QUALITY ASSURANCE OF REAL-TIME OCEAN DATA: EVOLVING INFRASTRUCTURE AND INCREASING DATA MANAGEMENT TO MONITOR THE WORLD'S ENVIRONMENT

William Burnett(1), Richard Crout(1), Mark Bushnell(2), Julie Thomas(3), Janet Fredericks(4), Julie Bosch(5), Christoph Waldmann(6)

(1) *NOAA (National Oceanic and Atmospheric Administration), NDBC* (National Data Buoy Center), 1007 Balch Blvd., Stennis Space Center, MS 39529, USA, Email: Email: Bill.Burnett@noaa.gov

(2) *NOAA (National Oceanic and Atmospheric Administration), CO-OPS (Center for Operational Oceanographic Products & Services)*, 672 Independence Parkway, Chesapeake, VA 23320, USA, Email: Mark.Bushnell@noaa.gov

(3) *Scripps Institute of Oceanography, CDIP (Coastal Data Information Program)*, 9500 Gilman Dr, La Jolla, CA 92093, USA, Email: jot@splash.ucsd.edu

(4) *Woods Hole Oceanographic Institution*, MS #9, Woods Hole, MA 02543, USA, Email: jfredericks@whoi.edu

(5) *NOAA (National Oceanic and Atmospheric Administration), NCDDC (National Coastal Data Development Center)*, 1100 Balch Blvd, Stennis Space Center, MS 39529, USA, Email: Julie.Bosch@noaa.gov

(6) *University of Bremen, Leobener Strasse*, P.O. Box 330440, 28334 Bremen, Germany, Email: Waldmann@marum.de

OVERVIEW

At the OceanObs'09 Conference, there will be numerous papers and many discussions describing the intense effort by the international community to completely observe the world's oceans. New technologies, new techniques, better ocean vessels, improved sensors and faster data collection – all of these items will be used to observe in real-time, and understand, the ocean more than at any time in our history.

Yet, with all the observations being collected, and all the new technology being developed – who, and better yet, how will these data be properly quality controlled, maintained, disseminated and archived? The next ten years will bring many challenges related to the distribution and description of real-time ocean data. One of the primary challenges facing the community will be the fast and accurate assessment of the quality of the data streaming in from new observing systems. Quality control and quality assurance of ocean observations must be a priority for data collectors and observation providers to ensure that the real-time users of the observations, as well as the climate community understand the value of the observation. This White Paper will describe how data managers can properly prepare for, and manage, the incoming wave of ocean observations that will arrive in the next few years.

1. QARTOD

High-quality, long-term observations of the global environment are essential for understanding the Earth's environment and its variability. The United States contributes to the development and operation of many ocean observation systems – some of which have been in operation for many years. To ensure that data providers, managers and users understand the value of the large amount of ocean observations that will be available in the near future will require more robust quality control and quality assurance systems and procedures.

The Quality Assurance of Real-Time Oceanographic Data (QARTOD) group is a continuing, U.S. multi-organizational effort formed to address the quality assurance and quality control of oceanographic data collected by the Integrated Ocean Observing System (IOOS®) community. The first workshop was held at the National Oceanic and Atmospheric Administration (NOAA) National Data Buoy Center (NDBC) office in Bay St. Louis, Mississippi, USA in the winter of 2003. Over 80 participants from federal agencies, universities, oceanographic institutions and private industry attended the meeting with the primary task of developing minimum standards for calibration, quality assurance (QA) and quality control (QC) methods, and metadata.

The first workshop [1] resulted in some monumental decisions for an ocean community struggling to understand the challenges related to the distribution and description of data from the IOOS®. First, the workshop agreed that every real-time observation distributed to the ocean community **must** be accompanied by a quality descriptor (Was the data quality controlled? Was the data quality questionable?). Second, all observations should be subject to some level of automated real-time quality tests. Third, quality flags and quality test descriptions **must** be sufficiently described in the accompanying metadata. Fourth, observers should independently verify or calibrate a sensor before deployment. Fifth, observers should describe their method of verification/calibration in the real-time metadata. Sixth, observers should quantify the level of calibration accuracy and the associated expected error bounds. Finally, manual checks on the automated procedures, the real-time data collected, and the status of the observing system **must** be provided by the observer on a time-scale appropriate to ensure the integrity of the observing system. Though the primary focus of the workshop was on real-time QA/QC, it was understood that some methods and requirements for the real-time data are easily extendable to "delayed mode" QA/QC and that the real-time and retrospective processing are both linked and ultimately required.

