2019 International Quality-Controlled Ocean Database (IQuOD) – 6th IQuOD Annual Workshop

6th CLIVAR GSOP IQuOD
4th IODE SG-IQuOD
3rd SCOR WG 148

29-31 October, 2019. IFREMER,
Brest, France
Belgium

Editors

Rebecca Cowley, CSIRO, Australia; **Alison Macdonald**, WHOI, USA; **Tim Boyer**, NCEI, USA; **Guilherme Castelão**, Scripps Institute of Oceanography, USA; **Matthew Palmer**, Met Office Hadley Centre, UK; **Steve Diggs,** Scripps Institution of Oceanography, USA; **Marlos Goes,** AOML, USA.

Bibliographic Citation

IQuOD (International Quality-controlled Ocean Database) 6th Annual Workshop Report, November, 2019.

Sponsorship

We greatly acknowledge financial and logistical support from:

- The International Council for Science, Scientific Committee on Oceanic Research (SCOR) via Working Group 148 (http://www.scor-int.org/SCOR_WGs_WG148.htm).

- IODE Project for the SG IQuOD project established by IODE-XXIII (2015) through Recommendation IODE-XXIII.3 and the project is managed by the IODE Steering Group for the IODE-IQuOD project (SG-IQuOD).

Organizers

Organizing Committee

- Rebecca Cowley (CSIRO, Australia)
- Loïc Petit de la Villéon (IFREMER, France)
- Catia Domingues (University of Tasmania, Australia / co-chair of IQuOD Steering Team)
- Matthew Palmer (Met Office, UK / co-chair of IQuOD Steering Team)

Local Organizers/IFREMER

- Loïc Petit de la Villéon
- Virginie le Saout

Goals of the workshop

- Obtain agreement for the AutoQC process for application to the next product version.
- Plan for the next IQuOD product: what will it consist of and what do we need to complete to get to the next release?

Workshop Documents

All workshop documents and presentations can be found in the Google Drive below.
https://drive.google.com/open?id=1H1YKYi0a9Yzg4H0ELSCHjYWvrKCiWHuj

IODE Portal: SG-IQuOD

Please refer to the IODE website as a direct link to IODE portal is not available at the time of publication.
https://www.iode.org/index.php?option=com_content&view=featured&Itemid=89

Table of contents

1. Introduction Session

1.1. Welcome: Loïc Petit de la Villéon

Summary:
Loic opened the meeting with a welcome, mentioning how 40K jobs in the west Brittany are linked to the sea, and all the different researchers, programs, Universities, institutions and agencies. There are 80 Infrastructures and science platforms and >30 "innovative" SMEs. Ifremer itself (https://wwz.ifremer.fr/en/) has 5 centers, 19 coastal stations,1500 people & 11 vessels used to "explore, understand and predict" the ocean. Data management at Ifremer (SISMER) has 55 staff including 38 permanent positions and runs large computing facilities for data management accredited by NODC (IODE-XXIV) - oceanography, geophysics, geology, sea floor mapping etc. There is the goal to develop cooperation and have Ifremer as a core partner in ocean data management (including French NODC, Argo, SeaDataNet, Copernicus-CMES).

Discussion:

a) Catia: How does Ifremer work with universities? There are joint Ifremer/University institutes.

1.2. Invited Talk: CMEMS: Loïc Petit de la Villéon

Summary:
CMEMS (https://www.copernicus.eu/en, Copernicus Marine Environment Monitoring Service has both thematic (TAC) and monitoring centers and structure of these was discussed. TAC has been operational since 2015. Its focus is the aggregation of various datasets - bringing together existing datasets from various programs, to provide similar format,synchronization and cataloguing for 7 components (Global+6 regions) - NRT near-real_time and REP reprocessed from European and International in Situ programs (takes data from JComm, EuroGOOS). Close collaboration with and SeaSDataNet on QC. http://marine.institu.edu Focus onT&S but other properties too. CORA dataset = an objective analysis rerun every 6 months based on previous datasets UK and US counterparts (WOD). Ifremr video.

Discussion:

a) Discussion about the number of standard levels, how a newer version reduces the size and how there will be a simplified product that only provides the "good" data rather than all the flags. Response: CORA is moving to netCDF4 to reduce the file size. It's also considered to provide a simplified product only with good data, and regional products.
b) Reprocessing is it done for the whole period? Response: No, every 6 months new data are included and some updates are performed. The job is made easier when products come with QC. There are 2 types products - monitoring data and reanalyses

1.3. IQuOD Overview and Workshop Objectives: Matthew Palmer

Summary:
IQuOD's Mission is to maximize the quality, consistency and completeness of the long-term global subsurface ocean temperature data. There are 7 task teams: GDAC, Uncertainty, Auto QC/Duplicates, Formats, Intelligent Metadata, Expert QC/ Machine Learning, Metrics.
The idea is to take Input data sets -> Auto QC -> expert QC -> IQuOD (with version control): all steps are intended to include training and best practices. Progress since previous meeting includes v0.1 March 16 2018: Intelligent Metadata, and Assignment of uncertainties.
Now we want a v1.0 with: Auto QC , Duplicate checks, Additional documentation, and potentially incremental improvement for uncertainty estimates (UE), and incremental improvement for intelligent metadata (IM).
So we will discuss: formats & uncertainties, IM, Metrics today; Tomorrow: Auto QC and dups, Expert QC, global data acquisition center; Thursday next steps and timeline.

Discussion:

a) Note, machine learning QC input is much less than the 40% of all data originally thought and is more likely <%10 and perhaps even as low as 1%. The original number was larger because it methods considered were quite different. Now we know to avoid the edges of the Gaussian curves where data quality is low. We know and have always known that we cannot afford to have expert QC on everything. Action item: It would be worth sketching out a new conceptual model for QC flagging workflows.
b) It is important that the data produced by and processes performed by others are consistent and interoperable, so Matt is keen to learn how QC that has been done in other places and how it might apply here. Response: We should get to some of this during this meeting.
c) Are we planning to continue to salinity? Matt feels that IQuOD's expansion to salinity is a few years off, but others are already working on it as there is a lot of power in working in density space. IQuOD has been limited by lack of available support (not interest). The automatic QC procedures should work for salinity. Viktor says it can be applied to O2, nutrient etc using the experience gained from the temperature QC effort. There was also the need to continue to work on the historical data set to make it consistent with the modern data sets. Salinity is the logical next step, but the conversation/discussion to get the word out about what IQuOD has been done. Action item: Getting good Best Practices (part of the SCOR proposal) out there is as important as getting the dataset out there. How much would be needed in additional resources to expand to salinity and coordinating/approaching other communities.

2. Formats & Uncertainties Session

2.1. Machine Learning and Ocean Data Quality Control: Guillaume Maze

Summary:

Machine Learning (ML) is fitting mathematical model parameters on data to make predictions. Why and why now? It has never been so easy to do and we want the experts to focus on the harder problems. Can ML improve Argo QC? Input Data -> QC ->90/99% = validated which means only 10/1% input data that needs to be looked at more closely by experts. ISAS (https://www.umr-lops.fr/SNO-Argo/Products/ISAS-T-S-fields) is both software and product and it is an interpolation of Argo data. - training is local physical structure of the data to capture spikes and drift, predict alarm status for each data point. There were negative results: the accuracy of the alarm status predictor was too sensitive to the training set. There were no difficulties in identifying bad data, but there are N different kinds of good data (i.e. N different good models). It was determined that there is a need to train as much bad and good data and noted that structures are different in different parts of the ocean. There were also positive results: to effectively reduce operator workload - create a new QC test (MOCCA), improve/optimize parameters of existing QC tests, between combine sequence of existing QC test(IQuOD). Now, we need to: better understand what experts do, how do experts make decisions; to form a better definition of good vs bad data; and to better understand possible errors. Creating a new QC test - most of the tests look at a single data point at a time, but experts look up and down a profile and at the surrounding profiles - should develop tests for profiles and platforms rather than single data points (BAD profile can have as few as 1 BAD data point). Need: met-data, per-profile metrics (climatologies, deepest observation, surrounding profiles). Just using raw data/met data -> raise a lot of alarms, but if one uses all "features" + delayed and grid search the number of alarms can be reduced.

