Report from the COL-NASA Data QA/QC Workshop
6-8 June 2012
University of Maine Ira C. Darling Marine Center

Summarized by Emmanuel Boss[1], Merrie Beth Neely[2] and Jeremy Werdell[3] and reviewed by all the participants

[1] **University of Maine, School of Marine Sciences, Orono, ME**
[2] **Consortium for Ocean Leadership, OOI, Washington, DC**
[3] **NASA Goddard Space Flight Center, Greenbelt, MD**

Contents

The Consortium for Ocean Leadership (COL, the entity contracted to administer the NSF-funded Ocean Observatories Initiative, OOI) and NASA co-sponsored an expert workshop on data QA/QC convened from 6-8 June 2012 at the University of Maine, Ira C. Darling Marine Center.

In this document, we present summaries of the discussions conducted during the workshop. While we do not believe this document to be exhaustive, we feel it presents an honest attempt to provide the sponsors with the state-of-the-art knowledge associated with quality control of observatory data.

I. Introduction to the workshop

A team of subject matter experts (SMEs) was assembled (Appendix 2) with broad experience with bio-optical sensors, observing systems, various oceanographic sensors and sampling platforms, observatory science, and automated and human-in-the-loop (HITL) data QA/QC procedures. The SMEs met in both small and large groups for two and half days to discuss a variety of specific questions associated with the workshop goals. The workshop goals were divided into those of interest for each co-sponsor:

NASA Goals:

- Interact with OOI and subject matter experts in the community to improve upon our collective ability to QA/QC in-situ bio-optical data. NASA acquires data from many sources, making high quality QA/QC metrics essential.
- Consider OOI as a future source of in-situ data for NASA ocean color calibration, validation, and bio-optical algorithm development activities. NASA needs (will need) an abundance of high quality data from diverse regions in support of current (upcoming) missions (e.g., VIIRS and PACE).

OOI Goals:

- Review the OOI Data QA/QC Plan.
- Discuss state-of-the-art approaches for QA/QC and emerging QA/QC methods.
- Recommend modifications to the existing QA/QC Plan that will ensure high quality ocean observations for the planned life of the system.

The co-sponsors also presented questions to be addressed during the event:

NASA Questions:

- What are the similarities/differences in QA/QC approaches for each measurement type? Do the similarities imply best practices and/or endorsement of the protocols? What are the lessons learned, particularly with regards to the differences in approaches?
- On an instrument, measurement, or infrastructure level, what is the state of the art? What are the outstanding issues that need to be resolved? Is any new technological development required? Are the issues technologically or manpower/resources limited or due to individual historical practices, without community consensus?

- With regards to synergy between OOI and its end users (e.g., NASA), what needs to be done to ensure a seamless transfer of information and high quality data?

OOI Questions:

- What are the similarities/differences in approaches and the best practices and lessons learned?
- What are the problem areas that the observatory community needs to concentrate on solving? List of must do's in data QA/QC and should do's (given sufficient resources). What additional steps, not yet considered for the OOI, should be added to the QA/QC procedures?

II. Recommendations for data QA/QC

The following recommendations do not represent an exhaustive survey of all that is known on QA/QC of in-situ oceanographic data records, but rather, provide a distilled summary of the knowledge of the assembled experts. Note that the recommendations are not inclusive of all OOI sensors and are skewed towards bio-optical sensors. This skew resulted from both the role of NASA as a co-sponsor and the fact that data from present-day bio-optical sensors present significant QA/QC challenges. Many of the data QA/QC issues addressed, however, can be extrapolated to other sensors and instruments in terms of protocols and lessons learned. The recommendations provided below represent lessons learned in a variety of observatories and observing systems, spanning different scales, among them: IOOS, MVCO, GoMOOS, ARGO, IMOS, QARTOD, CDIP, BOUSSOLE, NASA, and LOBO.

The assembled SMEs divided the tasks associated with data QA/QC into three periods associated with the data collection: (A) *pre-deployment* – actions taken before instrument deployment; (B) *real time* – diagnostic actions taken during data collection; and, (C) *post-deployment* – retrospective actions taken following data collection.

Climate quality data

The recommendations provided here are designed to ensure that, where required, OOI assets collect *climate quality data*. Reliable climate monitoring and assessment of climate change are exacting disciplines requiring a high level of accuracy and reliability in their data sets. In order to detect climate change with certainty, careful data quality control is essential. For example, one of the more pressing needs from a climate monitoring perspective is the identification and correction of time-dependent data biases in observation systems as early as possible. This is a fundamental aspect of scientific data stewardship that depends on a high level of training and expertise of the data analyst. These biases are sometimes due to straightforward sensor degradation, but just as frequently result from changes in algorithms or spatial and temporal sampling methods that yield unintended consequences. Although the concept of *climate quality data* is frequently noted in one form or another (e.g., WOCE, 1988; CLIVAR, 2012], a simple all-encompassing definition is almost impossible to prescribe because the accuracy of measurement of any specific parameter will depend on the science question, on whether the measurements are of physical, chemical or biological properties, on instrument technology, and on location and depth. Here, the term *climate quality data* is used for those data that constitute a time series of measurements of sufficient length, consistency, continuity, and accuracy to determine climate variability and change, given the known or expected variability of the parameter being measured (NRC, 2004).

A. Pre-deployment

Gathering high quality data starts with pre-deployment efforts. As technicians dealing with sensors are often not experts and manufacturer user manuals are often incomplete in describing the latest best management practices, highly detailed manuals must be developed which spell out the specifics of every step associated with sensors, their calibration and field tests, best practices and standardization across the program, and the rationale for such actions. Manuals should be approved by SMEs and include laboratory and field protocols.

Calibration

Throughout the OOI, calibration practices should employ best practices, minimize inter-laboratory inconsistencies, and ensure that calibrated sensors have known accuracies. This is especially important for relatively novel sensors (e.g., some bio-optical and chemical sensors). On-site calibrations are required for instruments that can be affected by shipping (e.g. beam transmissometers) and whose dark currents might be affected by their deployment systems/platforms (e.g. backscattering sensors). Cross-checking with multiple sensors of the same type during deployments will add confidence regarding the accuracy of the sensors. Inter-calibration with well characterized and calibrated sensors from manufacturers other than those used by OOI, that measure the same or similar parameters as the OOI sensor, will provide additional verification of data quality and accuracy.

Biofouling prevention or minimization

Techniques exist for mitigation of fouling that require specific actions prior to deployment; Delauney et al. (2010) provide a recent review of conventional methods and Manov et al. (2004) discuss specific techniques for bio-optical sensors. For example, they suggest that open-face sensors be positioned to minimize particles settling on the face (i.e., not act as particle collectors) and that upward-facing sensors (e.g., radiometers sensing downward light) be positioned such that the CTD pump outflow washes off the optics windows (via water jetting on their surfaces, being careful that the radiometers are not shaded by other instruments or parts of the platform) or have a wiper.

B. Realtime QC

Flagging

Both automated and human-in-the-loop (HITL) methods can be used to flag data in quasi-real time to alert potential users that the accuracy is suspect. This process is referred to as ‘flagging’ of the data.

Cross calibration

Inter-comparisons of sensors with similar sensors on gliders and ships visiting the site mid-deployment provide an opportunity to verify data consistency across OOI platforms (e.g. when the OOI gliders are used) and provide an independent source for assessment of data quality (when freshly calibrated non-OOI sensors are used for cross-check). These cross calibrations should be carried out for more than a single profile.

Biofouling recognition

Sensor fouling affects optical data by causing an exponential or other rapid increase in signal (e.g. in fluorometer, transmissometers, e.g. see Delauney et al., 2010). Immediate recognition of such trends in real-time data streams will result in these values being flagged as suspect until HITL evaluation overrides this designation or provides corrected values. Differential changes in duplicate sensors would indicate biofouling (or sensor drift). However, detecting such trends against the natural background variability remains a technical challenge.

Biofouling mitigation

Replacement/calibration/diver-assisted cleaning of bio-optical sensors during deployment may be necessary to ensure data quality from sensors that are otherwise unlikely to provide usable data throughout the deployment. This requires abundant pre-cleaned and post-cleaned measurements be collected for comparison and quantification of before any corrections can be applied.

