

AOOS

Alaska Ocean Observing System
Arctic Animal Telemetry Network
March 25th, 2013, Anchorage

In Attendance: Molly McCammon (AOOS) Rosa Meehan (ArcticTurn) George Divoky (Friends of Cooper Island) Rob Bochenek (AXIOM) Mike Rassur (BOEM) Cathy Coon (BOEM) Will Koeppen (AXIOM) Ellen Tyler (AOOS) Jeff Denton (BOEM) Russ Andrews (UAF, ASLC) Andy Sietz (UAF)

On the phone: Hassan Mustahfid (IOOS), Phillip Goldstein (OBIS USA), Bruce Mate (Oregon State U), Dan Costa (UC Santa Cruz), Joel Garlich-Miller (USFWS), Josh London (NOAA/NMML), Amy Kennedy (NOAA/NMML), Mike Weise (ONR), Samantha Simmons (MMC)

Introduction – Molly McCammon

Molly gave a brief overview of the Alaska Ocean Observing System and how facilitation of an Arctic Animal Tagging Network (AATN) fits into the goals, mission and philosophy of AOOS. Her PPT is linked below.

- Goals: To provide easy access, integration and translation of marine observations into information products and decision support tools
- Mission: Congressionally directed, stakeholder driven and science based
- Philosophy: Map once, use many times
- Primary Activities to date: Hosting and integrating data via web-based open-source tools like the real-time sensor map, model explorer, and asset maps; supporting some observations, primarily by fostering collaboration and leveraging resources; providing tools and support for stakeholders
- Future activities: Arctic Observing System Build-out, a statewide modeling plan and greater attention to complementary system of marine observations.
- Things AOOS does not do: archive data, conduct hypothesis-driven science

Molly described National IOOS efforts to date, both with telemetry data specifically and broader data management including development of: protocols, standards and interoperability services. She mentioned two unique aspects of this Arctic network: there are very few existing instruments or platforms, and there is a big international component.

Molly framed the goals of an AATN effort in two phases:

- Phase I – to explore interest amongst stakeholders and funders, describe data sharing repositories, and support research and management efforts
- Phase II – Identify potential integrated/synthesized projects

Discussion

Currently there are several factors driving interest in data synthesis projects. Scientists are interested in greater understanding and discovery of patterns and trends, and data accessibility standards are increasingly being levied on researchers and data collectors as a requirement for (federal, and even some foundation) funding.

- One concern is that several separate data integration projects are spinning up and ultimately creating redundant systems.
- Redundancy is not bad, so long as data is interoperable and comprehensive.
- There may be an appearance of redundancy, but each portal is responding to distinct and diverse needs.
- Goal is for data systems to be complementary, which is challenging in practice and requires communication.
- There are lots of portals being built, but most of them are just repositories—most don't focus on the needs of stakeholders collectively and do what is needed in terms of integration and facilitating analysis across different scales.
- If we can recover resources from what redundancy exists now, we could repurpose and leverage those efforts towards a systems approach to telemetry data integration.
- Ideally, stakeholder needs and desired end products could inform how (and what) data is collected and managed.

Agency Needs and Desired End Products

USFWS (Joel Garlich-Miller)

- Quick/easy visualization of spatial information (habitat needs) to evaluate potential impacts and make agency decisions (pulling this out of scientific publications is time-consuming and difficult and there can be a significant lag time between data gathering and publication).
- Spatial and temporal use by species
- Ability to overlay various stressors
- Increased transparency within research community so that data collectors can cater collection strategies to researcher needs
- Information for contingency planning and disaster response

Navy ONR (Mike Weise)

- Navy is interested in marine mammal behavior and wellbeing based on federal requirement to monitor marine mammals in relation to their navy exercises
 - Particularly interested in physiological and behavioral responses to sound
- Arctic and Ocean Environment program has goals of “Safe, Stable, Arctic”

- Interested in investigating new approaches for measuring Arctic environment such as animal telemetry
- Naval Oceanographic Office (NAVO), IOOS and TOPP successfully collaborated on a tagging project with distribution of real-time and historic data to NAVO modelers to demonstrate capability, but they had trouble finding funding or leveraging resources.
- Need to collaborate more closely with oceanographers
- Need one-stop-shop for tag data

BOEM (Jeff Denton)

- Need science/observations to inform Environmental Impact Assessments and management decisions - driven by NEPA analysis framework
- Data must be highly scalable because BOEM is responsible for a wide range of activities from issuing site-specific permits all the way through development of multi-year leasing program planning
- Challenges: integrating across information types/sources (local knowledge, industry data, interagency)
- Interested in the potential to understand cumulative environmental stressors via animal telemetry
- Need consensus among PIs as analysis occurs
- Need derived products such as NMML's habitat polygons

Discussion

It is important to distinguish between raw data and synthesized products. Raw data can be released earlier if PIs are willing, but it can be messy and agencies are often interested in the synthesized products. Need to look at how physical oceanographers deal with this – they often release raw data without any QA/QC with a “user-beware” disclaimer. Synthesized products are usually developed for a specific question or audience and may be static in one space/time. In the Arctic, time is a proxy for changing conditions. Could the AATN contribute to modeling efforts to hind-cast physical oceanographic parameters based on greater understanding of animal behavior?