New approaches, two families of procedures (detecting bad measurements, detecting drifting sensors) have been developed. In theory, a bad measurement is when the sensor doesn't work. Requires a model of functioning sensor. Also, if you're using more than one sensor, one needs to know how/when they agree and/or disagree. To this end a deep Argo with 3 different CTDs is deployed to do such an intercomparison (check that data are compatible with a priori ocean knowledge: QC fails when reference is inappropriate, which occurs when local conditions change, e.g. sides of a front or island or temporally variable locations or when "if" test parameters are inappropriate).

a) How to select an appropriate data set (Maze et al., PO 2017 - selecting "classes" of data)
b) Use neural networks to combine all references and predict reference for a space/ltime location

When is a model good enough? If we compare to human performance, how do we quantify "human performance". Argo community QC challenge to see the spread in human performance. Noted that: validating measurements with a priori knowledge is the best way to bias the result.

Discussion:

a) How can we bring the operational machine learning techniques together? Auto detection vs/complements neural nets. Training a machine to look at all possible reasons for problems. Again the question of whether a group is working with just T or T & S. For machine learning you can have training data sets which represent different classes of problems. There should be an analytical model for each type of "bad" data (e.g. a spike equation in the conductivity cell) rather than trying to find "bad" data based on another set of parameters.
b) Viktor: Can reduce false flags by tuning for each grid cell.
c) Catia wondering if the meta data can also be QC'd - response yes.
d) Question about complexity - response that this one is quite simple.
e) What about coastal data sets - would this technique work - response no, it is more complex

2.2. Format Task Team Overview: Christine Coatanoan

Summary:
At the last IQuOD meeting, the work was to review and comment on the NetCDF files of the release 0.1, to review the Argo formats, to begin a draft document on format and create a repository on github. Some requirements have been looked (TS profiles, metadata, adjusted data, etc). From this meeting, the actions were to get input from IQuOD members and potential users on the files, to write a detailed draft document to be circulated for more feedback and produce drat data files to test usability. Unfortunately, little work has been realized. Nevertheless, the presentation has described in detail how the Argo data are managed, with a clear description of the different format files available on the GDAC. Moreover, an example of each Argo and IQuOD format have been presented showing the differences in the dimension of the data and variables definition.
The first question to answer is do we go with the current ragged array format served by WOD, which is a 1-dimensional ragged array format, or do we think we should change to an Argo-style/CCHDO-style structure, which is a 2-dimensional array representation? The easiest route might be to stay with the ragged array format because: it is already in use, it is supported and developed by WOD and presumably there are tools/documentation being built by NCEI for the ragged array format.
Putting a format out there, we will get feedback from the users. We just have to accept that in terms of file structure, it is quite different from the Argo format. If we go for a different format: the question is who has time and resources to develop it. Regardless of the format chosen, it will be necessary to describe the format more precisely (User's Manual as for Argo) and to propose broader tools for reading files (help to read file : not only fortran as for WOD but Matlab/Python/R/ODV). Tools and documentation are required. We have started to work on a IQuOD user's manual based on what has been done for Argo.
Another concern is what kind of level do we want : are we planning on serving the originator and the best copy data ('raw', 'adjusted' when available, both ?) Maybe we should just serve a 'best'/'adjusted' version with only IQuOD flags. This would keep the files smaller and simpler, ideal for users who just want to use the IQuOD product. Originator data is traceable through the WOD. Regarding the flag, we need to have better link with the AutoQC/EXpert QC teams, we also need to think about the error, the metadata (more). Do we need to have mandatory fields, in

this case we need a format checker. More for the future, another concern is about the future when we will be able to provide biogeochemistry data, do we need to go as Argo for separate B format?
The github is open for questions to be discussed all around the year:
https://github.com/IQuOD/Formats/issues

Discussion:

a) Formats - discussed Argo formats in relation to WOD formats -> for IQuOD would be easiest to use the ragged array format (Thomas Moore at CSIRO has also looked at this issue) but have to accept that the Argo and WOD formats are different..
b) What about the flags (link with AutoQC/ExpertQC, error, …)
c) Discussion concerning ragged arrays and the software for handling them. Depends on the end goal and particularly how many parameters are being discussed. Bill Mills is writing a reader in Python to help read the ragged arrays. Seems easy to handle for some people, but not others depending on understanding and tools.
d) Is Argo likely to use a ragged array format? Synthesised files contain all the data. But synthesised files are a gridded product, not observational levels.
e) Could have a layer on top that can provide the files in the format style that suits the user. Considered whether this should be an action item. Should think about a service layer that provides various formats like WOD - rather than changing formats.

2.3. OceanObs19 Uncertainties Session: Alison Macdonald

Summary:
This presentation reported on the OceanObs19 U.S. CLIVAR Phenomena, Observation and Synthesis (POS) Panel Breakout Session on Uncertainty Quantification (UQ) lead by Shane Elipot and Kyla Drushka. The POS panel sees UQ as a priority now due the current acceleration in the expansion of observing networks and data integration efforts, the proliferation of definitions for the term "uncertainty", and the lack of broadly accepted best practices. There were 4 panel members: Mark Bushnell (NOAA Center for Operational Oceanographic Products and Services/US IOOS) who focused on quality control of real-time oceanographic data and their best-practices; Chelle Gentemann (Earth and Space Research) who focused on aspects of satellite data integration, uncertainty estimation and reduction of redundancy of efforts; Robert Sabia (European Space Agency) who presented the Pilot Mission Exploitation Platform for Salinity (Pi-MEP/SMOS) as platform for satellite salinity validation, monitoring and integration with non-satellite data; and Patrick Heimbach (University of Texas/Odin Institute) who discussed UQ in the context of data assimilation. Each panelist provided a set of recommendations and the panel produced a broader set of overarching recommendations which included: a) training of ocean observers and modelers in statistical terminology and techniques for the purpose of uncertainty quantification; b) building on existing efforts to produce a series of peer-reviewed and open-access documents that define and recommend strategies and best practices for uncertainty quantification in ocean observing; c) pushing for research programs to require and fund routine uncertainty estimates on ocean observations and derived products, and should fund dedicated efforts to develop freely available resources (software & databases for UQ. It is important for IQuOD to get its results out there along with the IQuOD understanding/definition of uncertainty.

Discussion:

a) Discussion at Ocean Obs - are uncertainties replacing flags all together? We have had this discussion in IQuOD as well. No decisions made at Ocean Obs.
b) Any concrete plans out of Ocean Obs for uncertainties? Not likely - too high level.
c) US Clivar POS panel (https://usclivar.org/panels/pos) is very interested in facilitating this - also recognised at higher levels. IQuOD has experience to offer them at the moment.
d) From the user point of view, we want uncertainty with an explanation of how it was calculated. Then ultimately, we can get rid of the flags (*perhaps we should have BOTH*?). Information about the data is not attached to the data, it is within cruise reports etc (Sylvie)
e) Confusion about what uncertainty means. Need to be specific. Is the representivity error actually what is wanted? This was also discussed at Ocean Obs.
f) Marlos - The Ocean Obs section did not discuss the importance of the communication of uncertainty between peers and for societal needs. They only focused on quantification. Good communication of uncertainty is the most important part for decision making.
g) Ocean Best Practice is where we can provide a recommendation. Action Item: review paper looking at what is the best way to represent uncertainty - confidence, errors etc
h) Simona - modellers do not care about a specific number on a particular datapoint. In SeaDataNet, the expert provides the guidance on what to use. Note that the expert may only know about their data and not have knowledge of their data in relation to other observations. Which begs the question - which is the expert needed?
i) The XBT data corrections have many options still - which one to use? WOD should have the recommendation attached to it.
j) Matt - could have a range of options available to the user, representivity, error, confidence etc. Coffee-machine options - allow the user to download the data for their particular application. The download will only contain data of acceptable quality for the use at hand. The end user doesn't often understand the flags/uncertainties and we should make it easy for them. But we should also have the more complex options for the knowledgeable user.

LUNCH

2.4. Uncertainties task team overview & discussion: Rebecca Cowley

Summary:

The Uncertainties task team has not achieved much in the last few years. The goals therefore for the team remain:

1. Publish the source of uncertainties for IQuOD v0.1.

2. Publish Matlab code (in Github) (done). Usefulness of the code at this stage is questionable, but it is there if required, for future development for checking the application of uncertainties to IQuOD.
3. Update the uncertainties for the next release (or future releases) of IQuOD.
4. Seek user feedback (via github or the website?)
5. Investigate the relationship between QC flags and uncertainties.

Is there a statistical method or techniques we can use to estimate uncertainties for future releases.