Use of non-OOI assets to enhance the ability of OOI to recognize suspect data

Remotely sensed data products (e.g., SSH, SST, Ocean color, SAR, winds) and assets associated with nearby observations (e.g., Argo floats) should be provided in parallel to OOI data streams (e.g., on the same website) to allow for rapid HITL and/or automated comparisons (of both absolute magnitudes and trends). Similarly, products from operational global circulation models (e.g., FNOC, MERCATOR) near OOI asset locations should be provided with OOI data streams to provide boundaries for common magnitudes and trends of physical properties. Studies of these assets used should be performed *a priori*, as should studies of temporal and spatial scales over which the assets are expected in order to provide constraints on likely in-situ values.

C. Post deployment QC

Calibration

Pre-recovery evaluation of sensors is critical to assess the potential for sensor drift during deployment and to provide avenues for correcting sensor drift (when possible). Corrections should be applied following this series of steps, which is designed to establish whether the sensor drifted and/or became fouled during its deployment. Sensors should be cross-calibrated against a freshly calibrated, unfouled sensor while still collecting data on the OOI asset (that is, before retrieval of the sensor) or through a controlled procedure that ensures that the cross-calibration occurs while the sensor is still in the same state as it was at recovery) or against physical samples collected at a range of depths at the site (The physical samples are especially recommended for chemical sensors. For example, the ISUS offset term is determined from physical samples of NO3 collected at deployment and retrieval). After the sensor is recovered, cleaned and disassembled, but prior to shipping it back to the calibration facility, additional laboratory calibrations should be conducted. This is particularly relevant for bio-optical sensors, in order to quantify and separate instrument drift from biofouling and to identify errors that might result from shipment to the manufacturer. Finally, the manufacturer should conduct similar or identical tests prior to recalibrating the instrument. The latter should occur at the same facility in which the initial calibration was completed. For instruments other than bio-optical ones, it may suffice to combine the latter two steps into one.

Human-in-the-loop (HITL) data QC

Whenever expanded sensor networks are deployed in the ocean, especially ones which are able to measure unexplored domains of the time-space variability of the ocean, new features and phenomena are seen. It should be expected that data collected by OOI will reveal details and events never seen before, and it is important to discriminate truly bad data from unusual phenomena never previously observed. This will require HITL procedures until enough experience has been accumulated to automate the procedures. In addition, simple algorithms like historical-range checks are unable to detect small drifts or outliers, which may be severe enough to adversely impact science analyses. Once data is flagged 'bad' during the data QC process, it is rare that a scientist will ever look at those data again. It is therefore critical that personnel involved in HITL data QC have experience in both engineering and observational oceanography.

OOI should anticipate the need for inter-comparison exercises to ensure the HITL component of delayed-mode data quality control is standardized across different operators. Somewhat by definition, the requirement of HITL QC is inherently subjective. Operators with different expertise and experience may make different decisions regarding data quality. Additionally, different types of scientific studies/calculations may have different demands in terms of what qualifies as good data, making documentation of technician training protocols and instrument accuracy and precision metrics important metadata to provide to users. Periodic meetings are necessary to ensure that all the QC technicians are making similar decisions. A recommended protocol to model would be the Delayed-mode Argo Data Comparison (Wijffels and Tchen, 2007, and the Argo Quality Control Manual).

Use of nearby OOI assets

Some data measured with different sensors are expected to co-vary with the relationship between them being bounded (e.g. backscattering, beam attenuation). Deviation from such relationships should help recognize suspect data.

Use of non-OOI assets to enhance the ability of OOI to recognize suspect data and be used in support of specific mitigation strategies

Remotely sensed data products (e.g., SSH, SST, ocean color, SAR, winds) and observations from nearby assets (e.g., Argo floats) could be used to identify potentially bad values and/or to adjust these values post deployment, so long as the accuracies of these assets are sufficient for this task. Data products from operational global circulation (FNOC, MERCATOR) models near an OOI asset could be used in a similar way. Prior to the use of such assets to flag/correct/adjust data, studies of the accuracy and spatial resolution provided by these assets should be performed, as should studies of temporal and spatial scales over which the assets are expected to provide constraints on likely in-situ values.

III. Additional recommendations and comments

Science teams and SME roles and responsibilities

Workshop participants recommended the formation of science teams for each OOI sensor type and measurement suite. They further suggested that SMEs be actively involved in data QC/QA to ensure the formation of the best possible data streams. Funding mechanisms for small-scale studies and science team formation will ensure high level of involvement by external SMEs in protocol development and addressing specific problems (e.g. improved method for biofouling, testing of different QC approaches) in a timely manner. Involvement of SMEs from outside the OOI is encouraged, to enlarge the circle of people in the community who contribute and oversee the processing of OOI data.

Operational budgets for QA/QC

During the workshop several participants associated with operational observatories shared their experience regarding data QA/QC efforts. In the Australian IMOS, an observatory spanning the coastal waters of Australia, about 30% of the salary budget (~10-15% of overall budget) is devoted to salaries of staff associated with data QA/QC. This large percentage is partly due to a series of laboratory activities designed to characterize commercial sensors during the current early phase of the observatory. In the US-Argo float network, a relatively mature (~10 years of operation) and standardized observatory, approximately five full-time equivalents (FTEs) deal with the real-time QA/QC of 150 CTD profiles per day (~1,500 US floats total) with an additional four FTEs devoted to delayed-mode QC and calibration. This accounts for about 12% of the overall budget, but is difficult to isolate the cost of QA/QC because the same technicians are also responsible for the bulk of data management. Approximately 5% of the CDIP total budget encompasses the steps from calibration through data processing (e.g., looking at the signal and archiving the signal), while the remaining 10% includes the temporal and geospatial QA/QC of processed data. Four FTE programmers and one half-time FTE data analyst perform HITL data QA/QC for CDIP. Finally, for the NASA SeaWiFS Bio-optical Archive and Storage System (SeaBASS; ~15 years of operation), one full-time FTE performs data QA/QC on in-situ data collected under the auspices of the NASA Ocean Biology and Biogeochemistry Program.

Based on the budget allocation for data QA/QC from these existing observatories, the participating SMEs concluded that an investment on the order of 20 FTEs would be needed for OOI QA/QC activities (771 individual sensors of 47 different kinds) during the first five years of the project. The sensors that measure hydrographic properties and currents, which are mature data streams, will require approximately one-third of the allotted FTEs. The remainder will be equally divided between chemical sensors and bio-optical sensors, for which the oceanographic community has significantly less experience for long-term deployments. As the observatory matures, it is likely that the number of FTEs will decrease, as automation of data QC procedures will replace HITL procedures for certain activities. It is clear to all participants, given the experience with mature observatories, that HITL procedures will always be part of the OOI data QC effort.

IV. References

Argo quality control manual. Version 2.7. Argo data management, 3 January 2012. Available online at: http://www.argodatamgt.org/content/download/341/2650/file/argo-quality-control-manual-V2.7.pdf

CLIVAR, 2012: CLIVAR Data Policy. Available online at: http://www.clivar.org/resources/data/clivar-data-policy

Delauney, L., Compère, C., and Lehaitre, M., 2010. Biofouling protection for marine environmental sensors, Ocean Sci., 6, 503-511, doi:10.5194/os-6-503-2010.

National Research Council, 2004. Climate Data Records from Environmental Satellites: Interim Report. Washington, D.C. National Academies Press.

Manov, D. V., Chang, G. C., and Dickey, T. D., 2004. Methods for reducing biofouling of moored optical sensors, J. Atmos. Ocean. Tech., 21, 6, 958–968.

Wijffels, S., and T. Tchen, 2007: Delayed-mode data intercomparison. CSIRO Marine and Atmospheric Research Prepared for AST-8, 59 pp. Available online at http://prelude.ocean.washington.edu/dmqc3/pub/wijffels_dmqc_intercomp_report.pdf and http://www.argodatamgt.org/Media/Argo-Data-Management/Argo-documentation/General-documentation/Reports/2006-DMQC-exercise]

WOCE, 1988: World Ocean Circulation Implementation Plan, Vol. I, Detailed Requirements. WCRP-11 (WMO/TD - No. 242), July, 1988.

