INTEGRATED OCEAN OBSERVING SYSTEM

# Data Integration Framework (DIF)
# Final Assessment Report

November 22, 2010

# Reference Documents

| # | Document | Date/Version | Author(s) |
|---|---|---|---|
| 1 | Quantifying HAB-FS Improvements Using IOOS DIF Compliant Data Streams | October 28, 2009 | Richard Patchen, et al |
| 2 | IOOS Hurricane Intensity Project Report | June 16, 2010 | Gustavo Goni, et al |
| 3 | IOOS GML Online Survey Results | January 19, 2010 | Alex Birger |
| 4 | DIF Customer Projects Performance Assessment Summary | November 2, 2009 | Carmel Ortiz |
| 5 | IOOS GML XML Schema and Examples | http://ioos.gov/dif/schema.html | IOOS WSDE |
| 6 | IOOS DIF Functional Requirements Document | Version 1.0, November 19, 2007 | IOOS Operations Division, DIF IPT |
| 7 | IOOS DIF Concept of Operations | Version 1.0, April 25, 2008 | IOOS Operations Division, DIF IPT |
| 8 | DIF Design Solution Overview | November, 2008 | IOOS Operations Division, DIF IPT |
| 9 | Testing of Early DIF SOS Implementations | December, 2008 | John Ulmer, NOAA Coastal Services Center |
| 10 | Model Data Interoperability for the United States Integrated Ocean Observing System (IOOS) | March 17, 2010 | Richard P Signell, USGS / IOOS Program Office |

# Acknowledgments

The IOOS Data Integration Framework project could not have been accomplished without the collaboration of many individuals inside and outside of NOAA; the IOOS Program Office thanks the following persons for their contributions:

| Contributor | Organization |
|---|---|
| Lance Acree | NOAA/NESDIS |
| Steve Baig | NOAA/NWS/TPC |
| Anne Ball | NOAA/NOS/CSC |
| Russ Beard | NOAA/NESDIS/NCDDC |
| Luis Bermudez | SURA |
| Rick Blair | NANOOS |
| Nathaniel Booth | USGS |
| Julie Bosch | NOAA/NESDIS/NCDDC |
| Jen Bowen | SCCOOS |
| Tess Brandon | NOAA/NESDIS/NODC |
| Eric Bridger | NERACOOS |
| Bill Burnett | NOAA/NWS/NDBC |
| Grant Cameron | CDIP |
| John Cartwright | NOAA/NESDIS/NGDC |
| Ken Casey | NOAA/NESDIS/NODC |
| Rob Cermak | AOOS |
| Mohamed Chaouchi | NOAA/NOS/CO-OPS |
| Joseph Chen | SCOOS |
| John Christensen | NOAA/NOS/NCCOS |
| Jeremy Cothran | SECOORA |
| Tom Crowley | LMI |
| Paul Daisey | Image Matters, LLC |
| John Dandy | NOAA/NOS/MBO/IMD |
| Cindy Diehl | NOAA/CIO/EAO |
| Paul DiGiacomo | NOAA/NESDIS/ORA/ORAD |
| Darrell Duncan | NOAA/NWS/NDBC |
| Lori Fenstermacher | NOAA/NOS/CO-OPS |
| Kathleen Fisher | NOAA/NOS/CO-OPS/PSD |
| Dave Foley | NOAA/NMFS/SWFSC/ERD |
| David Froehlich | NOAA/NESDIS/NGDC |
| Jennifer Frye | NOAA/NESDIS |
| Charlton Galvarino | SECOORA |
| Mike Garcia | NOAA/NWS/NDBC |
| Felimon Gayanilo | GCOOS |
| Gustavo Goni | NOAA/OAR/AOML |
| John Graybeal | OOI-CI Project, UC San Diego |
| Heng Gu | NOAA/NESDIS/ORA |
| Ted Habermann | NOAA/NESDIS/NGDC |
| Sara Haines | SECOORA |
| Steve Hankin | NOAA/OAR/PMEL |
| Andrea Hardy | NOAA/NOS/CO-OPS |
| Lisa Hazard | SCCOOS |
| Dan Henderson | NOAA/NWS/NDBC |
| Tim Hopkins | NOAA/NWS/SEC |

| Contributor | Organization |
|---|---|
| Matt Howard | GCOOS |
| Bill Howe | CMOP |
| Eoin Howlett | MACOORA |
| Kent Hughes | NESDIS/ORA/ORAD |
| Hugh Johnson | NOAA/NOS/CIO |
| Mike Johnson | NOAA/OAR/OCO |
| Phil Keegstra | NOAA/NESDISCoastWatch |
| Mike Kosiara | NOAA/NWS/WFO Slidell |
| Nicole Kurkowski | NOAA/NWS/MDL |
| Lyon Lanerolle | NOAA/NOS/CSDL |
| David Layton | NOAA/CIO/EAO |
| Steven Le | CeNCOOS |
| Rebecca Love | NOAA/NOS/CSC |
| Stuart Maclean | UW APL |
| Daniel Martin | NOAA/NOS/CSC |
| Emilio Mayorga | NANOOS |
| David McClure | NOAA/CIO |
| Lewis McCulloch | NOAA/NESDIS/OSD |
| Ken McDonald | NOAA/NESDIS/OSD |
| Avichal Mehra | NOAA/NWS/EMC/MMAB |
| Roy Mendelssohn | NOAA/NMFS/SWFSC |
| Kurt Nelson | US Navy |
| David Neufeld | NOAA/NESDIS/NGDC |
| John Parker | NOAA/NOS/MBO/IMD |
| Rost Parsons | NOAA/NESDIS/NCDDC |
| Richard Patchen | NOAA/NOS/OCS/CSDL |
| Josh Pederson | NOAA/Ntl Marine Sanctuaries |
| Jim Potemra | University of Hawaii, PacIOOS |
| Armin Pruessner | NOAA/NOS/CO-OPS |
| Justin Rabin | NOAA/NWS/NDBC |
| Ed Rappaport | NOAA/NWS/TPC |
| Paul Reuter | SCCOOS, UC San Diego |
| Frank Revitte | NOAA/NWS/WFO Slidell |
| Craig M. Risien | Oregon State University |
| Eric Roby | NOAA/NOS/NCDDC |
| Bill Sammler | NOAA/NWS/WFO Wakefield |
| Jim Sargent | NOAA/NMFS/OAA/SOCIO |
| Whitley Saumweber | NOAA/NOS/OCRM/ERD |
| Franklin Schwing | NOAA/NMFS/SWFSC |
| Tony Siebers | NOAA/NWS/WFO Wakefield |
| Richard Signell | USGS |
| Bob Simons | NOAA/NMFS/SWFSC |
| Walter Smith | NOAA/NWS |
| Derrick Snowden | NOAA/OAR/OCO |

| Contributor | Organization |
|---|---|
| Richard Stumpf | NOAA/NOS/NCCOS |
| Sudha Sundar | NOAA/NOS/CO-OPS |
| Arthur Taylor | NOAA/NWS/MDL |
| Christopher Taylor | NOAA/NWS/NDBC |
| Julie Thomas | CDIP |
| Hendrik Tolman | NOAA/NWS/EMC/MAB |
| Michelle Tomlinson | NOAA/NOS/NCCOS |
| Stephen Uczekaj | NANOOS/Boeing |
| John Ulmer | NOAA/NOS/CSC |
| Sam Walker | SECOORA |
| Kyle Wilcox | ASA / MARACOOS |
| Darren Wright | CDIP |
| Chung Wu | NOAA/NWS/SEC |
| Tim Wynne | NOAA/NOS/NCCOS |
| Martin Yapur | NOAA/NESDIS/OSD |

| DIF Project Team – NOAA IOOS Program |
|---|
| ***NOAA STAFF*** |
| Charles Alexander |
| Jeff de LaBeaujardiere |
| Jack Harlan |
| CAPT Christopher Moore* |
| Hassan Moustahfid |
| Kurt Schnebele* |
| Rebecca Shuford* |
| Marcia Weaks* |
| ***CONTRACT SUPPORT*** |
| Alexander Birger – Skjei Telecom |
| Jack Oliva - LMI |
| Carmel Ortiz – Skjei Telecom |
| Robert Ragsdale – ERT |

*no longer with NOAA IOOS Program

# Table of Contents

## TABLE OF FIGURES

## TABLE OF TABLES

# 1.Executive Summary

**Overview**

In December, 2006, the NOAA Executive Council and NOAA Executive Panel approved the formation of NOAA's Integrated Ocean Observing System (IOOS®) Program within the National Ocean Service. This included approval for the IOOS Program to develop a Data Integration Framework (DIF) project with a nominal duration of three years, from February 1, 2007 to February 1, 2010. The DIF was designed as a limited-scope, risk reduction project to measure the value of integrated data on four (4) ocean decision-support tools that span multiple NOAA mission goals. Following the nominal three year project duration, experience gained through the DIF would be used to inform long-term IOOS data management planning and implementation.

Interoperability tests, sponsored by IOOS during the spring of 2007 revealed gaps with interoperability among potential IOOS data providers: "In general, that testing showed the provisioning of temperature, salinity, water level, currents, and ocean color data to be quite variable in format and content as they expressed time, position, and parameter names differently. This prevented direct integration of the data sets and sources with each other." (Ulmer, December 2008)

The DIF project intended to address those gaps for four selected NOAA decision-support tools, using data from sources of five (5) core IOOS variables (the original 5 core variables was subsequently expanded to 7), such that improvements in accuracy and/or efficiencies in time and production costs of model outputs, assessments and forecasts could be achieved and measured. The project premise was that data integration and improved management and dissemination of mission-critical ocean-related data can increase the value and effectiveness of these presently disparate data in supporting decision-making tools/models.

Initially, three major NOAA Data Providers were selected for the DIF: NWS' National Data Buoy Center (NDBC), NOS' Center for Operational Oceanographic Products and Services (CO-OPS), and NESDIS' CoastWatch. The selection was based on the fact that these providers together represent a significant amount of data for the selected core variables. Subsequently, all IOOS Regional Associations (RAs) have implemented at least one of the DIF-adopted standards to provide their data; the DIF customer projects also established OAR's Atlantic Oceanographic and Meteorological Lab (AOML) as a DIF-compliant data provider for synthetic temperature and salinity profiles.

Data Customers consisted of selected decision-support tools in the areas of **Coastal Inundation (CI)**, **Hurricane Intensity (HI)**, **Integrated Ecosystem Assessments (IEAs)**, and **Harmful Algal Blooms (HAB)**. These four decision-support areas were selected because they address critical environmental issues aligned with NOAA mission goals.

**Decision Support Tools/Models (Customer Projects)**

The DIF Customer Implementation Projects (CIPs) were identified in 4 decision-support areas. The CIPs were initially designed to support one of DIF's primary goals and objectives: to validate the DIF premise that data integration and improved access via standardized interfaces has a value that can be measured in specific models/decision support tools. The assessment of these projects against a set of *a priori* metrics is the primary subject of this report. A brief summary of the assessments appears at the end of this Executive Summary.

**Impact on IOOS Regional Associations**

The IOOS RAs were not originally intended to be data providers in the DIF project (other than regional data that is served by NDBC). However, several factors contributed to the RAs implementation of DIF standards and services and general advancement in the homogeneity of data from the RAs. Factors that contributed to this progress include:

- Establishment of regular and formal coordination/collaboration between RA data managers and IOOS DIF team members
- DIF adoption of specific web services and encoding schemes that could be broadly implemented

- Availability of an SOS code base, due to NDBC implementation of DIF SOS services, that formed the basis of many RA implementations
- The DIF Model Data Interoperability Project (see Section 5.2) which resulted in interoperability between a set of previously heterogeneous structured grid model outputs. The project showed how existing model outputs can be homogenized and made accessible via standard desktop tools.

The DIF activity served as a catalyst to guide the RAs in implementing standard methods of serving their data. As a result they transitioned from a group of nascent, fairly independent data providers to a more integrated data management group. Across the RAs there is now more uniformity of services and data formats. Section 5 contains details on the RA implementation of DIF standards.

**Summary of Results**

A summary of the DIF customer projects against their metrics is provided in Table 1-1. In the case where assessment was possible, the impact of the project on the decision support tool was overwhelmingly positive. Unfortunately, half of the metrics could not be effectively evaluated due to insufficient data (either no baseline data or inability to measure the enhanced performance).

| **Item** | **CI** | **HI** | **HABs** | **IEA** | **TOTAL** | **% Total** |
|---|---|---|---|---|---|---|
| **# of Metrics** | **5** | **2** | **7** | **2** | **16** | **100%** |
| Positive Impact | 1 | 1 | 4 | 2 | **8** | **50%** |
| Negative Impact | 0 | 0 | 0 | 0 | **0** | **0%** |
| Insufficient data for evaluation | 4 | 1 | 3 | 0 | **8** | **50%** |

**Table 1-1 DIF Customer Project Metrics Summary**

In addition to the anticipated benefits which were the formal metrics for the projects, numerous additional benefits were identified during the course of the execution of the customer projects. These "additional benefits" are detailed in Sections 3.2.5.2, 3.3.5.2, 3.4.5.2 and 3.5.5.2 but generally fall into one of the following major categories:

- Additional scientific or technical knowledge regarding the potential for particular data sets or processes to improve decision-support tools
- Increased understanding of the benefit of standardized data formats and access methods on IT investments
- Demonstration of successful collaborations across organizations
- Overall expansion in the availability of easily accessible oceanographic data in standard formats that can be used by a wide range of scientists and application developers

The assessment contained in this report is based primarily upon an examination of project performance against the original DIF goals and objectives: to measure the value of integrated data on four (4) ocean decision-support tools. The assessment does not thoroughly evaluate whether or not the initial goals were appropriately selected or were, in retrospect, the most meaningful. Section 3.3.5.2 provides anecdotal data regarding how the DIF integrated data services allowed a programmer to quickly incorporate multiple datasets into their client applications; a result that would benefit a broad range of users. Going forward, a more meaningful measure of success might be the ability of IOOS data to easily reach a broad range of users in the ocean data analysis community; users that analyze data with popular commercial-off-the-shelf client software such as Excel, ArcGIS, Matlab, R, etc. Ultimately, the widespread use of IOOS data by the ocean community will depend on its relative ease of access for users with limited specialized knowledge in data access techniques.

