

FY16 PROJECT PROPOSAL SUMMARY PAGE

Continuing, Multi-Year Projects

Project Title: Collaborative Data Management and Holistic Synthesis of Impacts and Recovery Status Associated with the Exxon Valdez Oil Spill

Project Period: February 1, 2016 – January 31, 2017

Primary Investigator(s): Matthew B. Jones, National Center for Ecological Analysis and Synthesis (NCEAS)

Study Location: Entire Exxon Valdez oil spill affected area

Project Website (if applicable): www.gulfwatchalaska.org, <https://goa.nceas.ucsb.edu/>

Abstract*: The AOOS-led Long-Term Monitoring (LTM) and the PWSSC-led Herring Research and Monitoring (HRM) programs propose an ambitious monitoring and research agenda over the next five years. These efforts could facilitate a more thorough understanding of the effects of the oil spill if the new data and information on the spill-affected ecosystems are effectively managed and collated along with historical data on these systems, and then used in a comprehensive synthesis effort. We propose a collaboration among NCEAS and the AOOS LTM and HRM teams to help build an effective data management cyberinfrastructure for proposed monitoring efforts and organize these data with historical data, including previous EVOSTC-funded efforts, to prepare for synthesis and ensure all data are organized, documented and available to be used by a wide array of technical and non-technical users. Building on the LTM and HRM syntheses and modeling efforts and the 20-year historical data from EVOSTC projects and any available current data, NCEAS would convene two cross-cutting synthesis working groups to do a full-systems analysis of the effects of the 1989 oil spill on Prince William Sound and the state of recovery of the affected ecosystems.

Estimated Budget:

EVOSTC Funding Requested* (must include 9% GA):

FY12	FY13	FY14	FY15	FY16	TOTAL
\$416.8	\$464.7	\$372.1	\$379.2	\$73.9	\$1,706.7

Non-EVOSTC Funds to be used:

FY12	FY13	FY14	FY15	FY16	TOTAL
\$0.0	\$0.0	\$0.0	\$0.0	\$0.0	\$0.0

* Funds expressed in \$1000 increments

Date: September 1, 2015

I. EXECUTIVE SUMMARY

In the two decades following the *Exxon Valdez* oil spill (EVOS), and after extensive restoration, research, and monitoring efforts, it has been recognized that full recovery from the spill will take decades and requires long-term monitoring of both the injured resources and factors other than residual oil that may continue to inhibit recovery or adversely impact resources that have recovered. Monitoring information is valuable for assessing recovery of injured species, managing those resources and the services they provide, and informing the communities who depend on the resources. In addition, long-term, consistent, scientific data is critical to allow us to detect and understand ecosystem changes and shifts that directly or indirectly (e.g. through food web relationships) influence the species and services injured by the spill.

An integrated monitoring program requires information on environmental drivers and pelagic and benthic components of the marine ecosystem. Additionally, while extensive monitoring data has been collected thus far through EVOS Trustee Council-funded projects as well as from other sources and made publicly available, much of that information needs to be collated and assessed holistically to understand factors affecting individual species and the ecosystem as a whole. Interdisciplinary syntheses of historical and ongoing monitoring data are needed to answer remaining questions about the recovery of injured resources and impacts of ecosystem change.

Data collected prior to and in response to the Exxon Valdez oil spill are profoundly heterogeneous. They range from long-term, automated sensing of oceanographic and atmospheric conditions, to short-term, experimental, monitoring, and behavioral studies of biological components of the system. The scientific data to be collected in these studies includes data on population trends, behavior, physiology, disease, and genetics of many species, as well as oceanographic and meteorological data at both regional and local scales. This diversity of data and data collection protocols substantially complicates data management by EVOSTC long-term monitoring projects. In addition, investigators on both the long-term monitoring and herring population studies are affiliated with many different institutions and agencies, each currently collecting data from many sites within the spill region and managing it within the frameworks dispersed among these agencies. Any data management system will necessarily need to accommodate this heterogeneity and dispersion by preserving the original data and providing mechanisms to access, integrate, and analyze the data for crosscutting synthesis. Data management activities for oceanographic information occur in isolated, physically distributed agencies, leading to low cross-agency utilization of data. Technical barriers, complex data formats, a lack of standardization and missing metadata have limited access to data and made the utilization of available scientific information cumbersome and daunting. As a consequence, existing data is underutilized and often has not undergone quality assurance.