Appendix A

Data QC Workshop Agenda
June 6-8, 2012
Darling Center, University of Maine

Wednesday, June 6

8:00 AM Continental Breakfast

8:45 AM Introduction, overview of workshop objectives (Cowles, Neely and Boss)

9:00 AM Overview of OOI (Daly)

9:15 AM Presentation/Discussion of NASA needs (Werdell)

9:45 AM Presentation/Discussion of OOI Data QC Plan (Lankhorst)

10:00 AM Overview of IOOS, CDIP approach (Thomas)

10:30 AM Break

11:00 AM Overview of Argo approach (Robbins)

11:30 AM Overview of IMOS approach (Lynch)

12:00 PM Working lunch

1:00 PM Overview of BOUSSOLE approach (Antoine)

1:30 PM Field Examples and Instrument-specific considerations (Perry/Boss)

2:30 PM Break

3:00 PM Plenary discussion

- What are similarities/discrepancies between OOI and other approaches?
    - Are there areas where OOI plans are **not** in alignment with community expectations?
- Identify the problem areas, based on families of sensors, that the OOI, NASA and the entire ocean observation community need to concentrate on

5:00 PM Adjourn for day

5:30 PM Travel to working dinner offsite, assemble for shuttle at designated spot

Thursday, June 7

8:00 AM Continental Breakfast

8:45 AM Charge for the day (Neely and Boss)

9:00 AM Technical Breakouts

- System-level QC approaches (as opposed to sensor-level)
- Sensor inter-calibration/verification
- Utilizing external (field or remote) observations
- Detection of slow drifts or small biases against large background variability

10:30 AM Break

11:00 AM Report out and discussion from breakout sessions (15 minutes each)

12:00 PM Working Lunch

1:00 PM Technical/Management Breakouts

- At-sea and instrument inter- calibration
- Bio-fouling: susceptibility, mitigation strategies, data QC checks to identify bio-fouling effects, possible corrections
- Centralized versus distributed approaches
- Levels of QC and associated effort (workforce) required

2:30 PM Break

3:00 PM Report out and discussion from all breakout sessions

5:00 PM Adjourn for day

5:30 PM Working dinner onsite

After dinner Reps for OOI and NASA meet to summarize key points needed for white paper

Friday, June 8

8:00 AM Continental Breakfast

8:45 AM Charge for the morning (Neely and Boss)

9:00 AM OOI and NASA summary of key points and recommendations of workshop

9:30 AM Lessons learned and whitepaper discussion

10:30 AM Break

11:00 AM White paper assignments

12:00 PM Lunch (box lunch to go available)

After meeting White paper discussion will continue for those who can stay longer

Appendix B

Data QC Workshop Participants
June 6-8, 2012
Darline Marine Center, Walpole ME

Steve Ackleson
Consortium for Ocean Leadership
sackleson@oceanleadership.org

David Antoine
Laboratoire d'Océanographie de Villefranche
antoine@obs-vlfr.fr

Andrew Barnard
WET Labs
andrew@wetlabs.com

Bill Bergen (OOI)
Ocean Observatories Initiative
wbergen@oceanleadership.org

Bill Boicourt
University of Maryland Horn Point
boicourt@umces.edu

Emmanuel Boss
University of Maine
emmanuel.boss@maine.edu

Kendra Daly (OOI)
University of South Florida
kdaly@marine.usf.edu

Grace Chang
Sea Engineering
gchang@seaengineering.com

Ken Johnson
Monterey Bay Aquarium Research Institute
johnson@mbari.org

Matthias Lankhorst (OOI)
Scripps Institution of Oceanography
mlankhorst@ucsd.edu

Marlon Lewis
Dalhousie University
Marlon.Lewis@dal.ca

Tim Lynch
CSIRO IMOS
tim.lynch@csiro.au

Merrie Beth Neely (OOI)
Consortium for Ocean Leadership
mneely@oceanleadership.org

Mary Jane Perry
University of Maine
perrymj@maine.edu

Chris Proctor
NASA
Christopher.W.Proctor@nasa.gov

Paul Robbins
Woods Hole Oceanographic Institution
probbins@whoi.edu

Collin Roesler
Bowdoin College
croesler@bowdoin.edu

Oscar Schofield (OOI)
Rutgers University
oscar@marine.rutgers.edu

Wayne Slade
Sequoia Scientific
wslade@sequoiasci.com

Daniel Smith
CSIRO IMOS
Daniel.V.Smith@csiro.au

Heidi Sosik
Woods Hole Oceanographic Institution
hsosik@whoi.edu

Julie Thomas (OOI)
Executive Director of SCOOS,
OOI-CI Observatory Manager
jot@cdip.ucsd.edu

Jeremy Werdell
NASA Goddard Space Flight Center
jeremy.werdell@nasa.gov

Ivona Cetinic
University of Maine
ivona.cetinic@maine.edu

Nathan Briggs
University of Maine
natebriggs@gmail.com

APPENDIX C

Day 1 Discussion of OOI and NASA Workshop Questions

Disclaimer: this is a summary of the verbal discussions and should not be considered as reproducing statements verbatim or a comprehensive representation of all statements presented.

What are the similarities and differences between OOI and other approaches?

- There are real- time data similarities between the ARGO and what will be necessary in OOI QC Flags. (e.g. range check, spike check)
- Some possible differences are that the collection and analysis of field samples vary among other time-series projects (e.g. BATS, HOT, CARIACO).

Should the community attempt to have a common approach? Not all scientists agree on what are the 'best' analytical approaches or data analysis approaches. However, it is recognized that results obtained by different time-series sites must be intercomparable. To address this issue, an international workshop funded by the International Ocean Carbon Coordination Project **(**IOCCP) will be held this coming fall to discuss best practices for time-series sites.

Other time-series projects (like those mentioned above) employ more human-in-the-loop practices rather than automated data QC approaches. Experience with the Bermuda Testbed Mooring over the 12+ years of the time series program reveal the human-in-the-loop data QC approaches varied greatly over time due to changes in personnel.

How do we standardize in the future? What are the steps we need to take?

The global collection of various parameters (e.g., optics, DO, pH) would benefit from improved standardization. To accomplish this, it is important to have good cross-section of representative groups, with international participation, as in the QARTOD waves buoy data example.

Example: JCOMM – International effort funded for wave data evaluation by co-location of instrumentation.

Also NASA or OOI can identify groups that have apparently done a good job executing data QC, and perhaps they have written about it – or would be willing to. The OOI approach will become stronger when more partners can be brought in. Choose one instrument or one data product at a time and examine them one at a time. We need participation from the experts to gain their knowledge.

How would that participation by experts be orchestrated?

QARTOD has set up a governing structure. Write a white paper? Suggest that it would require a director and a board of advisors. Many of the workshop participants have connections with national/international professional societies or groups that may have expertise among members. CALCOFI at Scripps has years of biological data and personnel experienced in data QC. Get the word out, get people together in an open, not a closed, group.

Who leads the effort? If this work is funded, what is the process?

Some feel there are enough interested users in the community that want to participate and will volunteer their expertise in an area. Other scientists feel they do not want to volunteer their time and such a concerted effort requires paid expertise. Users are crying for measurement and metadata standards. Philosophically, perhaps a group like the IOC body (at the international level), or JCOMM/JCOMMOPS could serve as the coordinating body. This would serve as a central library to house technical documents (e.g. how to sample oxygen correctly). Realistically, salaried technical coordinators will be required.

We need to standardize what 'data QC' means. Suggest annual meetings of working groups for different sensor families (bio-optics, CTDs) to discuss and develop strategies for those instruments. Prepare an updated report on annual progress to the community. Have enough funding to cover participants travel to this meeting.

What about OOI data file format standards? Include metadata, user generated information. How does that get standardized? Otherwise standard QC can never be agreed upon (e.g. ARGO started discussions very early on in its development). OOI CI has experience with this; looked at other structures, the data managers are in communication with them. Regarding data format, CI has its own internal data model and it will be transparent to end users, because they will be able to access data in several formats, e.g. NetCDF. OOI metadata will follow the standard format.

If someone queries the OOI database, does data collected at different times get merged? e.g. if a CTD collects samples at 1 hr. intervals, and another instrument at 30 minutes, will OOI merge them? This is user specified (i.e. merging of files) – short answer is all data comes with a timestamp. OOI will aggregate data files and there will be a number of formats users can select to receive data. MOS has had to deal with data merging problems. Their system does not aggregate files. Users must download and splice them together.