**Enduring DIF Outcomes**

Many DIF benefits and outcomes have endured beyond the initial DIF project timeframe. Specific DIF “leave-behinds” include:

1. SOS services operating at NDBC and CO-OPS, serving six IOOS core variables (Currents, Water Level, Sea Surface Temperature, Salinity, Winds and Waves)
2. OPeNDAP services operating at CoastWatch serving one IOOS core variable (Ocean Color)
3. Synthetic temperature and salinity profiles for three (3) hurricane case-studies, available via OPeNDAP from AOML
4. Successfully integrated DIF data into NCEP’s operational data “tanks”. This is a major technical achievement for establishing protocols/procedures to accept additional IOOS data into NCEP operational data streams. The experimental Temperature and Salinity profile data are now available to all NCEP models - NCEP’s coupled hurricane forecast model, the global ocean model, the Global Forecast System, and a lower resolution ocean model
5. An enhanced understanding of the effect of improved ocean state information on hurricane intensity forecasting which could lead to improved forecasts
6. An enhanced Sea, Lake and Overland Surges from Hurricanes (SLOSH) Display Program (SDP) which integrates IOOS-formatted water level and wind observations with SLOSH model output. This version of the SDP has been deployed for operational use at coastal NWS WFOs for use in advance of a hurricane landfall.
7. Enhanced operational HAB-FS bulletin software which integrates IOOS in-situ currents observations to assist analysts in HAB forecasting; this allows the analysts to enhance nowcast predictions for harmful blooms. Without in-situ currents data the nowcasts were only possible using remotely-sensed chlorophyll data and observed wind data.
8. Demonstration of a potential capability of producing chlorophyll anomaly transport and extent forecasts, derived from SeaWiFS imagery. The forecasts, in conjunction with cell count confirmation, could serve as additional guidance to the HAB analysts when developing forecasts.
9. Demonstration of the potential improvement in the spatial and temporal resolution of HAB Extent forecasts. The forecast used in the DIF HAB project resulted in a spatial error did not exceed 18 km in a 7 day forecast. The existing HAB-FS cannot be resolved at scales finer than 30 km and forecasts are limited to 3 to 4 days.
10. Reinforced understanding of the importance of subsurface chlorophyll and cell count observations to HAB forecasting
11. NOAA subject matter experts have a better understanding of the limitations of a physical model in predicting bloom transport and extent, the importance of biological factors in bloom transport, extent and intensification, and the limitations of a 2-D particle tracker versus 3-D. The project results demonstrated that the use of a three-dimensional trajectory model can produce viable forecasts for up to one week, while the use of a two-dimensional tracker is insufficient to provide such forecasts because the nature of the HAB processes was proved to be essentially three-dimensional.
12. An increased number of datasets of interest to IEA developers, available using IOOS standard web services and well-defined formats
13. Significant technical refinements to the NMFS/ERD ERDDAP software (a tool that can read from a variety of the most common data transport standards, and can output the data in a wide variety of formats used by analysis and visualization applications) to provide enhanced integration with selected IOOS data services and, in collaboration with the Ecosystem Goal Team, to prototype the implementation of these services into the IEA model for the Gulf of Mexico and California Current Regions.

14. Various software tools, code bases and reference implementations that can be referenced, modified and/or re-used by other data providers or data consumers to expedite subsequent implementation efforts
15. The recognition that a significant benefit of the standardization of the data formats and methods for requesting and receiving the data across multiple providers is significantly reduced software coding complexity for client-side application developers. A standardized access and data format across multiple data providers simplifies the task for the client-side software developer. This increases the value of the initial software development investment, and reduces long-term costs of software maintenance enhancements. In one instance client development time of 5 days was reduced to 2 days when adding a data provider; in another instance the software modification effort was reduced from 5 days to 1 hour.
16. Establishment of a core set of initial standards and conventions to enable data interoperability (documented in *Guidance for Implementation of the Integrated Ocean Observing System (IOOS®) Data Management and Communications (DMAC) Subsystem* http://ioos.gov/library/dmac_implementation_2010.pdf)
17. Identification of an initial approach to data provider compliance testing
18. Significant momentum for moving forward with IOOS DMAC planning and implementation
19. DIF-compliant SOS services operating experimentally at 9 of the 11 RAs (see Table 5-1)
20. OGC WCS or OPeNDAP services operating at all 11 RAs
21. Interoperability of model data output across all 11 RAs
22. The GCOOS Data Portal that consolidates data from 10 local data nodes into a single aggregated SOS according to DIF standards. All GCOOS regional data nodes offer their data via a standardized IOOS interoperable interface; this is a direct result of the IOOS DIF activity. They nodes provide an up-to-the-minute accurate representation of what parameters are currently available via the continually updated XML files they host for the IOOS Regional Observation Registry. The local data nodes have adopted a common vocabulary (that was recently submitted to MMI - See the "IOOS Parameter Vocabulary V1.0 http://mmisw.org/orr/#http://mmisw.org/ont/ioos/parameter). This common vocabulary was developed with input from the IOOS DIF effort.

# 2. Introduction

## 2.1. Document Purpose

The purpose of this document is to:

- Assess how the DIF project performed against its original objective,
- Assess technical decisions and identify lessons-learned,
- Describe the DIF project benefits to the Integrated Ocean Observing System (IOOS®) Regional Associations,
- Document additional unanticipated benefits and "leave-behinds" that will persist beyond the end of the DIF,
- Make specific recommendations for IOOS DMAC.

## 2.2. Organization

This document is organized around the following major sections:

- Section 1 – Contains an Executive Summary of the DIF project outcomes and performance against the original objectives
- Section 2 – Provides background information and an overview of the DIF project and its technical approach
- Section 3 – Provides detailed information on each of the four DIF Customer Implementation Projects, including a performance assessment of each project against the metrics
- Section 4 – Provides an assessment of the technical decisions that were made during the course of the DIF project execution. It also provides recommendations for evolving the DIF infrastructure for the benefit of IOOS DMAC
- Section 5 – Describes how the DIF project benefitted the IOOS Regional Associations
- Appendix A – Provides analysis and results of the SOS Encoding Survey undertaken at the end of the DIF Project.
- Appendix B – Contains acronyms used in this report

## 2.3. Background

In December, 2006, the NOAA Executive Council and NOAA Executive Panel approved the formation of NOAA's IOOS Program within the National Ocean Service. This included approval for the IOOS Program to develop a Data Integration Framework (DIF) project with a nominal duration of three years, from February 1, 2007 to February 1, 2010. The DIF was designed as a limited-scope, risk reduction project to assess the value of integrated and interoperable data on four (4) ocean decision-support tools or themes that span multiple NOAA mission goals. Following the nominal three year project duration, experience gained through the DIF would be used to inform long-term IOOS data management planning and implementation.

The DIF project was conceived to address a lack of overarching operational data management system or capability within NOAA that could assemble data from diverse sources, and meet the geographic coverage, vertical and horizontal resolution, accuracy, timeliness, and data processing requirements and needs of multiple NOAA decision-support tools: ocean models, assessments and forecasts, and other products.

Interoperability tests, sponsored by IOOS during spring of 2007 revealed gaps with interoperability among potential IOOS data providers: "In general, that testing showed the provisioning of temperature, salinity, water level, currents, and ocean color data to be quite variable in format and content as they expressed time, position, and parameter names differently. This prevented direct integration of the data sets and sources with each other." (Ulmer, December 2008)

The DIF goal was to address those gaps for four selected NOAA decision-support tools, using data from NOAA sources of the initial five (5) core IOOS variables, such that improvements in accuracy and/or efficiencies in time and production costs of model outputs, assessments and forecasts, could be achieved and measured. The project premise was that data integration and improved management and dissemination of mission-critical ocean-related data can increase the value and effectiveness of these presently disparate data in supporting decision-making tools/models.

## 2.4. Technical Approach and DIF Scope

The DIF focused on the integration of data from selected sources of core IOOS variables to enhance the efficiency and/or effectiveness of four (4) selected decision-support tools/models. Five (5) initial core variables were selected for the DIF: **Sea Temperature**, **Water Level**, **Ocean Color**, **Currents** and **Salinity** (Section 3 contains details on the specific variables/datasets used in each project). It was anticipated that these initial variables would be of high value for the selected Data Customers and could be made readily available from a limited number of Data Providers. Subsequently, **Winds** and **Waves** were added to the list of core variables for a total of seven.

Integration within the DIF means improving the way the selected sources of the variables are made available to the decision-support tools through the consistent application of community-based standards and protocols, such as for data format and access methods. Utilizing the principles of IOOS Data Management and Communications (DMAC), the DIF developed methodologies to improve upon existing ocean data integration efforts that could facilitate flexibility and extensibility to other variables, systems and decision-support tools. By adopting, adapting, or expanding existing standards and other capacities and capabilities for data management services, the DIF formalized a standards-based common data sharing infrastructure that is expected to facilitate and improve data integration of ocean variables across NOAA Line Offices and beyond.

Initially, three major NOAA Data Providers were selected for the DIF, NDBC, CO-OPS, and CoastWatch, based on the fact that these providers together represent a significant amount of data for the selected core variables. Subsequently, all Regional Associations (RAs) have implemented at least one of the DIF-adopted standards to provide their data. The DIF customer projects also established Atlantic Oceanographic and Meteorological Lab (AOML) as a DIF-compliant data provider for synthetic temperature and salinity profiles in support of the Hurricane Intensity project. Table 2-1 summarizes DIF data providers and the variables offered.

| | Currents | Water Level | Sea Temp | Salinity | Surface Winds | Waves | Ocean color (chlorophyll) |
|---|---|---|---|---|---|---|---|
| **Single station** | NDBC, CO-OPS | NDBC, CO-OPS | NDBC, CO-OPS | NDBC, CO-OPS | NDBC, CO-OPS | NDBC | n/a |
| **Group of stations** | NDBC | NDBC, CO-OPS* | NDBC | NDBC | NDBC | NDBC | n/a |
| **Profile** | NDBC, CO-OPS | n/a | NDBC, CO-OPS, AOML | NDBC, CO-OPS, AOML | n/a | n/a | n/a |
| **2D grid** | NDBC, CSDL | | | | | | CoastWatch |

* NDBC temporarily serves groups of stations data (collections) for CO-OPS;
CO-OPS plans to start serving collections on its own in FY2010

**Table 2-1 DIF Data providers and Variables**

Data Customers consisted of selected decision-support tools in the areas of:

1. Harmful Algal Bloom Forecast System (HAB-FS)
2. Coastal Inundation (CI)
3. Hurricane Intensity (HI)
4. Integrated Ecosystem Assessments (IEA)

These four decision-support areas were selected because they address critical environmental issues aligned with NOAA mission goals.

Customer Implementation Projects (CIPs) were identified for each of the 4 decision-support areas. They were initially designed to support one of DIF's primary goals and objectives: to validate the DIF premise that data integration and improved access via standardized interfaces has a value that can be measured in specific models/decision support tools.

The CIPs were expected to provide visible and meaningful impacts within a relatively short timeframe and limited funding. Independent metrics were established for each CIP against which performance could be measured. A number of candidate projects were investigated and evaluated for use as DIF CIPs. One project was selected in each of the four model/decision-support areas.

The DIF data customer projects were selected in the context of various DIF project constraints. The constraints drove the decisions as to which project was selected, although in some cases the constraints were re-evaluated to determine if any flexibility existed to enable implementation of a project that might not otherwise have been achievable.

Specific constraints include:

- Cost – During the course of project identification, cost constraints limited the domain of potential projects to those that could reasonably be achieved within the budget.
- Schedule – the DIF project schedule required that customer projects be identified, executed and evaluated before the end of FY10. Many of the decision-support areas addressed seasonal conditions, such as hurricanes or harmful algal blooms that have specific seasons of high likelihood of occurrence. This further constrained the schedule; project implementation needed to be completed before the start of the season in order to maximize the ability to measure impact.
- Scope – the objective of the project was to achieve the highest level of integration possible – integrating as many core variables as possible from the maximum number of data providers. However, because the DIF scope was limited to seven variables (expanded from the initial 5) and four initial data providers, the projects were constrained in scope.

Briefly, the following are the customer projects that were selected in each decision-support area (refer to Section 3 for a detailed description of each project):

- Harmful Algal Blooms (HAB) – two independent projects were identified:
    - Enhancement of the operational Gulf of Mexico HAB-Forecast System with real-time current data from NDBC to assist analysts with HAB transport and extent forecasts
    - Study into the feasibility of using satellite ocean color data, modeled currents and a particle tracker to predict HAB transport and extent
- Coastal Inundation (CI) – Enhancement of the SLOSH Display Program (SDP) to include real-time water levels and wind data from NDBC and CO-OPs. The enhanced SDP could improve the ability of coastal Weather Forecast Offices and the Tropical Prediction Center to forecast inundation in advance of a hurricane land fall. It could also improve visual communication with key stakeholders such as the emergency management community and media.
- Hurricane Intensity (HI) – A study that evaluates if the inclusion of synthetic temperature profiles, derived mainly from satellite altimetry observations, in the HYCOM model produces a reduction of error in hurricane intensity forecast in HWRF runs coupled to HYCOM. As a byproduct, this project would also allow NCEP models to work with new datasets and a new type of data, and opened a gate for IOOS data into NCEP data tanks – the source of all data for NCEP operational model runs.