In this proposal, we outline the collaboration between the National Center for Ecological Analysis and Synthesis (NCEAS), the Alaska Ocean Observing System (AOOS) and their partner Axiom Consulting, and the investigators of the pending Long Term Monitoring (LTM - proposal submitted by McCammon et al.) and Herring Research and Monitoring (HRM – proposal submitted by Pegau et al.) programs. This project will augment the expertise in data management and synthesis of these groups to maximize the efficiency of data collection and management for the LTM and HRM programs and expand access to these data, collate additional historical data that are useful for synthesis from the EVOS affected area, and conduct a broad-ranging synthesis of twenty years of EVOSTC funded research data to generate a comprehensive assessment of ecosystem impacts and recovery status for the spill affected area.

This collaboration document augments the data management, infrastructure development, and synthesis activities previously proposed by the AOOS partners with additional objectives that introduce new

technologies from NCEAS to jointly improve the data management infrastructure available to researchers, broaden the scope of data collation and integration, and embark on an ambitious synthesis plan (Figure 1). During the first three years, NCEAS has focused on mining historical data and contributing to development of both the AOOS cyberinfrastructure and the DataONE Federation infrastructure in order to create the necessary data resources for synthesis; during years 3-5, NCEAS has been conducting a multi-year working group effort to synthesize what is known about spill effects and recovery of ecosystems. These activities will be interwoven with the complementary but distinct data management, technology development, and analysis activities previously proposed by Axiom and AOOS and which are referenced in the objectives below.

II. COORDINATION AND COLLABORATION

A. Within an EVOTC-Funded Program

- Alaska Ocean Observing System (AOOS), Axiom Consulting: NCEAS project participants have coordinated cyberinfrastructure development to link the AOOS/Axiom data management system to broader community infrastructure, including the DataONE project. This work will be completed in FY16 with the full integration of the historical data portal to AOOS to operate as a DataONE member node.
- Gulf Watch Alaska and Herring Research and Monitoring (HRM): NCEAS has worked with all current Gulf Watch Alaska and Herring Research and Monitoring researchers to collate, document, and archive data collected under these programs, and to make this data widely available to the research community.

B. With Other EVOSTC-funded Projects

Numerous historically funded EVOS projects: NCEAS has worked with researchers from historically funded EVOSTC projects to collate, document, and archive data collected under these programs, and to make this data widely available to the research community. This collaboration has allowed us to collate a rich set of data arising from past Trustee project funding.

C. With Trustee or Management Agencies

This proposal was originally in response to a request for involvement in data management activities from Elise Hsieh, with the stated goal to connect the data management activities of the project to other activities in which we are involved, including DataONE and related efforts.

Conduct of the synthesis working groups described below is done in collaboration with the National Center for Ecological Analysis and Synthesis, as well as many participating scientists that volunteer their time to work on integrative synthesis. These collaborations are highly productive because of the diversity and intensity of the participating scientists. NCEAS working groups under this project are led by community researchers, including:

- Predicting CGoA dynamics Working Group
 - Thomas A. Okey, School of Environmental Studies, University of Victoria
 - Terrie Klinger, School of Marine and Environmental Affairs, University of Washington
 - James J. Ruzicka, Cooperative Institute for Marine Resources Studies, Oregon State University
- Gulf of Alaska portfolio effects Working Group
 - Kristin Marshall, School of Aquatic and Fishery Sciences, University of Washington
 - Anne Beaudreau, School of Fisheries and Ocean Sciences, University of Alaska Fairbanks

- Richard Brenner, Alaska Department of Fish and Game Division of Commercial Fisheries
- Mary Hunsicker, National Center for Ecological Analysis and Synthesis
- Eric Ward, NOAA Northwest Fishery Science Center
- A. Ole Shelton, NOAA Northwest Fishery Science Center

III. PROJECT DESIGN – PLAN FOR FY16

A. Objectives for FY16

Objective 1. Provide data management oversight and services for EVOS LTM project team data centric activities that include data structure optimization, metadata generation, and transfer of data between project teams.