Data QC: Data quality must be mostly good to begin with, or you can't QC it to make it much better. That's the grand challenge -e.g. surface mooring data, where fouling is an issue, might not be able to be corrected with QC.

It depends on the data-product and sensor, but some QC solutions might come from working with instrument manufacturers. IMOS started off their detailed examination of processes with salinity and temperature, moving onto bio-optics.

What is needed to improve dissolved oxygen data? Better QC procedures, better hardware, better deployment processes? Clark membrane type sensors may have sensor degradation problems and there is quite a lot of work remaining to do on fully identifying and solving that problem. This is an active area of research and sensor development.

What happened to the proposal to calibrate ARGO floats? There were two types of oxygen sensors that have been deployed on ARGO floats, but there are issues with drift and manufacturer calibrations. There are solutions to solve these issues and scientists are working with the manufacturers to develop more robust sensors and improve calibration methods. The ARGO program has a much defined target. Currently the data from one manufacturer's sensor is acceptable 90% of the time, but there is a need to resolve the last 10% of unacceptable instances to meet ARGO performance standards. There is a proposal still in review on this matter.

What's the best solution to the problem of good quality data - to focus on the sensors or the data product you produce? Do we develop oxygen sensors or the 'XYZ' sensor that is a proxy and is currently available? Are sensors selected because they measure backscatter (a property), or POC (a desired product)? Things to think about include the complexities of sensor issues, and the conversion of measured values to final products.

The OOI sensors chosen to be the core sensors were determined by the high level science questions (i.e., the Science Traceability Matrices: what parameters were needed to address the high level science questions). In addition, there were several criteria for sensors that included whether they were COTS, proven to be durable and reliable, and could be deployed for relatively long periods of time, etc.

In the case of some instruments in the OOI program, rather than offering certain biogeochemical parameters as data products, they specifically named a particular proxy that the user can convert into a relevant data product themselves. This may not be the best approach for every program. In some cases the OOI is deploying a handful of unique or non-COTS sensors and in 10 years, those sensors might be idiot proof as a result of the programs efforts. But it is best to start with deploying those COTS sensors that community experts have well characterized. The OOI developed a list of the best available (COTS) sensors that will be installed on the OOI. OOI selection of instruments and where to deploy them was crafted to address high level science drivers, the 6 over-arching scientific questions and sub-questions that define the OOI. When starting your own program define what parameters you need to measure to answer your science questions. Oxygen, POC, etc. [See Traceability Matrix example in Daly's OOI Overview talk]

The question facing the workshop participants is:

> **How does OOI take the list of sensors the program will deploy, and apply the best QC measures possible?** Instrument level responses should be solved in collaboration with instrument experts.

Who gets these people committed to working with the data and providing information? It is an inter-agency effort for the national level, IOC for global (UN types of channels). But the bigger the community you engage, the harder it is to work on 1 sensor for 1 measurement. If you engage a smaller community, you might be able to focus on something specific - as a hypothetical example, dissecting QC needs for a WETLabs FLNTU would be much harder to do with a large scale group assembled (whether US or international). The latter might be better for more general issues (Phytoplankton sensors, carbon sensors). Working groups must scale properly as well. If too many people are involved they break down, so an ideal size is 2-6 people. Task them to produce a tightly constrained brief in a certain amount of time. IMOS has had good success with the model of constraining time and size of working groups. It is important to gather the experts together to develop (or revisit) best practices. However, it is essential to include people who are collecting data in the working groups. Groups of theoretical experts who don't have experience collecting data often yield unrealistic advice.

Pay attention to what was the original intent of the instrument (i.e. not all sensors are designed to be moored and left unattended for long time periods). Consider QA/QC of the sensor output because maintenance is also critical! Consider budget issues and constraints when selecting instruments and in determining your maintenance schedule. In order to get high quality data from certain sensors, an operator might have to perform onsite service every 2 weeks – but you can't unless that is in the operational budget.

Back in early 2000s, NASA drafted technical manuals on Ocean Color satellites. The manual contents were the accumulation of the community and vendor recommended best practices available at that time for instruments and satellites. The manuals were drafted for a specific purpose – to 'match-up' the data, but they haven't been updated lately. OOI should obtain copies of the NASA documents and convene a similar group to emphasize the best features of the OOI and identify best practices for all instruments.

Regarding bio-optics and ocean color, the IOCGG exists and an effort to restart something similar to SIMBIOS is underway and might be a good source of information. There is a white paper coming out soon on this topic – the idea being to make recommendations to agencies, with working groups for revisions of protocols. This approach might be useful in coming years. Task a working group with advocating with the international community. Funding must come from multiple agencies (NASA, NOAA, NSF, and EPA) to accomplish this. One recommendation that might be made from that working group is what % of the operating budget should be devoted to the characterization of sensor data QC? 20%? 30%? The user community redefines what level of time and effort needs to go to data QC.

How would the solicitation happen for who wants to be the head of X or Y team to produce something akin to the NASA manuals? There's no OOI budget for this effort. One possibility discussed was that the community might decide to develop a community consensus statement and work with federal agency program managers to implement NOPP funding. The Consortium for Ocean Leadership also could help put that suggestion in front of the appropriate funding agencies.

NRC – provides an infrastructure – largest number of agencies funding a study in over a decade. Build a wide range – stress the broad importance of what OOI oceanographic data can do for health and safety, etc.

Can we identify problem families that OOI and NASA need to focus on?

Of concern is that OOI instrument data will be flowing starting this summer and the IO's may be swamped with engineering issues and may not have time to be working on data QC protocols and standards. OOI needs to address this NOW, won't have the luxury of waiting 2 years for a NOPP panel to come out with recommendations. For the IMOS program, data QC is essential to ensure data quality and is budgeted at 30-40% of overall funding – mostly in salaries.

In the short term, OOI solutions may need to be sensor specific with groups of 4-5 assembled to inform OOI procedures in the near term. This effort has to begin immediately as the user community deals with longer term solution(s). The OOI has to come to the oceanographic community to request this though. OI and ROSES E2, National Academies may have funding mechanisms. For small NASA working groups, ROSES (Appendix E2). i.e. get 6-8 people to work on the ultrapath measurement of the dissolved absorption coefficient, or CDOM measurement. Another example is the coastal carbon synthesis project - important to get community members 'charged up to write'. The user community needs a volunteer program manager and community member as a near-term solution to hit the ground running on this issue.

OOI Project Scientists (global array) are working with OceanSITES on some activities. However, this is a voluntary PI driven effort to collect the open ocean time-series measurements. BATS and HOT are part of this organization too. OceanSITES are part of GOOS. They deal with best practices, standards, QC. OOI and NASA could connect/discuss with them, as they are a very active group.

How do we get the word out and get the community entrained in all OOI efforts?

The oceanographic community lacks knowledge about OOI activities. The OOI began as a community wide effort with large participation in workshops held in Puerto Rico and Salt Lake City. Following the SLC meeting, many of the OOI decisions that have been made were not as broadly communicated as before and publicity dropped off. Community apathy toward OOI has ensued except at the institutions involved in the program. Town Hall Meetings at Ocean Sciences are ineffective, and so is referral to the OOI webpage (www.oceanobservatories.org). Webinars were of interest to some, but not all participants. Colleague-to-colleague marketing has been somewhat effective to date for OOI and other programs, but is obviously limited in scope.

APPENDIX D

Workshop Breakout Group Summaries

Data QA/QC Instrument Drift Working Group Leads Collin Roesler and Oscar Schofield

Automated ocean networks, such as the OOI, will deploy a diverse array of sensors at sea for extended periods of time with the goal to detect long-term trends in the ocean. This will require correcting the data for shifts related to factors that do not reflect actual changes in the environment. One such factor that will be present in the data sets collected by instruments deployed for long periods of time (months to year) is instruments drift. The drift in instrument baselines can be abrupt or can be a slow linear change. Drift can reflect a myriad of factors within the instrument. For example in bio-optical sensors containing internal light sources and filter wheels, the drift can be driven by changes associated with aging in either the light source and/or the filter system, as well as detectors. The drift is often slow and thus over short time scales can be difficult to recognize in the presence of natural variability. Indeed, it may only be upon redeployment of a new, well calibrated and characterized sensor that an offset related to drift is observed or when post-recovery calibrations are performed on the sensor that changes in the instrument baseline are observed. This working group focused on identifying the steps that are critical to detecting and will assist in correcting for instrument drift. Specific recommendations are provided below.