Discussion:

a) Matt - would be good publish something on the v0.1 method (action item), even if short and based on the existing tables. Maybe a section on each instrument. The first thing should be an instrument error, there can be discussion of other components of uncertainty. Alison, John Gould and Viktor can assist with different instrument types. The idea is to get something out so other communities can see and provide input. And interesting things that have been found, e.g. poor data quality during WWII or the change in data quality after the introduction of standard seawater. It would be good to go beyond just a technical description of the data and explore some of the scientific issues - and capture some of the useful discussion that has been held on this topic over successive IQuOD workshops
b) Flags and uncertainties need to be consistent. Some time in the future the link/relationship between them needs to be defined.
c) Who is the user community and what do they need? Catia suggests talking with Patrick Heimbach from ECCO model reanalysis for feedback on uncertainty needs. Alison: This is a great idea because it underlines a connection between the US CLIVAR POS panel and IQuOD.

3. Imetadata Session

3.1. Imetadata Task Team Update and Discussion: Matt Palmer

Summary:

The IQuOD v0.1 data product was delivered in March 2018. This release included intelligent metadata probe and manufacturer assignments for individual profiles based on a decision tree approach that incorporates information on country code, maximum recorded depth and record date, as documented in Palmer et al (2018). This v0.1 iMeta algorithm has an average success rate of 77% based on evaluation of known XBT probes for the period 1966-2015. Leahy et al (2018) have applied a machine learning method to this classification problem using a neural network. They report a substantial improvement over the Palmer et al (2018) iMeta algorithm, with an average success rate of 90%, and substantial differences in the distribution of XBT probe assignments. Ongoing work at the UK Met Office has demonstrated further improvements over Leahy et al (2018) using a number of different machine learning approaches and the highest skill may be achieved using an ensemble approach. Ultimately, the plan for this work is to extract probabilistic information from the machine learning methods that could be used to develop

ensemble XBT bias corrections. Other outstanding areas of intelligent metadata that could usefully be pursued include estimating the XBT launch height and the data acquisition system – both of which are factors that affect XBT biases.

https://doi.org/10.1175/JTECH-D-18-0012.1 Leahy et al, 2018
https://journals.ametsoc.org/doi/10.1175/JTECH-D-17-0129.1 Palmer et al, 2018

Discussion:

a) Matt - there could be some work to be done on the distribution of data over the years and how it impacts on the heat content uncertainties (eg, during the early 1990 period where there is very little XBT data). Not sure if we should update the imeta for V1.0 with ML estimate.
b) Gui - likes the monte-carlo idea to give a probability, then that can help with applying fall rates and measuring impact on heat content estimates.
c) Viktor - the height of probe launch could be a feasible option for ML. Would be an interesting experiment that could lead to improved bias correction. Also talked about recorder types that might be a candidate for ML.
d) Matt would like to present his work to the XBT Science meeting next year as a means to starting dialogue/interactions with that group. Or work offline with XBT Science team. **(**action)
e) Do we need to get any of these ideas into the next release? Matt is happy to leave the imetadata as it is for now, but look at developing these ideas.It would be feasible to get something in on a 12 month timescale so this question should be revisited.
f) Rachel - have we done any sanity checks on the imeta that was applied? Not really, would be good to do this.
g) Bec asked about data rescue of XBT paper traces. Alison to send Bec a contact that might be able to use this as a project ideas. There is also software available (eg Un-Scan-It).
h) Steve Diggs leads data rescue group for CODATA (IDAR-TG) - happy to assist with promoting projects and even funding. Usually a small funding pool, but good to get into larger pools of funds.

4. Metrics session & XBT invited talks

4.1. Model evaluation and initialization: which datasets do modellers dream of?: Anne-Marie Treguier

Summary:

The numerical models considered here range from global forecast models (grid size of a few km, timescales of a week to a few years) to climate models (grid size from 20 to 200 km, decadal to centennial time scales). First, we note that initialization does not require new developments for recent ocean datasets. A good knowledge of ocean properties in 1850 would be useful to validate preindustrial control climate runs, but these simulations are very long (1000 years) so that their

initial conditions are forgotten. For forecast models, assimilation methods are generally more critical than dataset quality.
Regarding model evaluation, good climatologies are necessary and often sufficient to reveal model systematic errors. Ocean reanalyses suffer from model biases and are less suitable than observation-based products to evaluate numerical models. Modeller need derived quantities such as mixed layer depths or large-scale potential vorticity. A consistent global analysis of bottom properties, similar to the hydrobase analysis in the North Atlantic, would be useful because models tend to misrepresent dense overflows and bottom water masses. Modellers also need collocated datasets from different sources (satellite and in situ) for process based validation. Key examples are processes of ocean ventilation, carbon uptake, oceanic transports, or atmosphere-ice-ocean interactions.
As climate models enter the eddying regime, model evaluation faces new challenges. A number of recent studies emphasize the large imprint on ocean mesoscales on low-frequency, large scale variability in many regions of the world ocean (Penduff et al, Oceanography, 2018). In-situ observation of mesoscale dynamics is sketchy in the interior. Modellers need observation-based knowledge to identify relevant statistics or target regions for the validation of eddying climate models.

Discussion

a) Model evaluations are done with observational data and there are not many datasets available for these evaluations.The models are usually not good enough to worry about the detail of good/bad observations (i.e. model biases dominate in model-observation discrepancy).
b) Alison - if there are big biases in the observational atlases, doesn't this impact on the biases that you see in the model evaluations? Anne-Marie - the models are the problem, the impact from observations is not considered to be the problem.
c) Density measurements within 70m of the bottom from observations is very valuable. Would be a nice product to have. Not achievable from products like the WOD. WOCE climatology might be useful for this. https://icdc.cen.uni-hamburg.de/1/daten/ocean/woce-climatology.html
d) Modelers would like documentation of the variability at different time/space scales. Need to identify target regions for model validation.
e) Modelers need the data in a simple format. This is to create easy synergies between datasets of different sources.
f) Matt mentioned the possibility of characterising the phase and amplitude of the seasonal cycle from observations for a number of key variables. Ann Marie said that this could be useful.
g) (from later subsequent discussion) uncertainty estimates on ocean data could prove a useful basis for perturbing initial conditions and generating the forecast ensembles?

4.2. Metrics task team overview & discussion: Lijing Cheng & Mauro Cirano

Summary:
Actions not done yet, but will be attempted soon.

a) The CH14 XBT correction will be updated with the imetadata outcomes.

b) IQuOD data flags were not present in the IQuOD v0.1 release, Tim has updated with the WOD flags for now.

Steve - questions sent to Lijing and Mauro via email:

- **Question #1**: This endorsement from the GSOP. Can it be found online? Would you be so kind as to forward the official statement to the IQuOD Team?

Here is the report for the 10th session of GSOP:
http://www.clivar.org/sites/default/files/documents/GSOP10%20Report.pdf

- **Question #2:** So, the Brazilian Navy released all of their hydrographic data? What exactly does this mean? For instance, is there an inventory somewhere?

Regarding the Brazilian navy data, it is available for everyone. Mauro will put the data in a cloud storage, so that Tim can copy it and check which cruises are in WOD.

4.3. Invited talk: Corrections for systematic errors in temperature profiles from mechanical bathythermographs and satellite relayed data loggers: Viktor Gouretski & Lijing Cheng

Summary:
Temperature profiles from the World Ocean Database obtained by means of mechanical bathythermopgraphs (MBT) and satellite relayed data loggers (SRDL) are checked for possible systematic errors. The profiles are compared with the collocated reference unbiased temperature profiles from CTDs, Nansen casts, and Argo profiling floats to derive temperature and depth biases. The reference profiles were checked for consistency and showed a good agreement between the different instruments.

Three bias correction models were tested: D-Model: the time-dependent depth bias model; DT-Model: the time- and depth-dependent depth bias with a time-dependent thermal bias, and T-Model: the time- and depth-dependent temperature bias. Four metrics are used to characterize the performance of the bias models, with the best score achieved by the DT-Model for the MBT profiles and the T-Model for the SRDL profiles. The three existing bias corrections schemes by Ishii&Kimoto (2009), Gouretski&Reseghetti (2010), and Levitus et al. (2009) were shown to be less successful in bias reduction compared to the models suggested in this study.

The derived corrections undergone the robustness test, with half of the data being used as a training data set and the other half as a validation dataset. The test confirmed the robustness of the methods.

The MBT data are characterized by a positive temperature offset which is composed of the pure thermal bias and the temperature bias due to the prevailing depth overestimation. The biases are derived separately for five main contributing countries.