- Integrated Ecosystem Assessments (IEA) – Modification of the ERDDAP software (a tool that can read from a variety of the most common data transport standards, and can output the data in a wide variety of formats used by analysis and visualization applications as well as in scripts) to provide enhanced integration with selected IOOS DIF data services and, in collaboration with the Ecosystem Goal Team, to prototype the implementation of these services into the IEA model for the Gulf of Mexico and California Current Regions. This ERDDAP application will be an underpinning element of the data management infrastructure for NOAA's IEA program.

The DIF customer projects were assessed after completion, based on a series of *a priori* project-specific metrics. Assessment was more challenging than originally anticipated for a number of reasons:

- Assessment can take several months and may be dependent on occurrence of weather or other events (HAB, land falling hurricane, etc)
- Lack of baseline data against which to evaluate prevents quantifiable results
- Metrics that were identified were often more qualitative than quantitative
- Schedule and cost constraints that restricted the scope and depth of assessment.

Section 3 provides a detailed description of the selected Customer Implementation Projects, metrics and performance.

## 2.5. IOOS Regional Associations

IOOS is based on the coordinated development of observing and prediction systems at the global, national, regional, and local scales. In addition NOAA and other federal observing and prediction systems, IOOS Regional Associations (RAs) and Regional Coastal Ocean Observing Systems (RCOOSs) provide a vital and vast network to identify and address regional priorities and expand the coverage of IOOS. There are eleven RCOOSs, each of which is comprised of many, smaller observing systems, as well as a regional management structure responsible for collaboration and coordination within the region. RAs provide the primary framework to coordinate ocean observing activities and are responsible for the design and coordinated operation of sub-regional coastal ocean observing systems within their respective geographies. RAs work with user communities at the state and local levels to ensure that the regional system is designed to be as useful as possible.[1]

The initial DIF scope did not include use of data from the IOOS RAs. However, the DIF activity eventually provided the technical guidance that led to strong harmonization in the way the RAs offer data: using standards and formats in concert with the IOOS vision. Section 5 provides a detailed description of how the DIF activity benefited the RAs and helped to advance data interoperability in the regions.

[1] http://ioos.gov/partners/regional.html

# 3.Customer Implementation Projects

This section provides a detailed description of each of the four DIF Customer Implementation Projects (CIPs). Project objectives, architecture, datasets and outcomes are discussed. The results of each project are evaluated according to the metrics defined, and additional benefits are described.

It is important to note that although the DIF was a limited scope, three-year project, much of the infrastructure, tools and services that were developed during the course of the project will continue. Key "leave-behinds" of the DIF Customer Implementation Projects include:

1. SOS services operating at NDBC and CO-OPS, serving six IOOS core variables (Currents, Water Level, Sea Surface Temperature, Salinity, Winds and Waves)
2. OPeNDAP services operating at CoastWatch serving one IOOS core variable (Ocean Color)
3. Synthetic temperature and salinity profiles for three (3) hurricane case-studies, available via OPeNDAP from AOML
4. An enhanced understanding of the effect of improved ocean state information on hurricane intensity forecasting which could lead to improved forecasts
5. An enhanced SLOSH Display Program (SDP) which integrates water level and wind observations with SLOSH model output. This version of the SDP has been deployed for operational use at coastal NWS WFOs for use in advance of a hurricane landfall.
6. Enhanced operational HAB-FS bulletin software which integrates in-situ currents observations to assist analysts in HAB forecasting
7. An increased number of datasets of interest to IEA developers, available using IOOS standard web services and well-defined formats
8. Various software tools, code bases and reference implementations that can be referenced, modified and/or re-used by other data providers or data consumers to expedite subsequent implementation efforts (reference project-specific assessments for more information)
9. The recognition that an important benefit of the standardization of the data formats and methods for requesting and receiving the data across multiple providers is significantly reduced software coding complexity for client-side application developers. A standardized access and data format across multiple data providers simplifies the task for the client-side software developer. This increases the value of the initial software development investment, and reduces long-term costs of software maintenance enhancements.

## 3.1. Performance Assessment Approach

While Performance Assessment specifics vary by project, in general the objective was to measure the value of enhanced operating performance (benefit) gained by each project due to the data integration. The evaluation process in general consists of four steps:

- Identify key ***Metrics*** for performance
- ***Baseline*** the forecast/prediction performance (skill) of the decision support tool for the selected metrics
- Assess the ***Enhanced*** forecast/prediction performance (skill) due to DIF data
- Compare the ***Baseline*** versus ***Enhanced*** performance to determine ***DIF Impact***

For performance assessment purposes, all CIP benefits have been divided into three main categories in regard to metric and benchmark availability. Table 3-1 provides the metric category labels and corresponding description.

| Category | Category Description |
|---|---|
| A. | Benefit has a quantifiable value, and a historic benchmark is available; |
| B. | Benefit has a quantifiable value but no benchmark is available; |
| C. | Benefit has qualifiable but not quantifiable value, i.e. the improvement definitely brings the tool/model to a qualitatively new and higher level but it cannot be expressed in terms of numbers (e.g. percents skill improvement, etc.), and there is no feasible benchmark. |

**Table 3-1 Customer Project Metric Categories**

Each project provides its own distinctive combination of benefit categories, and thus the assessment methodology as well as benchmarks and metrics are also CIP-specific. The benefits, assessment methodology, metrics and benchmarks are explained in more details in the succeeding sections devoted to each CIP.

## 3.2. Harmful Algal Blooms (HABs)

### *3.2.1.* Decision Tool/Model to Be Enhanced

This project was based on the existing operational model for the Eastern Gulf of Mexico – Harmful Algal Bloom Forecast System (HAB-FS), and the experimental model for the Western Gulf of Mexico. The existing operational HAB-FS output is in the form of a HAB Forecast Bulletin, which contains an operational forecast of bloom extent, transport, intensification, and impact over 0-3 days. This bulletin is used by local decision makers and the general public to inform regional public health and safety decisions.

### *3.2.2.* Scope, Architecture, and Data

The primary objectives of the project were to:

(a) Ingest and display surface currents data from CO-OPS and NDBC to the existing HAB-FS bulletin software, providing additional data for analysts to use in bulletin creation; determine whether the surface currents observations could improve the analysts' ability to provide accurate guidance;

(b) Determine whether HAB nowcast and forecast quality would benefit from an objective and spatially-articulated transport model. The transport model input was fed by modeled surface currents data in DIF standard formats/services as well as winds (in indirect fashion via the currents model), and ocean color data was used to augment the forecast;

(c) Assess the operational requirements that DIF-formatted data would need to meet, should this enhanced transport model become operational.

The project was divided into two phases. Phase 1 focused on the capability of the existing HAB-FS bulletin software to ingest surface currents data provided by CO-OPS and NDBC using data standards and protocols identified by the DIF. Phase 1 was developed by collaborative efforts of NOAA/NOS' Coastal Services Center (CSC) and CO-OPS teams. Phase 1 resulted in reliable access to DIF-formatted surface currents data served by NDBC and CO-OPS, and automatic ingest and display of that data into the operational HAB-Forecast System bulletin generation application for the western Florida region of the Gulf of Mexico. In the existing operational HABS-FS, analysts considered winds but not surface currents when forecasting a bloom's transport.

Phase 2 of the project augmented the bloom transport forecast with the help of an enhanced transport model driven by HF Radar observations, and forecasted marine winds and surface currents provided by CSDL, although winds forecast data and HF Radar observation data have been excluded later due to the limited resources and data availability. The main objective of Phase 2 of the project was to illustrate that using wind field forecasts and forecasts of currents from a three-dimensional numerical circulation model, and a multi-dimensional particle tracking trajectory model that would ingest the aforementioned fields, would address limitations in the operational method presently employed in the NOAA's HAB-FS. The initial project plan provided for the use of a two-dimensional particle tracking trajectory model; as the project evolved, it became evident that a three-dimensional model produced better results.

Because of the project time constraints, Phase 2 was performed in hindcast mode for the time period from November 1, 2004 through January 31, 2006. Within that fifteen month time domain, two shorter periods in 2005 were selected to perform intensive model to data inter-comparisons, i.e. January 18, 2005 through February 1, 2005, and September 26, 2005 through September 30, 2005. These time periods were selected based on the availability of relatively clear (cloud-free) satellite imagery. For the convenience of the experiments, the January, 2005 period was divided into two sub-periods; the resulting experiment time domains are identified as follows:

- January 18 through 25, 2005
- January 25 through February 1, 2005
- September 26 through September 30, 2005.

The geo-spatial domain was set from just north of Tampa Bay to south of Charlotte Harbor on the west Florida shelf. This domain was used because it matches the domain of the high-resolution surface currents forecast model from CSDL.

In Phase 1 of the project, the surface current observations for the project geographical area were provided by NDBC and CO-OPS in IOOS DIF format for gridded data, i.e. netCDF/CF. In the Phase 2, the surface currents modeled by CSDL were used instead of observations. The CSDL model output data was provided by CSDL in the same netCDF/CF format, and was accessible via IOOS DIF standard access service, i.e. OPeNDAP/WCS. However, the netCDF files to feed the particle tracker were retrieved via FTP because of limitations on the client side. The project architecture and data flow diagram is presented in Figure 3-1.

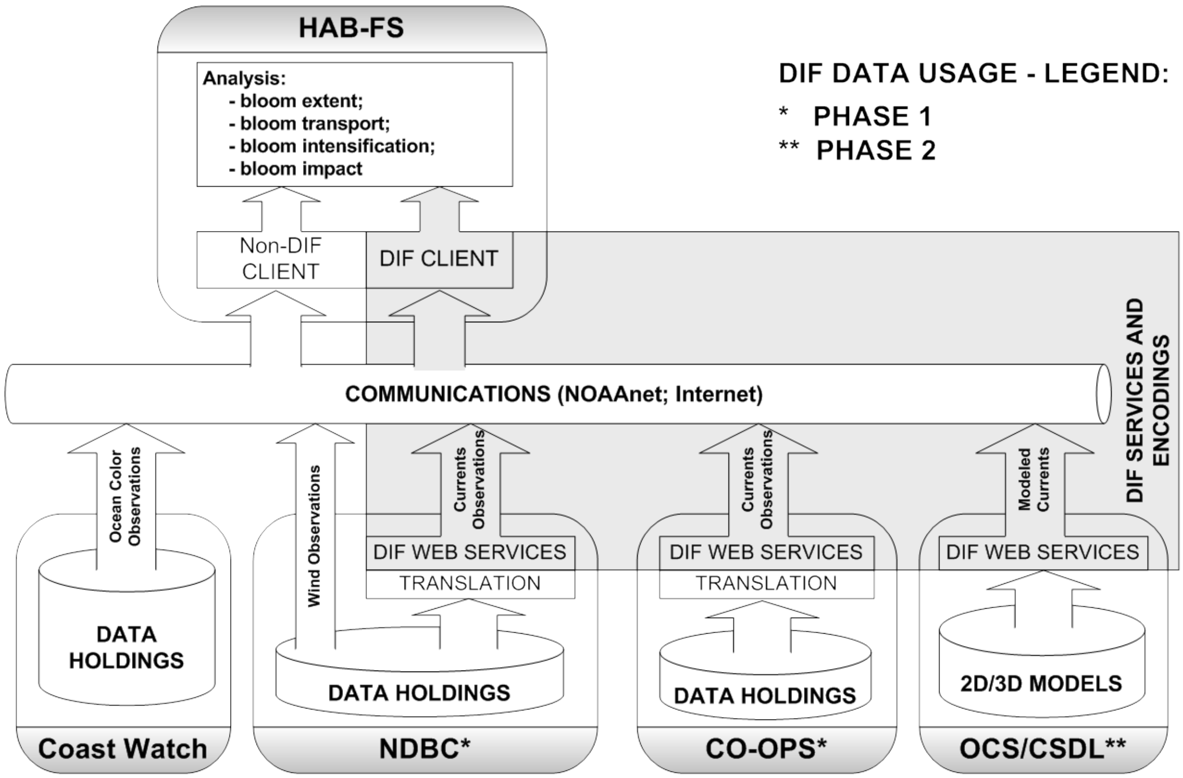


**Figure 3-1 HAB-FS Data Flow Diagram**

### *3.2.3.* Anticipated Benefits

A number of benefits were expected as a result of this project; categorized benefits are shown in Table 3-2.

| # | Anticipated Benefit | Category[1] |
|---|---|---|
| **1** | A measurable increase in the spatial and temporal accuracy (skill) and precision (reproducibility) of HAB forecasts | A |
| **2** | Increased probability of providing an accurate nowcast for the time periods when satellite imagery is not available due to clouds | A |
| **3** | Extension of forecast spatial range to the areas where the forecast has been previously unavailable, e.g., Tampa Bay | A |
| **4** | Extension of forecast temporal range from 3 days to 5 days and more | A |
| **5** | Increased HAB forecast objectivity | C |

[1]See Table 3-1

**Table 3-2 HAB-FS Benefits**

## *3.2.4.* Performance Assessment Methodology

### *3.2.4.1.* Assessment Methods and Benchmarks

The assessment method was planned to be a comparison of the performance of the existing HAB-FS bulletins against the performance of the enhanced modeling process (for the specific, selected periods). The spatial and temporal accuracy of the existing HAB forecasts was intended to serve as the baseline performance. Unfortunately, it was determined that, although the bulletins provided textual predictions/forecasts of the HAB behavior, these predictions often were not detailed enough to be directly compared with the enhanced model-based performance.