1) Often there are several sensors available that are commercially available that fit specifications for procurement. However, sensors made by different manufacturers are all subject to different amounts of drift and different sensors built by the same manufacturer have different amounts of drift. Information on manufacturer-specific drift characteristics should be given a large weight in the decision making process prior to procurement. Information is often best obtained by requiring input from SME. Doing so will decrease overall cost over the program and reduce risks to the program, as monitoring and correcting for the drift will increase the O&M needed to produce data that the science community will trust. **Given that purchase is a onetime cost while O&M are ongoing costs, choosing stable and well tested sensors yield long term cost savings and higher quality data products. This is the most important step in the effort to combat drift.**
2) Once procurement has proceeded, and the manufacturers have delivered instruments, it is recommended that all the sensors be tested together in batch in the laboratory. If possible, the batch should consist of sensor from both the first and second deployments so that intra-deployment and inter-deployment comparisons can be conducted. The tests should monitor stability, and depending on the sensor should also be tested across a range of temperatures, light regimes and/or water calibrations. This data along with the instrument calibration information must be recorded and maintained with the instrument metadata records allowing users to access the information. Specific tests vary with sensor and advice from the broader community of how those bench top tests should be conducted would be extremely useful. For example, the baseline or dark readings many sensors exhibit sensitivity to temperature and these readings are those subject to drift. However, sensors moored at a fixed depth and those profiling through a thermocline necessarily require different modes for assessing sensitivity to temperature; the former in a series of thermally stable readings (i.e. at thermal equilibrium), the latter in a mode to quantify the time course of thermal response.
3) Once the sensors have been taken to sea, it is recommended the sensors be tested on the bench to assess if there has been drift in the baseline readings during shipping.
4) Once the sensors have been mounted on the platforms a final assessment of the baseline readings are taken to ensure that no changes occurred in the mounting of the sensors. Optical sensors are very sensitive to alignment and orientation and thus this final pre-deployment assessment should be conducted in the deployment configuration.
5) Once sensors are deployed in the ocean, a series of calibration and comparison tests are required. A ship-board CTD system outfitted with the same sensor suite as that mounted on the platform should be profiled as close as possible to the deployed platform. Multiple profiles must be conducted (at minimum 4 profiles should be conducted). These comparisons provide the time zero point for assessing drift in the baseline for the data correction.

6) During the deployment multiple comparisons of deployed assets should be conducted between the platforms. For example mobile assets should profile beside fixed and profiling sensors on mooring on a routine basis as much as possible given the prevailing science needs.
7) Prior to recovery, the ships should conduct once again a series of 4 ship CTD profiles with the suite of freshly calibrated and characterized sensors to compare to the sensors that have been deployed. **This step in conjunction with the next step is critical in order to quantify and separate instrument drift from biofouling.** Instrument drift is generally linear from deployment to recovery while biofouling exhibits a lag followed by an exponential or hyperbolic pattern. Thus separating them for delayed correction is vital data quality.
8) After the sensors are recovered and disassembled **but prior to** shipping back to the manufacturers, the bench top tests of the recovered and cleaned sensors should be conducted. **This is a critical step in being able to quantify and separate instrument drift from biofouling.**
9) The science community should work with sensor manufacturers to allow continuous recording of the instrument dark current, as this can be a useful means to monitor the instrument baseline. The difficulty in providing this information is likely to vary with the instrument class.
10) Real-time identification of drift may be achieved by inter-sensor comparisons, although real-time drift and biofouling signals may be confounded. Regardless, these comparisons could enable real-time flagging of suspicious observations. Often these comparisons are being developed by SMEs and therefore entraining the broader community will be necessary.
11) Some sensors have a high risk for drift; however they have the potential for high science reward. The OOI should consider reducing risk by considering a phased deployment of these sensors and perhaps deploying such systems on platforms that will be subjected to more frequent revisit schedules such as the pioneer and endurance arrays. Another strategy would be to mount several of the sensors on a platform and leave several hibernating and unexposed to seawater until the first sensor has experienced significant drift or bio-fouling.
12) *Ultimately, for ocean sciences, there needs to be a sustained commitment to metrology across agencies.* This is an inter-agency challenge and currently is not being addressed in any systematic manner.

Utilizing external (field or remote) observations: Kendra Daly, Emmanuel Boss, Bill Bergen, David Antoine, Bill Boicourt, Nathan Briggs, Wayne Slade

(1) Using satellites as an external tool to test data consistency

Examples:

<u>SST -Temp</u> – Average satellite SST values over 8 days. Satellite images cover larger scale areas than in situ sensors would and can act as a gross check for consistency.
<u>Sea surface height and flow</u> – integrate over water column
<u>Winds</u>
<u>Ocean color</u> –Signals timing of the spring bloom, but magnitude may vary spatially.

Limitations of this approach: Satellites measurements only measure at specific times, e.g. ocean color remote sensing is only available when cloud cover is low, during daylight hours, and at the sea surface. When phytoplankton cells are exposed to high light, measurements may be biased if their fluorescence (used to estimate Chlorophyll) is quenched. Possible recommendation would be to use nighttime values of in-situ fluorescence.

Benefits of this approach: Puts single point (in situ observations) in context of larger region (satellite). If sensor values are highly variable, satellite data can show if the reason for this variability may be because the instrument is located in a front or an eddy, etc. Use satellite images as a check. Now being done for ARGO biogeochemical floats, grab associated NASA satellite image and post it next to the data allowing a comparison of sensor trends with satellite trends. OOI could provide this data stream as separate data product, so try to put comparable data products on the same page as this image appears to ensure optimal availability to the data user.

NASA (GIOVANNI) is building tools to provide all their data streams for reference areas.

(2) Other remote sensors:

- **Data from other process research cruises**
- **Station P Canadian cruise results especially can be implemented by summer 2013**
- **ARGO floats -** use this data when the floats are within the de-correlation scale of an OOI platform
- **Instruments on other ships of opportunity** – ferry lines, fishing boats, etc. But use with cautions as the QC of instruments on these platforms may not be as high a quality as OOI or the user community would prefer.

NASA has great tools already, to provide all their data streams for each relevant site. But the user community needs confidence in the products OOI or NASA use for such comparisons too**.**

(3) Gliders can provide internal cross validation for mooring sensors.

(4) European OceanSites mooring in Irminger Sea is another sensor check opportunity.
NASA helped solve their QC problems by making data easily available to the user community. Getting people to look at data right away alerted NASA to problem areas and the people who helped perform the QC and inform NASA of those issues felt empowered.

Programs need to automate the reprocessing of data in order to add new data products. In the ESA ENVISAT mission (which carried the MERIS ocean color sensor), the reprocessing capabilities were inadequate at the start of the mission, so improvements in vicarious calibration and algorithms took a very long time before being incorporated in the MERIS processing. This situation led the community to be suspicious about the quality of these data, and it took a long time before they eventually come back to a better feeling (and start to use the data). OOI should avoid making the same mistake right at the beginning, because it may have adverse effects on the perception of the user community, which are subsequently difficult to counterbalance.

(5) Use of Predictive models as a QC for OOI sensors.
This is still a research question, but would provide a longer term use in QC for context. Would models that assimilate all Argo data be useful? Models could ID eddy motion and tracks. For some small amounts of funding could improve model development in OOI areas.

Compare in situ data with predictive models that incorporate satellite data – use of the coarser scale Navy models are best – could be a research subject for some variables (surface fields of temperature and surface currents, dynamics height, etc.), but are not good for interior measurements, and most models don’t get the mixed layer depth correct.

Models can provide de-correlation scales to give a threshold value, and perhaps longer for physical properties. For chlorophyll values they would only be useful for shorter scales of variability, but again the satellite data can provide reference surface values, on kilometer scales, but do suffer from cloudiness. The Navy models are the best operational models currently.

OOI needs to put the best data available on the website. OOI may need to reprocess the original data to improve data quality and correct for biofouling. If someone in community identifies a data quality issue and a solution is available, OOI needs to reprocess the data, as done in ARGO. All other programs have mechanism for reprocessing. If OOI doesn’t do this, it is not following community best practices.