The SRDL exhibit different biases in the northern and southern hemispheres. The northern subset is characterized by a positive temperature bias, whereas the bias for the southern hemisphere profiles is rather small.

The corrected data will be used for the calculations of the ocean heat content

---

END OF DAY 1

Wednesday

4.4. Invited talk: AOML's XBT science contribution: Marlos Goes

Summary:
AOML/SIO XBT flow - XBT transect data to monitor Meridional Overturning Circulation (MOC) and boundary currents.

- Goal is to use update hydrography to infer salinity from temperature (Goes et al 2019)
- Using CORA3.4 (1990-2011) + NODC Argo(2012-2015) use 4 methods to create T-S Lookup produces velocity errors of about 6 cm/s and now has a global project - sees a future using machine learning.
- Using the XBT data see impacts of Gulf Stream on sea level on east coast US (Dong et al 2019) - no impact seal level increase from GS strength, SLR due to larg-scale warming

Goes et al (in prep) compares the density Argo with XBT profiles along 6 XBT transects by looking at different mapping in time and space. In addition, the estimates of boundary currents, meridional overturning circulation and meridional heat transport are investigated. For this comparison, the TS-lookup table is used to estimate salinity for XBTs, and the WOA13 climatology of temperature and salinity is used below 800m for both Argo and XBT.

The results from the comparison between Argo and XBT were: (Goes et al in prep)

- Looks at different mapping in time and space (GS example)
    - 2.5 days there are far many more XBT profiles than Argo
    - By 30 days, 3-5 degrees Argo catches up
    - Argo sees a blurred Gulf Stream and Brazil Current at 0.25-0.5 degree spatial mapping, and these currents practically disappear at 3-5 degree mapping.
    - At smaller scales XBT is necessary
    - XBT & Argo get different transport estimates because Argo misses the gradients
    - (AMOC/MHT) at lower resolution they match - 4 Sv, 0.5 PW error at the scales where they match because the estimates have lost the meso-scale signal.

In summary:
- XBT Data are unique and cannot be easily replaced by Argo
- Best quality, long-term time series for climate signals
- Reliable NRT XBT data with visual and auto-QC
- High quality controlled salinity data
- Accessible to users, research centers

Discussion:
a) Have you used the corrections for height for the different launchers? The deployment was from a Research Vessel, so all deployments were about 3 m height, therefore not important. Have corrected for the fall rate? All biases were estimated, but only shown depth offset here for comparison to Ann Thresher (2014) work. The Devil launcher

shown here is an early version which was not the same as the more recent one, but comparison is still valuable.

b) Is the lookup as good across all locations? It has really been tested along the transects. Salinity data would be limited along the land boundaries.
c) What effect does that have on the estimate of velocity? It varies with region.

4.5. Invited Talk: Uncertainties in XBT measurements and recording systems - work in progress: Franco Reseghetti

Summary:
There is a long history of the recording systems (RS) used - from analogic strip charts to LMC.16. The first popular digital RS were the MK9 and BathySystem SA 810. Those doing QC tend to consider mainly depth & temperature biases, but temperature bias also includes RS bias, which is important because there has been an improvement in RS over time. Moreover, equation that transforms resistance to a temperature has also changed over time and the difference among the results from different equations is larger at higher temperatures. Starting from middle of 90's, Sippican used Bennett's equation and its application to all earlier profiles from side-by-side comparisons improves the agreement as well as we found that the RS type influences the time dependency of the T bias. In a short time, we should have final results.
A comparison was performed between XBT profiles from SOOP after May 2004 vs. Argo in the Western Mediterranean. Specific constraints were applied to find matching pairs of XBT & Argo (avg. distance 9.4±3.1 km) and ± 1m in the vertical (within 1- and 7-day windows), with separated analyses to values < 100m and >100m depth. After July 2010, all XBT profiles were recorded after a check on the RS with a test canister: different RS showed slightly different results. Upper layer shows an even great disagreement between XBT and Argo values, mainly due to the summer thermocline, but below 100m the mean and median of (XBT-Argo) T differences are nearly coincident and small, (0.05 ± 0.12) °C.
Last, a comparison between Argo vs. CTD profiles from the Mediterranean was performed (only for profiles flagged as GOOD from both WOD and providers) using the same geographical constraints as for XBT and a 1-day window. There are two classes of Argo profiles: 199 Not-adjusted and 161 Adjusted profiles. Below 600 m, mean difference (Argo-CTD) both in T and S converges to very small values, close to -0.006 °C and -0.005 PSU, respectively. If a general uncertainty on T and S is applied to CTD values, Argo profiles are well consistent with CTD below that depth. The drift in time was also analysed (unfortunately, the most part of data is close to the Argo deployment): only non-adjusted Argo T profiles are statistically constant in time, otherwise there is a drift at a level of more than 10^{-6}/day.

Discussion:

a) Reprocessed the XBT data where available. He had the raw Sippican files with the resistance.
b) Shoichi: we have no information of the thermistor changes, so there is no reason to simply believe that the changes in the R-T conversion equations can be an explanation of T bias and its temporal change. Franco says the combination of thermistor and recording changes to produce changes over time.
c) Viktor - says their Argo-CTD differences are smaller because they have many 1000s of pairs.

d) Shoichi: What is the cause of T bias in the tester experiment? Self-heating? Seemingly, Franco had no idea about that.

4.6. Invited Talk: Leakage, revisited: a possible cause of pure warm-bias in XBT: Shoichi Kizu

Summary:
Many statistical studies showed that there is a "pure" temperature bias in the XBT profiles throughout its history, separated from depth-bias-oriented T bias. But no studies showed how it was caused, so the problem is not solved at all from an instrumental viewpoint. The presenter believes that the leakage from wire can be a good reason for such warm bias in the XBT temperature readings.
TOGA Lessons: A certain type of recorder, SEAS II (whose main part is the BathySystems SA-810 controller), was found to show sizable warm bias, often linearly increases with depth. More than 70% of total drops (over 1700) showed >0.05K positive bias, and more than 34% showed >0.15K positive bias, in excess of the claimed accuracy for XBT temperature. The task team investigated the problem and reached a conclusion that the cause of such T bias is in low-current applied by that system, and they recommended modification to the system to increase the current according to the Sippican's design. But the developer had a different view: low current is advantageous to prevent self-heating of the thermistor, and whether or not the distributed units were properly replaced or modified is not known. Similar phenomenon is found in Japanese system (Murayama Denki's DAS) combined with TSK-T7, when compared against CTD (plots showing a bias that often increases linearly with depth). Mixed Layer analysis revealed that the problem is confined to Murayama and not with TSK's MK-130.
Circuit of the XBT. Signal (change of resistance of thermistor) is measured by the potential difference between two legs in the Wheatstone bridge. Once the probe is sent into sea, two of the four legs are soaked. As far as the insulation is working well, this makes no problem. However, when the leakage happens, some portion of the soaked wire is exposed to seawater and makes parallel path to the soaked legs. The seawater is a good conductor, hence the seawater path makes the readings of thermistor resistance lower than true. Because the thermistor has negative-temperature-coefficient, the lower readings yield positive temperature bias (i.e. warm bias).
Quick visual inspection suggests that the problem is seemingly common in many SOOP XBT line transects (e.g. AX32, PX40, PX34, PX38, Mediterranean) realized by Sippican's XBT. Hard to identify without good reference, but caution is needed.

Discussion:

a) Matt: Do spikes indicate leakage? Response: Not always. There are various causes for spikes. The direction and/or "looking" of spikes often have mechanical meaning. CSIRO's "XBT Cookbook" is a good guide for learning how to interpret errors in temperature readings.
b) Viktor: How many profiles are affected by the leakage? Response: It is impossible to tell the statistics because so far the mixed layer is only the clue for making judges. In regions where salinity is effective in density stratification, like high latitudes and tropics where freshwater input is effective, for example, visual inspection is helpless.

c) Tim: Can we improve the XBT correction based on this information? Response: Not clear how to improve except for improving the wire quality. The clear message is that the problem is still with us. The presenter asked how the data originators think about such "suspicious" profiles, but generally their response is not good, and seemingly not even interested. Bec says there is a labor intensive way to find much of it. Matt says that it is a non-correctable error in some profiles (either so large you toss the profile or so small that it cannot be found). Franco - perhaps a check could be made and warning could be provided.

