

Collaborative Oceanography and Virtual Experiments

James G. Bellingham
Monterey Bay Aquarium Research Institute
7700 Sandholdt Rd
Moss Landing, CA 95039
phone: (831) 775-1731 fax: (831) 775-1646 email: jgb@mbari.org

Grant Number: N00014-09-1-0405
<http://itop.org>

LONG-TERM GOALS

Our long term goal is to develop online collaboration environments that support cooperative operation of complex environmental field programs by distributed interdisciplinary teams. The need for such collaborative environments is growing. Modern field experiments are challenging to manage, in part because new autonomous and unmanned platforms increase both the complexity and the number of options available to participants. Consequently, managing a field program demands a diverse project team. Traditionally, the tight cooperation of the team is enabled by collocating participants for the duration of the experiment. Our work focuses on developing a collaboration environment that would allow a team to manage a field program collectively, without leaving their home institutions.

OBJECTIVES

The backbone of the collaborative environment is a data management system that aggregates and presents observations and models to experiment participants. Thus the objective of our work was to collect data from ITOP observational assets and other relevant platforms and models, and to make that information available in a useful form in near real-time. Although our initial proposal was to also continue development of an existing online collaboration environment, early ITOP planning sessions made clear that this would be less useful. In effect, the rapid tempo of the meteorological components, especially planning for the aircraft flights, drove the ITOP team towards a conference call format. Consequently, our efforts narrowed to capturing oceanographic observation and modeling elements of the program, and focused on data aggregation and information presentation.

APPROACH

Our team developed an online portal for the ITOP effort (<http://itop.org>) that aggregated data from diverse sources, produced data products, supported a variety of visualization approaches, and hosted online discussions. The work built on a data collection and collaboration system originally developed for the Adaptive Sampling and Prediction (ASAP) component of the Monterey Bay 2006 Experiment. The ASAP portal and data toolsⁱ, specifically COOPⁱⁱ (the Collaborative Ocean Observatory Portal) and MOQuAⁱⁱⁱ (Metadata-Oriented Query and Access) provided the departure point for a portal customized around the needs of the ITOP team.

A central challenge for experiment management for oceanographic field programs, especially large oceanographic field programs, is the decentralized nature of observational capabilities. A diverse suite of systems, including in situ autonomous platforms, moorings, ships, aircraft, satellites and real-time models all provide information important to participants. While the data management for each system maybe be expertly conceived and executed, they are neither integrated nor do they, for the most part, conform to common standards. Much of the program effort focused on tools to take highly unstructured data and make it available to experiment participants through a single portal with a consistent and useful presentation. The decentralized nature of data production led us to develop tools that allowing individual investigators to identify and describe new data streams so they can be integrated into the system. Investigators from both the AOSN and ITOP consider periodic standard visualizations to be one of the most desirable features to include in an experiment's data system. However, many expressed a desire to perform ad-hoc visualizations within the context of the data system. The approach for ITOP will included both standard, periodic visualizations and ad-hoc visualization capabilities.