****Recommendation: create NASA–type science (product) teams.**
The team members would offer community service by developing protocols for debugging problems that may arise, but also teams would consist of subject-matter experts (SME) that write proposals that include salary to develop the protocols for these corrections, or new algorithms to use on the raw data. Form teams of experts for families of sensors or each

sensor (bio-optics, biogeochemical, ADCPs, CTDs). For example, fluorometers need repeated dark field measurements of backscatter, but this is not called out in the user manual. Yet an SME will recognize this need and offer a solution. SME's would also make sure that sensors are calibrated, validated, installed and deployed properly, etc.). The team(s) would: define tasks, recommend certain studies that could be funded to provide guidance, would provide good information on intermediate and long term data QC, entrains expertise of community, builds the user base and achieves 'buy in' from the oceanographic community. Need to bring these teams into the OOI structure now (they are already part of NASA approach). OOI needs community buy-in, and trust of quality assurance of OOI data. How to get NSF buy-in to this approach? Investigators currently can write proposals related to OOI, such as for model development, but that pathway is part of peer review process. NSF and NASA need to provide a mechanism to fund investigators for short periods of time to provide specific data products.

Whether for OOI or NASA, the workshop participants feel funding is needed for this kind of team participation, or people won't commit much time to this assistance even though the community wants to get involved. If NASA were to consider implementing working groups they require a workshop report recommendation – and NASA wants to do so for broader international purposes and new missions.

OOI's Program Advisory Committee thinks the oceanographic community needs to be brought in using other mechanisms beyond traditional town meetings**. These science teams could serve several purposes, but we need to consider what can be implemented by summer 2013.**

Sensor inter-calibration/verification: Steve Ackleson, Andrew Barnard, Tim Lynch, Ivona Cetinic, Merrie Beth Neely

Must develop a highly detailed ***Manual for Dummies*** spelling out the specifics of every step and detailing why it is important to do them. Don't assume anything is understood – referring a technician to a published method is not enough. Even if this is 4 inches thick and regularly revised – it will be worth it. Must revise at least annually in the beginning, possibly will have longer shelf life as O&M matures but should never be shelved – continually evolving and updating is crucial. Explaining why it is important to follow all these steps and not cut corners can be emphasized to technicians by using an overall error equation example: Σerrors = Procedure A (up to 10% accuracy lost if not followed) + Procedure B (up to 50% precision lost if not followed) +Procedures C-F (up to 4 units annual drift lost if not followed). Track sensor serial numbers, not just instrument serial numbers – can change out a detector portion of the instrument (e.g.) but give you back the same instrument – this information varies by vendor and is not provided by every vendor – you may get an instrument back from service and not know this has been done – yet may explain weird results.

You must not sit back and wait for problems to be anecdotally identified, be proactive in identifying problems through a regular lab/procedures QC inspection process and utilize round-robin exercises for the evaluation of adherence to inter-laboratory procedures – have consequences and corrective actions for failing the test in order to fix problems. Invest in training of your staff, not just once but continually, and cross-train among laboratories to ensure consistent applications of procedures.

<u>Calibration centers:</u> have a centralized calibration center – for redundancy (natural disasters), in order to minimize errors, and maximize efficiencies have 2 centers (E coast W coast). Such facilities must have flowing seawater, temperature controlled environments, bench space, access to cultures and chemicals reagents/standards for the instruments used by the program and serviced there, appropriate safety devices and lab setup (fume hoods, DI system, etc.). Independent evaluation of the procedures

and facilities (NIST/NELAC certification by a trade association or gov't accreditation program) – in absence of these could be an ad hoc panel of oceanography experts. Business case to utilize vendor warranties and take advantage of their service agreements vs. doing it all in-house (requires expertise, facilities, voids warranty, but may be preferable option for some locations or instruments).

Golden instruments: each calibration center (marine IO) should have an instrument highly calibrated for expected communities (opticals) or in situ conditions the program will encounter and 'treated well' and that never goes to sea (sits on a lab shelf) and is regularly recalibrated (every 6 months).

Optical instruments example: Multipoint 'secondary calibration' against a fluorescein dilution series (accuracy and precision of instrument) followed by characterization of the instrument with expected field communities from culture. Try to get the vendor to use the programs preferred calibration method at the factory testing and service stage – even if it differs from their method – this enables the program to actually validate the vendor calibration (primary calibration) otherwise the best practice would be to verify operations, do reference standards and blank tests and then 'characterize' the instrument with a suitable culture that approximates the field conditions – not validating calibration, just optimizing your results.

Lab to field protocols: multi-sensor (fleet) testing is important to do on all sensors and spares slated for the next deployment. Do simple checks of operability of the instrument at regular intervals (when unpacked from service, prior to deployment and assembly of mooring on land, immediately pre-deployment on ship, immediately post deployment in water, immediately post-recovery on ship, when unpacked from cruise prior to shelving or sending off for annual service). This SOP can be developed by an instrument (or data product) SME, and with rigorous oversight, carried out by well trained technicians (again cross training among facilities is important here). This leads to a pass/fail that instrument is 'working' (pass move to next phase in prep for deployment, fail send to vendor for repairs). Next phase is doing serial blank and reagent standard testing on all instruments and spares that are slated for the next deployment cruise with same pass/fail outcome. Then do serial tests on all instruments against gold standard instruments to characterize – can verify accuracy and precision of the calibration and adjust slopes here (scale factors in opticals). Ready to go to the field!

Field protocols: Have formal lab assembly and field pre-deployment checklists prepared. 2 people must sign off on each checklist items as completed, one person only holds the clipboard doesn't put it down and do something then pick it back up again, etc. Also can use flags (attach flagging tape at pre-deployment assembly stage, remove each flag on ship – count that all flags are there prior to deployment.

> Turn instrument on, ensure operating normally, ensure communications are working, and evaluate a few 'samples' on deck or at the surface on the hydrowire to ensure data output looks reasonable. May need to spend some time purging the system of bubbles or cycling through some reagent, or equilibrating at depth before data is acceptable (housing floods may be immediate, when lowered to depth or otherwise fail at depth but work on the surface – better to recovery and replace now than find this out when back at the dock – spend the money to stay on station a bit more time to save money and preserve data long term). If failures at any point program can utilize spares on board that were also put through same lab to field procedures outlined above. Suggest a minimum number of spares to carry on all cruises (can be based upon a percentage of those instruments you are deployments but never less than 1 of everything) – default is to take as many spares as possible (or practical) on cruise. Perform similar pass/fail test for instrument working, etc. as above – if pass can deploy the buoy string.

Deployment testing: Do at least 1 profile up and down adjacent to mooring using all 'reference' instruments deployed on a rosette – use spares for this purpose (field golden instrument?). This should be viewed as the gold standard, but may use alternative protocols if field conditions or deployment depth location dictate varying from this (gliders swing by, do cast some distance away instead of adjacent to mooring, etc.). Alternative or in addition to this: can have both new and old moorings in the water (co-located) for several hours on station taking overlapping time series to aid characterization. This may do much for biofouling corrections as disturbance during recovery will introduce substantial bias if this testing occurs only after recovery – not as good as if done before. Alternative; take pre- and post-diver-cleaning readings then recovery them – also minimizes disturbance. In fact corrections for biofouling may be characterized in the metadata as 'we followed gold standard protocol', we followed silver standard, we followed bronze standard, etc.

ACTION ITEM: add details to clarify what is meant by primary calibration definition in the OOI data QC document (see calibration standards section above).

System-level data QC approaches: Matthias Lankhorst, Jeremy Werdell, Marlon Lewis, Julie Thomas, Paul Robbins, and Daniel Smith

Interpretation of three tiers of data QAQC success (from plenary discussion):
OOI is currently developing semi-automated QA/QC methods in place to identify and flag questionable data. This includes a series of ~19 binary (pass/fail) tests to be applied autonomously to their incoming data streams. The exact number of tests applied to each stream is specific to the data product. Our understanding was that all data are to be retained, but failure of any test will result in a data point being flagged as "bad". This breakout group reviewed the technical application of these data QA/QC tests and made recommendations for future enhancements.

