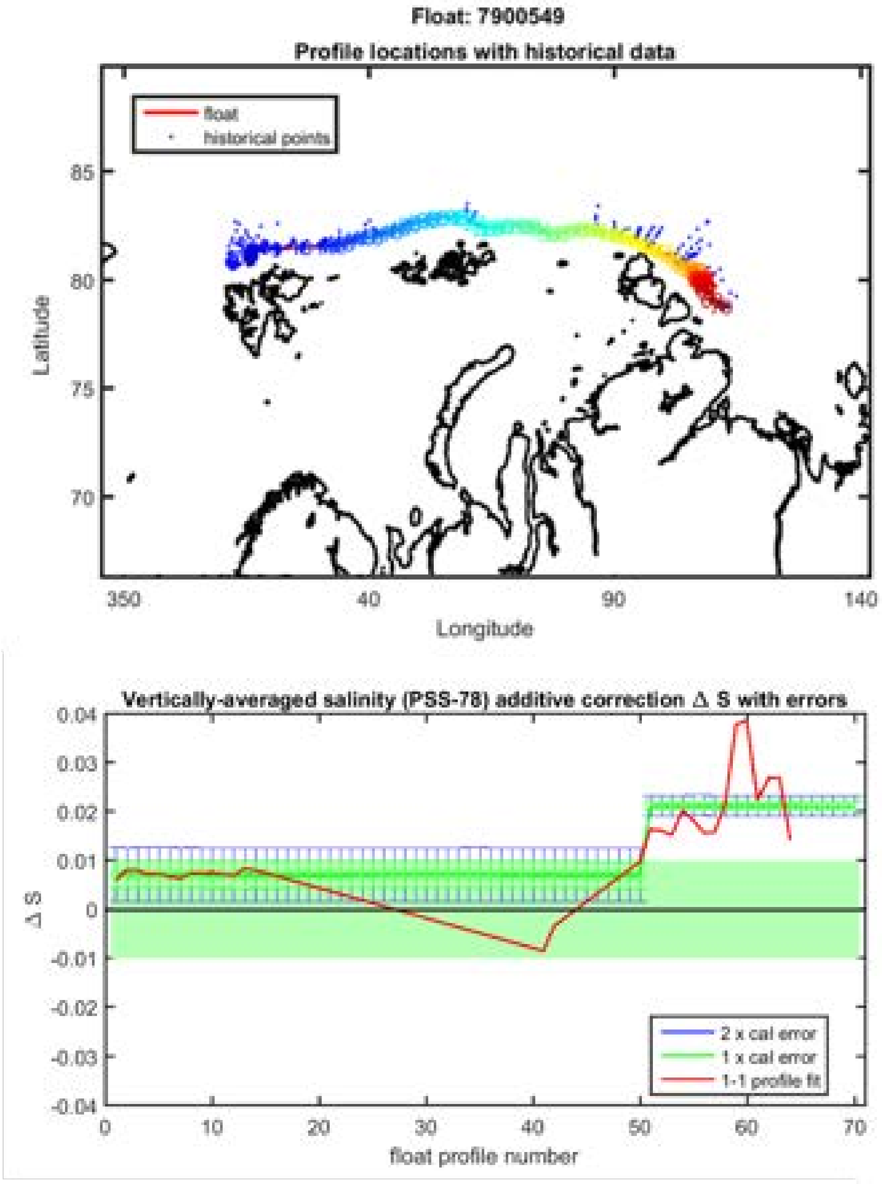


Figure 4.12. Trajectory of float 7900549 (top panel), with profile positions shown as colour coded dots, and available reference profiles shown in blue. Suggested additive salinity correction for each float cycle calculated by OWC (bottom panel) and extracted from its diagnostic plot number 3. Reference data used was CTD_for_Argo_2021V02_plus.

Float 7900550 travelled a similar trajectory as 7900549 but has not resurfaced after it went under ice still in the Sophia Basin (Figure 4.13). Reference data are available along its entire trajectory, with the similar caveat that all are situated north of the float trajectory in deeper water. According to the OWC diagnostics the float shows a similar large fresh offset as 7900549 and would need to be corrected by 0.007.

