

IMOS QC Summit

CSIRO, Hobart, 29 & 30 August 2012

Actions, Recommendations and Minutes

Actions

- Send a USB with all talks to Marian Wiltshire to load onto ANMN IMOS webpage (**Tim Lynch**)
- Post all Calibration data to Marty Hidas at eMii (**Robert Kay**)
- Post all Calibration data to deployment data bases (**Robert Kay**)
- Send FLNTU Cap method to Marian W to load onto ANMN IMOS web pages (**Robert Kay**)
- Send link to technical forum to Marian Wiltshire (**Daniel Smith**)
- Send pre-deployment check lists send to Marian Wiltshire for posting onto the IMOS ANMN webpage (**Craig Steinberg and Dave Hughes**)
- Develop scoping document for Tim M for QC procedures methods/technical papers/manuals - includes regional range table (**Tim Lynch, Guillaume Galibert, Lesley Clementson**)
- Develop an ANFOG – ANMN regional range table Chl a (**Christine Hanson, Lesley Clementson**)
- Automatic QC needs testing (especially Chl a) hence provide a Maria Island NRS (QC'ed) dataset for Guillaume Galibert to model (**Tim Lynch, Lesley Clementson, Daniel Smith, Ken Ridgway, Karen Evans, Bee Morello**)
- Collect rules for expert/scientific QC cross validation method (**Ken Ridgway**)
- Manifesto for data labelling easily available and obvious at eMii site (**Marty Hidas**)
- Flags are sorted and reported (flags sorted, flags passed, flags failed) (**Guillaume Galibert**)
- Work with central processing labs to roll out QC of BGC metadata system (**Claire Davies**)

Recommendations

- Calibrate instruments every 12 months
- Rationalise QC teams by parameter - Develop joint facility focus on QC of particular parameters e.g. FLNTU: ANFOG and ANMN, S & T: ARGO, ADCP: ABOS – Rebecca Cowley
- Meta data server for deployment database – web accessible
- Bio-optics test for spike data (myocean) is not appropriately scaled and should be re-parameterised based on the regional ranges
- Additional step for QC possible in naming convention e.g. QC0 – raw, QC1 – automatic/technical, QC3 – expert/scientific
- Also note reprocessing of data in naming convention
- Data streams should be aggregated to detect sensor calibration issues – each new deployment is added to the previous by the facility.
- Automatic QC is developed as a consistent policy around the 9 tests developed for the Matlab toolbox. This policy is explained in Morello et al 2011 and will be expanded upon in the following journal publication.
- We accepted that the general rule: “the greater period between collection and QC more likely data will degrade”.
- Cross validate at all sites.

Minutes

Comments from the Chair

- It was apparent that we have made much progress since the last summit and, from our work with international collaborators, it seems we are on the leading edge of this field.
- QC is labour intensive and difficult. To get the most efficient use of our resources it would be ideal if we work with sister facilities (ANFOG, ABOS, ARGO) to rationalise our QC systems based on common parameters.
- We need to be completely transparent in what we have done for QC and make this easy for people to understand our processes and steps. We also need to provide completely unprocessed data so peers can review our work, undertake their own QC as they see fit and to assure that we do not ever remove or 'throw out' data.

Automatic QC

- Both AIMS and ICT appear to have developed similar systems for real time QC. In the ICT case this has been developed around a maths framework while the AIMS system is intuitively rule based. Interestingly, this seems to have produced similar weighted system for assessing data. AIMS has developed this QC system based on their metrological data while ICT has based theirs on in water sensor data.
- A further potential test for Automatic QC that was suggested was a thrashing test. This test detects wide changes in values across small time scales. Based on the known failure mode of the FLNTU of macro-algae fouling of frames that hold the instrument and passing in front of the sensor, this may be a useful test to develop.

Sensor drift and aggregation

- How and where is sensor drift accounted for? While this issue is often raised it may not be of high consequence for ANMN data streams. Sensors deployed for ANMN are rotated through the calibration laboratory which will detect sensor drift. This is a more pressing issue for ARGO, where sensors are deployed until they, or the platform, fails.
- A larger issue for ANMN data is differences in precision or accuracy between individual sensors that are deployed at the same location. Aggregation of data is a good method for detecting these issues and should be a priority for eMii and the facility.
- Aggregation is also an issue with the NRS BGC data. At the moment it is not user friendly, for instance 280 downloads are required to collect data from all NRS stations.

Cross validation – expert/scientific QC

- The heart of the expert or scientific QC was cross validation of data from independent sources. Ken Ridgeway provided useful examples of cross validation of the Maria Island NRS dataset.
- A potential issue with cross validation using Niskin bottles samples adjacent to the WQM was expressed. Density or temperature are better parameters to line up rather than pressure due to thermocline heave. Really cross validation needs smart Niskin bottles that know the temp/density at the moment of the sample.

Metadata

- Meta data is critical for QC, allowing for time calculations and cross validation. However metadata itself also needs to be QC. For instance tests such as impossible date and location need to be conducted as these issues have been identified by users of the data. Metadata is also not easily searchable and a central store for metadata would be ideal.
- The 9 x QC tests developed for QC of sensor data need to be applied were applicable to the BGC metadata.