We identified three events for application data QA/QC metrics:

(1) fabrication/calibration/delivery/acceptance/installation of instrument: very engineering-centric, acceptance testing of equipment and proper equipment handling during preparations following documented procedures.
(2) Deployment of instrument and beginning of time-series ("real-time"). Uses automated data processing stream.
(3) Historical re-evaluation of time-series ("delayed-mode"). Post-processing of data, heavy involvement of subject-matter expertise in judging data quality and possibly making adjustments. Key step to achieve "climate-quality" data stream.

We assumed that each individual data record will report:

- Station metadata (date/time/latitude/longitude/depth/etc.)
- the geophysical uncertainty for the geophysical variable
- the individual binary pass/fail for each test applied
- the overall binary good/bad flag assigned to the geophysical variable

Recommendations from our technical evaluation of current OOI methods:

- adopt community-vetted instrument acceptance tests that are standardized in collaboration with instrument manufacturer; ensure commonality throughout OOI

- ensure coordination of (cross training of / consistency between) delayed mode operators for flag and data QA/QC assignment; organize periodic gatherings of human operators who make these subjective decisions (along with experts on the measurements);

- limit the subjectivity of each test to the best extent possible (preference for quantified tests);

- develop automated data QA/QC tests (e.g. thresholds/ranges) internally within OOI (taking advantage of established quality control efforts all in place). Further vet qc checks with community/industry experts;

- quantify spatial and temporal correlation/covariance between appropriate variables for use as semi-automated data QA/QC tool (this is well described in sensor network literature);

- consider the review of multiple, variable temporal scales (daily-weekly-monthly-annual, etc.) of delayed mode data; there should be value in the automated comparison of these multiple scales;

- develop the capacity for reprocessing; evaluate the implications of data QA/QC metric changes with time (e.g., to avoid introduction of step functions in long-term time-series); evaluate de-coupling from climatologies used in data QA/QC; and,

- Build standing science teams to assess quality of data for addressing 6 OOI science themes; enable community involvement and develop feedback loops.

Methods exist that utilize data from multiple sensors to assess quality of data from one sensor (Lynch's IMOS presentation showed an example and Lankhorst's OOI presentation showed a proposal to that effect). This is based on "sensor network" theory, and published material is readily available. These methods have the potential to narrow the detection limits of automated tests, and to reduce the level of effort needed for human inspection of data. Argo's 'delayed-mode procedure' is, in effect, an example of this.

At-sea and instrument inter-calibration: Six step inter-comparison for sensors on moorings and gliders. Mary Jane Perry, Oscar Schofield, Marlon Lewis, Steve Ackleson, Nathan Briggs

At factory: Replicate calibration of 15%- 20% of each sensor type, done at different times, and by different technicians at the factory.
At institution: a) bench test for inter-comparison baseline, stability, dark currents. b) In situ pre-deployment batch inter-comparison of each sensor type in local waters (cf. David Antoine and Eric D'Asaro); frame include the 'golden' sensor. Each operator has a gold sensor / reference sensor. If sensor varies (by how much? within spec), return to vendor for repairs. c) For some sensors, like beam c, recalibrate in water (or in dry air) because alignment can be compromised during transit.
At deployment:
Deep-sea mooring, at deployment: For some sensors, like beam c, recalibrate in water (or in dry air). Launch gliders, deep CTD cast, aggregate gliders for simultaneous CTD-glider profiles (target five sets with gliders as close as possible (< 1 km) before departing station), deploy central mooring, CTD-glider calibration, shallow CTD near central mooring with appropriate sensors and water samples, and flanking moorings.
Shallow water – CTD profiles by mooring, CTD – glider profiles at deployment, at retrieval, and opportunistically. Autonomous cross calibration of glider/glider and glider/CTD sensors.

Autonomously: Data for cross calibration collected from co-located glider-to-glider profiles and from transits near moorings by gliders.

At recovery:

Deploy new gliders, do aggregated glider profiles with new and old gliders.

Multiple CTDs and water samples near moorings.

Final inter-calibration: Return sensors to factory for pre-cleaning calibration and post-cleaning calibration.

Biofouling: Kendra Daly, Collin Roesler, Grace Chang, David Antoine, Bill Boicourt, Heidi Sosik, Ken Johnson, Wayne Slade

(1) Preventative strategies,

- TBT (tri-butyl-tin)
- Copper
- Put electrical tape on everything first then copper tape for easy removal
- PVC tape also good –cheaper, more durable, wider
- Tubed or pumped system, buy lengths of plumber cooper tubing (Manov et al. 2004 - JAOT) JMSTAC.
- for long deployments, technicians must be careful to not create anoxic zones, anaerobic bacteria may feed on the copper. Specifically, be careful when wrapping copper tape around areas of the sensor that may create anoxic zones, for example at the top of a WETLabs ECO sensor, there are socket cap screw holes – be sure these are ultra clean and sealed. Also, ensure the outside of the instrument housing is very clean before wrapping. Do not breed organisms under your anti-biofouling measures!
- Use wipers on optical faceplates, especially wipers that don't get hard and scratch the windows. If power goes off for extended time, the user may not elect to turn the instruments back on, because in these cases the wiper could scratch the window and destroy instrument. Also, copper faceplates with wipers are essential if power is off for an extended time – copper faceplates ensure that large organisms do not grow on the faceplate that may prohibit the movement of the wipers, e.g. barnacles.
- Satlantic's ISUS and other sensors and some instrument manufacturers have shifted away from using copper, 'too much copper will kill it'. For other pumped sensor systems, like conductivity, the WETLabs ac-s, etc., the operator can use copper-nickel or plumbing tubing for the inflow and outflow lines – it works better because it is stronger, and deployment evidence has shown its use doubled the biofouling mitigation period (need to service moorings to remove biofouling half as often)
- Silver tubing works in freshwater deployments, but in saltwater it exchanges with the chloride quickly, doesn't work
- Application of UV light can kill biofilms, but you need to use deep UV wavelengths. This method is power hungry. The use of UV lamps could be a good solution for cable systems, but further research on the effectiveness of this method would be necessary prior to implementation – perhaps a funded research project?
- Silicone grease applied to the ADCP transducer face works well. It is also recommended to wrap the ADCP body with PVC tape or **copper-nickel tape that is commercially available.**
- There are some new CTD/water quality sondes that have sleeves in which you can deploy your instrument. The sensors then slip easily out of the sleeve upon recovery for "instant" cleaning.
- To protect the structure on which the instrument is deployed it is recommended to tape the cages or frames, because if organisms colonize the frame, they move to the sensors more rapidly. U Md wraps frames with plastic foodwrap, and paint frames with boat antifouling paint or E-paint (ecominder paint). But be careful of the paint's composition, because some components can cause corrosion or interactions among instruments

- Use cross-correlation checks with other sensors on gliders, etc.
- For long-term deployment an operator could put multiple sensors out at the same deployment depth and location (example WETLabs ac-s) that come 'on line' at different periods with some overlap (encased in an anti-fouling capsule and remotely triggered to open and turn on).
- Could minimize biofouling by using a chlorine generator on cabled system with high power, as used on Sea Keeper flow through systems. This uses a small battery – this is an area where more research is needed to learn what, if any, interference may occur with some sensors?

(2) Real-time strategies while sensors under water

- For near surface fixed sensors such as the WETLabs ac-s (fouls within a few weeks to month), so an operator must either replace sensors frequently or send divers out on boats to manually wipe off the optical windows with lens paper and to exchange tubing in situ. For the OOI, this could be reasonably accomplished on the same trips as when gliders are being recovered and deployed
- Could take an in situ reference check by using a solid chlorophyll standard to get single point reading on WETLabs ac-s.

(3) Post-processing to recover data

- Need fly-by in situ measurements to assess biofouling. Need data from non-biofouled instruments to compare to biofouled instrument data. Before every recovery conduct 4 CTD profiles but collect representative water samples only once. Then recover the mooring. Clean and recalibrate the instruments in air (dark counts), and repeat the same procedures in water to assess instrument drift component of the error. Then subtract the instrument drift from the overall error and attribute the remainder to biofouling (ship board-in water test)
- Level 0
- Flagged data
- Professional in the loop assessment is necessary prior to correcting data
- Develop education tools to allow students to go on interactive website and create biofouling corrections- may not yield correctable data but would be a good exercise for educators to use online data
- Digital fishers – Neptune Canada ROV and camera images, utilize citizen scientists to analyze large amount of video data

Centralized vs. distributed approaches– human in the loop efforts applied to delayed-mode data:
Merrie Beth Neely Tim Lynch, Jeremy Werdell, Bill Bergen, Ivona Cetinic.