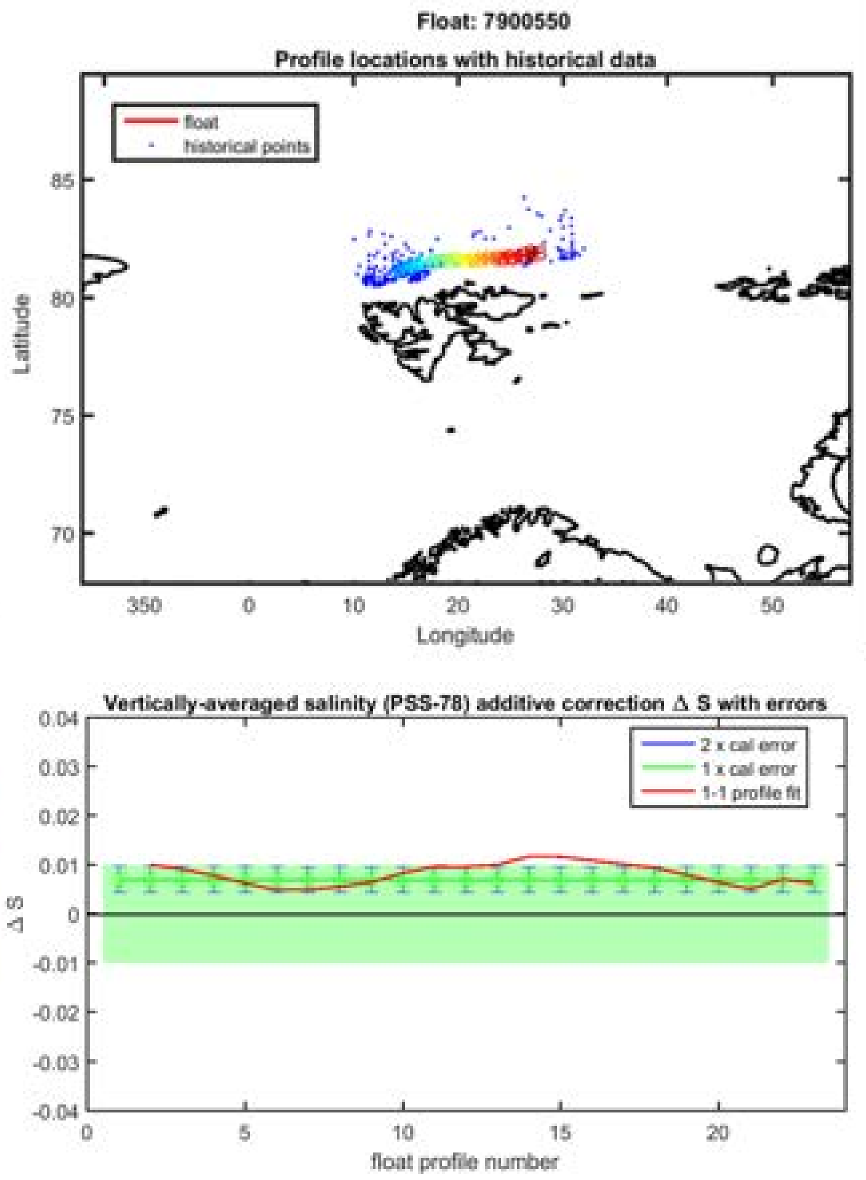


Figure 4.13. Trajectory of float 7900550 (top panel), with profile positions shown as colour coded dots, and available reference profiles shown in blue. Suggested additive salinity correction for each float cycle calculated by OWC (bottom panel) and extracted from its diagnostic plot number 3.

4.2.6 Alternative quality control check

Since it is unlikely to increase the amount of reference CTD data in the area in the near future, the quality assessments need to address potential drift and offsets by checking on the stability of a respective float record over time and the consistency with data from other floats deployed in the area. If more and more floats are D-moded this will build up a substantial Argo contribution to the reference database over time. This situation is similar to the Caribbean, where the ratio of available reference data from CTD and ARGO is already dominated by Argo by a factor of 2.3. It is also strongly recommended to perform a CTD cast at deployment to assure the quality of the Argo data set. Figure

4.13 shows the intercomparison of the existing CTD reference data and 11 floats which have operated in the Nansen Basin along the continental shelf. Except for the three floats addressed above, all others have already been through quality control (Table 4.1).

