**Alaska Ocean Observing System**

**DMAC Advisory Committee Meeting Minutes**

**Jan 24st, 2011 Anchorage**

*In attendance*: Phil Mundy (Chair, NOAA/NMFS/AFSC), Rob Bochenek (AOOS data team lead), Warren Horowitz (BOEMRE), Molly McCammon (AOOS), Peter Olsson (AEFF/State Climatologist), Scott Pegau (OSRI), Carl Schoch (AOOS), Tom Heinrichs (GINA), Igor Katrayev (NPRB), Angel Corona (NOAA/NWS), Darcy Dugan (AOOS), Allison Gaylord (BASC/Nuna Technologies), Steve Lewis (NOAA/AKRO), Stan Smith (USGS).

*By phone*: Emily Fergusson, NOAA/NMFS/ Auke Bay Lab)

**SUMMARY**

**Updates and Discussion Topics:**

* USGS and NMFS are switching from FGDC metadata standard to the new ISO 19115 metadata standard to be consistent with the international community. This involves a lot of difficulties, so the FGDC standard is highly likely to be retained until the FGDC to ISO conversion program is operational
* The AOOS data team has expanded its model explorer to include more models and allow users to drop virtual sensors and view/download time series data from models.
* The data team will be expanding these applications so that users can be notified when sensors reach specific thresholds of their choosing. The team anticipates cell phone apps of this nature in the next year.
* A question remains of how much real-time data should AOOS archive; what is our role in that; and how much do we make accessible?
* Either AOOS or GINA or some other entity needs to deal with the enormous issue of archiving shorter term project data. It is not easy dealing with project level data because there are so many considerations and constraints. AOOS is hesitant to take this on fully.
* It’s important to find leveraging opportunities with partners, but these can steer AOOS away from delivering the basics. A good topic for the next meeting is leveraging vs. staying the course.

**Recommendations:**

* The national IOOS office should provide instructions to standardize the visualization of IOOS core variables in order to make them uniform across regions.
* AOOS needs to pay attention to the guidance memorandum from national DMAC officials, and participate in the development of protocols that might eventually be mandatory.
* The committee, with help from AOOS staff and stakeholders, needs to prioritize key datasets for AOOS to provide. Categories include: models, remote sensing, real-time sensors, time series/GIS, and project data.
* More real-time sensors need to be added to the sensor map. A working group should be established to indentify and prioritize these sensors.
* To better understand how users interact with the maps and what their specific needs are, the data team should contact them by phone for a short interview rather than use online surveys.
* When deciding what datasets to archive, the data team should focus first on core variables that are not being archived. Avoid duplication.

**Requests:**

* A disclaimer needs to be added to the AOOS website regarding reliability of data and conditions for use of data.
* Rob Bochenek was asked to update the committee on the DIF at the next meeting.
* The NAM-WRF in the Model Explorer needs to be changed to WRF-ARW.
* The outline of the coast should overlay the data layers so the coastline is visible.
* The committee would like Axiom to develop a system for responding to the feedback provider about how the issue they identified was responded to.
* The committee would like a log-in to look at the feedback database.

**Action Items:**

* The committee accepted the minutes from the last meeting and asked for the addition of Rob Ragsdale’s presentation from the national IOOS DMAC to be included.
* The committee determined it will hold 2 meetings per year (Jan/Feb and June) except for this year when there will be three (Jan/June/Sept) as the new system ramps up.

**MINUTES**

**Director’s Update (Molly McCammon)**

AOOS staff have been busy working on proposals this fall and winter. After completing in October the 5-year funding proposal to NOAA for FY11-15, AOOS also submitted a proposal to fund its High Performance Computing Cluster (HPCC) through NOAA. In December, they also worked quickly to lead a last-minute partnership applying for NOAA’s National Ocean Partnership (NOP) funding opportunity focusing on Coastal & Marine Spatial Planning (CMSP). Other partners on the proposal included ACCAP, the ASLC, The Nature Conservancy, ISER, and Axiom Consulting. The total proposal was for around $2 million, and decisions will be made this spring or early summer.

McCammon reported on the regional ocean observing systems’ national meeting in November in DC. One of the outcomes from the meeting was an effort to synchronize regional websites with three minor additions which would be easy to implement. The regions are waiting to hear about funding for the upcoming fiscal year. There is a continuing resolution in Congress until March 4th. The rumor is that NOAA will be given 2010 funding with a 5% cut as well as no funding for earmarks. AOOS’s current grants end on May 1st and July 1st.