Two other factors limited the ability to quantify the performance of either the baseline or enhanced process:

- Satellite imagery (chlorophyll anomaly) of the HAB hindcast events was unavailable for some portions of the hindcast time period due to clouds
- Even when the imagery was available, the chlorophyll anomaly in the image could not be confirmed as a harmful algal bloom without cell count data; such cell count data was not available.

### *3.2.4.2.* Metrics

Each benefit described above has a corresponding metric. For consistency with previous HAB-FS skill assessments, the same basic HAB parameters will be used as some of the metrics for the performance assessment. The primary metrics traditionally used for HAB-FS skill assessment are:

1. **TRANSPORT**: the direction (north, south, offshore, onshore, etc.) in which the bloom is likely to migrate.
2. **EXTENT**: increase or decrease in bloom area.
3. **INTENSIFICATION**: expected change in HAB concentration (increase, decrease, no change).
4. **IMPACT**: presence of adverse coastal conditions, including respiratory irritation and presence of dead fish (from no expected impacts to very low, low, moderate or high impact).

It was proven in the process of the initial model testing that INTENSIFICATION metric cannot be calculated reliably, and will likely be dropped as a project metric. IMPACT requires in-situ observations, and therefore has been rejected as well.

It was suggested in the initial testing that the best way to calculate the TRANSPORT metric would be to define a trajectory of the bloom’s center. Nevertheless, later on the scope has moved to more sophisticated methods of tracking the bloom edges.

EXTENT metric was suggested to be calculated as a probability of the single model particle to be in certain place at certain time. It was assumed that the probabilistic approach would result in a meaningful guidance for a HAB area expansion based on certain probability level.

The metrics and benefits cross-reference is shown in Table 3-3.

| Benefit | Metric | Measurement Unit |
|---|---|---|
| **1** | Gain in TRANSPORT forecast accuracy / precision | Percentage skill increase |
| | Gain in EXTENT forecast accuracy | Kilometers (miles) and/or percentage skill increase |
| | Gain in EXTENT forecast precision | Percentage skill increase |
| **2** | Gain in TRANSPORT forecast availability | Number of cases when TRANSPORT nowcast was available |
| | Gain in EXTENT forecast availability | Number of cases when EXTENT nowcast was available |
| **3** | Forecast availability area growth | Square kilometers (miles) or percentage skill increase |
| **4** | Forecast length increase | Days or percentage skill increase |
| **6** | Degree of objectivity in HAB-FS | TBD. This may be a subjective or anecdotal result. |

**Table 3-3 HAB-FS Metrics**

## *3.2.5.* Project Performance Assessment

### *3.2.5.1.* Project Results and Anticipated Benefits

As the project consisted of two independent phases, the performance of each phase was assessed separately. The results of Phase 1 and Phase 2 are also very different – while Phase 1 focused on augmentation of the current operational HAB forecast system by allowing HAB analysts to get access to in-situ currents observations via the existing HAB-FS software, Phase 2 was more of a research of a possible benefits of the adoption of modeled currents in conjunction with particle tracking models.

#### 3.2.5.1.1. **Phase 1**

As a result of the Phase 1 development, the in-situ currents observations provided by CO-OPS and NDBC are now being ingested into the HAB-FS bulletin software. The data is then displayed for bloom analysis in the HAB-FS software using an integrated Google Maps application. The Google Maps tool in Figure 3-2 shows the CO-OPS and NDBC stations in the Gulf of Mexico along southwest Florida shore (where the HAB-FS is presently operational), which provide surface currents observation data. The visual drifting paths shown in Figure 3-3 are calculated from the observations of current direction and speed; observations are obtained through NDBC's and CO-OPS' Web services.

All station data, drifting paths, and observations, which are displayed along the virtual drifting paths, are now available for HAB-FS analysts to enhance the resulting forecast. The analysts use the IOOS currents data to determine the potential transport and extent of a confirmed or potential harmful algal bloom over the past 72 hours when satellite chlorophyll imagery is unavailable. When a confirmed or potential HAB is identified in close proximity to one of the stations, the analysts use the in-situ currents data in conjunction with the marine forecasted winds to form a nowcast prediction of the relative direction the bloom may have transported in over the past 72 hours. This is helpful when satellite imagery for the previous 72 hours is not available; the analysts can predict the bloom's location so that a forecast can be developed. The in-situ currents data and forecasted winds are also used to predict potential intensification of a bloom over the last 72 hours. Transport and intensification nowcasts are included in the text of the HAB-FS bulletin whenever possible in order to give all interested parties an idea of possible bloom location, and also provide an indication of whether the bloom may have intensified.

Integration of the IOOS in-situ observations allows the analysts to enhance nowcast predictions for harmful blooms; without in-situ currents data the nowcasts were only possible using remote sensing chlorophyll data and observed wind data. However, the observations can only be used if a bloom is located in the vicinity of a station that can measure surface currents. The challenge at present is the limited geographic extent of the existing currents stations in the regions where HABs occur (see Figure 3-2); this limitation could be reduced through the addition of HF Radar currents observations and/or modeled currents for the areas where in-situ measurements are not available.

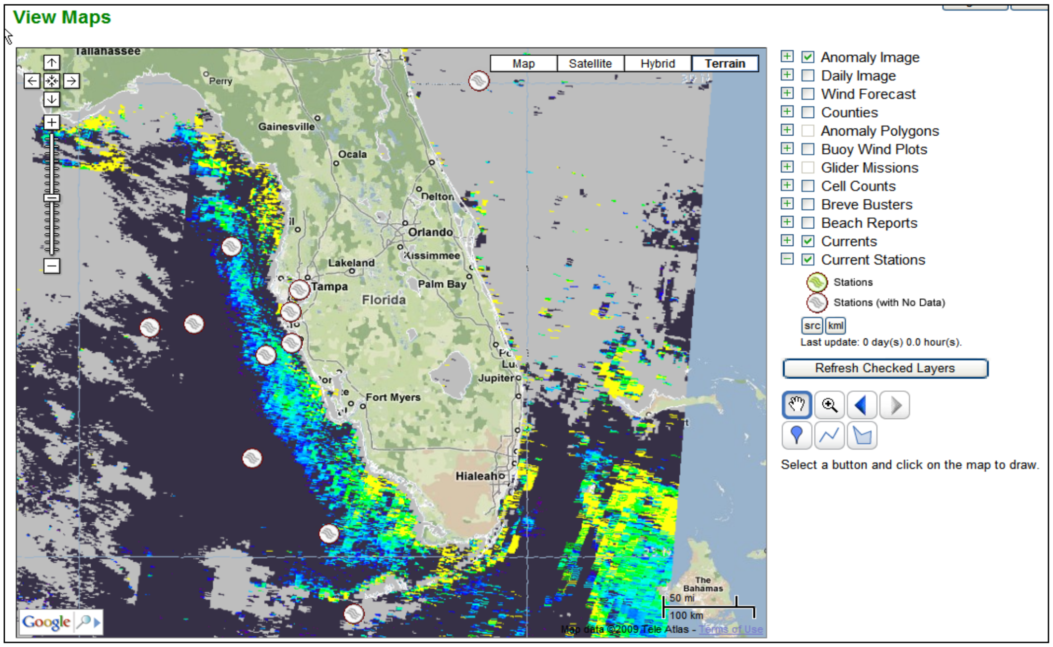


**Figure 3-2 HAB-FS iGoogle tool showing CO-OPS and NDBC currents stations in operational area of the Gulf of Mexico (overlay on chlorophyll image)**

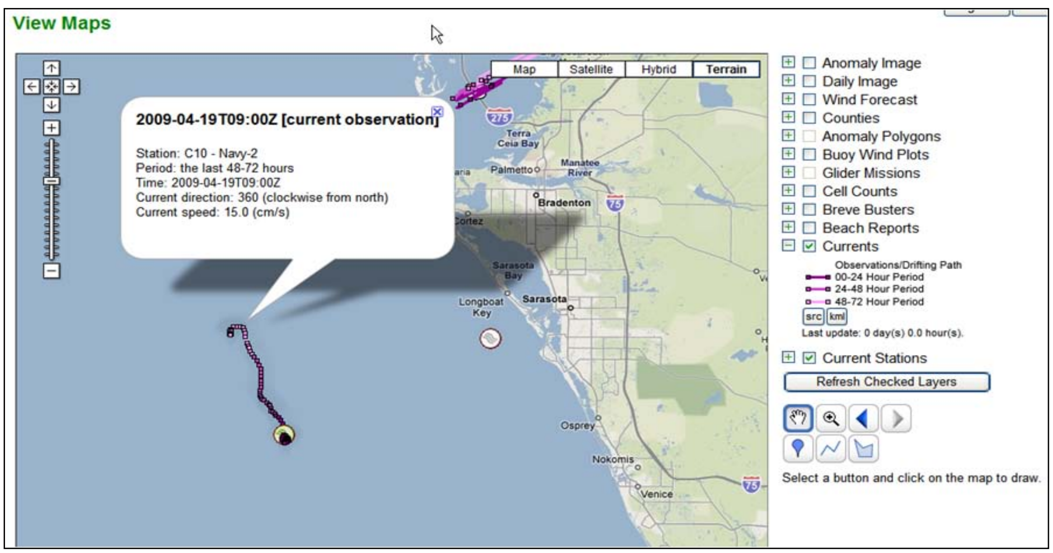


**Figure 3-3 Surface current observations for the past 72 hours**

Although HAB-FS analysts have subjectively concurred that Phase 1 of the project had a positive effect, there was no way to assess the Phase 1 performance in terms of the established methodology (comparison of baseline performance with enhanced performance). Skill assessments for HAB forecasts for the last several years are not available, so no skill baseline could be established against which enhanced performance could be compared. Therefore, the decision was made not to assess Phase 1, but to focus on Phase 2.

#### 3.2.5.1.2. **Phase 2**

As stated above, two particle tracking trajectory models were used in the Phase 2 of the project: (1) the particle tracking trajectory model that is a component of NOAA's GNOME (General NOAA Operational Modeling Environment), and (2) a trajectory model that supports Chesapeake Bay Oyster Larvae Tracker (CBOLT).

GNOME's tracking component is two-dimensional; whereas the trajectory model used by CBOLT employs a fully three-dimensional approach.

**GNOME** trajectory model was developed by the Emergency Response Division (ERD) of NOAA's Office of Response and Restoration (OR&R), and is available at http://response.restoration.noaa.gov/software/gnome/gnome.html. GNOME trajectory model has been used by OR&R ERD responders during an oil spill to:

- estimate the trajectory of spills by processing information about wind and weather conditions, circulation patterns, and river flow;
- predict the trajectories that can result from the uncertainty in current and wind observations and forecasts;
- provide trajectory output (including uncertainty estimates) in a geo-referenced format that can be used as input to geographic information systems.

A GNOME user just has to describe a spill scenario by entering information into the program; GNOME then creates and displays an oil spill "movie" showing the predicted trajectory of the oil spilled in the scenario.

**GNOME** trajectory model is essentially two-dimensional; it was designed to accurately trace the path of nearly inert particles transported horizontally by the velocity fields of a hydrodynamic model, i.e. modeled or observed currents. Despite the 2D limitations, the GNOME trajectory model provided accurate HAB transport and extent predictions in the Great Lakes where it has been used for the last several years. However, when using the 2D model in the Gulf of Mexico, the impact on the HAB due to a significant coastal upwelling was not well modeled, so it was determined that a 3D trajectory model like CBOLT would be more effective.

**CBOLT** trajectory model was developed by CSDL to meet the needs of CSC Chesapeake oyster project and other remote sensing applications. Unlike GNOME or any other 2D tracker, it simulates an active particle swimming behavior, allowing it to control its vertical position in the water column; the behavior depends on the salinity, temperature and depth, as well as the life history of the particle. The CBOLT model allows simulating an impact of 3D environment, e.g. coastal upwelling, making the trace much more accurate.

During the two time periods selected, the remotely sensed chlorophyll anomaly product was digitized and used to delineate *K. brevis* blooms. The digital imagery was used both to provide the initial conditions for the trajectory models, and for validation of the model results.

The project performance metrics depended on the information provided in the HAB Bulletin. Unfortunately, forecast information in the HAB Bulletin was more verbose/verbal and qualitative rather than numerical, and that did not permit a quantitative assessment to be made. The following statements are typical for the Bulletins:

*"Strong northerly winds from Saturday night through Monday make intensification and continued southern transport of the bloom likely."*

*"Imagery is not available for the last 48 hours, the bloom may have moved slightly south."*

*"Current extent of bloom cannot be estimated."*

*"Recent and forecasted northerly winds may continue southerly transport and promote upwelling favorable conditions through Tuesday."*

*"...harmful algal bloom in the Tampa Bay region has expanded southward, trailing offshore of North Captiva..."*

*"Variable winds will maintain bloom location or southward current may cause slightly southerly movement..."*

*"Much of the recent satellite imagery has been obscured by clouds...."*

*"Variable winds and southward currents will likely cause southward and offshore expansion..."*

These qualitative forecast statements, used as the performance baseline, prevented an objective performance analysis for the project skill assessment.