Existing programs: and overview of TOPP (Dan Costa)

Dan shared his experience building and implementing the Tagging of Pacific Predators (TOPP) research program, which was a technology-driven component of the Census of Marine Life, funded by the Sloan Foundation. Ultimately the program deployed >4,000 tags on 23 species. This led to important scientific discoveries about how different animals occupy different niches within an ecosystem, and showed some environmental basis for animal migratory behaviors in the North Pacific. His PPT contains many examples of spatially explicit research findings and is linked below.

Dan talked a bit about technical differences across tags- such as measurement period; research length and seasonality; and frequency of data collection. They were lucky that the Sloan, Moore and Packard Foundations were interested in

funding many of the tags. Additionally each of the collaborators secured some additional grants and through Census of Marine Life, the community worked together to leverage limited resources, agreeing to contribute to one central data integration archive. He stressed that one of the best things about the project was having this human network to bounce ideas off of and go through the mechanics of putting physical data together with biological information. Synergy came from having funding initially. The program was structured with 4 PIs, science oversight group, and then working groups by species guilds. There has not been any formal quantification of the benefits of this collaboration, but none of the multi-species papers referenced in his slides would have happened otherwise.

Existing Programs: AOS Research Workspace (Rob Bochenek)

Rob talked about his experience managing a wide variety of data types collected under various techniques throughout the AOS system. His efforts to date have focused on integrating and upscaling some data sets while also co-hosting “odd man out” datasets that don’t line up exactly with other datasets. Looking at the spectrum of telemetry data collection protocols and platforms, he thought it could be useful to set up an Arctic Animal Tagging Network to better understand what actual data exist, and what steps would be needed to transition to a more structured system for data integration. He drew a distinction between existing systems like TOPP, which are focused on making data available as final products, and the AOS system, which is serving up raw data and tools for analysis.

Rob walked the group through an existing “Research Workspace” tool on the AOS website which allows research collaborators to document and stage data. Here, someone with tag data can:

1. Create a project: first by describing the taxa, geographic area, and research abstract; then by creating folders to drag and drop data, formal metadata and informal workflow processes
2. Invite collaborators: invite new members to the project. New members can access and share data and files, compiling a spectrum of available data, sorting through data sets that can be integrated into particular research questions, describe what and why some data sets may not fit in initially, and review what project collaborators are doing to manipulate raw data.
3. The Workspace automatically generates metadata. AOS is currently establishing mechanisms to publish data from the workspace and automatically archive it to NODC and any number of archival systems.

The Workspace may be an ideal location to build a network to develop shared understanding within a secure data-sharing environment. There was some discussion about options for controlling access and securely sharing data within a community of experts. The TOPP system established firewalls both because they could (being funded largely by foundations and before mandates for open data access were so pervasive) and because of the nature of the data synthesis tasks they were performing.

Existing Programs: OBIS USA Data Enrollment Process (Phillip Goldstein)

Phil walked through the US node of the Ocean Biographic Information System (OBIS) federal data archive and analytics. His slides go through a step-by-step process for data enrollment and are attached below. He made three salient points to keep in mind as this network comes together:

- It takes resources to be able to support and integrate different data types.
- If you are reusing existing data types, don't re-import them; integrate with them remotely on a standards-interface basis.
- If you are innovating new data types or extending features of existing data types, get funded for it. It will require effort and this effort will contribute value to your community and others.

Hassan Moustahfid stressed the importance of learning from other systems and identified a need for an inventory of tools for operational use. There was discussion about a potential role for tag manufacturers to help with this or even getting someone like Rob to help develop approach to make the data gathered and compiled properly upfront so that it is easy to capture and share, and so that it is compliant with existing and developing data standards.

Data Gatherer's Perspective (Josh London)

Josh began by acknowledging that researchers who collect raw data need assistance getting that raw data into useable forms. But just because tools exist doesn't mean that researchers are going to find them- or know how to use them. He suggested that a key priority for the Arctic Animal Tagging Network be a clear and concise write-up of data management methods, establishing a clear process for scientists to communicate and develop systems.