Rob Bochenek, the AOOS data team lead, has been staying connected with the other regions and the national IOOS program through conference calls every other week. They have been discussing interoperability, and sometimes guest speakers talk about specific products. Bochenek reported there seems to be disagreement between regional and national levels regarding specific data integration protocols. It is difficult to share code between regions for projects and tools because the tools aren’t always designed in ways that are easy to share. Bochenek is the only one using Adobe Flex while the other regions use JavaScript. They were surprised by AOOS’ use of Flex, but Bochenek explained they had expanded beyond the capacity of what JavaScript could handle. Bochenek has been actively participating during these meetings, and felt he was beginning to win over technical staff from other RAs by showing that AOOS is contributing. Bochenek plans to attend the annual NFRA technical conference in DC in April.

Carl Schoch voiced the need for a conversation at the IOOS level that establishes commonality in how the regions visualize core data. There is a set of core data that all RAs are required to produce but it is hard to navigate between regional portals since each is handled differently. McCammon noted that HF radar is already handled in a coordinated way through a common radar site.

*Committee recommendation*: Another step towards regional standardization should be a standardized protocol across the regions on how to visualize core data– common color palate, common template, wind icons, etc.

Chair Mundy asked for McCammon’s advice. She said this should be brought up through NFRA, and possibly at the next NFRA product workshop. Bochenek noted he is already receiving legend attributes through web services. Using THREDDS, the coloring for the relative magnitude for an attribute like waves is automatically in the same order, but the colors might correspond to different values depending on the typical size of waves in the region. He noted his wind visualization has strayed from the standard wind rose because he believes it does a better job of recording specific temporal data and allowing the user to track timing of winds instead of just looking at averages.

**Review of Previous DMAC Meeting**

Following up from last meeting, a disclaimer still needs to be put on the data portal. McCammon sent Bochenek several options, and Bochenek said he would implement one this week. Chair Mundy advised it should be something simple saying we are providing data developed by other entities and although we make every effort to reflect the data they gave us, it is up to user to do the due diligence of verifying quality of data with the original sources.

*Request: Implement a disclaimer on the AOOS website*

*Action Item: The committee approved the minutes from the last meeting and asked for the summary of Rob Ragsdale’s presentation from the national DMAC to be included.*

There was a brief discussion about the DIF. The DIF is a document developed by the national DMAC summarizing how core ocean variables should be collected and provides general high-level directives on how to correctly implement an interoperable data system. The group acknowledged we don’t really know what the DIF is or what it does. The previous person to work on this project was Rob Cermak, and it will now be handled by Bochenek.

*Committee Request: Chair Mundy asked Bochenek if he would make a presentation to the group about the status and functionality of the DIF at the next DMAC meeting.*

Chair Mundy flagged the IOOS DMAC implementation guidance document and noted that the group needed to keep an eye on it. There are a lot of directives that may one day be mandatory for AOOS to comply with.

**Update on Metadata from Emily Ferguson – NOAA Auke Bay**

Emily Ferguson (NMFS Alaska metadata lead) gave a brief presentation about the NMFS/USGS effort to transition from FGDC metadata to ISO metadata. NMFS and USGS have been using the FGDC standard for writing metadata which has worked well; there are a lot of tools for creating metadata and the interface is user friendly. However, there is a push to switch to the ISO 19115 metadata standard which is an international standard that Canada and many European nations are using. It makes sense to adopt this standard so all metadata is discoverable at the same level worldwide.

Ferguson was asked to participate in a 2-day workshop to develop an approach to training and creating materials for ISO metadata so the NOAA community can adopt and use standard curriculum to transition from FGDC metadata. Ferguson didn’t initially think it would be problematic, but got to the workshop and saw the differences in HGDC and ISO and realized it would be a lot of work. The workshop was led by National Coastal Data Development Center who was creating the workbook the group will use to write the ISO metadata. There were representatives from NOS, NGDC, OAR, Ocean Exploration and Research, USGS, BII, NESDIS, and a big group of metadata trainers and experts.