5. AutoQC & Duplicates Session

5.1. Invited Talk: ODV latest developments and QC procedures: Reiner Schlitzer

Summary:
Ocean Data View (ODV), the popular software for analysis and visualization of marine and other environmental data, is now available as version 5.2.0 (https://odv.awi.de/software/download/). Important new features include: (a) import support for WOD ragged-array netCDF files, (b) importer for CORA netCDF files, and (c) import support for Argo synthetic profile (s-profile) files. In addition, the Argo and WOD importers now extract many more meta data. ODV has a flexible duplicate station checker, and provides a variety of procedures for resolving duplicate station issues. ODV also provides a number of effective interactive workflows for automated and visual identification and flagging of outliers, ranging from single sample flagging to flagging of large number of samples in given stations or data windows in one go. Usage of derived variables, such as vertical derivatives or ratios, facilitates detection and flagging of spikes or anomalous biogeochemical data. Action item: The ODV team is interested in collaborating with IQuOD and will contribute resources if needed.

Discussion:
a) What tests do you use? Spike and range check, both trivial. Spike test is derivative-based.
b) Action item: Gui would like to work with Reiner's group on machine learning. Would be good to work together and develop tools
c) Simona: We would like to learn what the human impacts are on the machine learning outcomes. Reiner is skeptical, concluded that we need to incorporate the ML with the human. SeaDataNet has machine learning followed by visual inspection, would be good to do an assessment of the sensitivity of the impact of the 'expert' check. Does a different person impact the results of QC? They use already quality controlled profiles before the ML step.
d) Tim: How do we get the work that SeaDataNet has done on QC, and how can we use this in IQuOD and get the QC into IQuOD? Would the tools be useful for IQuOD? Could export ODV QC results into IQuOD in WOD.
e) The export format from ODV was built to suit the user. Might need to create a new output format to suit. Expert QC is valued very highly.

5.2. Task Team Update: AutoQC - IQuOD Duplicate Checking: Edward King and Rebecca Cowley

Summary:
The presentation given by Ann Thresher and Edward King in 2015 was presented. The WOD 2015 release was tested for duplicates using the following method, on HPC cloud and using Python scripting:

- Download the 10 large tarred and zipped netcdf files
- Decompress and split into smaller tarred and zipped files for easy access
- While decompressing, read the metadata and record to text files ordered by the number of points in the profile. Also exclude unrealistic points with a crude filter and sum the temperatures, depths, salinities. Record this information with the metadata for each file.
- Using the text files, look for pairs where the sums of the temperatures and depths are identical or within a set threshold.
- For identified potential pairs, open each file individually and sum the differences in the temperatures and depths to identify true duplicates

Next step requires that we look at the duplicates and make a decision about the next steps. It is difficult to make the decision for some pairs as they might have widely varying metadata.

Discussion:

a) Viktor: Where are most of the issues? Ann suggested in her talk that at Npts=1577 there were 800,000 duplicates. It would be good to analyse the pairs and figure out the percentages of different pair types (eg, exact, near in space/time, different data types etc).
b) Tim: although computing time is costly, the people time for handling the detail is more costly. How do we get rid of them within IQuOD when there are little differences until we determine the correct answer? Bec - for IQuOD we could choose and flag as bad, but include pointers back to the one chosen and why. Matt: we don't need to solve the problem, we just need a tractable reasonable step forward. Tim: thinks there are bad data that look like duplicates, but are not. Bec wants to redownload the data and see what WOD has already found. The are CTDs & bottle data which are duplicates, as well as MBTs and bottle data, which are meant to be duplicates.

Action item: to develop a set of procedures to apply to identified duplicates, including flagging system between identified duplicates. Check WOD (or IQuOD) for exact duplicates.

5.3. Task team update AutoQC: Simon Good, Bill Mills

Summary:
The current focus of the automatic quality control task team is to benchmark quality control tests in use around the world and determine which are the best to use. The resulting set of quality control tests will be used to supply the quality control decisions for the next release of IQuOD data.
A subset of the QuOTA dataset (Gronell and Wijffels, 2008), which has been subjected to scientific quality control as part of that study, is used as a reference dataset in this work. The quality control flags supplied with the data are assumed to be accurate and usable for benchmarking the performance of automatic tests. The quality control checks that are benchmarked (57 in total) have been rewritten in Python based on their original code or

documentation and are available from Github under an open source license. Software runs each of the quality control checks on each profile and benchmarking statistics are produced. Algorithms are run to select three groups of quality control tests that work in combination to give 1) decisions with a low probability of incorrectly flagging a clean profile, 2) decisions with a low probably of missing a profile with bad data in it, and 3) a compromise between the two. It is proposed that these three cases are used to generate quality control decisions and a different quality code assigned for each so that users can choose what level of flagging they require.

The results have been tested by subsetting the QuOTA data and looking at the variance in results. This suggests that the methods used are robust. The quality control has also been applied to four validation datasets. The results for some of these datasets are not consistent with the results for QuOTA, but this seems to be a feature of the datasets rather than an issue with the quality control tests combinations derived from the QuOTA data.

Discussion:

a) EN4 Also has a duplicate check (based on what Ann and Ed have done). Looking at both profiles flagged as but are actually good, and profiles flagged as good, but are actually bad. Simon has come up with some decisions about selecting the preferred profile in an identified duplicate. He can share this information (Action Item: Bec to follow up with Simon).
b) Gui: the idea has always been that everything (data, flags, software…) should be public. How do we make the links between the expert QC and auto QC? Matt: The three collections of tests would be applied and there would be flags 1 to 4. Good to bad. Shouldn't be a situation where there has been no QC.
c) Are there things that haven't been incorporated that should be in v1.0? Matt: feels that there will always be new tests but we cannot mark progress until get something out for the users. Tim: IQuOD should be state of the art, but we are missing Christine's climatology anomaly scheme and Jerome's min/max scheme. These should be incorporated into v1.0. Bill: if this is just a couple more qc tests, he can code them up quickly. Rachel has already been coding up some of the tests. Bill, Rachel, Jerome will talk about it and get the coding into the IQuOD GitHub repository.
d) Tim: another point about the flagging scheme, he thought we were going to include a reason for a data point failing (ie, the test that failed a data point). Simon says the tests are not run in any order, so we could definitely include this information about each test.
e) Now let's hear about the tests. The IGOSS flagging scheme should be used (Tim), as it gives a higher level flag, followed by a second level that indicates why a data point failed. Simona suggests having a list of failures/passes for each test - either a separate file or a byte for the failure. Some discussion on this topic - many opinions on bits and bytes.
f) The order of the tests being applied can give different results. At the moment, all the tests are applied independently. Therefore, the order is not important for this current method.

Action: to review the IGOSS flagging scheme and figure out how it can be implemented. Are there already existing maps between WOD and IGOSS flags and other schemes (eg GTSPP/Argo and IGOSS)?

LUNCH

6. Expert QC Session

6.1. Invited Talk: Requirements for Reanalyses: Andrea Storto

Summary:
Reanalyses have reached a good degree of maturity to be used for climate change and process-oriented multi-decadal investigations, and prove a valuable complementary tool to objective analyses and model simulations. In particular, it is argued that the multi-model ensemble approach is able to provide reliable and robust estimates for what concerns the ocean heat content and thermosteric sea level centennial evolution and variability. In order to advance reanalyses, a few recommendations/wishes for future potential consideration from the IQuOD community are given: i) consider the possibility of providing multiple (ensemble) realizations of the same observing dataset, to span the observation processing uncertainty and for use in ensemble reanalyses; ii) consider the possibility to provide the uncertainty linked to the processing of the observations (calibration, bias correction, fall rate correction, reliability of metadata, etc.); iii) consider the possibility to extend the dataset in near real time (around one month delay) to allow for seamless production of reanalyses and long-range predictions.

Discussion:

a) How can IQuOD benefit from the reanalysis community?
b) Simona - What is your definition of processing "error" uncertainty? The error due to the processing/metadata that happens between the observation and whatever ends up in IQuOD. Tim: How does it help to include that uncertainty as IQuOD also has an error assigned to it? Andrea thought that history had already been quantified - but if not, it may not be useful.
c) Matt: How big a difference does it make to the reanalysis? To Andrea - are you able to use the depth correction information (response - yes).

6.2. Invited Talk: QC based on local min/max values: Jerome Guourrion OceanScope

Summary:
Jerome uses Argo data to determine the min/max alpha values. The details are in Gourrion et al 2019 (submitted to JAOT).