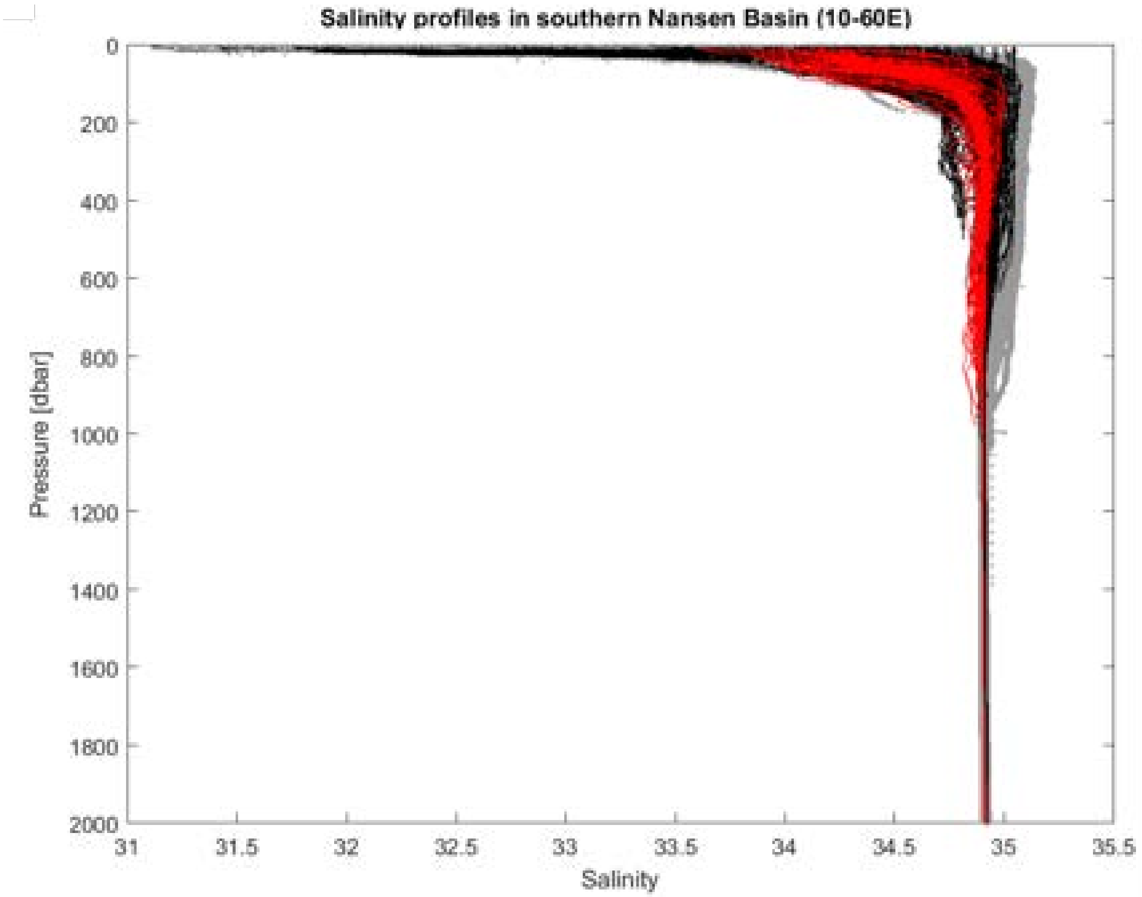


Figure 4.14. Comparison of argo profiles from the 11 floats listed in table Arctic 1. The comparisons are made for all profiles in the Nansen Basin in the longitude range 10-60 °E. Only data in the depth range down to 2000 m are shown, which is covered by Argo floats. The profiles from floats 7900549 and 7900550 discussed above are highlighted in red and show uncorrected salinities.