The take home messages:

1. It is not simple to move metadata from FGDC to ISO standard.
   * Someone has written transformations to transform FGDC into ISO standard but will take lots of proof reading.
   * An entirely different metadata set is needed for biological data.
   * All the files at Auke Bay will require 2 metadata files instead of one.
   * Many categories in the metadata are named slightly differently so metadata readers will have to re-teach themselves how to read a metadata file.
2. ISO does not have a nice user interface for creating new metadata.
   * One is an xml editor. Metadata authors will need to know how to write in xml which is a learning hurdle in itself. When the workbook and curriculum come out, there will be a big learning curve.
3. There are three programs, and two are not free.
   * XML Spy and Oxygen cost money. XML spy has good graphics.
   * Cat MD8 is free and has an interface that is good for writing new metadata but doesn’t allow for transformation. It also doesn’t support biological extensions yet.
   * You have to pay to get a copy of the ISO standard. The workbook being created will be free, and act similarly to the FGDC workbook which walks you through each element and what’s required. It will be user friendly once it is done and available, but is not yet available.

Scott Pegau asked when is this going to be an issue for AOOS -- are we writing metadata for records written by someone else or just pulling from them? Is this our problem or “their” problem? Bochenek explained there are different types of metadata. One is project level data; another is an interoperability service. By default, the interoperability service creates its own metadata. Bochenek thought it doesn’t come into play with some of the data we are exposing. If there are metadata protocols for IMS feeds, AOOS should probably comply with those.

McCammon mentioned that under the RCOOS authorizing legislation, there is a requirement for “certification” and there is a lot of ongoing discussion about what “certification” means. It was initially included in the legislation as a way to remove regional observing system management teams if they weren’t performing. Certification standards are still being developed (enumerating what’s required for users, board, stakeholder input); however, now the IOOS staff is going a step further and talking about data certification also. Is it data provider certification or data source certification? The regions feel that having to certify data discourages data sharing since there is a big difference between data that IOOS is producing vs. serving up data from others.

Chair Mundy noted he worked hard in his own agency to encourage scientists to add metadata records. NPRB requires scientists to do so. Mundy would not want to face a room full of scientists who are producing data and tell them they had to implement metadata to a new protocol such as ISO without being able to give them an easy way to do so. However, it was noted that it would be 10 times more effective for the PI to write the metadata than the person grabbing data through interoperability. Peter Olsson made the comment he wasn’t going to write metadata until someone mandated he do it. It is likely many others feel the same since most agencies are minimally compliant right now.

The conversation concluded in agreement that FGDC is an accepted usable standard and ISO is still experimental. The committee offered the following advice:

*Committee recommendation*: *It will be important to pay attention to guidance memoranda from national DMAC officials, and engage in its development. Metadata is just one example of the technical guidelines AOOS needs to be vigilant about. Chair Mundy asked that any time there is a movement on this issue that Bochenek notify the DMAC committee.*

There was brief discussion about additionally recommending that AOOS push back on the ISO standard, and that certification for metadata should accommodate FGDC and ISO in tandem. (current standard and experimental standard). This idea was floated by Tom Heinrichs. Heinrichs said GINA is investing a huge amount of effort in documenting metadata right now because they want the data used in 50 years still to be good. There was no firm commitment from the group on this item.

A question for AOOS is whether we want to work hard to serve data in compliance with the standard, or only serve a subset with the standard which are easy to do so, and maintain the goal of making as much data available as possible. A lot of ocean observing regions are taking the second approach: get as much data out there even if it is not interoperable and doesn’t meet standards. It was agreed that right now, AOOS is focusing on the data that will most help our users. At present, the AOOS system is providing mostly real-time data, but other time series and historical datasets will be added in the near future.

**Setting Annual DMAC Meeting Schedule**

Chair Mundy suggested developing an annual schedule that corresponds with the AOOS funding cycle, and serves the information needs of the Board so we can keep them informed about data matters in a timely way.

*Action Item: It was decided that the AOOS DMAC committee will have two meetings per year, synchronized with the data team’s 6 month development cycle in June and January, since having feedback around the two deployment dates would be useful. One meeting will follow the AMSS every year, and be held the last week of January or first week of February. The other will be in June. Pegau mentioned M-W-F works best for getting in and out of Cordova. This year, there will be three meetings since the system is in its early stages. Darcy Dugan will send out a Doodle poll to schedule a meeting for the week of June 6th or 13th and one for the week of Sept 12th or 19th.*

**Meeting User Needs**

*Discussion Question:* Who are the constituents for the website and how do we find them?

There are a lot of people who don’t access physical and atmospheric data now because they don’t know it exists. Chair Mundy highlighted an example about NOAA fisheries and ADFG biologists trying to predict run times and seasonality for salmon and how useful it would be to include AOOS data when making these predictions. Bochenek noted that they have devoted part of their budget to ADFG to get some of the commercial harvest data to be served up through interoperability protocols.