Discussion:

a) Viktor - the ICDC uses a similar check (already included in the IQuOD benchmarking), but calls it something else and uses a climatology to get a longer comparison.
b) The reference fields are delayed mode data and the tests are applied to RT data. Some manual methods to set up the reference fields (some hand-QC required for 'bullets').
c) Uday: is there a seasonal mode taken into account? The fields are independent of time.

7. Expert QC Session

7.1. Expert QC Task Team update, Community quality control: Guilherme Castelão

Summary:
Given some questions raised earlier in the meeting, discussions from previous meetings were reviewed. To quality control oceanographic data is a classification type of problem in the Machine Learning perspective, which typically requires sufficient samples of each class being classified, but bad samples are typically scarce compared with good measurements. Such unbalanced problem requires special procedures for most of machine learning techniques, and can result in bad predictors even if good results were obtained from the training/testing datasets. Deep neural networks is an example of those sensitive techniques. In contrast to that, Anomaly Detection is based on the characteristics of the good data and the contrast with the bad data, therefore, an appropriate technique for an unbalanced problem. Another common limitation of QC techniques is on the independence of the tests, which is not how a human brain operates. An expert quality controller takes advantage of multiple information in a complex case. To mimic that, Anomaly Detection uses a multi-dimensional criteria, combining multiple aspects for an overall decision. A single dimension criterion is forced between minimize false positives or false negatives, while a multi-dimensional criteria allows more flexibility in the decision.
The web application used to collect experts evaluation was improved as well as the data pipeline. It now includes the official Argo data together with XBT, CTD, profilers, and gliders from WOD. As decided in the previous meeting, it has been running with ~30 oceanographers, where ~10 are actual 'professional' experts. A new version of CoTeDe with updated parameters will be released soon.

Discussion:

a) New users registered in the web application are exposed with cases considered easy, and as the system learns with the expert it moves to more complex cases.
b) Anyone can download the current version of CoTeDe, which is engine doing the data decision, for use with your own data and system, which has implicitly what was learned from the iteration through the web application.
c) The data manually flagged will be transmitted to Tim with the proper WOD id so it can be easily ingested back into WOD. Also, all the flagged measurements and flags themselves will be available for public use.
d) Guillaume Maze presented the intention to release next year a similar web application, and since IQuOD already has this operational, it was offered for the community to use to optimize the efforts.
e) Another way to move forward with the Machine Learning is by taking advantage of the already manually evaluated data, like from SeaDataNet, and this should be a goal for the near future.

7.2. Expert QC: Sensitivity Experiments for Polygon (Convex Hulls for Profiles) Checks: Udaya Bhaskar

Summary:
Initially the method was devised for picking up outlier in individual standard depth based on the convex hulls built using patterns observed when plotting temperature/salinity against longitude and latitude. The methodology was published in MethodsX journal. Even though this method is found to be good, its tedious when implemented for enormous amounts of data as multiple depths levels need to be catered to. In view this an extension to the data was tried out in the form of building Alpha convex hulls, where in the entire profile is used for building the convex hulls so that profiles can be directly classified as good or bad. The only hitch found in this extension is that of fixing the alpha values which is not unique for all months, seasons and annual data. Accordingly sensitivity experiments were done to arrive at a look up table of alpha values for each month, season and annual data which were used for performing quality control. The method an extension to the already published work is found to be worthy as whole profiles are being quality controlled. Results from the test cases (XCTD and Argo data) showed that the method is worthy of using by expert QC team for performing outlier analysis. Some more sensitivity analysis can be done to see if an optimal alpha which is common for all the months, season and annual data be obtained so that, the number of rejections of good data and inclusion of bad data be minimized.

Discussion:

a) Could be combined with Gui's approach. Use this to classify broad good/bad data before sending to the machine learning or auto qc.

END OF DAY 2

Thursday

8. Future View Session

8.1. Invited Talk: SeaDataNet: A Distributed marine data Infrastructure for in situ marine data management: Michele Fichaut (Project Coordinator)

Summary:
www.SeaDataNet.org is a distributed coalition of European data centers with specific interests in European Seas. SeaDataNet maintains a metadata directory with information about projects, organizations, cruises, etc. for European countries - with more than 11 data centers connected to this catalogue. The metadata formats are catalogued under ISO, and use NERC-BODC vocabularies. SeaDataNet provides software tools to generate metadata and data files, including file compliance checking. Beyond data and metadata management, SeaDataNet creates products based on the shared data. Ocean Data View (ODV) is utilized by SeaDataNet as a quality control tool. DIVA (Data Interpreting Variational Analysis; http://gher-diva.phys.ulg.ac.be/) is used to interpolate data to a regular grid. SeaDataNet has specific software tools for data centers and for users. Seatdatanet is moving to cloud HPC computing with EUDat (with variable update frequency). SeaDataNet follows FAIR principles (https://www.go-fair.org/fair-principles/) where possible.

Discussion:

a) T. Boyer asked about the SeaDataNet associated Caspian Sea portal. The portal is no longer maintained but the data should still be within the SeaDataNet.: G. Castelao asked whether all data have DOIs? They do not, but there are DOI's on aggregated data and planned to have on cruises. People should contact M. Fichaut for more information.

8.2. Invited Talk: SeaDataNet Products and IQuOD: Simona Simonocelli and WP11 team

Summary:
SeaDataNet, EMODnet, and Copernicus work in tandem to ensure European oceanographic data and metadata are gathered, aggregated, distributed, and incorporated into products. SeaDataNet performs quality control on the oceanographic data, reports problems back to the individual data centers and asks them to make a final determination and amend their data files if necessary. SeaDataNet employs a number of different methods for quality control, and uses Ocean Data View (ODV) as a visual quality control tool. Duplicate checking from different source data is done by SeaDataNet using ODV (by station) and DIVAnd (point-by-point). A detailed duplicate check between CORA (Coriolis Ocean Dataset for Reanalysis;

https://www.seanoe.org/data/00351/46219/), WOD (World Ocean Database; https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:NCEI-WOD), and direct SeaDataNet sources was performed with SeaDataNet the main source everywhere except the Arctic Ocean. SeaDataNet performs data and metadata rescue activities to restore data to a scientifically usable state. With the combined data set (see above) SeaDataNet uses DIVA to calculate climatological mean fields on decadal time scales on the same standard levels as the World Ocean Atlas (WOA; https://accession.nodc.noaa.gov/0176314). Comparison between SeaDataNet results and WOA help show the utility of regional climatology and quality control work. The DIVA user interface is presented in Jupyter notebooks. SeaDataNet products are issued DOIs, as are documented in Product Information Documentation (PIDoc) which are also given DOIs.

Discussion:

a) TVS U. Bhaskar - asked about the interpolation. DIVA has the capability to stabilize in the vertical and this may be applied in the future, but with care, as inversions in places like the Baltic Sea are not uncommon. J. Gurrion asked about residuals: There is ongoing work on trying to better understand residuals and to create fields for initial conditions.
b) C. Domingues asked about duplicate checking in SeaDataNet. Duplicates are sought with a space-time criteria using ODV algorithm and flagging. Decisions on duplicates have been made for the climatology only.
c) F. Reseghetti was concerned that Sippican MK12 have 4 different versions, but there is only label available through the vocabulary. In the Med, the height of the platform is also provided. T. Boyer commented that the work done comparing WOD and CORA and SeaDataNet should be loaded into IQuOD so it doesn't have to be done it again. There was discussion about the particulars of SeaDataNet licenses and DOIs and the wider issue of acknowledgement and downstream use of the data. G. Castelao noted that we need to plan for the future using a cascading DOI schema.