The internal consistency of the salinity profiles delivered by the 11 floats in the Arctic appears good. But all profiles are fresher than the climatological data in the upper 1000 m of the water column and agree better with climatology below that depth. The freshening of the upper layer salinities is most likely reflecting temporal changes in the upper water mass due to climate change. The two floats discussed above which showed fresh biases in the OWC diagnostics are also fresh compared to their fellow floats, but would obtain a much better match with their buddy floats if corrected by +0.007. It has to be noted that there is nearly no reference data available between 30 E to 60 E to verify and quantify the fresh water masses at the depth between 650-1000 in the float profiles from 7900549 collected close to Franz Josef Land. The continental shelf east of Franz Josef Land towards Severnaja Semlja is totally void of reference data and floats reaching this area can not be verified independently. The apparent rise in suggested corrections of float 7900549 to +0.020 is difficult to justify, the reference data only go back to 2018 and are mostly much older. The water mass characteristics in the reference data show a massive freshening and warming in the float data compared to the reference data which reaches down to 2000 m.

Recommendations for the DMQC in the Arctic

- Use OWC with similar settings as in Nordic Seas, but longer time scales (Table 4.3)
- Carefully inspect OWC proposed fits for unreliable data due to gaps in reference data distribution
- Compare float data with buddy dataset to check for internal consistency
- Make efforts to engage with the Arctic observational community to expand the CTD reference database with new data and in a timely manner.
- Perform a thorough detailed quality control of the profiles in the reference database with assistance of the experts in Arctic oceanography.
- Perform a CTD cast at deployment whenever possible to assist the DMQC
- Continue to test the interpolation methods for the estimation of under-ice positions.

4.3 Summary and future work

The use of OWC to perform salinity DMQC is possible in the deep basins of the Arctic without modifications. However, the lack of recent and well distributed reference data in the region make the outputs of the procedure somehow unreliable. In the future, the efforts to obtain more reference data from scientists should be intensified and a CTD cast should be done at deployment. However, it is likely that even adding all existing CTD casts would not solve the problem of large regions void of reference data. As in the Caribbean sea, the quality of the float data can still be checked by comparing it with that of other Argo floats in the region, slowly generating a reliable Argo reference database that would improve the reliability of the OWC results.

Obtaining good estimates of the under-ice profile positions is important not only for the OWC method but also for the scientific value of the Argo data in the Arctic. Although the definition of a standard method for a method more refined that linear interpolation has yet to be discussed in the ADMT, an easy to use method (terrain-following) has been made available to the DMQC operators via the Euroargodev Github, and easy to use versions of it are underway.

For the DMQC in the shallow Barents Sea, we recommend that the future efforts emulate the work done for the Baltic during this project, namely: establishment of a reference dataset and calculation of min/max climatologies (see subsection 2.2.1) that can be used in real-time quality control, and the use of the proposed alternative DMQC procedure for shallow waters (see subsection 2.2.5).

References

Allen, J.T., Munoz, C., Gardiner, J., Reeve , K.A., Alou-Font, E., and Zarokanellos, N. (2020) Near-Automatic Routine Field Calibration/Correction of Glider Salinity Data Using Whitespace Maximization Image Analysis of Theta/S Data. Frontiers in Marine Science. 7:398. DOI: 10.3389/fmars.2020.00398.

Allen, J.T., Fuda,J-L., Perivoliotis, L., Munoz-Mas, C., Alou, E. and Reeve, K. (2018) Guidelines for the delayed mode scientific correction of glider data. WP 5, Task 5.7, D5.15. Version 4.1. Palma de Mallorca, Spain, SOCIB - Balearic Islands Coastal Observing and Forecasting System for JERICO-NEXT, 20pp. (JERICO-NEXT-WP5-D5.15-140818-V4.1). DOI: http://dx.doi.org/10.25607/OBP-430

Angel-Benavides, I.M., Klein, B., and Coatanaon, C., (2020). D4.4.5 Report on the update of the CTD reference database for Salinity DMQC in the Nordic Seas. MOCCA report. https://www.euro-argo.eu/content/download/142333/file/D4.4.5%20Report%20on%20the%20update%20of%20the%20CTD%20reference%20database%20for%20Salinity%20DMQC%20in%20the%20Nordic%20Seas_v1.0.pdf

Behrendt, A., Sumata, H., Rabe, B., and Schauer, U. (2018). UDASH– Unified Database for Arctic and Subarctic Hydrography. Earth System Science Data, 10(2):1119–1138. https://doi.org/10.5194/essd-10-1119-2018