Status: Completed during years 1-3, but with a small emphasis during years 4-5 on ongoing consultation. Dependent upon continued collaboration by other project participants.

Objective 2. Consolidate, standardize and provide access to related and historic data sets that are critical for retrospective analysis, synthesis and model development within the LTM and HRM programs.

Status: This activity is complete from years 1-3 (recoverable data sets have been collated, but many remain that are non-recoverable), but will continue to gather specific data sets that target specific data needs for synthesis working group activities in years 4-5. A new data recovery effort is described in Changes to Project Design.

Objective 3. Develop tools for user groups to access, analyze and visualize information produced or processed by the LTM and HRM efforts.

Status: Completed, in that we have developed and released analysis and visualization tools in R that provide direct access to the historical data collected and deposited in the AOOS member node of DataONE.

Objective 4. Organize, integrate, analyze, and model the 20-year historical data from EVOSTC-funded projects and other monitoring in the spill area in preparation for LTM and HRM program and NCEAS working group synthesis efforts

Status: Historical data are now documented and organized through the single GoA Data Portal, and are accessible via analytical scripting systems such as R. The Projects Data Coordinator has worked on integration of key data sets, including ADCP data across the study area, seabird data, and others. Results of these integrations are also being pushed to the GoA member node as derived data products that can be reused in synthesis efforts. Some of these products are being further developed by the synthesis working groups that started work in year 4 and continue through year 5.

Objective 5. Integrate all data, metadata and information products produced from this effort into the AOOS data management system for long-term storage and public use.

Status: All historical that have been collated to date have been archived in the AOOS data systems, and additional data from synthesis activities are added as they are identified,

collated, and documented. The NCEAS Data Coordinator (Couture) continues to work with groups to produce syntheses of the various historical and contemporary data sets. Emphasis has shifted in year 4 and 5 to focus on data integration needed for the two NCEAS synthesis working groups.

Objective 6. Augment AOOS/IOOS preservation and interoperability system with other data systems through integration of DataONE services.

Status: The Gulf of Alaska Member node has been developed, tested, and deployed (<https://goa.nceas.ucsb.edu>). All historical data that have been collated to date have been archived in DataONE and replicated across multiple nodes, and additional data are added as they are identified, collated, and documented. Historical data can be accessed at <https://goa.nceas.ucsb.edu> in addition to the AOOS site and the DataONE data portal (<https://cn.dataone.org>).

Objective 7. Conduct broad synthesis activities on spill impacts and recovery as part of whole-ecosystem analysis through NCEAS working groups.

Status: Two working groups have been selected and developed in year 3, and are conducting synthesis activities in years 4-5. Two postdocs, Rachael Blake and Colette Ward, were recruited and started in January 2015 to work with these working groups.

B. Changes to Project Design

While the overall objectives have not changed, some of the activities have shifted in time due to the delayed selection of working groups and postdocs, and to reconsideration of project priorities. These changes come in three principal areas: 1) delayed synthesis working groups and postdocs; 2) renewed effort to integrate DataONE support in AOOS, and 3) additional historical data recovery.

1) *Delayed synthesis groups and postdocs.* Synthesis groups were originally scheduled to hold two meetings in each of project years 3 and 4, but due to an extended selection and peer review process, these groups did not convene until the beginning of year 4. Each of the synthesis working groups held their inaugural meeting in the first half of year 4, and has their second meeting is scheduled for September and October of year 4. The third and fourth meetings are scheduled for year 5. Similarly, both postdoctoral fellow positions were shifted to years 4 and 5 to support and expand working group synthesis activities. While the schedule has shifted, the deliverables in terms of synthesis papers has not changed. The budget for the year 5 request has not changed, but we will plan to roll funds forward from years 3 and 4 to cover these additional synthesis and postdoctoral costs in year 5.