As it was mentioned above, the availability of clear and unbiased remotely sensed chlorophyll imagery for the test period was another key factor of the project skill assessment. The imagery is extremely important as it allows an accurate benchmarking of the proposed algorithms against the established baseline. The imagery is usually taken by satellite; therefore, when the area of a HAB is covered with clouds or the satellite is not right over it, the image quality is affected. In these cases imagery is either completely unavailable, or the available imagery is biased and thus unreliable. The excerpts from the Bulletins provided above proved that problems with imagery have been a frequent issue.

The availability of the imagery was one of the main reasons for the selection of the hindcast time domain: the most clear and unbiased anomaly images were obtained for these periods. In addition, it turned out that some skill assessment of the operational HAB-FS had been done for approximately same period of time, which provided a relatively good baseline for the project performance metrics.

The performance of the Phase 2 of the project was assessed in accordance with the suggested methodologies, baseline and metrics. The assessment results are tabulated in Table 3-4. Some of the metrics had insufficient amount of data for direct and exhaustive comparison with the baseline; for some metrics, the project covered just a portion of the baseline time period. Due to these constraints, the enhanced performance was quantifiable not for the whole set of metrics; for some metrics just indirect, qualitative and anecdotal evaluation was possible.

However, with all reservations and limitations, the project results undisputedly demonstrated that the use of a three-dimensional trajectory model can produce viable forecasts for up to one week, while the use of a two-dimensional tracker is insufficient to provide such forecasts because the nature of the HAB processes was proved to be essentially three-dimensional. Although the full-scale operational use of the three-dimensional tracker still needs verification, the process can be used to augment existing HAB-FS.

| # | Metric | Baseline Performance | Enhanced Performance | DIF IMPACT | |
|---|---|---|---|---|---|
| **1** | Increased HAB intensification forecast availability and accuracy (# forecasts, % improvement) | Intensification forecast availability: 38% Intensification accuracy: 73% | Physical model used does not adequately support prediction of bloom Intensification. | Insufficient data for evaluation | Project model does not forecast intensification |
| **2** | HAB transport and extent forecast availability increase (# forecasts) | Transport forecast availability: 83% Extent forecast availability: 21% | Enhanced model forecast availability is >85% (although skill decreases when imagery is cloudy or unavailable). | ***Positive*** | Results are encouraging. See Additional Benefits |
| **3** | Improved HAB transport and extent forecast accuracy (% improvement) | Transport forecast accuracy: 90% Extent forecast accuracy: 77% | Project study period somewhat different from baseline period; baseline and enhanced data not directly comparable. | Insufficient data for evaluation | Results are encouraging. See Additional Benefits |
| **4** | Extent precision | Not less than 30 km | Less than 20 km | ***Positive*** | 30% gain |
| **5** | Increase in geographic coverage area for operational forecast (miles or additional features) | Entire eastern gulf of Mexico (from the FL panhandle through to the Keys) [1] | Enhanced process has limited domain, existing process would be used outside domain. | Insufficient data for evaluation | Result seems to be neutral |
| **6** | Forecast length increase (% improvement) | 3 or 4 day forecasts when a bloom is present | Forecast length unrestricted; skill persists up to 7 days | ***Positive*** | About 100% gain |
| **7** | Degree of objectivity in HAB-FS (qualitative assessment) | Existing bulletin based on analyst synthesis of various data | Bloom particles transported using 3-D model; graphical forecast more objective | ***Positive*** | See Additional Benefits |

[1] A bloom originated in FL and moves into AL and MS, is tracked but HAB-FS is not fully operational there.

**Table 3-4 HAB-FS Performance**

#### *3.2.5.2.* Additional Results and Benefits

In addition to the anticipated benefits which were the formal metrics for the project, a number of additional benefits were identified during the course of the Phase 2 project execution:

- Forecasting: The project demonstrated a potential capability of producing chlorophyll anomaly transport and extent forecasts, derived from SeaWiFS imagery. The forecasts, in conjunction with cell count confirmation, could serve as additional guidance to the HAB analyst.
- Resolution: The project demonstrated potential improvement in the spatial and temporal resolution of HAB Extent forecasts. The enhanced process spatial error did not exceed 18 km in 7 day period, the existing HAB-FS cannot be resolved at scales finer than 30 km and forecast cannot be longer than 3 to 4 days.
- Observations: The importance of subsurface chlorophyll and cell count observations to HAB forecasting was highlighted
- Other benefits: As a result of the project, NOAA subject matter experts have a better understanding of the limitations of a physical model in predicting bloom transport and extent, the importance of biological factors in bloom transport, extent and intensification, and the limitations of a 2-D particle tracker versus 3-D

## 3.3. Coastal Inundation

### *3.3.1.* Decision Tool/Model to Be Enhanced

The SLOSH Display Program (SDP) application was developed by NWS' Meteorological Development Lab (MDL) and is used to assist forecasters in evaluating the threat from storm surge. The SDP is used by NWS Weather Forecast Offices (WFOs) to inform their Hurricane Local Statements (HLS'), and by WFOs and the Tropical Prediction Center (TPC) to communicate storm surge information to state and local Emergency Managers to help inform their public safety decisions, including determination of which areas to evacuate. Some Emergency Managers themselves use the SDP during storm preparations.

### *3.3.2.* Scope, Architecture, and Data

The SDP enhancement project integrated DIF water level, winds, tide predictions, and Low Astronomical Tide/High Astronomical Tide (LAT/HAT) products into the SDP to provide a much more rich and integrated set of surge-related data for users.

For the TPC team, the project provided a display of IOOS DIF water level observations, tide predictions and other water level products for user selectable locations within the SLOSH basin. TPC users would benefit from an enhanced visualization of the data for use with media and other users.

For the WFO forecasters, who upload the TPC operational storm surge forecast into the SDP, this project's enhancements integrated the IOOS DIF wind, water level observations, tidal predictions, and other water level products with the surge forecast. This allowed forecaster to compare the observations to the forecasts, and to have additional information to determine total water level, which in turn assisted them with their Hurricane Local Statements. They also used the enhanced visualization to communicate with local emergency managers during a tropical event.

The data for the project is provided by CO-OPS and NDBC in IOOS DIF format for in-situ observations, i.e. IOOS GML, via SOS Web service. CO-OPS provided real-time observations of water level and winds, as well as tidal predictions and datum. NDBC delivered real-time wind observations and collections of water level observations (requests for data from series of observing stations rather than individual stations) on behalf of CO-OPS. The project architecture and data flow diagram is presented in Figure 3-4.

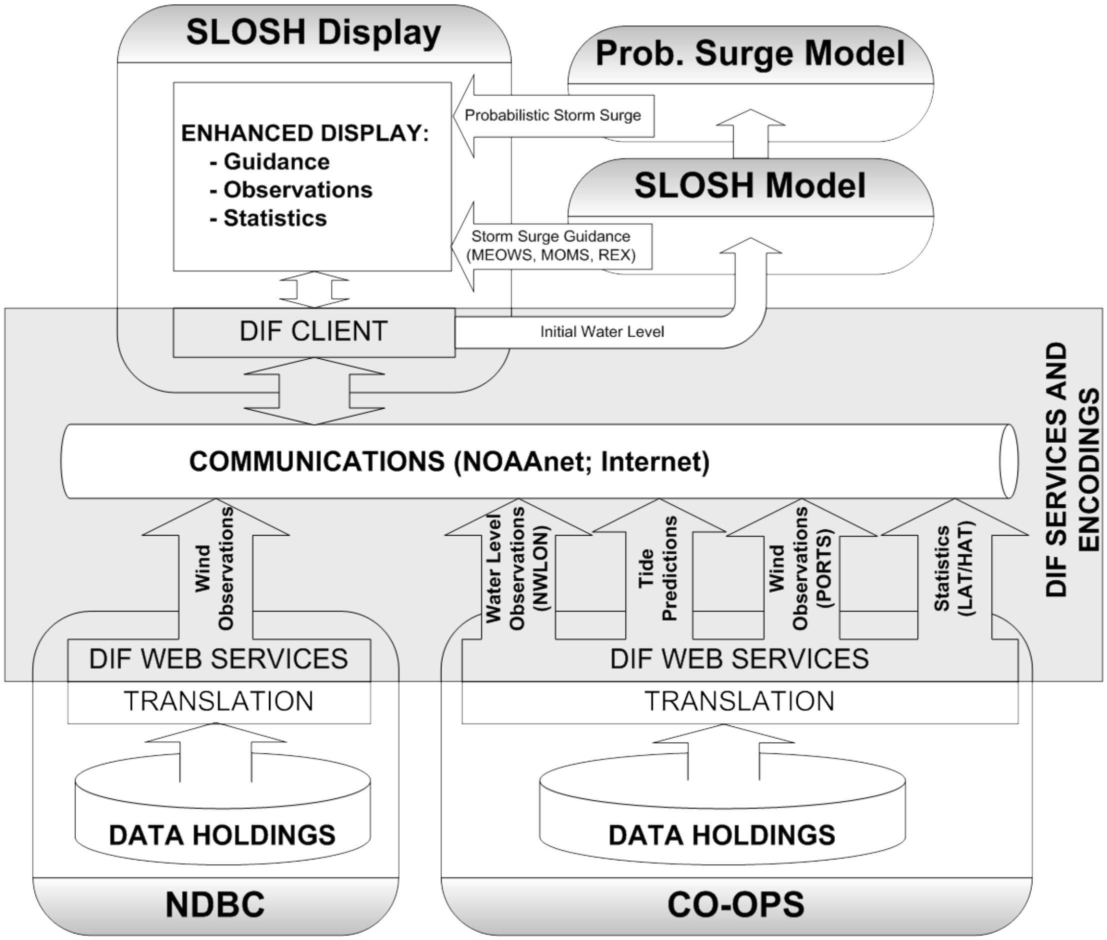


**Figure 3-4 CI Data Flow Diagram**

### *3.3.3.* Anticipated Benefits

TPC and all coastal WFOs are both expected to benefit from the project, however, due to different goals and objectives, the anticipated benefits are somewhat different so are described separately below.

#### *3.3.3.1.* Benefits to TPC

The list of categorized benefits to TPC is shown in Table 3-5.

| # | Benefit | Category[1] |
|---|---|---|
| **1** | Increased efficiency and reliability of access to real-time water level data for SLOSH model initialization (using DIF services to acquire the needed data from either CO-OPS or NDBC source instead of “web scraping” from sites). | A |
| **2** | Ability to display real-time water levels in the SLOSH Display program, particularly useful during TPC briefings to external users during the 24 hours before landfall. This will help communicate the actual situation to emergency managers and general public as the storm progresses (currently, TPC does not communicate real-time water levels to external audiences). | B |
| **3** | Visualization enhancement to the SDP to facilitate communication will dramatically improve the presentation of the information provided, to better inform emergency managers and external users. | B |

[1]See Table 3-1

**Table 3-5 CI Benefits to TPC**

#### *3.3.3.2.* Benefits to WFOs

The list of categorized benefits to WFOs is shown in Table 3-6.

| # | Anticipated Benefit | Category[1] |
|---|---|---|
| **4** | The enhanced display and graphics will help better communicate the potential coastal inundation impact to local emergency managers. | C |
| **5** | The integration of real-time water-level observations into the SLOSH display will allow forecasters to adjust the surge portion of their HLS (Hurricane Local Statement), providing more up-to-date and accurate guidance. | C |

[1]See Table 3-1

**Table 3-6 CI Benefits to WFOs**

### *3.3.4.* Performance Assessment Methodology

#### *3.3.4.1.* Assessment Methods and Benchmarks

Benefit #1 is the only anticipated benefit that can be quantifiably benchmarked and assessed (Category A). The existing method of collecting real-time observation data, tide predictions, and statistics (typically by going to multiple web sites and performing manual calculations) takes a significant amount of time. This is expected to decrease with the integration of all the data into the SLOSH Display. The existing procedure, specifically amount of time it currently takes to obtain and process this data, served as a baseline. The value of benefit is a measure of a direct comparison of the demonstrated improvement using the enhanced SLOSH Display with the baseline.

The Category B benefits, #2 and #3, arisen from completely new functionality that has never existed before. The numerical baseline for such results could not be determined since TPC had never displayed water levels in their media briefings with the legacy SDP. As for the enhanced SDP performance, the number of times in a FY09 hurricane season that the water level observations were presented in briefings with a help of the enhanced SDP, was suggested as a quantifiable measure of the benefit of #2. Similarly, the visual enhancements of the SDP interface are new, and the suggested quantifiable measure for #3 is the number of times in a season that the visual enhancements are used in media briefings.

Category C benefits, #4 and #5, are largely subjective and have no real benchmarks. Therefore, benefits #4 and #5 were expected to be assessed only in a qualitative fashion. In addition to the benchmarks and methodologies described above, the anecdotal feedback regarding the value of the enhancements was anticipated.

#### *3.3.4.2.* Metrics

The proposed set of metrics along with the measurement units in reference to the benefits is shown in Table 3-7.

| Benefit | Metric | Measurement Unit |
|---|---|---|
| **TPC #1** | Improved availability of real-time observations and tide predictions | Percentage improvement |
| | Decrease of time required for data collection and processing for model initialization | Hours or percentage improvement |
| **TPC #2** | Presentation of real-time observation data, tide predictions, and statistics to the external audience through the season. | Number of events |
| **TPC #3** | Presentation of enhanced display to the external audience through the season | Number of events |
| **WFOs #4** | Improvement in HLS forecast due to additional SLOSH Display data | Qualitative assessment |
| **WFOs #5** | Improved communication with local emergency managers in pre-storm time frame | Qualitative assessment |

**Table 3-7 CI Metrics**

## *3.3.5.* Project Performance Assessment

### *3.3.5.1.* Project Results and Anticipated Benefits

The project performance was assessed in accordance with the suggested methodologies, baseline and metrics. The assessment results are tabulated in Table 3-8.