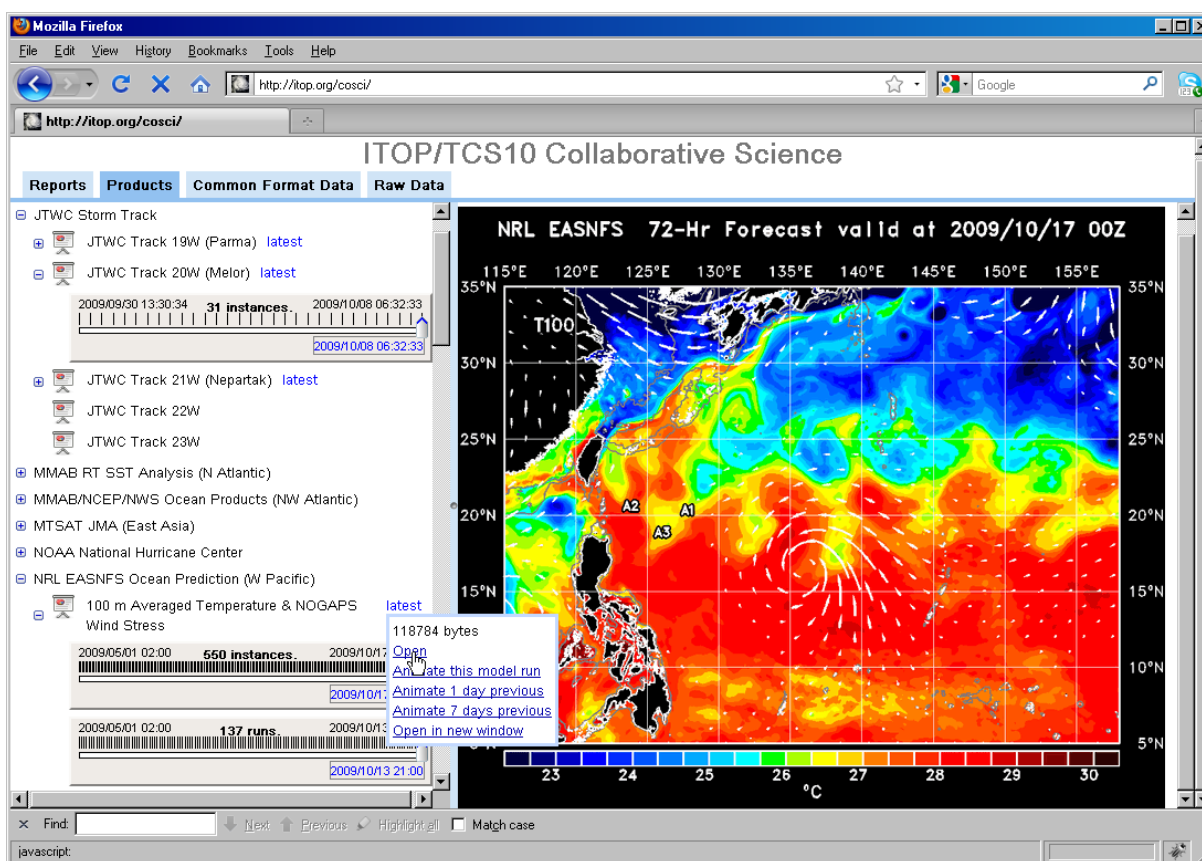


Figure 2, ITOP Collaborative Science (CoSci) portal prototype, showing latest instance of Ko's EASNFS model forecast product.

WORK COMPLETED

The first versions of the Collaborative Science portal (CoSci) were released in September 2009 at <http://itop.org/cosci> (see Figure 1). As requested by ITOP PIs, this system duplicated many of the features of NCAR's Earth Observing Laboratory (EOL), but also contains an internal architecture which will allow it to evolve into a collaborative communication tool. During September 2009, the

prototype system collected over 100,000 model and observational products from repositories around the world.

By the time of the ITOP field program in 2010, we developed a portal and supporting tools that allowed non-experts to add new data sources, edit data, add scripts, and schedule operations. Specific components of the data system included:

Data Collection and Storage: Data was collected by a "bot" running on itop.org and stored redundantly during the experiment in the Amazon cloud.