Pegau said this assumes the constituency is defined (which may or may not be true.) Users like easy interfaces. Pegau noted he personally prefers Weather Underground to the NWS because it’s easy to find things. The group discussed how to provide the missing pieces of information people most likely want: (1) identify the gaps and do straight value-added with existing data; then use specific projects to create value added products for defined constituencies; (2) instill confidence in the products – good metadata and reliability/consistency; (3) allow users to download their data in addition to visualizing it on the AOOS site.

Peter Olsson noted that people frequently come to him asking for historic wind data. He said it is very tedious to populate historical databases with weather data. It’s possible to do certain places for certain parameters, but it is piecemeal due to relocated or terminated observing platforms. Olsson would love to be able to find historical wind data on the AOOS website but questioned whether that was really feasible. Peter said he would be willing to work on behalf of the NOAA climate center and AOOS to rescue wind data, if funded.

It was agreed that it is hard to find historic weather data but easy to find the real-time data. If AOOS starts archiving the real-time data now, then in 20 years it will be useful as historical data. Many datasets are on an accessibility escalator and after a given time (like a year), they are no longer available from their source website. *Big question: how much do we archive; what is our role in that; how much do we make accessible?*

**Update from Data Team Lead (Rob Bochenek)**

Bochenek gave an update on recent developments from the data team. The biggest enhancement since the last meeting is the installation of a High Performance Computing Cluster which greatly increased the computing capacity of the AOOS system. The data team updated the Model Explorer application by loading more models and allowing the user to search by parameter, look at future and past forecasts with time slider, and drop virtual sensors to compare to actual sensors.

*Request: During the demo, Olsson noted that the name NAM-WRF should be changed to WRF-ARW. He also noted it’s hard to tell where the land is with the current color and overlay scheme, and suggested putting the coastline layer on top of the model data instead of the other way around.*

The committee discussed the need to be able to see longer time series of real time data (like a week) instead of just one day in order to compare real time sensors to virtual sensors. This brought up the question of whether AOOS should be storing and distributing the time series data, or whether that would be competing with entities like Snotel. There are pros and cons to both. To keep a good quality archive, it’s necessary to backfill monthly with QA/QC’d data. Not all entities store good data, and if it goes into NCDC, it can be so hard to get back out that it is often considered “lost”. The group discussed archiving the data that are not being archived elsewhere in an accessible format.

The same methodology used for models will be applied to satellite data. When loaded into the system, the Model Explorer will allow people to visualize satellite data to see primary productivity and other parameters. Satellite images will be assimilated into the AOOS system every 5 minutes, and Bochenek’s team will be able to providing batched information to users. Users will also be able to create their own virtual sensors and ask to be notified every 12 hours as they change. Similarly, they can ask to be notified when certain thresholds are reached, such as winds exceeding 20 m/s at specific location. The data team plans to develop iPhone apps which can track your geographic location and tell you the forecasted parameters based on where you are and what you have indicated you are interested in. Bochenek thought these capabilities could be launched within a year since the technology already exists.

Chair Mundy asked the group to focus on process -- how and why do we develop the tools? These questions should be approached from the user standpoint. Bochenek said he needs feedback regarding the new model explorer since they were developing it rapidly without outside input. This first application is just a teaser. Now they would like to figure out how to make it more useful.

The group highlighted five categories and discussed which were the priorities for AOOS, and what priority data should go into each category. (they are not in order of priority here).

1. **Models** - The model explorer is up and running which provides an excellent base to visualize additional models efficiently. Bochenek is interested in adding more oceanographic and atmospheric models -- forecasts, climatologies, and hindcasts. The list of forecasted variables can also be expanded. The lingering question posed by McCammon and Chair Mundy is to identify *who* wants them and *why do we want them*.
2. **Remote Sensing** - Remote sensing data is in the same format as models so it will be straight forward to plug remote sensing data into the existing Model Explorer framework. This will allow comparison between hindcasts and satellite images.
3. **Real-time sensors** – The system needs to continue adding available real-time sensors. A working group should be established to identify and prioritize these sensors.
4. **Long time series monitoring or GIS** –there is a serious need for AOOS to provide bathymetry, bottom character, climatological data, temporal series, ecological classification, species distribution and abundance, as well as large vector GIS datasets that are long term data syntheses. These data would be useful for a wide array of users. The data team has not tackled this category yet. Axiom has other projects on large GIS vector style datasets and there could be some good leveraging and synergy. In the next six months, Rob would particularly like to explore some of the oceanographic data. Adding a large volume of data is likely to complicate the user interface, but Bochenek said they have a design in mind to minimize that.
5. **Project data** – this is not the priority from the standpoint of the data management team, since it burns staff time, is difficult to automate, and will not be accessed by a broad user group. In counter, McCammon noted the need to accommodate data from PIs who don’t have anywhere else to put their project data. Pegau also brought up the EVOS herring program, noting that that would require data from many small projects and AOOS was submitting a proposal to do precisely that. The group decided to table the project-level data discussion for now.