Given the rather lofty goals set by the first workshop, attendees agreed that future workshops should work piecemeal on each of the goals. NOAA's Center for Operational Oceanographic Products and Services (CO-OPS) funded and hosted the QARTOD II (the second workshop) was held February 28-March 2, 2005 in Norfolk, VA [2]. This meeting focused on QA/QC issues in HF radar measurements and wave and current measurements' unique calibration and metadata requirements. This workshop attempted to develop the quality descriptors for each system, set the level of automated (and manual) quality control for observations and determine the type of real-time metadata pertinent to each observation. QARTOD III, held on November 2-4, 2005 at the Scripps Institution of Oceanography, La Jolla, CA., continued the work on High Frequency (HF) Radar, waves and *in-situ* current measurements, and initiated work on CTD (Conductivity-Temperature-Depth) measurements. QARTOD IV was held at the Woods Hole Oceanographic Institution, June 21 - 23, 2006, added QA/QC for dissolved oxygen into the agenda and also began the engagement of the international community [3]. Related materials are posted on the QARTOD website: http://qartod.org.

Previous QARTOD meetings worked on the qualitative and quantitative specifications for various ocean parameters like temperature, surface waves, surface/subsurface currents, salinity and dissolved oxygen. Developing these specifications requires an "all hands meeting" of ocean sensor, ocean science and data management experts, sharing quality control algorithms, quality assurance techniques and real-world experiences. The meeting begins with all participants gathered together in an auditorium to receive direction from the QARTOD Organization Committee – consisting of volunteers. Briefs on the outcomes of previous QARTDOD meetings are presented and the goals for the current meeting are discussed. Then participants "break out" into different ocean parameter groups to work on their respected areas. A facilitator for each parameter break-out group provides questionnaires to the participants and gathers information to obtain the answers. After a day and a half, the participants join together to discuss the outcome and brief the participants. When the meeting ends, the organization committee compiles all the information and delivers a report to all the participants.

At the QARTOD IV meeting, quality control recommendations for two parameters, waves and ocean currents, were approved and forwarded to the U.S. IOOS® Data Management and Communication (DMAC) organization. With the approval of these QC specifications, the U.S. IOOS® community will be able to quality control in-situ, real-time wave and current observations in a community-accepted, standardized method approved by the U.S IOOS® Office. Standardization will enable interoperability of the data. Quality control flags will be assessed for those observations providing the user with a valuable understanding of the accuracy of the observation. Future efforts will focus on how to graphically display the observation so the users will not have to look at text information to assess the accuracy of the observation, while still enabling machine-to-machine interoperability.

The first four workshops were a success, from the standpoint that disparate groups from federal, state, academic and private organizations worked together to develop data management standards. These groups agreed to a minimum level of quality control for surface wave observations and in-situ currents collected by a specific manufacture. These groups also developed quality flags and test descriptions that are actually in place at some operational data centers. However, much work remains to meet the seven "goals" set during the first workshop – and the U.S. group realized that a global ocean observing system would require the participation of the international community.

2. INTERAGENCY AND INTERNATIONAL CHALLENGES

QARTOD addresses issues relating to the collection, distribution and description of real-time oceanographic data. One of the primary challenges facing the oceanographic community will be the fast and accurate assessment of the quality of data streaming from the IOOS® partner systems. Operational data aggregation and assembly from distributed data sources will be essential to the ability to adequately describe and predict the physical, chemical and biological state of the coastal ocean. These activities demand a trustworthy and consistent quality description for every observation distributed as part of IOOS®. Significant progress has been accomplished in previous workshops towards the definition of requirements both for data evaluation and relevant data flags for real-time QC. The intent of future QARTOD workshops is to report on the recommended QC descriptions for parameters such as waves and currents, expand the work to additional parameters and evolving sensor systems, and develop guides for best practices to assure data quality.