Unfortunately, due to the lack of hurricanes through the season of FY09, there was no opportunity for extensive use of SDP in an operational environment.

| # | Metric | Baseline Performance | Enhanced Performance | DIF impact | |
|---|---|---|---|---|---|
| **1** | Increased efficiency and reliability of access to real-time water level data for SLOSH model initialization (% improvement) | No estimate provided. | Easy access to water level data during pre-storm landfall. | ***Positive*** | More efficient access to data for TPC operators. |
| **2** | Graphic display of real-time observation data, tide predictions, and statistics to external audiences via SDP. (# events) | # of events = 0[1] | # of events = 5[2] | ***Positive*** | New capabilities were used operationally. |
| **3** | Enhanced GIS display for briefings to external audiences via SDP. (# events) | # of events = 0[1] | # of events = 7[2] | ***Positive*** | New capabilities were used operationally. |
| **4** | Improved communication with local emergency managers in pre-storm time frame (qualitative assessment) | Baseline data not available. | Visual improvements anticipated to be positive | ***Neutral*** | Very little opportunity to test; seems positive but addl enhancements desired. |
| **5** | Improved HLS forecast due to DIF data integration into SDP (qualitative assessment) | No baseline data. In general, no skill assessment on HLS'. | No DIF data was used for HLS during the season. | Insufficient data for evaluation | Not enough storms during test period for evaluation |

[1] - Existing SDP does not include this data

[2] - TPC: during hurricanes Bill and Ida; tropical storms Claudette and Denny
WFOs: hurricane Ida exercise

**Table 3-8 CI Performance**

### *3.3.5.2.* Additional Results and Benefits

In addition to the anticipated benefits which were the formal metrics for the project, a number of additional benefits were identified during the course of the SDP project execution:

- The standardization of the data formats and methods for requesting and receiving the data across multiple providers significantly reduced the software coding complexity for client-side application developers. With previous legacy formats and access methods, separate and custom client-side software modules were required for each data provider. Each module would be customized to match the data provider's legacy request/response protocol and data format. Each time a client application wants to add a data provider, a new custom module is required. By contrast, the standardized access and data formats adopted by the DIF, across multiple data providers, simplified the task for the client-side software developer. A single request/response module is compatible with any DIF-compliant data provider (if they are using the same data type and access method); only minor changes are required to add a new data source. This increases the value of the initial software development investment, and reduces long-term costs of software maintenance enhancements.

  The SDP experience illustrates this benefit:

  - The client-side SDP application developer spent approximately 5 days developing the module to harvest and graphing wind data from CO-OPS. Repeating the process to harvest wind data from NDBC took only an additional 2 days of effort.
  - The software module to ingest collections of water level data from NDBC took approximately 5 days to develop. Repeating the process to ingest collections of water level data from CO-OPS (hosted by NDBC) took only 1 hour.

- The SDP with its enhanced functionality was extremely well received by forecasters at the two test WFOs. So much so that, at their request, it has been operationally deployed to all coastal WFOs instead of just the original two test WFOs.
- The enhanced SDP is now web-downloadable for easy updates and installations by users. This is a significant improvement to the previous process of mailing DVDs to users each time a new version was released.
- TPC had intended to use SDP in media briefings and to initialize the SLOSH model, and in practice they have also found it useful to monitor real-time tide levels. It was used by TPC in this manner during Tropical Storm Claudette in August, 2009 because they felt the tide level plot in SDP was easier to read than web-based sources of the data.
- Some of the client-side software modules developed for SDP can be used as reference software implementations for other application developers. The code used to parse the IOOS GML data is available on the DIF website (http://ioos.gov/dif, Software) for use by any other application that would ingest point data in the IOOS GML format.
- The project was successful collaboration that crossed lines offices (NOS/NWS) and produced a deliverable aimed at meeting specific user needs. The enhanced SDP received very favorable reviews from many offices. Below are testimonials from some key stakeholders:
  - "Many of the enhancements to the SDP were based specifically on meeting NWS's evolving needs for producing storm surge forecasts. We are especially pleased to now have the ability to integrate water level observations now made available in one location versus the old method of collecting the data via the internet. Such an enhancement will streamline our data collection process during hurricane landfall situations. Additionally, we are pleased to see improved functionality and plotting options which greatly increases the usability of the SDP for briefings." Bill Read, Director of NHC
  - "NHC is very pleased with the enhancements to the SDP and the tool will now better serve our needs during land falling cyclone events. We are especially pleased to see the increased capabilities (Inundation, GIS layering, NOS tide stations, etc) and the increased access (web download/update)." Jamie Rhome, Storm Surge Team Lead, National Hurricane Center

## 3.4. Hurricane Intensity

The Hurricane Intensity project expected to prove the critical importance of ocean data in tropical cyclone intensity prediction, as existing data does not describe mesoscale features, which are vital for hurricane intensification modeling.

Key conclusions reached as a result of this project, that will continue to influence hurricane intensity prediction, include:

- New synthetic temperature profiles introduced mesoscale features to the model thus providing a new capacity that has not been available before.
- Synthetic temperature profiles derived from satellite altimetry appear to have a very small error with depth (approximately 1ºC) when compared against actual in-situ observations.
- A working environment can be implemented to create and distribute these synthetic profiles using state-of-the-art technologies.
- When these profiles are included in HYCOM runs, the output of the model provides an improved upper ocean thermal structure field, which results in a small and consistent reduction of intensity error in the coupled HWRF-HYCOM runs.
- Tools: In order to ingest the profiles into NCEP's models, they needed to be converted from netCDF (the DIF-adopted format for gridded data) to BUFR (used by NCEP models). A general purpose netCDF-to-BUFR translator is available (http://ioos.gov/dif, Software) for any other users that might need it

### *3.4.1.* Decision Tool/Model to Be Enhanced

This project was focused on improving a NOAA/NWS's National Centers for Environmental Prediction (NCEP) operational hurricane prediction model, which contains two components: the atmosphere component – Hurricane Weather Research and Forecasting (HWRF), coupled to the ocean component – Hybrid Coordinate Ocean Model (HYCOM). The HWRF ingests the HYCOM output, while HYCOM is initialized and forced on the boundaries with data from the operational Real-time Ocean Forecast System (RTOFS)–Atlantic model, which is also based on HYCOM. RTOFS-Atlantic is capable of assimilating additional observational or synthetic data sets from NCEP Local Data Repositories.

### *3.4.2.* Scope, Architecture, and Data

Research indicates that tropical cyclones intensify over warm mesoscale features located in the open ocean. The two-fold purpose of this project was to evaluate this effect by ingesting new ocean temperature data into air-sea numerical model and to assess the operational requirements of DIF-formatted data for these purposes.

The goal of this project was to evaluate the benefits of integration of ocean data in DIF standards into an air-sea numerical model, to aid the scientific and operational community to reduce the error in Atlantic hurricane intensity forecasts. Although tropical cyclones are formed in several basins, this work was concentrated in the Gulf of Mexico and Caribbean Sea. The objective of this work was to produce prompt results and analysis to aid the scientific and operational community in their effort to reduce the error in the forecast of intensity of tropical cyclones. If successful, the same tools and methodologies applied in this effort could be applied to the forecast of tropical cyclone intensity in all basins where tropical cyclones occur.

NOAA/NWS's NCEP collaborated with NOAA/OAR's Atlantic Oceanographic Marine Laboratory (AOML) to conduct a series of hindcasts for three Category 4-5 hurricanes from 2005 – Rita and Wilma in the Gulf of Mexico and Emily in the Caribbean Sea – to evaluate forecast accuracies with and without the new temperature data (Figure 3-5).

For the purpose of this project, AOML obtained temperature data from various sources, combined and translated data into IOOS DIF format. The resulting data sets were made available from the CoastWatch THREDDS server in netCDF/CF format. The data sets were also converted into BUFR, and delivered to NCEP via FTP, because currently NCEP can assimilate data only in that format, and only by FTP. The temperature data provided by AOML consisted of temperature profiles obtained from regional XBTs, AXBTs, profiling floats, thermistor chains, and moorings. Some of these data has been archived at NODC, while other data are available from AOML.

In addition, AOML provided synthetic sea temperature profiles for Emily (2005-07-06; 2005-07-13; 2005-07-20; 2005-07-27), Rita (2005-09-14; 2005-09-21; 2005-09-28) and Wilma (2005-10-12; 2005-10-19; 2005-10-26). The profiles were developed within the area shown in Figure 3-6 using historical statistical relationships between sea surface level change and deep-sea isotherm distribution. The sea surface level data was obtained from all available satellite-based altimeter sources, including NASA's Jason-1, ESA's Envisat, and US Navy's GFO. Depth of isotherms obtained from eXpendable BathyThermograph (XBT) probes, Connectivity-Temperature-Depth (CEB) sensors and floats were calculated for the area in Figure 3-6; the depths were then linearly regressed onto corresponding Sea Surface Height Anomaly (SSHA) values obtained from AVISO in a 0.5° x 0.5° grid. The synthetic temperature profiles were derived from the regression parameters and the daily real-time SSHA fields, and combined with the real time SST observations.

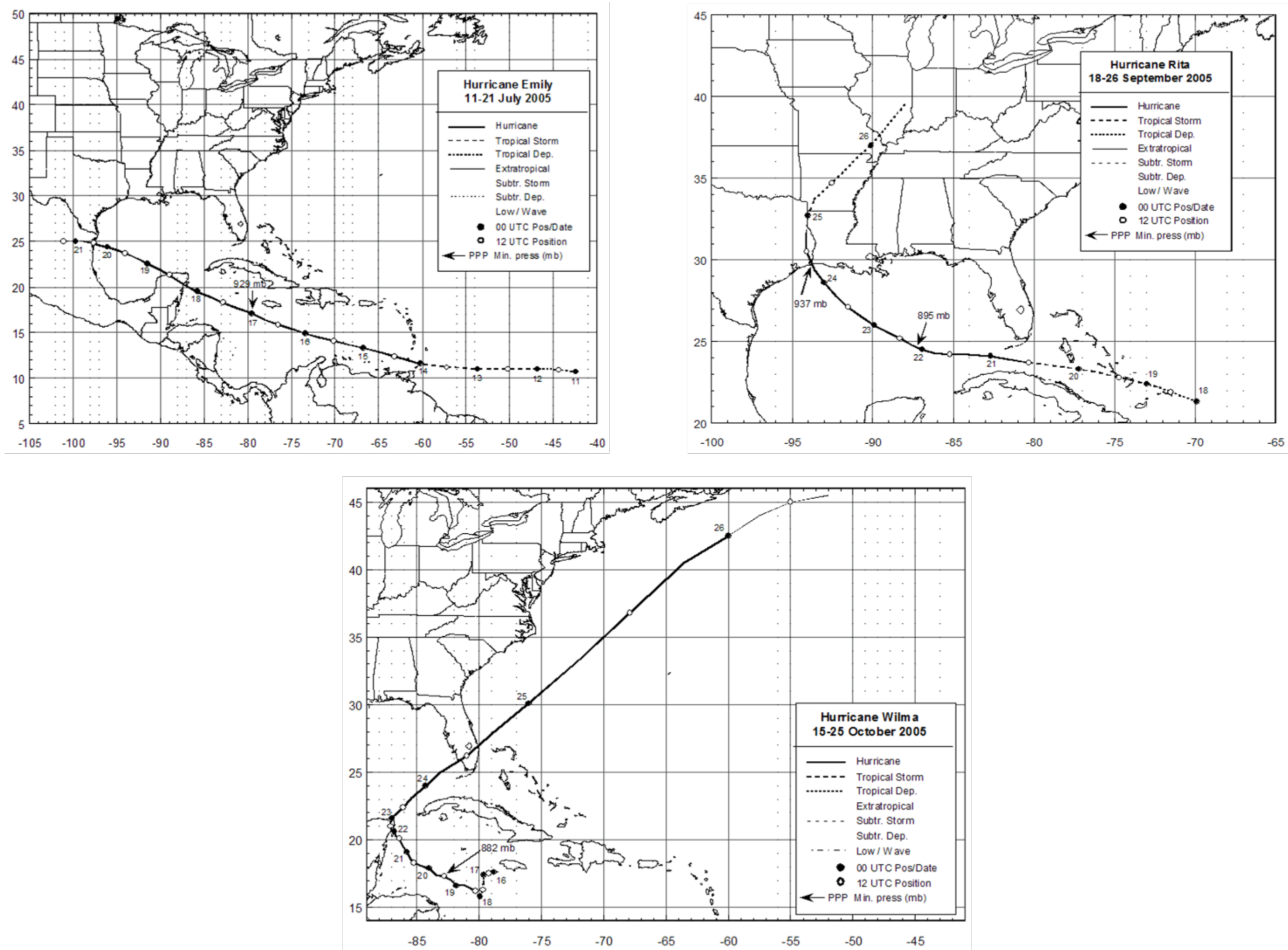


**Figure 3-5 Hurricanes Emily, Rita and Wilma**

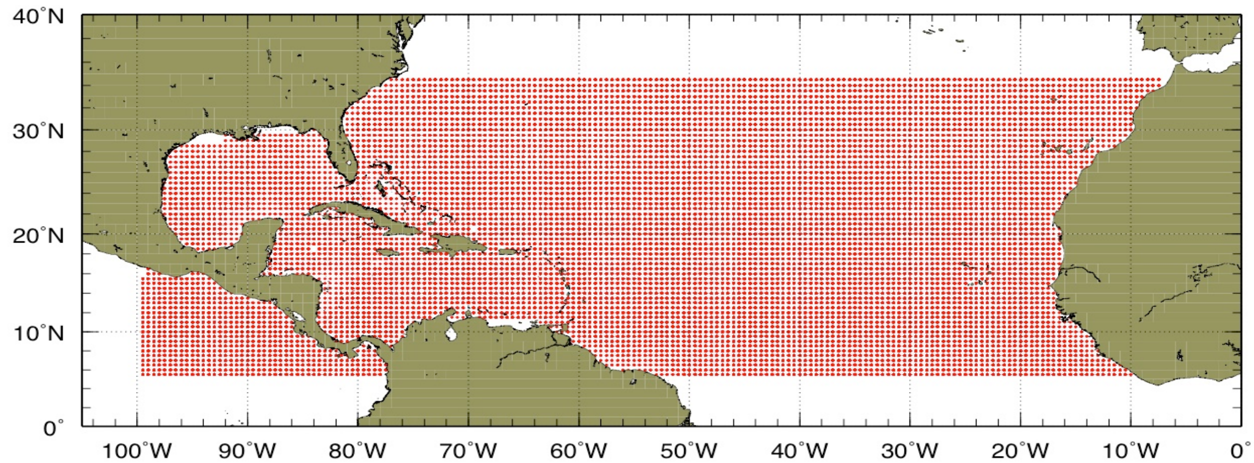


**Figure 3-6 Synthetic profiles region**

The synthetic temperature profiles were compared against the in-situ observations, and demonstrated very good correlation. In fact, the mean temperature difference was smaller than 1ºC, and the maximum discrepancy was about 3ºC.