9. GDAC & Final Session

9.1. GDAC task team overview & discussion: Tim Boyer

Summary:
Version 0.1 of the IQuOD dataset was released in March, 2018. A DOI: https:/doi.org/10.7289/V51R6NSF was assigned. Landing page at the National Centers for Environmental Information (NCEI), https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:IQuODv0.1 has access to the IQuOD data set through ftp, https, and THREDDS. IQuOD data are also accessible through data.gov (https://catalog.data.gov/dataset/international-quality-controlled-ocean-database-iquod-version-0-1-aggregated-and-community-quaQuOD). IQuOD data files contain an uncertainty estimate on each profile measurement and iMeta algorithm v0.1 metatdata. In lieu of IQuOD Auto QC flags, World Ocean Database (WOD) flags are used. There was discussion at the previous meetings of using Coriolis objective analysis flags, but these were not included in IQuOD v0.1 (and probably will not be in the future as they have been superseded). The IQuOD dataset can be subsetted and downloaded through the WODselect tool

(https://www.nodc.noaa.gov/OC5/SELECT/dbsearch/dbsearch.html and http://wod.iode.org/SELECT/dbsearch/dbsearch.html). The IQuOD dataset is expected to be mirrored in France (T. Carval) and Japan (T. Suzuki). The IQuOD dataset is available in Climate-Forecast (CF) compliant netCDF ragged array format (https://www.nodc.noaa.gov/OC5/WOD/netcdf_descr.html). Specific readers are available in the latest versions of Ocean Data View (ODV, R. Schlitzer) and ERDDAP (B. Simmons). A reader is also available in FORTRAN and a Matlab reader is being tested and should be available soon. A python module is also in the works (B. Mills). ODV and ERDDAP have the ability to output the data in a variety of formats.

Discussion:

a) IQuOD has metadata space for position uncertainty, but this is not yet used. IQuOD can develop a position uncertainty based on measurement system (dead reckoning to Global Positioning System; GPS).
b) AutoQC modules need to be installed and operated at NCEI for loading and dissemination of IQuOD quality flags. A system for ingesting expert flag information needs to be standardized so the different groups doing expert qc can communicate their results to the main IQuOD data set.
c) S. Good suggests that IQuOD data be available in NEMO format.
d) T. Boyer suggested that IQuOD could extend intelligent metadata beyond XBT probe type; drop rate to other variables such as wire stretch.

9.2. Discussion on Roadmap for v1.0.

Summary:

Discuss and provide target dates for all activities below (see action list for resultant actions and due dates):

- Decide on and implement the duplicate check scheme
- Submit + Publish paper on the duplicate check scheme
- Decide on and finalise benchmarking the AutoQC tests (if new tests added)
- Agree flag definitions and secondary flag format
- Submit (+ publish?) paper on AutoQC and Duplicate algorithm
- Implement the AutoQC suite for use at NCEI
- Generate QC flags at NCEI and cross-check (perhaps subset of data?)
- Additional documentation needed for NCEI website (?)
- Publish the data! :)

Discussion:

a) A decision was made to use the AutoQC as defined in S. Good paper instead of attempting to add more tests to the benchmarking exercise. A concern that the min/max test is an improvement on the standard deviation test, but if they both exist at the same time a conflict arises. J. Gurrion noted that the min/max test does not cover all grid cells - where it does not the standard deviation check is necessary. There is a resourcing issue with adding more tests. There was further discussion about how adjusting the AutoQC tests shouldn't change too often as then every time it changes the expert quality control will be affected (it was agreed that while this is true, effect would be minimal).

b) There was discussion about the process of data passing and flagging by different stages of the QC process. It was agreed that expert QC is the best QC and does not get replaced by any new autoQC process subsequently run on the dataset.
c) It was agreed to hold quarterly meetings - short online progress meetings, and then in-depth discussions for particular topics.
d) Regarding next workshop timing, there is no more funding for any more meetings from SCOR, though there may still be some IODE funding. C. Domingues stated that IQuOD might also be able to build another SCOR working group proposal for Expert QC. Funding options will become apparent within 18 months. Hosting options in China (Lijing) and San Diego (Steve Diggs).
e) We are reminded to add the SCOR, IODE and CLIVAR acknowledgments to IQuOD publications and work.
f) There was a short discussion on the IQuOD website - idea from G. Castelao to list problems that IQuOD can address (eg, application of XBT fall rates). M. Palmer suggests that we put the website review on the regular meetings as a standing item. R. Cowley got no response when asking for help with website maintenance.
g) Steering team roles/expectations could be clarified- perhaps this needs review? C. Domingues noted there are two different definitions of the steering group for IQuOD (SCOR and IODE), with SCOR more restrictive. Some people have been added to the IODE steering group but not the SCOR steering group.
h) There was a discussion about CCHDO data - might need some care when taking flags from CCHDO data as 'expert QC' flags as not all data at CCHDO has received final primary investigator quality control. Also in order to incorporate CCHDO and other QC into IQuOD, we need to look at mapping between flag schemes (ask M. Hida at IMOS). It was suggested that maybe we need levels of expert QC flags (eg, some groups/people might be considered to produce higher quality QC than others). How do we decide this and do we keep a list of a priority? G. Castelao does not keep a list of the people's expertise. But he does have levels of QC interface (one for the experts). T.Suzuki: Primary level flag should be simple following IOC Manual and Guides No.54(3), similar ODV flag. Secondary flag can be designed by results of auto QC and expert QC and others.
i) The question was asked as to how to entrain experts to do the QC? Perhaps a presentation at EGU? S. Simonocelli is offering to take the IQuOD and compare it to the expert QC she has done in the Mediterranean Sea.
j) S. Simonocelli asked: Should an expert QC be forever? It was suggested that probably an iterative approach to revisiting already expert QC'd data is preferable.
k) S. Diggs, (remote) Can QC be harmonized with GLODAP QC?
l) A question was asked about training and community effort to get the data QCd (citizen science). No definitive answer was offered.
m) TDS U. Bhaskar asked about outreach - can we provide a few slides so that everyone can see. Action Item: provide a repository of slides for use by IQuOD representatives.

The final decision: the next meeting will be scheduled for approximately 18 months time - nominally in May 2021. IQuOD activities will continue to be coordinated through quarterly teleconferences and opportunistic interim meetings (e.g. XBT Science Team meeting in France during May 2020).

END OF DAY 3

10. Action Items

Task Team	Priority	Action	Who	Goal date	Status
Uncertainties	1	Write and publish paper for uncertainty/acccuracy/precision estimates released with IQuOD v0.1	Bec Cowley to lead. Alison Macdonald, Shoichi Kizu, Viktor Gouretski, John Gould and people in their networks. Others?	End 2020	
AutoQC, ExpertQC, GDAC	1	Develop workflows for AutoQC/Machine Learning/ExpertQC flagging	Tim, Bec, Gui, Uday, Simona, Christine, Simon	ASAP	
Duplicates	1	Re-run duplicate checking on WOD (or IQuOD?) latest release. Classify duplicate types and figure out percentages.	Edward, Bec	Early 2020	
Duplicates	2	Develop a QC flagging scheme for identifying duplicates in IQuOD v1.0. Develop a decision tree for different classifications of duplicates.	Edward, Bec, with input from Simon, Simona, Christine	Mid - 2020	
Duplicates	2	Submit + Publish paper on the duplicate algorithm	Edward, Bec, Simona	End 2020	
Duplicates	3	Investigate/develop the duplicate algorithm in conjunction with SeaDataNet procedures	Edward, Bec, Simona/Reiner, Christine, Franco		
Duplicates	1	Publish the code to github repository	Edward, Bill, Bec	Mid 2020	
AutoQC	1	Finalise benchmarking the AutoQC tests	Bill, Tim, Simon	End 2019(?)	
AutoQC	1	Submit (+ publish?) paper on AutoQC	Bill, Tim, Simon	Draft in January 2020 for comments	
AutoQC	1	Implement the AutoQC suite for use at NCEI	Bill, Tim, Simon	Mid-2020	
AutoQC	1	Generate QC flags at NCEI and cross-check (perhaps subset of data?)	Bill, Tim, Simon	Mid-2020	
AutoQC	1	Additional documentation needed for NCEI website (?)	Matt, Catia	End of 2020	
AutoQC	1	Publish the data!	Bill, Tim, Simon	End of 2020	

All	1	Overview paper on IQuOD and describes the data (eg ESSI, Nature Scientific Data)	Matt, Catia	End of 2020	
AutoQC/ExpertQC/Formats	1	Agree flag definitions and secondary flag format. Investigate using the IGOSS/IODE flag scheme and identify if there are any mapping tables available between different existing schema	Marty, Bec, Christine, Tim, Gui - others?	Mid-2020	
AutoQC	2	Code up the Ifremer min/max and improved spike tests	Bill, Rachel, Christine, Jerome	Mid-2020	
All	1	Set 3-monthly online meetings	All	Today	
Imetadata	2	Matt to attend XBT Science meeting if possible to liaise with team for ML uses in XBT metadata decisions	Matt, (Bec to try and assist with meeting timings)		
ExpertQC	2	ODV team to collaborate with IQuOD through the ExpertQC team.	Guilherme and Reiner	Next workshop	
All	2	Review and update the iquod.org webpages	All attendees at regular online meetings	3-monthly	
All	2	Review the roles of the IQuOD members and better define them.	All attendees at regular online meetings	3-monthly	
All	2	Present at US CLIVAR webinar	To be discussed	In 3-monthly meeting	
All	3	Rescue expert QCd databases and import the flags into IQuOD (eg Hydrobase, Quota, SeaDataNet)	To be discussed	In 3-monthly meeting	
All	2	Identify datasets in WOD that the originator flags can be imported as 'expert QC' level flags for v1.0	Bec (Australian data), Steve (CCHDO), others??	End 2020	
All	2	Identify the resources the task teams need to get the tasks completed. This is to help with funding proposals.	All task team leaders	In 3-monthly meeting	
All	2	Provide a stock set of IQuOD slides for use for outreach	Uday (?)	In 3-monthly meeting	