Cabanes Cecile, Angel-Benavides Ingrid, Buck Justin, Coatanoan Christine, Dobler Delphine, Herbert Gaelle, Klein Birgit, Maze Guillaume, Notarstefano Guilio, Owens Breck, Thierry Virginie, Walicka Kamila, Wong Annie (2021). DMQC Cookbook for Core Argo parameters. https://doi.org/10.13155/78994

Cabanes, C., Thierry, V., and Lagadec,C. (2016). Improvement of bias detection in Argo float conductivity sensors and its application in the North Atlantic. Deep-Sea Research Part I, 114, 128-136. https://doi.org/10.1016/j.dsr.2016.05.007

Chamberlain, P. M., Talley, L. D., Mazloff, M. R., Riser, S. C., Speer, K., Gray, A. R., and Schwartzman, A. (2018). Observing the ice-covered weddell gyre with profiling floats: Position uncertainties and correlation statistics. Journal of Geophysical Research: Oceans, 123(11):8383–8410.https://doi.org/10.1029/2017JC012990

Detoc, J., Thepault, B., Carval, T., Mahoudo, P., Garo, M. (2021). Scoop-Argo : visual quality control for Argo NetCDF data files. SEANOE. https://doi.org/10.17882/48531

Fofonoff, N.P. and Millard Jr, R.C. (1983) Algorithms for the computation of fundamental properties of seawater. Paris, France, UNESCO, 53pp. (UNESCO Technical Papers in Marine Sciences; 44), DOI: https://doi.org/10.25607/OBP-1450

Gourrion, J., Dobler, D., Gallian, M., Szekely, T., Beaulieu, N. (2020). Reference files and python code to run Min/Max QC test. SEANOE. https://doi.org/10.17882/77199

Gourrion, J., Szekely, T., Killick, R., Owens, B., Reverdin, G., Chapron, B. (2020). Improved Statistical Method for Quality Control of Hydrographic Observations. Journal Of Atmospheric And Oceanic Technology, 37(5), 789-806. Publisher's official version : https://doi.org/10.1175/JTECH-D-18-0244.1 , Open Access version : https://archimer.ifremer.fr/doc/00628/74031/

Klein,B. Coatanaon, C., Purokoski, T., Siiriä, S.M, Tuomi, L. and Walcowski, W. (2020) Data management for floats in the Baltic. MOCCA Report 4.4.7. . https://www.euro-argo.eu/content/download/142348/file/D4.4.7_Data_management_for_floats_in_the_Baltic_v1.0.pdf

Notarstefano, G. (2019) D4.4.1 Report on the reference dataset for the delayed-mode quality control activity in the Mediterranean and Black Sea. MOCCA report. Restricted access.

Owens, W.B. and Wong, A.P.S. (2009). An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by Ɵ -S climatology. Deep-Sea Res.Part I, 56, 450–457. doi:10.1016/j.dsr.2008.09.008

Sea-Bird Application Note 31: "Correcting for Conductivity Drift Based on Pre- and Post-Cruise Laboratory Calibrations".

Siiria, S.M., Roiha, P., Tuomi, L., Purokoski, T., Haavisto, N., and Alenius, P. (2018) Applying area-locked, shallow water Argo floats in baltic sea monitoring. Journal of Operational Oceanography. 0 (0): 1-15. DOI: https://doi.org/10.1080/1755876X.2018.1544783

Tomczak, & Godfrey. (1994). *Regional oceanography: An introduction*. Oxford: Pergamon.

Wallace, L. O., van Wijk, E. M., Rintoul, S. R., and Hally, B. (2020). Bathymetry-constrained navigation of Argo floats under sea ice on the Antarctic continental shelf. Geophysical Research Letters,47(11):https://*doi*.org/10.1029/*2020GL087019*

Wong Annie, Keeley Robert, Carval Thierry, Argo Data Management Team (2022). Argo Quality Control Manual for CTD and Trajectory Data. https://archimer.ifremer.fr/doc/00228/33951/

Yamazaki, K., Aoki, S., Shimada, K., Kobayashi, T., and Kitade, Y. (2020). Structure of the subpolar gyre in the Australian-Antarctic basin derived from Argo floats. Journal of Geophysical Research: Oceans, 125(8). DOI: https://*doi*.org/10.1029/*2019JC015406*