The HYCOM model cannot assimilate temperature profiles alone; thus salinity profiles were also developed using monthly climatology values obtained from the NOAA's World Ocean Atlas (WOA2005). Finally, joint temperature and salinity profiles were created. Synthetic data output has been converted in DIF standard format for gridded data as well.

The project architecture and data flow diagram is presented in Figure 3-7.

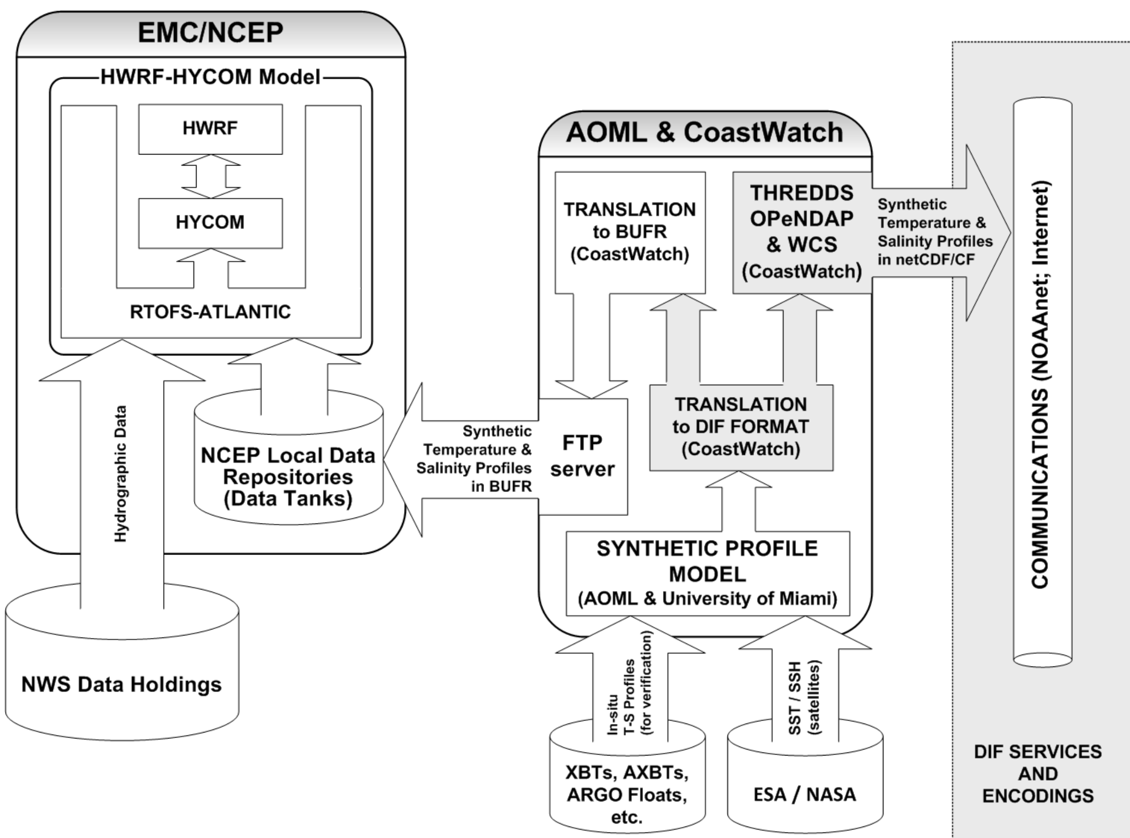


**Figure 3-7 HI Data Flow Diagram**

### *3.4.3.* Anticipated Benefits

The project has been expected to prove the critical importance of ocean data in tropical cyclone intensity prediction, as existing data does not describe mesoscale features, which are vital for hurricane intensification modeling. New synthetic temperature profiles introduced mesoscale features to the model thus providing a new capacity that has not been available before.

The list of categorized benefits is shown in Table 3-9.

| # | Anticipated Benefit | Category[1] |
|---|---|---|
| 1 | Reduction in the error of hurricane intensity and track forecasts. | A |
| 2 | Information on how to improve the ocean observing system for tropical cyclone prediction studies. | C |

[1]See Table 3-1

**Table 3-9 HI Benefits**

### *3.4.4.* Performance Assessment Methodology

#### *3.4.4.1.* Assessment Methods and Benchmarks

In regard to benefit #1, the error reduction of hurricane intensity and track forecast, the project the results of the HWRF-HYCOM runs for 3 major hurricanes from 2005 were evaluated with and without the integrated data. The numerical model results constituted a number of hindcasts, which were directly compared against the baseline. The historical forecasts issued by NHC for these 3 major events served as baseline for hurricane intensity and track.

The example of that approach is presented in Figure 3-8 and Figure 3-9, provided by Avichal Mehra (NOAA/NCEP). Figure 3-8 shows the results of one of the HWRF-HYCOM coupled model runs for Rita showing storm track and intensity. The “best” refers to the archived estimate provided by NHC for Rita, while “H209” is the operational HWRF (which uses POM instead of HYCOM).

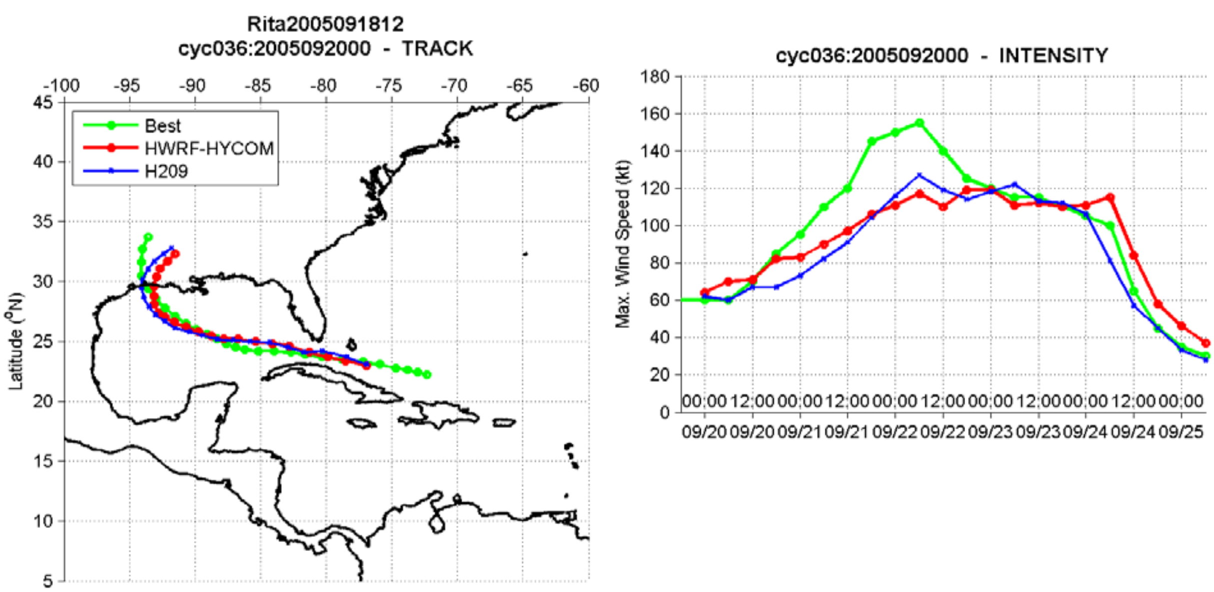


**Figure 3-8 Rita storm track and intensity (maximum wind speed) forecast**

In Figure 3-9, HWRF-HYCOM forecast intensity and track errors is presented relative to a best track (%) with variations (black line) over model runs (in parenthesis), as function of simulation hour; the same errors and variations for operational HWRF are shown in gray.

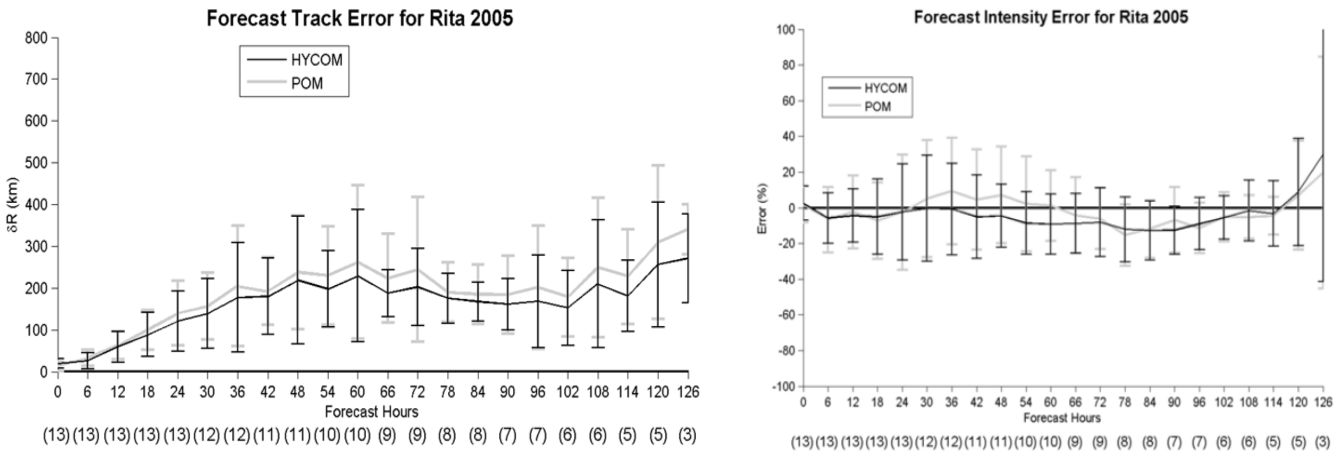


**Figure 3-9 Rita storm track and intensity forecast error**

In addition to forecast enhancement, the project provided valuable information on how to improve the ocean observing system to benefit tropical cyclone prediction studies. However, that benefit is primarily qualitative and does not have an established baseline. Thus, the benefit has been attributed to Category C.

#### *3.4.4.2.* Metrics

The proposed set of metrics along with the measurement units in reference to the benefits is shown in Table 3-10.

| Benefit | Metric | Measurement Unit |
|---|---|---|
| 1 | Error reduction in hurricane intensity and track forecast | Percentage error reduction |
| 2 | N/A | Qualitative assessment |

**Table 3-10 HI Metrics**

### *3.4.5.* Project Performance Assessment

#### *3.4.5.1.* Results and Anticipated Benefits

The operational ocean forecast system RTOFS-Atlantic was used in a hindcast mode for the 2005 Hurricane season. Two sets of runs were performed for each of the 3 storms of interest – with and without synthetic Temperature and Salinity profiles. The results of both sets were saved in NCEP's archives.

Then two sets of the coupled HWRF-HYCOM model runs were performed for the same 3 storms using the saved results of the RTOFS-Atlantic runs with and without synthetic profiles as the initial boundary conditions.

The results of the experiment showed some error reduction in the prediction of the hurricane intensity (HI), which is a measure of maximum sustained wind speeds during the certain storm. The error reduction is illustrated for the hurricanes Rita, Emily and Wilma by Figure 3-10, Figure 3-11 and Figure 3-12 respectively.

The plots in the figures represent mean error of HI in the forecast as compared to the National Hurricane Center's official HI estimates based on observations and existing operational prediction systems. These mean errors are plotted against the simulation hour with the number within the brackets below the hour indicating the number of forecasts (runs) used to compute the mean error. Such comparison is one of the preferred metrics for evaluating performance of a hurricane forecast system at NCEP. For all three plots, the green line is for runs without using DIF data while the red line is for those experiments where the impact of additional DIF synthetic profiles on the ocean state was included.