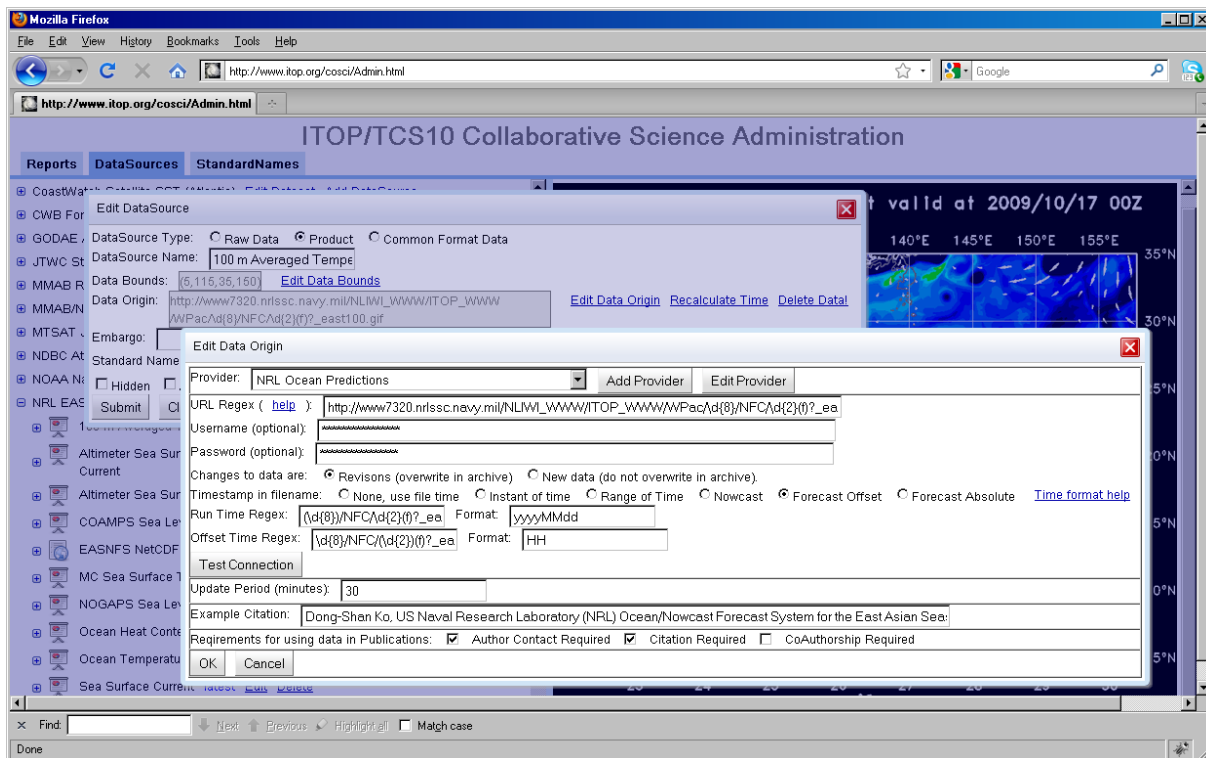


Figure 2, CoSci administrative tools, shown editing an entry for an EASNFS forecast product.

Data Stream Editor: This was the system that the site manager, used to interact with every data set. The editor allowed her to add and remove data sources, and to edit the categorization of data in the system (see figure 2).

Executive Scheduler/Planner: Like the Data Stream Editor, the executive allowed the site manager to add, remove, enable and disable processing scripts, and to adjust the data crawler repetition rate for each

Scripts: A large number of scripts were developed to handle specific processing steps and to produce custom visualizations during the experiment. The previously described executive allowed scripts to be automated.

Products: For any ocean measurement platform or ocean model, if possible, plots of SST, 24C isotherm, currents, mixed layer depth, T(100m), and T(100m)-SST were plotted, as well as plots of

change in these quantities since recent storms passed. Also, products for winds, assets trajectories, storm tracks, and tides were produced.

Visualization: Innovative visualization approaches included: allowing users to build-up plots using overlays and auto-generation of kml wrappers for displaying most plots in Google Earth, including some external plots that required sophisticated re-projection into Google Earth's simple cylindrical projection.

The implementation of CoSci web application runs entirely on compute cloud infrastructure, meaning that no hardware investment has been necessary, and no hardware maintenance will be required in the future. A small monthly bill covers the cost of compute cycles on the virtual server hosting the application, the cost of data transfer to and from the server, and the cost of data storage. Most importantly, running CoSci in the “cloud” means that duplicating the system for another field program would take hours, rather than months that would be required to procure, install, and configure a physical server.