Determining the approach to use starts with trusting the field protocols are adequate and properly followed by field personnel. One instrument (or data product) SME should control this within the program (can be vetted with several SMEs inside or outside the OOI but responsibility lies with one SME in both approaches. The SME must know the instrument, the data product (and expected results for all OOI regions), and the operational theory (pros and cons, sources of errors, bias, or confounding factors) well in order to be given this duty. If rigorous oversight procedures are followed, the SME can be decoupled from the technicians operating the reliable, high data volume instruments (distributed approach). But this approach will not work for complicated, unreliable, beta-testing or prototype units (low serial number), or unique-low user demand instruments (centralized approach). The SME may need to interact regularly with: a platform expert to discuss platform issues such as shading, aliasing, or cross-talk concerns; a regional expert looking at broader picture (multiple sensors) to evaluate results of a detected event or section of flagged data to determine if it is good or bad; the technicians doing their field work unless they have the luxury of going on all cruises to conduct the field work themselves.

Internal data 'workshops' should be held regularly to inform the OOI-wide data procedures as well as to evaluate regional phenomena (refine lookup tables for local regimes, early-warning system for blooms, *El Nino*, etc.) - How often? At least quarterly, and as needed otherwise.

When using the distributed approach, in order to avoid bottlenecking, designate a person as being responsible, and they must have consequences for not doing this job in a timely fashion, and structure it so this person is not a gatekeeper to ensure that **not** doing this job doesn't prevent updated data from getting to CyberInfrastructure (CI) if an Implementing Organization (IO) is doing the updates and just waiting on the SME approval – maybe the data just isn't released – but need a mechanism to tell when updates aren't happening.

As described above and shown in Figure 1 below, we recommend two models:

- Model 1 (ex. CTD): reliable and widely used instruments that produce a high volume of data either because there are a lot of them or because they are on a lot. May be in high demand by users. 1 SME is responsible for instrument or data product data QC oversight, but they can rely upon trained technicians at Marine IOs to evaluate and implement the HITL aspects of field verification protocol data for these instruments and to execute repeatable data corrections (can be automated – mean shift, linear extrapolation over time of deployment, etc.).
- Model 2 (ex. Direct CO_2 Covariance Flux [OOI's FDCHP/FDCLP instruments]): unique, low serial number, low user demand, fussy or complicated instruments that require a lot of HITL effort to ensure high quality data. The community may be more skeptical about this data – assuming the necessary HITL data QC/QA will not be done or what is don't won't be useful – may require longer time or more outreach efforts to get community buy-in to trust and use the data. Such instrumentation may be subject to extra scrutiny after some period of initial deployment that if the cost to benefit ratio of the instrument is too high it may be deprecated. This length of this period of time should take into consideration parallel efforts to inform the community skeptics and garner community buy in to use the data. Alternatively, the OOI may continue to operate the instrument but disclaim the data shall always remain raw or provisional as it is cost prohibitive to apply the necessary HITL required to meet community standards. A user may choose to invest their own time into applying QA/QC, post the updated data set and associated metadata, and share it freely through OOI website. OOI should not invest time and money into these instruments at the expense of doing all you can do with the Model I instruments.

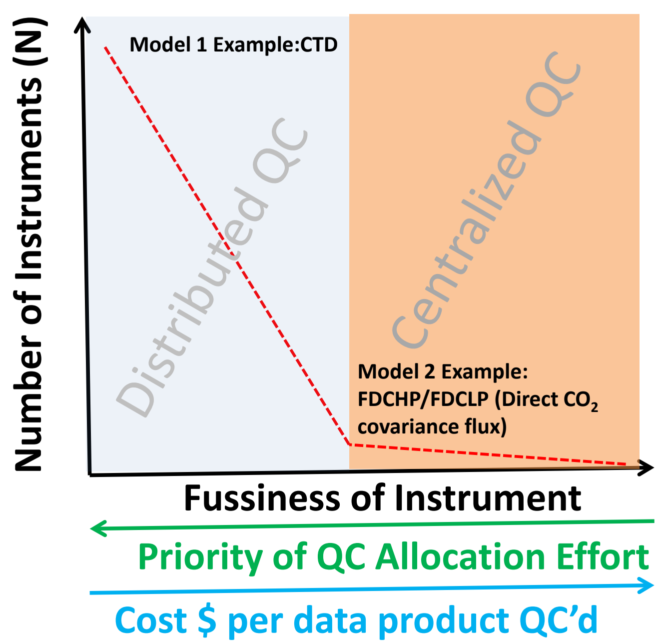

Figure 1. Human-in-the-loop (HITL) data QC effort models.

Levels of QC and associated effort required: Group Leads Matthias Lankhorst, Emmanuel Boss

The OOI is not yet clear about the level of human involvement that is needed for data QC or that can be provided given the budget and scope.

Long term accuracy requirements impose need for high human involvement, deriving e.g. from the science topic “climate”.

Australian IMOS: 30% of labor cost goes into QC development.

Argo US: Have 5-6 people working operationally, ca. 3 full-time equivalent, including developing tools. Only evaluating conductivity and pressure data. The technicians need to look at the data many times. Reprocessing occurs sometimes, due to finding issues with sensors. Generally they wait for an erroneous signal to show up in the climatology data and then back-correct. Changes in sensors always cause Argo to reprocess data. Initial number of FTE performing data QC significantly higher. Development of initial data QC system took (very approximately) five years, was an international effort.

MOVE (M. Lankhorst): 3 moorings in tropical Atlantic with ca. 40CTDs. Climate-quality accuracy needed. Moorings are recovered and replaced annually initially, but now they do it bi-annually. It takes about 1.5months to get data posted in ‘delayed mode’ (this is corrected, QCed data). They also do reprocess the data with adjusted coefficients.

IOOS wave buoys: Datawell wave sensor, 55 instruments in a mature and mostly automated system. Have ca. 4 fulltime programmers perform data management, incl. QC and archiving. One part-time person (0.5 FTE) evaluates data exceptions. Project is about 30 years old, initially a significantly higher portion of labor needed for data QC. Instruments return full spectrum of motions, i.e. plenty of parameters available to feed into automated algorithms. Still, will always need a human for data evaluation and would never be fully automated. They do utilize an automated messaging system for critical problems. They’ve had a handful of cases when they reprocess the data; sometimes they just reformat the data. Huge level of ‘people effort’ was necessary to build and implement a list of automated QC procedures.

SeaBASS: Data reprocessing occurs regularly in PI data.

Matthias Lankhorst presented one estimate of data management/QC needs for OOI: for global array – 40 instrument types, 250 total instruments. Only highlighted part is actually QC:

- Start data flow (data management) – 0.5 FTE (15 mooring, 12 gliders)
- Figure out & apply adjustment (Delayed mode) – 1.7FTE
- Visual QC and flagging (delayed mode) – 0.7FTE (tools are programmed by CI)
- Visual quick check (Real time) – 0.8FTE
- Resolve trouble cases –2.2 FTE
- Documentation – 1.0FTE
- Interface w/ CI -- 0.5FTE
- Science oversight -- 0.5FTE
- User office –1.0FTE

Sum is 9FTE for only the Global Array Portion.

Assuming steady state and using the calculation factors above, for the entire OOI with approximately 750 physical/biological/chemical instruments, perhaps 25FTEs would be needed (15-20FTE to perform QC tasks). Currently the O&M budget does not reflect this need. QC tasks are rather distributed among engineering/data management/CI personnel.

Recommendation: General consensus that the above numbers are reasonable. Plan for 18 FTE's for next 5 years while the OOI system matures, maybe fewer when mature and automations are achieved. They could be proportionally assigned by need, for example: 6 people for physics, 6 for chemical and 6 for biological.

OOI should identify data users group.

In construction phase, OOI should budget for the development of the initial set of QC procedures. Include some budget to write procedures brought from within and outside the program. Also include refresh of these QC procedures in the budget in O&M.