Fortunately, there are some specific data collection platforms that collect global ocean observations that are successfully quality controlled and calibrated. The Argo system collects salinity and temperature profiles using an array of robotic floats in oceans deeper than

2000 m. The Argo data are subjected to 19 quality checks at national data centers before being sent to the Global Telecommunication System (GTS). The data are disseminated in netCDF (Network Common Data Form) format that contains the profile, trajectory data, associated metadata and quality control flags. Similarly, the global drifter program quality controls observations from the enormous amount of surface ocean drifters that are providing location and SST (Sea Surface Temperature) observations in real-time.

Related quality control and assurance efforts are taking place around the world. Europe's implementation of ESONET (European Sea Floor Observatory Network) includes the standardization of hardware and software for interoperability, as well as the standardization of data quality, access and semantics. In regard to quality management, it has been recognized that an essential prerequisite for tracing the quality of ocean data is to describe the process of data acquisition in detail. This will allow for the definition of workflows for the measurement process, which forms the basis for intercomparison of the quality control and assurance procedures between different data providers. In the end, this forms the base for the recommendation on best practices for instrument preparation and deployment and the subsequent processing of the collected data. ESONET aims at harmonizing the newly developed recommendations with existing, analogous procedures such as those found in meteorology and with other initiatives in the field. The European Global Ocean Observing System (EuroGOOS) recommends ISO (International Organization for Standardization) 9001:2000 as a coherent quality management system for service providers. The implementation of ISO 9001:2000 standards would then define the mission, strategies and strategic aims of the data provider, and documentation for quality controls and quality assurance tests. This would allow for transparent quality management procedures, better comparison of processes and dynamic adaptation to future systems.

Seeing a need for a continued effort on establishing standards, the U.S. IOOS® Program Office worked jointly with the IOC's (International Oceanographic Commission) International Oceanographic Data and Information Exchange (IODE) and the WMO (World Meteorological Organization) Joint Commission for Oceanography and Marine Meteorology (JCOMM) to hold the first session of the Forum on Oceanographic Data Management and Exchange Standards. One objective of this first meeting, held in January 2008, was to gain broad agreement and commitment to adopt common standards related to ocean data management and exchange. Taking a lead from the QARTOD measurement types and five core variables identified by the IOOS® Program Office, this meeting began to address the QA/QC of surface waves, currents, temperature, salinity, and sea level data as well as compare QC flag sets and several vocabularies being used by a variety of international programs and National Data Centers. Included in this meeting were presentations on the QARTOD effort and the SeaDataNet (Pan-European infrastructure for Ocean & Marine Data management) vocabulary harmonization work. This forum provided a start to further discussions and actions for jointly updating QC practices. Additionally the Forum established an Ocean Data Standards Pilot Project (http://www.oceandatastandards.org/) under IODE and JCOMM and a process for vetting recommendations and establishing practices for international application. As data quality control protocols and quality flag scales continue to be key factors for successful data interoperability by the QARTOD/U.S. IOOS® and SeaDataNet, IODE, JCOMM communities, efforts must continue to engage each other throughout the development process. Upcoming activities including the planned QARTOD V workshop (http://qartod.org/) in the U.S. and the International Conference on Marine Data and Information Systems (IMDIS) 2010 provide venues for continuing joint work, demonstrating current capabilities and engaging the broader community.

The World Meteorological Organization (WMO) Integrated Global Observing Systems (WIGOS) Pilot Project is a WMO/Intergovernmental Oceanographic Commission (IOC) funded effort to establish a "comprehensive, coordinated and sustainable system of observing systems with assured access to data and products from the component observing systems through interoperability arrangements." WIGOS is the system of observing systems and the WMO Information System (WIS) provides data access through interoperability arrangements for collecting observations with various national Data Assembly Centers (DACs) providing the distribution mechanisms.