2) *Renewed effort to integrate DataONE support in AOOS.* The Gulf of Alaska Member Node was established as part of the DataONE federation and deployed in year 3, but is currently still a separate part of the infrastructure than other services that AOOS maintains. Thus, there are two separate data portals, one within AOOS for current monitoring, and one at NCEAS holding historically collated data and data from synthesis activities. The historical data was manually deposited in the AOOS Ocean Workspace, but this process does not allow automated synchronization between the data systems. NCEAS, AOOS, Gulf Watch Alaska, and Axiom met to discuss how to remediate this, and have established a two-stage plan to eliminate this redundancy. First, Axiom will work in the remaining months of Year 5 to automate the process of ingesting metadata and data from DataONE compatible data repositories, which will streamline the process of ingesting data from the NCEAS synthesis efforts, as well as new historical data that are rescued (see next section). NCEAS will provide a part-time software engineer to help design and implement this effort and consult on the DataONE system.

Second, in year 5, we will use remaining rollover funds for software engineering from years 4 and prior to help Axiom to develop a full implementation of DataONE member node within the AOOS-maintained systems. This will include: 1) supporting the REST APIs from DataONE directly within Axiom, and 2) both ingestion and production of standard metadata and data formats used in the DataONE federation (focused mainly on ISO 19139 and Ecological Metadata Language). Time for this joint design and development activity has been budgeted in year 5 by utilizing rollover funds from previous years. When complete, the current historical data portal will be shut down and replaced by an AOOS DataONE member node.

3) *Additional historical data recovery.* The original historical data recovery effort in years 1-3 identified data collected from all prior EVOSTC-funded projects, prioritized these, and attempted to recover and the data through a series of email and personal contacts with the original investigators. Currently, 27% of these identified historical data sets have been recovered and archived, with the remainder still unrecovered due to several reasons, including retirement of key personnel, destruction of the data, unresponsiveness of key personnel, and at times the amount of labor required to digitize original paper files and notebooks. We do not feel that additional data can be recovered using the approach to email and phone contacts that we employed in years 1-3. However, in discussions with Gulf Watch Alaska leadership, we have agreed that recovering additional high-priority data sets would be extremely valuable to synthesis efforts and to long-term, retrospective analysis. There was general agreement that GWA leadership could assist in the recovery effort by utilizing their personal networks at the various state, federal, and university organizations that hold the data. Thus, we plan to use rollover funds from years 1-4 to restart a data rescue effort to obtain, document, and publish additional high-priority data sets from prior funding and from related efforts in the region. This work will be accomplished in phases. First, the existing listing of historical data will be reviewed and organized, with explanations provided on the status of each data set and the potential effort involved and barriers to recovery. Second, this listing will then be reviewed and triaged by GWA leadership to identify the high priority targets, and those will be examined to determine which are both tractable and amenable to alternative approaches for recovery. For example, lack of time from agency personnel was a large barrier to recovery for many data sets, and we will utilize high-level contacts within those agencies to raise these items as priorities for agency staff. Third, the Projects Data Coordinator (Couture) and two student interns at NCEAS will correspond with these agency staff to identify, obtain, understand, document, and deposit the data in the Gulf Watch repositories. Upon review and approval of the original data creators, we will publish the data sets with citable identifiers and make them available to the broad community of researchers and the public. These data, like the other historical data recovered, will be deposited in a DataONE compatible repository, and replicated at multiple sites to ensure preservation and accessibility. Funds for these students and the additional time for the data coordinator will be drawn from rollover funds from prior project years.

IV. SCHEDULE

A. Project Milestones for FY 16

Objective 1. Provide data management oversight and services for EVOS LTM project team data centric activities that include data structure optimization, metadata generation, and transfer of data between project teams.

Completed during years 1-3.

Objective 2. Consolidate, standardize and provide access to related and historic data sets that are critical for retrospective analysis, synthesis and model development within the LTM and HRM programs.