11. Appendix A. Participants

Participants

Last Name	First Name	Institute	Country	Email
Boyer	Timothy	NOAA/National Centers for Environmental Information	USA	tim.boyer@noaa.gov
Castelão	Guilherme	Scripps Institution of Oceanography	US	castelao@ucsd.edu
Coatanoan	Christine	Ifremer	France	christine.coatanoan@ifremer.fr
Cowley	Rebecca	CSIRO Oceans and Atmosphere	Australia	rebecca.cowley@csiro.au
Diggs	Stephen	Scripps Institution of Oceanography / UC San Diego	United States	sdiggs@ucsd.edu
Domingues	Catia	National Oceanography Centre	UK	catiad@noc.ac.uk
Fichaut	Michèle	Ifremer	France	
Goes	Marlos	NOAA/AOML	United States	marlos.goes@noaa.gov
Good	Simon	Met Office	UK	simon.good@metoffice.gov.uk
Gouretski	Viktor	Institute of Atmospheric Sciences, Chinese Academy of Sciences	China	viktor.gouretski@uni-hamburg.de
Gourrion	Jérôme	OceanScope	France	jerome.gourrion@ocean-scope.com
Han*	Chunhua	National Marine Data and Information Service	China	hanchunhua2008@126.com
Kizu	Shoichi	Tohoku University	Japan	kizu@tohoku.ac.jp
Kramp	Martin	JCOMMOPS	France	mkramp@jcommops.org
Petit de la Villéon	Loïc	Ifremer	France	Loic.Petit.de.la.Villeon@ifremer.fr
MAZE	Guillaume	IFREMER	France	gmaze@ifremer.fr
MIAO	Qingsheng	National Marine Data and Information Service	China	mqs1986865@163.com

Macdonald	Alison	WHOI	United States	amacdonald@whoi.edu
Palmer	Matthew	Met Office Hadley Centre	United Kingdom	matthew.palmer@metoffice.gov.uk
Reseghetti	Franco	ENEA	Italy	franco.reseghetti@enea.it
Simoncelli	Simona	INGV	Italy	simona.simoncelli@ingv.it
Suzuki	Toru	Marine Information Research Center	Japan	suzuki@mirc.jha.jp
Udaya Bhaskar	TVS	INCOIS	India	uday@incois.gov.in
Treguier	Anne Marie	IFREMER	France	anne-marie.treguier@univ-brest.fr
YANG*	YANG	National Marine Data and Information Service	China	yang03034101@126.com

* Participant was unable to attend

Remote Participants

Last Name	First Name	Institute	Country	Email
Cirano	Mauro	Federal University of Rio de Janeiro	Brazil	mauro.cirano@igeo.ufrj.br
Cheng	Lijing	Institute of Atmospheric Physics, Chinese Academy of Sciences	China	chenglij@mail.iap.ac.cn
Chu	Peter	Naval Postgraduate School, Monterey	United States	pcchu@nps.edu
Hidas	Marty	IMOS	Australia	marty.hidas@utas.edu.au
Killick	Rachel	Met Office Hadley Centre	United Kingdom	rachel.killick@metoffice.gov.uk
Mills	Bill		United States	mills.wj@gmail.com
Repetti	Luca	Istituto Idrografico della Marina	Italy	luca_repetti@marina.difesa.it
Schlitzer	Reiner	AWI	Germany	Reiner.Schlitzer@awi.de
Storto	Andrea	Centre for Maritime Research and Experimentation (CMRE)	Italy	andrea.storto@cmre.nato.int

12. Appendix B. Agenda

Tuesday 29 October		
Time	**Title**	**Who**
	Introduction Session	
	Chair	**Steve Diggs**
9:00	Opening of the meeting, welcome, logistics.	Loic Petit de la Villeon
9:20	Invited talk: The Copernicus Marine Service (CMEMS) and it in situ component. The in situ component ensures a consistent and reliable access to a range of in situ data for operational oceanography (service production and model validation)	Loic Petit de la Villeon on behalf of Mrs Sylvie Pouliquen (coordinator of the CMEMS IN SITU activities)
9:40	IQuOD overview - where are we now, objectives for the workshop	Catia Domingues/Matt Palmer
10:00-10:30	**Coffee break**	
	Formats & Uncertainties session	
	Chair	**Shoichi Kizu**
10:30	Invited talk: Machine learning	Guilliaume Maze
10:50	Formats task team overview & Discussions	Christine Coatanoan/Marty Hidas (remote)
11:10	Invited talk: Uncertainties and highlights from Ocean Obs	Alison Macdonald
12:30 - 2:00	**Lunch break**	
	Imetadata session	
	Chair	**Marlos Goes**
2:50	Uncertainties task team overview & discussion	Bec Cowley
2:00	Imetadata task team overview	Matt Palmer/Shoichi Kizu
2:20	Imetadata discussions	
3:00 - 3:30	**Coffee break**	
	Metrics session & XBT invited talks	
	Chair	**Alison Macdonald**
3:30	Invited talk: Model evaluation and initialization: which datasets do modellers dream of	Anne-Marie Treguier

3:50	Metrics task team overview & discussion	Lijing Cheng and Mauro Cirano
4:30	Invited talks: XBT science team - "Correction for systematic errors in temperature profiles from the mechanical bathythermographs and satellite-relayed data loggers"	Viktor Gouretski & Lijing Cheng
5:10	**Close**	

Wednesday 30 October		
Time	**Title**	**Who**
	XBT invited Talks (Continued)	
	Chair	**Steve Diggs**
9:00	Invited talks: XBT science team	Marlos Goes
9:30	Invited talks: XBT recorder biases	Franco Reseghetti
10:20 - 10:40	**Coffee break**	
	Auto QC & Duplicates	
	Chair	**Steve Diggs/Matt Palmer**
10:40	Invited Talks: XBT Science Team	Shoichi Kizu
11:00	Invited talk: ODV	Reiner Schlitzer (remote)
11:30	Task team update Duplicate Checking	Bec Cowley
11:50	Task team update AutoQC	Simon Good/ Bill Mills
1:00 - 2:00	**Lunch break**	
	Expert QC session	
	Chair	**Matt Palmer**
2:00	Invited talk: CLIVAR GSOP end user perspective - reanalyes community	Andrea Storto (remote)
2:20	Auto QC & duplicates discussions	
2:50	OceanScope QC test	Jerome Gourrion
3:20 - 3:40	**Coffee break**	
	Expert QC session continued	
	Chair	**Matt Palmer**

3:40	Expert QC task team overview & discussions	Gui Castelao
5:10	Sensitivity Experiments for Polygon (Convex Hulls for Profiles) Checks	Uday Bhaskar
5:40	**Close**	

Thursday 31 October		
Time	**Title**	**Who**
	Future view Session	
	Chair	**Bec Cowley**
9:00	Invited talk: SeaDataNet : a distributed Marine Data Infrastructure for the management of in situ marine data SeaDataNet is a distributed Marine Data Infrastructure for the management of large and diverse sets of data deriving from in situ of the seas and oceans.	Mrs Michèle Fichaut -Project coordinator
9:20	Invited talk: Seadatanet and IQuOD	Simona Simonocelli
10:00-10:30	**Coffee break**	
	GDAC & Final session	
	Chair	**Guilherme Castelão**
10:30	GDAC task team overview & discussion	Tim Boyer
11:00	Goals to v1.0 release	
	Review of current meeting actions	
	Wrap up, next meeting	
12:30 - 2:00	**Close + Lunch**	
2:00 - 17:00	IODE report and Workshop Report writing (for those staying and willing to help)	Tim, Bec, Alison and Gui