While the WIS will enhance distribution of observations and products, it will not impact existing services like the GTS. WIGOS, on the other hand, will integrate WMO/IOC management and governance, increase interoperability between systems and ensure broader governance frameworks (similar to the U.S. IOOS® objectives).

One WIGOS objective is to develop, document and integrate best practices and standards for oceanography, using similar frameworks that are in place for marine meteorology. The practices used for making meteorological observations have been standardized by WMO through its Commission for Instruments and Methods of Observation (CIMO). The WIGOS Concept of Operations recommends that all WIGOS observational data and metadata adhere to WIGOS standards through the promotion of instrument centers dedicated to marine and other appropriate calibration procedures, ability to provide assistance to inter-comparisons of instruments and use of training facilities located at instrument centers.

3. QUALITY CONTROL OF OBSERVATIONS

So, what makes a real-time oceanographic observation "good?" Take four different observation platforms providing real-time ocean wave observations in the same location. One system is a recently deployed spherical, 1-meter diameter, moored buoy using a calibrated wave motion sensor. The next system is a spherical, 3-meter diameter, moored buoy that has been deployed for over two years in waters that contain seals. The third system is a 287-meter-long, container vessel, Voluntary Observing Ship (VOS) platform that takes waves observations from a bridge that sits 20 meters off the water. The fourth system is a private yacht that operates a one-of-a-kind wave observation system that only the owner knows how to use and calibrate.

The first system reports a 2-meter wave height and a dominant wave period of 8 seconds. The second system reports a 1.8-meter wave height and a dominant wave period of 9 seconds. The third system reports a 3.5-meter wave height and a dominant wave period of 8 seconds. The fourth system reports a 5-meter wave height and a dominant wave period of 6 seconds.

The first two observations are transmitted in real-time to a DAC, and forwarded to the GTS after being validated by quality control processes that includes both automated algorithms and manual (i.e. human) verification. The third observation is transmitted via e-mail to a local weather office where the observation is sent to a server where the header/format is validated and then forward to the GTS. The forth observation is forwarded to a popular surfer website which also displays observations from the other three systems.

It is left to the user of the data – a surfer, boat operator, Coast Guard, or wave modeler – to determine the "correct" wave height and period by looking only at the data without knowing the observation system behind the data. With more knowledge, the user might have more confidence in the first two observation systems since the wave measuring systems are tailored to the environment and the data quality is controlled by a DAC. With a little more knowledge, the user might have more confidence in the first system instead of the second system – and then conclude that the wave height was 2 meters with a period of 8 seconds. Unfortunately, none of the system information or procedures listed above is readily available to the user.

4. NEXT TEN YEARS

In the next ten years, the overarching goal for ocean data managers should be to provide quantitative and qualitative information about the ocean observation to the users in real-time. The qualitative information would be detailed quality assurance metadata regarding the ocean platform that took the observation. The platform location at the time of the observation. When the system was last calibrated. What is the accuracy of the sensor? What environmental conditions exist that might impact the sensor measurement?

Within the framework of Global Earth Observing System of Systems (GEOSS), the architecture for collecting and disseminating data has been defined. One of the main challenges will be to follow common standards and procedures in regard to the data collection and dissemination process. These also have to be reflected in the metadata description. Standards for quality management that, for instance, apply to instrument qualification and performance assessment, have to be included as well. Implementing these procedures will for transform today's ocean observation activities into an operational mode.

The quantitative information would be QCl flags that provide the user with confidence values in the observation. Was the data quality controlled to IOOS® specifications? Is the observation accuracy at a high, medium or low level? What is the rating level (Level 5 – outstanding, Level 3 – good, Level 1 – minimal) for the DAC that provided the observation?