Question: How do we prioritize data team effort in the categories above?

Pegau noted that if he wants to know current conditions, he doesn’t look for a model. He looks for real time sensors. Scott places much more emphasis on real time sensors than models or remote sensing. Mundy voiced that this shouldn’t be an either/or situation. The sensors are an important source of driving the models forward, and they are ephemeral. If we have a user base, we should archive the sensor data. Users should be able to archive the sensors they are interested in using a MyAOOS application that is under development. Pegau noted that this isn’t currently possible but there is demand for it – he said that many people ask him where they can find historical data for Cordova. He can point them to 12 months of data easily, but it’s very difficult to track down the data prior to that. The question resurfaced about “archiving” data, which would then require the data to be quality controlled. This is not likely for AOOS in its current capacity.

Dugan mentioned we are getting 60-70 hits a day on the AOOS website right now and asked if it would be useful to put out a survey of existing users to see what they are currently accessing, and why.

*Committee Recommendation:* Committee members suggested targeted conversations instead of online surveys.

Pegau noted that our users are changing and we can’t know right now who the users will be in 15 years. If we begin archiving real time sensors today, it will allow us to present the data in a more unified manner in the future. If information is already stored someplace accessible, we should simply point users there. The data that is not made accessible should be archived by AOOS.

*Committee Recommendation*: Focus on core variables that are not being archived, and begin archiving them. Avoid duplication of datasets that are already being well archived.

Bochenek wants users to have both broad and granular access to data. Nightly, their system could grab the last day’s data from all the sensors and load them into database. However, that would make Snotel and other sensors obsolete. Do we want to do that? Chair Mundy and some other committee members thought no. Olsson said Rick McClure with Snotel would be delighted for AOOS to serve the data. They just want the data to be used. Peter thinks backfilling for some reasonable amount of time once you get the quality-controlled data should be fairly simple. From the AOOS perspective, freshwater input (from the Snotel stations) is critical, and storing the data is important since it’s the best freshwater data we have.

🡪 Either AOOS or GINA or some other entity needs to deal with this huge data archive/project issue. It is not easy dealing with project level data because there are so many considerations and constraints. GeoSpatial One Stop is doing this partially. Chair Mundy said he submits all the projects annually and it’s all there. You get a little map with the coordinates and abstract, which says who has the data and a link to where to get it.

*Request*: The committee would like Axiom to develop a system for responding to feedback providers about whether the issue they identified was responded to and how.

*Request*: The committee would like a log-in site to look at the feedback database.

*ISSUE: It’s important to find leveraging opportunities but these can steer AOOS away from delivering the basics. This is a good topic for the next meeting: leveraging vs. staying the course.*

Bochenek’s examples of a positive leveraging opportunity is with the pelagic sea bird database. The database includes 400,000 pivoted observations and a grid of densities. Right now there is no option for serving up the data because it’s so huge, but could be done through a partnership with AOOS.

*Request*: Dugan will resend the security plan to the group. Members can provide comments to Mundy and Dugan.

*Request*: Dugan will send out a Google document where people can populate the best, most robust datasets for the core variables.

**Meeting adjourned at 2:45pm.**

A small group continued the conversation after 3pm, including Rob Bochenek, Molly McCammon, Phil Mundy, Scott Pegau, Carl Schoch, Steve Lewis, Allison Gaylord, Tom Heinrichs, and Darcy Dugan. Discussion focused around the AOOS data portal that is under development, and how to change the existing “access data” interface. The group felt that having the four applications tiled on the screen made it look like that was the extent of the AOOS data system, and was confusing to a viewer who might not know which application to enter. They suggested sending the user straight into the real-time sensor application with other applications available from that window. McCammon emphasized the need for a status update to be posted on the transition data portal, indicating when the next developments would take place and what they would include.