Completed during years 1-3. A new data recovery effort is described in Changes to Project Design for year 5. To be met by January 2017.

Objective 3. Develop tools for user groups to access, analyze and visualize information produced or processed by the LTM and HRM efforts.

Completed in years 1-3.

Objective 4. Organize, integrate, analyze, and model the 20-year historical data from EVOSTC-funded projects and other monitoring in the spill area in preparation for LTM and HRM program and NCEAS working group synthesis efforts.

Completed during year 4. Some of these products are being further developed by the synthesis working groups that started work in year 4 and continue through year 5.

Objective 5. Integrate all data, metadata and information products produced from this effort into the AOOS data management system for long-term storage and public use.

Completed for historical data in years 1-3. Additional data to be deposited in year 5 after second recovery effort. To be met by January, 2017.

Objective 6. Augment AOOS/IOOS preservation and interoperability system with other data systems through integration of DataONE services.

Completed deployment of historical data node for DataONE in years 1-3. Additional integration planned for AOOS systems in year 5. To be met by December, 2016.

Objective 7. Conduct broad synthesis activities on spill impacts and recovery as part of whole-ecosystem analysis through NCEAS working groups.

In progress during years 4-5. To be met by January, 2017.

B. Measurable Project Tasks for FY 16

FY 16, 1st quarter (February 1, 2016 - April 31, 2016)

*February: Assess year 1-4 datasets and metadata historical data collations, prioritize new collation
Completed consultation with Axiom on data and metadata ingest from DataONE*

March: Continue synthesis group analysis activities

April: Completed contact with priority historical data PIs and organizations

FY 16, 2nd quarter (May 1, 2016-July 30, 2016)

*July: Completed the 3rd working group meeting for each of two NCEAS synthesis groups
Collated data from high and medium priority historical data projects
Start metadata generation for historical data, and publish after each is reviewed*

FY 16, 3rd quarter (August 1, 2016 – October 31, 2016)

*October: Completed the 3rd working group meeting for each of two NCEAS synthesis groups
Complete metadata generation for historical data, and publish after each is reviewed*

FY 16, 4th quarter (November 1, 2016- January 31, 2017)

November: Participate in LTM program PI meeting

*December: Completed the 4th working group meeting for each of two NCEAS synthesis groups
Completed consulting with Axiom on design and deployment of an AOOS Member Node*

*January: Complete publication of any final historical data sets
Submit any remaining synthesis group papers and products for publication*

V. PROJECT PERSONNEL – CHANGES AND UPDATES

No changes to senior personnel.

VI. BUDGET

A. Budget Forms (Attached)

Attached.

B. Changes from Original Proposal

Our original budget request for FY16 has not changed, and remains at \$67,766 for the year 5 effort.

To accommodate the schedule shifts and changes to project priorities above, we plan to use the remaining funds from years 1-4 to complete the objectives as described above in section IV. As described in section III.B above, we plan to use the prior-year funding to:

1. Postdocs: Both postdocs will be employed through the end of year 5 (\$115,995)
2. Working groups
 - Both working groups will continue to meet in year 5 to complete their work (\$118,200)
 - Added honoraria for working group leads to compensate for their time in organizing and planning working groups (\$36,000)
3. Cyberinfrastructure design and development for AOOS DataONE integration (\$36,472)
4. Data rescue effort
 - Student interns for new data rescue effort (\$46,201)
 - Increase projects data coordinator time for data rescue to 60% for 12 months (\$41,771)
5. Additional travel to GWA PI meeting, and scientific conferences for postdocs (increase by \$1,271 in year 5)

Cumulatively, these shifts in the expenditure of prior year funds to year five will allow us to complete both the original objectives slated for year 5, plus the additional objectives from years 4 and previous that shifted due to changes in timing (such as the working groups) and increases in priority of the data rescue effort. In addition to the original \$67,766 for year 5, the total rollover expenditures from prior years will be \$447,625.

C. Sources of Additional Funding

N/A