Therefore, initial actions will be to consolidate existing efforts made by the U.S. QARTOD, similar international efforts and collaborative groups supporting global programs (i.e. Argo). The JCOMM and IODE will need to be more aggressive in taking a leadership role for recommending and establishing practices for international applications. Ocean data managers need to submit ocean instrumentation quality assurance measures and ocean quality control schemes to the world's largest developer of international standards, similar to the International Organization for Standardization (ISO). This non-governmental organization forms the bridge between the public and private sectors, ensuring a consensus to be reached on solutions that meet both the requirements of organizations and the broader needs of society.

At the same time, ocean data managers need to understand and review the output from the OceansObs'09 meeting to get a quantitative understanding of the amount of data that will need to be disseminated in real-time. Currently, time scales to provide real-time ocean information range from one hour to six minutes. With societies needs for increased (and faster delivery of) ocean observations, data managers might be asked to quality control and disseminate data every minute (or less) in the near future. This will require the purchase of more computer equipment, large bandwidth and more robust algorithms to ensure accurate and timely data delivery.

Real-time data centers will need to provide by an infrastructure that is stable and robust. Network infrastructure will need to be enhanced with hardware that is leading edge and as such, will not exceed its technological usefulness in a short period. Visualization tool suites and modular framework, over the next ten years, must utilize automated quality control modules and require less human intervention during the final stages of processing prior to data distribution. Finally, to allow for time efficient implementation of new marine observations, the introduction of repeatable and proven configuration management processes agreed to through the ISO will ensure all new observations are evaluated against requirements needs and those observations will follow a logical sequence of activities for incorporation into the data center enterprise.

The future data centers will be on the leading edge of technology. Robust servers and efficient communications pipelines will transmit data effortlessly and routinely. The quality control centers will be a visual showroom of screens displaying observation sources and the status of those sources in real-time. Contingency sites will be in place providing for load balancing and full-scale operations in case of a catastrophic event. Personnel will be aligned to meet the new quality control enterprises to maximize efficiency while minimizing costs.

5. SUMMARY

While there is still a tremendous amount of work that needs to be finished in the U.S IOOS® QARTOD efforts, OceanObs'09 provides an opportunity to expand the QARTOD philosophy to meet the needs of the broader ocean observation community in the next decade. Instrument developers, data providers and data managers will need to meet international standards to ensure real-time observations are properly maintained and disseminated. The grass roots effort of the U.S. QARTOD can and will expand into an international effort to ensure appropriate quality controls are in place for the rapidly expanding ocean observation effort. By 2019, QARTOD will morph from a local "grass roots" effort to a standard international body that oversees, manages and approves all oceanographic data disseminated in real-time.

Achieving this international QARTOD body will require a concerted effort between nations participating in the Global Earth Observing System of Systems (GEOSS). The Intergovernmental Oceanographic Commission (IOC) must provide governance and organizational structure/support. An international Data Management and Communication (DMAC) organization will need to be assigned the role of validating and approving QA techniques and QC algorithms. Nations will need to provide funding and travel for participants to attend meetings, write reports and develop/transition algorithms. Finally, ocean sensor technicians will need to work closely with their data management counterparts to ensure that the required sensor and platform metadata are provided.

The U.S. QARTOD effort, the WMO/IOC WIGOS effort and related efforts like EuroGOOS ISO9001:2000 implementation are excellent first steps to establishing a coordinated international quality control and assurance. OceanObs'09 and the discussion of data management for current and future systems provides the necessary foundations for nations to agree upon a QARTOD-like governance body which will ensure the accurate and reliable ocean observations for the next decade.

REFERENCES

1. Kearns, E., Woody, C., Bushnell, M. Eds., Final Report, Quality Assurance of Real-Time Ocean Data (QARTOD) First Workshop, National Data Buoy Center, Stennis Space Center, MS, 3-5 December 2003, 89 pp.

2. NOAA, Second Workshop on the Quality Assurance of Real-Time Ocean Data, Norfolk, VA, CCPO Technical Report, Series No. 05-01, 48 pp.

3. Final Report, Quality Assurance of Real-Time Ocean Data (QARTOD) Fourth Workshop, Woods Hole Oceanographic Institution, Woods Hole, MA, 21-23