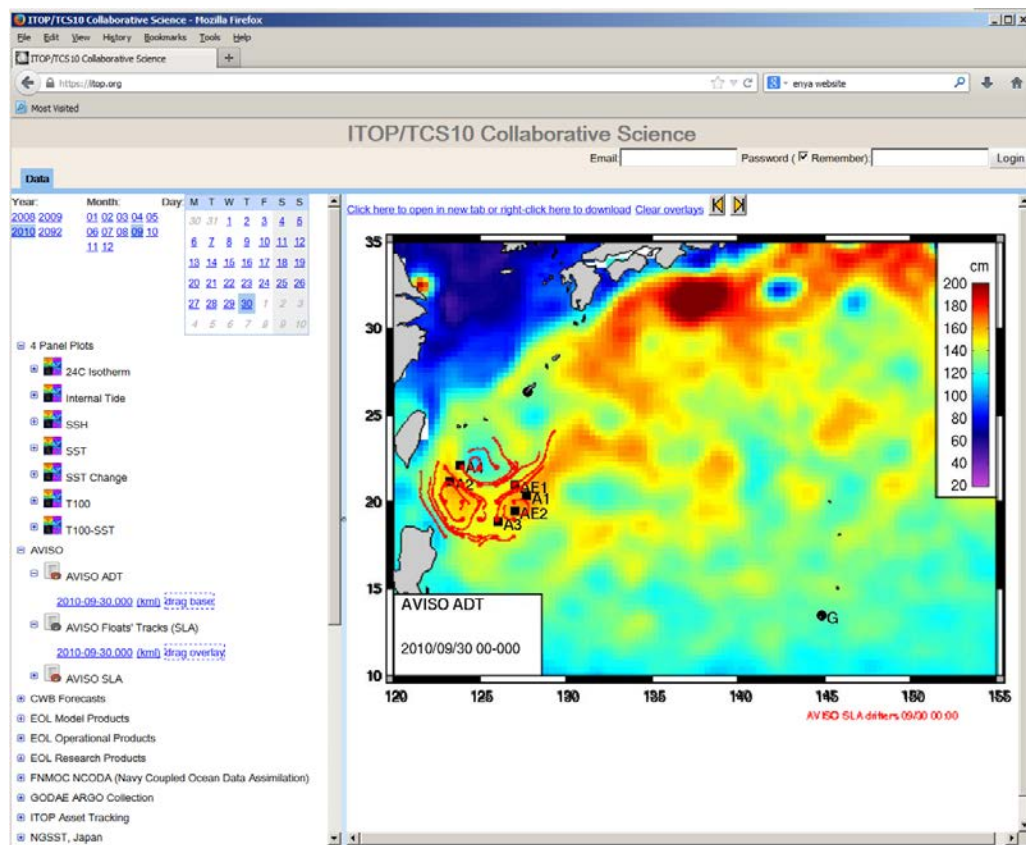


Figure 3: An example of a ‘built up’ graphic combining background and overlay plots.

RESULTS

The portal was operational from July 2010 through December 2010. The majority of activity was from August to October. The data crawler collected 25,854 observational and model data files from 98 sources (325 GB). From these files, using 219 different data processing scripts, an additional 50,988 "common" products were generated (241 plot types), along with 47,600 overlays (101 plot types).

From 52 non-EOL sources, 24,471 products were collected, and from 1486 EOL data collections, 643,263 "federated" products were indexed and made available through itop.org.

Collaborative portion of the site hosted 194 discussion messages by 27 of the 100 registered users. Eric D'Asaro was the top poster, followed by Peter Hawkins, Mike Godin, Peter Black, and Dong-Shan Ko.

Most demand from experiment participants was for specific visualizations. Consequently, some other activity such as data discovery was significantly curtailed. Also, in the implementation of the ITOP pilot experiments, a strong preference was shown for "live" web meetings, rather than the asynchronous "discussion forum" style communications that characterized the ASAP experiment. However, current technology (primarily screen sharing) is bandwidth intensive, making participation in live web meetings awkward for many remote participants. This is encouraging us to develop collaboration features that enhance the capability of users to participate in web meetings without requiring that participants have unlimited bandwidth.

IMPACT/APPLICATIONS

Naval relevance of our work stems both from the immediate impacts on environmental field programs, and for the more general contribution to the challenge of managing a complex distributed system with a geographically distributed team. The latter problem in particular is common to some multi-platform Naval activities.

REFERENCES

ⁱ Bellingham J. G. and Godin, M., "Exploring Ocean Data," *ACM Sigmod*, Jim Gray Tribute Issue, 37(2) 78-82, 2008.

ⁱⁱ Godin, M.A.; Bellingham, J.G.; Rajan, K.; Leonard, N.; Yi Chao, "A Collaborative Portal for Ocean Observatories," *OCEANS 2006*, vol., no., pp.1-5, Sept. 2006

ⁱⁱⁱ Godin, M.A.; Bellingham, J.G., "Data Exploration for Multidisciplinary Research," *Oceans 2007*, vol., no., pp.1-4, Sept. 29 2007-Oct. 4 2007