Biologists are not trained in large-scale data management. Overall, the field has limited programming skills, primarily relying on "mouse click" type interaction. Biologists are sometimes hesitant to even share data because of worries about data quality, and even errors, which are non-critical to their research, but could be critical under conditions of some other research question. Data collection and study design is sometimes so nuanced that researchers are hesitant to separate their specific study information from its data. Researchers see benefits of networking and sharing data, but note that there are a lot of operational barriers. Importantly, at the end of the day, time spent developing skills and training for effective, scalable database design and data management practices is time that biologists are not being biologists.

Agencies have limited capacity for data consultation and data base development either. Agencies are constrained by:

- Opportunistic or ad-hoc database design and management by the least skilled users
- Limited funds, especially as agency internal needs compete with agency mandates
- IT policies and investments vary across institutions making collaboration difficult
- No standards for sharing “raw” data
- Limited resources to qa/qc data; yet ultimate responsibility if posting for other users

Josh gave a personal example of the limitations of his current approach:

- Oracle framework in use (others are in excel)
- Single point of access failure (input limited to Josh)
- Limited documentation of design
- Design may or may not be the best.
- Main objective is to provide data to analyst

And he identified internal and external priority needs for improvement:

(internal)

- Data management and design consultation
- Data design and documentation
- Low tech exploration tools
- Database and application developers
- Training opportunities

(external)

- Meaningful public access
- Clear data sharing policy with best practices
- Encourage/facilitate scientific collaboration
- Meet agency needs and requirements
- Communicate scientific knowledge in the modern/digital age – sharing knowledge and expertise just as important as sharing data
- Publish data once

Phillip Goldstein commented that the info about tag model choice or programming should be captured in the Metadata – but that is not possible now with the current standards, so we need to raise this issue higher up to the Metadata standards decision makers.

Dan Costa pointed out that the Southern Ocean Observing System did propose to SCOR an “Essential Biological Variables” standard (actually it was for a SCOR working group to develop “Ecosystem Essential Ocean Variables (eEOV) for measuring change in the biological properties of marine ecosystems.”)

Science Questions – Russ Andrews and Rosa Meehan

Rosa and Russ facilitated a discussion about what sort of science question the network could start working on as a sort of “proof of concept” to begin sharing arctic tagging data. The purpose of this network is to:

- Increase coastal and ocean observation
- Facilitate data access
- Enhance collaboration
- Enable broadly synthetic studies

There was interest in the question of how animals respond to ice melt. An important goal will be to look at many species together to search for commonalities on how species are responding. While TOPP initially just focused on tagging and tracking data, ultimately they began integrating other information. Today, knowing what we know now, they would start with bigger questions, incorporating other types of animal observation data alongside the tag-collected data, right from the get-go. Russ and Rosa shared a list of preliminary questions in their PPT, posted below.

Within the principal interest of seasonal and long-term adaptation to changes in ice, the following criteria was identified to guide the research:

- Immediately opportunistic funding
- Animal-borne sensor data of value to oceanographers defines “telemetry” as primary source of data
- Multiple data collectors
- Multiple species – system type focus
- Addresses management focus
- Geographic focus – Beaufort and Chukchi

Approach

- Increase access and sharing of existing data
- Integrated research project – build in increased access to telemetry data
 - Derived product/s would be good to identify – then work backwards to determine how to get there
 - Can be pitched to funders

Possible pilot project - looking at bowheads, eiders, bearded seals and other benthic feeders would meet the interests of management, research and operational interests, and have added benefit that all of these animals have already been tagged.

Products

More than the scientific question itself, people seemed interested in becoming more efficient in their own efforts to share and integrate data. One really beneficial result of this network will be to identify what biological ocean observing data best meets the needs of multiple users.

Funding was identified as a need to provide an incentive for participation. Upcoming funding opportunities could include– MARS; NPRB/NSF; ONR; IOOS sensor development; AON

AON –

- Idea of what is a network
 - What is a network to include
 - Data management
 - Communication
 - Need group – white paper to push forward
 - Include data challenges Josh identified as justification
 - What resources are available for this?
 - Identify all of the groups out there and pull information together
 - Organize all the ways similar problems of data sharing were solved
 - Interoperability feeds
- Explore funding opportunity to pull together available data via Research Coordination Network proposal (NSF)

Phase II

- Outline for project
- Group and ID candidate species
- Cross discipline
- Need to be prepared to respond to funding opportunities

Next Steps:

- Develop a white paper/outline to use as a template
 - Scientific motivation
 - Bruce, Dan, Josh, Russ, Rosa

Come back together as a group/network when the MARES FFO comes out – likely in June.