Working Group 148: International Quality controlled Ocean Database (IQuOD)

1. Brief summary with the main highlights (200-300 words)

International Quality controlled Ocean Database (IQuOD) group continues to develop, implement and document new techniques to provide an unbiased climate dataset of the highest quality, consistency, and completeness. The community has developed a machine learning approach for the assignment of "intelligent" metadata for XBT data and assigned the uncertainty for each individual measurement. The group has also evaluated the most effective combination of automated quality control (AutoQC) procedures for temperature profile observations. With these developments, the group has freely disseminated an interim version of the IQuOD global temperature profile database through NOAA/NCEI. IQuOD has been very active with an online meeting every month and has established collaboration with OTGA. An in-person meeting has been organized on July 10-11 in Potsdam, Germany 2023, and the steering team has discussed and decided the main activities for the following 1-2 years.

2. Activities since previous report to SCOR (e.g., virtual or in-person meetings, email discussions, special sessions). Limit 1000 words

The paper describing work on benchmarking automatic quality control checks has been published (see below, Good et al. 2022).

IQuOD v0.1 dataset has been available to the public through NOAA/NCEI (https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0170893).

Activity in the task team focused on detecting duplicate profiles has made significant progress, with three years of data being checked and flagged. The group is preparing a benchmark dataset for duplicate checks, which could be useful for machine-learning approaches in the future. The team is preparing a scientific paper on duplicate checking.

A new paper has been published describing a new Auto-QC system (CODC-QC), with the collaboration of several members of IQuOD (Tan et al. 2023). This paper highlights proper treatments of the local climatology check, gradient check, and instrument-specific check in the Auto-QC system.

A paper published by several IQuOD members proposes a new bias correction scheme for historical Nansen Bottle data (Gouretski et al. 2022).

A paper has been published (Haddad, et al. 2022), which develops a process of estimating missing instrument type metadata (the combination of both model and manufacturer) systematically, constructing a machine learning pipeline based on thorough data exploration to inform these choices.

During the last year, there were online meetings (1.5 hours) every month, with presentations on QC-related topics either from IQuOD members or the broad community. The progresses of task teams were routinely reported at each meeting.

A collaboration has been established with OceanTeacher Global Academy (OTGA). Several members and co-chairs have an online meeting with Gregory Reed from OTGA group to discuss the potential collaboration. IQuOD group decided to start to design the training courses in 2023 and contribute to OTGA in 2024.

On July 10-11, 2023, IQuOD organized a steering group meeting to review the progress and discuss the activities in the following one or two years.

3. Documents published since previous report to SCOR (e.g., peer-reviewed journal articles, reports, Web pages) and should be limited to publications that resulted directly from WG activities and which acknowledge SCOR support

(1). Auto QC paper: Good S, Mills B, Boyer T, Bringas F, Castelão G, Cowley R, Goni G, Gouretski V and Domingues CM (2023) Benchmarking of automatic quality control checks for ocean temperature profiles and recommendations for optimal sets. Front. Mar. Sci. 9:1075510. doi: 10.3389/fmars.2022.1075510

(2). There are a number of updates for the IQuOD website (www.iquod.org), for instance the documents and papers from IQuOD group, updates the membership, and the support of IQuOD.

4. Progress toward achieving group's terms of reference. List each term of reference separately and describe progress on each one. Limit 1000 words

1. To develop, implement, and document algorithms for assignment of "intelligent" metadata – i.e. an informed guess as to likely values for missing information – for temperature profiles where crucial metadata is missing.
   - Achieved: a paper was previously published on this by IQuOD and the IQuOD dataset already includes the results of applying its intelligent metadata for expendable bathythermograph (XBT) data.
   - Several other papers have followed from the Met Office group (i.e. Haddad, et al. 2022).

2. To evaluate and document the most effective combination of automated quality control (AutoQC) procedures for temperature profile observations. International collaboration will be required for the design and coordination of benchmarking experiments using high-quality reference datasets.
   - Achieved: a paper has been published (Good et al 2023)

3. To establish and implement a set of optimal automated quality control procedures, by reaching international community consensus and using the knowledge gained in the benchmarking tests from ToR-2 (above); to produce and publish a reference guide for best practices in automated quality control of ocean temperature profiles; and to develop and freely distribute an open-source quality control software toolkit to promote wide and rapid adoption of best practices by the oceanographic community.
   - Achieved: the software described in ToR-2 is open source and is published under an open license (the MIT license). The AQC benchmarking paper has been published (Good et al, 2023).
   - Note: there will be some improvements in the coming months, which will be implemented in the IQuODv-2023 dataset.

4. To examine and document the feasibility of machine learning and other novel computational methods for enhanced quality control, to potentially minimize labor costs associated with human expert quality-control procedures.

- o Achieved: a cloud computing account to support this activity continues to be funded; we have transferred code to this cloud facility.
- o Achieved: Several papers have been published by MetOffice groups using the ML method to better identify the missing meta data of the XBT dataset.

5. To develop, implement and document internationally agreed best practice methods for the assignment of uncertainty estimates to each temperature observation.
   - o Achieved: a paper was previously published on this by IQuOD and the IQuOD dataset already includes those uncertainty estimates.

6. To freely disseminate (interim) versions of the IQuOD global temperature profile database (and added-value products) as it evolves over the next 3 years, in user-friendly file formats.
   - o Achieved: IQuOD v0.1 data are freely available from the US NCEI World Ocean Database website (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:0170893) in the widely used netCDF format.
   - o Note: the IQuOD group intends to release a 2023 version with AQC flags and subsequent versions will be numbered with the year of release.

7. To share knowledge and transfer skills in instrumentation, regional oceanography, quality control procedures, and data stewardship with international scientists in both developed and developing nations.
   - o Achieved: IQuOD virtual meetings are now being held on a monthly basis, in three alternating time slots across the day to increase the accessibility of the meetings for scientists across the globe.
   - o Note: there is an established collaboration with OTGA from 2023-2024, which will help to share knowledge and transfer skills worldwide.

5. WG activities planned for the coming year. Limit 500 words

Not applicable as WG will be finished this year.
There is a funded publication coming at the end of this year. The paper will be a data description paper for IQuODv-2023. The remaining SCOR funding will be used to pay for the publication fee of this paper.

6. Is the group having difficulties expected in achieving terms of reference or meeting original time schedule? If so, why, and what is being done to address the difficulties Limit 200 words

Not applicable

7. Any special comments or requests to SCOR. Limit 100 words.

Thank you for the multiple years of funding.
We request that the remaining SCOR funding be used to pay for the publication fee for an IQuOD data description paper, to be submitted and published by early 2024.

The group will continue the activities in the future and we hope to keep a connection with SCOR.

Additional information can be submitted and may be posted at the SCOR Annual Meeting webpage at the discretion of the SCOR Executive Committee Reporter for the WG and the SCOR Secretariat.