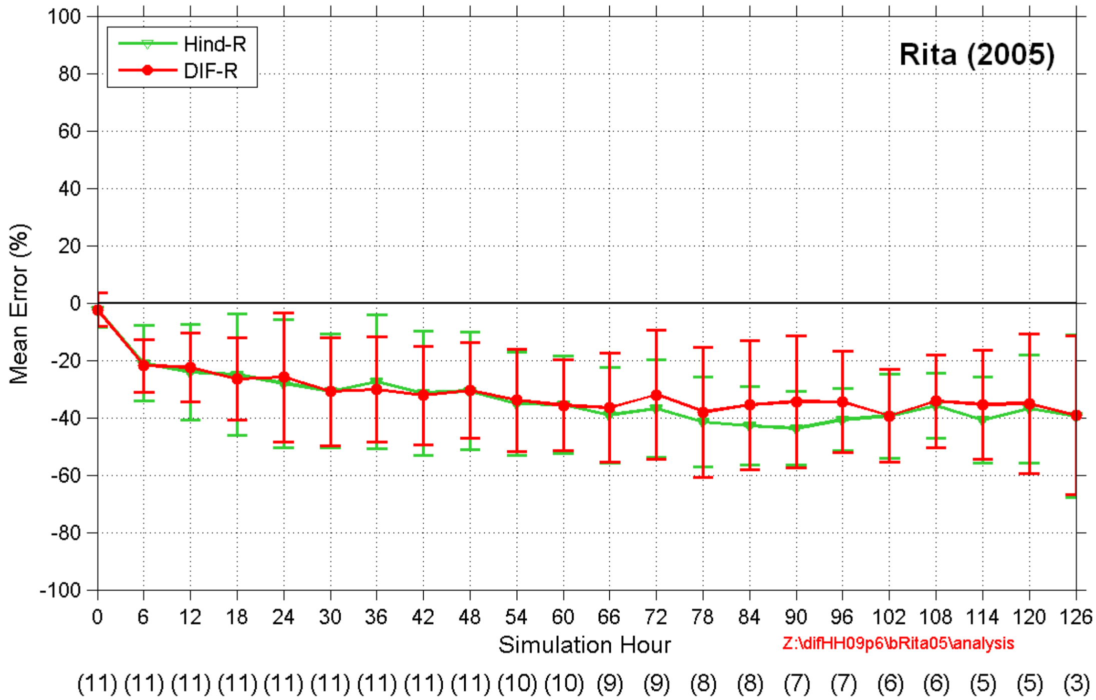


**Figure 3-10 Intensity error comparison for hurricane Rita**

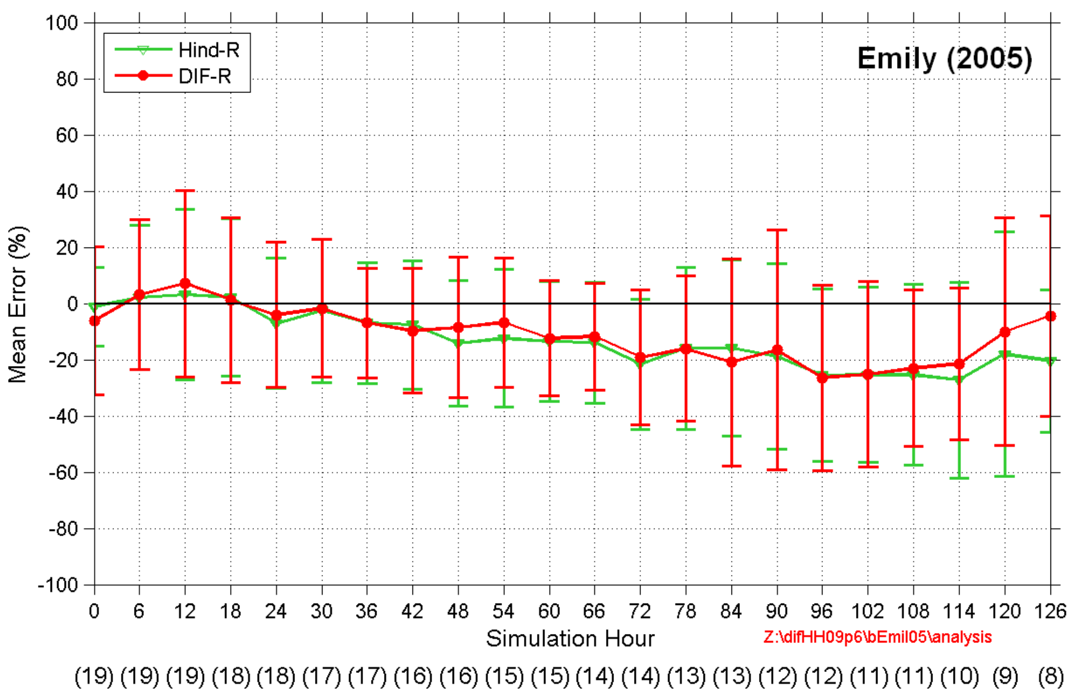


**Figure 3-11 Intensity error comparison for hurricane Emily**

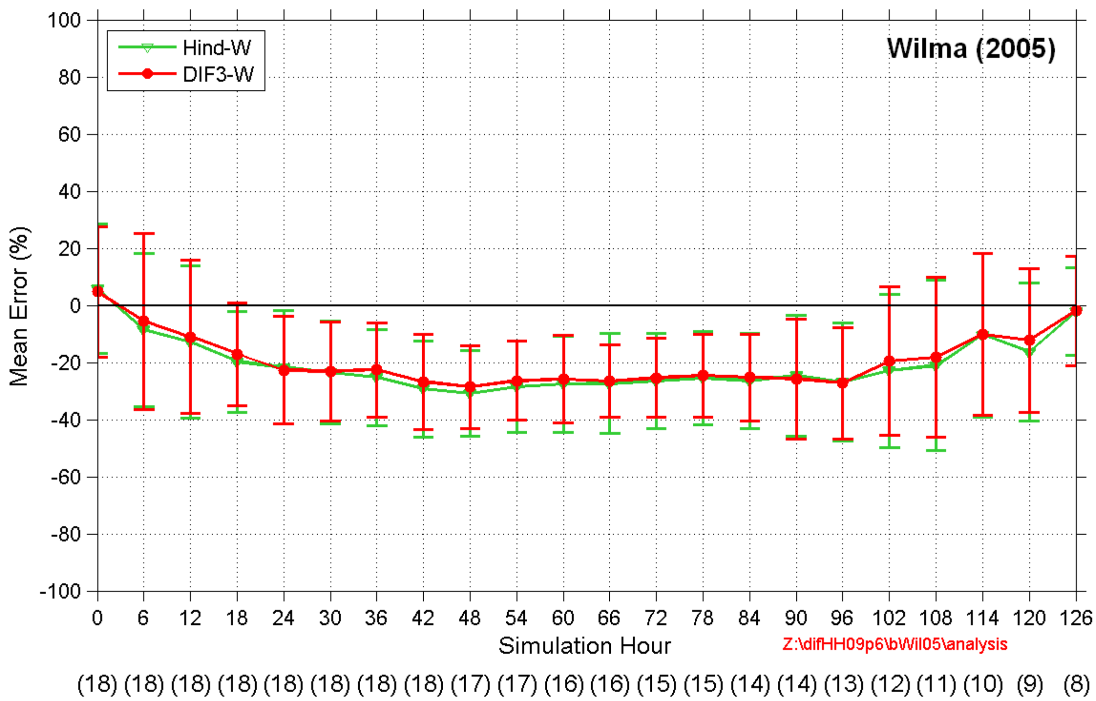


**Figure 3-12 Intensity error comparison for hurricane Wilma**

The prediction errors generally demonstrate steady, although not large, reduction; for all storms, the standard deviation of the intensity errors seems to be the same with or without the use of the synthetic profiles. The error reduction generally becomes most noticeable after 36-40 hours of prediction length. For hurricane Rita (Figure 3-10), the error reduction almost reached 10% between 72 and 102 hours. For Emily (Figure 3-11), the error reduction is about 5% in the 48 to 60 hours range and around 84 hours, and reaches almost 20% at 128 hours. For Wilma (Figure 3-12), the error reduction varies from 2% in the 6 to 60 hours range to 3.5% at 128 hours.

The table in **Error! Reference source not found.** shows the summary of the HI project performance assessment. Although the experiments demonstrated a noticeable (but sometimes just marginal) reduction in prediction error, these results are not sufficient for changing the data feeds of the operational model. Typically, such experiments would need to be repeated for an entire suite of past hurricanes (~ 300-400 storms) before a new hurricane forecast system is accepted for implementation for operational use at NCEP. The entire review process would be performed by an external committee consisting of domain experts.

| # | Metric | Baseline Performance[1] | Enhanced Performance | DIF impact |
|---|---|---|---|---|
| 1 | Error reduction in hurricane intensity forecast (% reduction) | Emily: 0 to 25%<br>Rita: 0 to 40%<br>Wilma: 0 to 30% | Emily: 5 to 20%<br>Rita: up to 10%<br>Wilma: 2 to 3.5% | Generally positive, although sometimes marginal |
| 2 | Better understanding of how to improve the ocean observing system for tropical cyclone prediction studies (qualitative assessment) | N/A – metric is a qualitative assessment of value of new information. | Insufficient data for evaluation[2] | Insufficient data for evaluation[2] |

[1] Absolute value of mean error

[2] Results from only three storms are not adequate to make a conclusion in this regard, but preliminary results indicate additional data helps in moving the water masses near the thermocline depths (> 50 m), towards the AXBT profiles, with less impact near the surface.

**Table 3-11 HI Performance**

#### *3.4.5.2.* Additional Results and Benefits

In addition to the anticipated benefits which were the formal metrics for the project, a number of additional benefits were identified during the course of the project execution:

- Modeling: The positive impact of the use of "synthetic profiles" on the upper ocean state is very promising for future use in both research and operations. Further refinement of the process for producing these profiles and improved assimilation algorithms are expected to lead to significant hurricane forecast error reduction
- Procedural: This project successfully integrated of DIF data into NCEP's operational data "tanks". This is a major technical achievement for establishing protocols/procedures to accept additional IOOS data into NCEP operational data streams. The experimental Temperature and Salinity profile data are now available to all NCEP models - NCEP's coupled hurricane forecast model, the global ocean model, the Global Forecast System, and a lower resolution ocean model
- Collaboration: The project was a successful collaboration between IOOS, NWS and OAR, highlighting the value and relative ease of such collective efforts.
- Tools: In order to ingest the profiles into NCEP's models, they needed to be converted from netCDF (the DIF-adopted format for gridded data) to BUFR (used by NCEP models). A general purpose netCDF-to-BUFR translator is available (http://ioos.gov/dif, Software) for any other users that might need it.

## 3.5. Integrated Ecosystem Assessments (IEAs)

### *3.5.1.* Decision Tool/Model to Be Enhanced

This project focuses on enhancing tools to aggregate and transform various data to and from DIF standards and by achieving improved integration of selected data sets to serve NOAA's Integrated Ecosystem Assessment (IEA) program. An IEA is a formal synthesis and quantitative analysis of existing information on relevant natural and socio-economic factors in relation to specified ecosystem management objectives.

This project enhances the ERDDAP tool. ERDDAP is the tool designed by Environmental Research Division (ERD) of the National Marine Fisheries Service's Southwest Fisheries Science Center that can read from a variety of the most common data transport standards, and can output the data in a wide variety of formats used by a number of analysis and visualization applications. It acts as a data collector and translator; providing unified access to data of interest to IEAs in standardized formats. It serves as an underpinning element to the IEA data management infrastructure.

### *3.5.2.* Scope, Architecture, and Data

The project goal is to augment the ERDDAP software to provide enhanced integration with selected IOOS DIF data services and, in collaboration with the Ecosystem Goal Team, to prototype the implementation of these services into the IEA model for the Gulf of Mexico and California Current Regions. The proposed enhanced version of ERDDAP will allow it to use IOOS DIF data formats as both a source and an output, and to integrate these capabilities with existing tools developed to support IEAs.

NOAA's National Coastal Data Development Center (NCDDC) is developing Regional Ecosystem Data Management (REDM) system as the underlying data management and discovery system for IEA efforts under the Ecosystem Observation Program. The REDM architecture contains several services to collect, transform and provide access to data, and ERDDAP will become a component within the REDM architecture.

The project better integrates data using the IOOS DIF standards into ERDDAP as well as better integrating with the REDM architecture. More specifically:

1. ERDDAP was updated to remain operational as a client with OOSTethys and DIF SOS services.
2. ERDDAP was updated to serve as an OGC Web Mapping Service (WMS) service, for greater interoperability with the GIS world and so that the graphical output can be accessed by the WMS clients.
3. ERDDAP was updated to output data in NetCDF4/HDF5 format following the "Common Data Model" being developed by Unidata.
4. ERDDAP was updated to translate data from other services into files consistent with DIF XML schemas in order to also become an IOOS DIF SOS server.
5. ERDDAP was integrated into the REDM architecture to act as "middle-ware" between the ecosystem data and the REDM catalog. This allowed users to transparently integrate data across multiple notes and in time/space.
6. Tools were developed for a user customizable dynamic IEA in order to
    a. use ERDDAP for the data delivery mechanism, similar to what can be done with iGoogle;
    b. integrate ERDDAP with the subscription-service based architecture underlying the REDM system.

The data set that the project operates with is pre-defined by the IEA requirements and DIF boundaries. IEAs operate with all 7 core DIF variables, which are served by 3 IOOS data providers – NDBC, CO-OPS and Coast Watch – in standardized DIF formats. In addition to those data providers, the data can be served by a number of non-NOAA providers such as WMO GTS, USGS, EPA STORET, etc. All data from different sources should be converted by ERDDAP into formats appropriate to IEAs.

The project data flow diagram is presented in Figure 3-13.

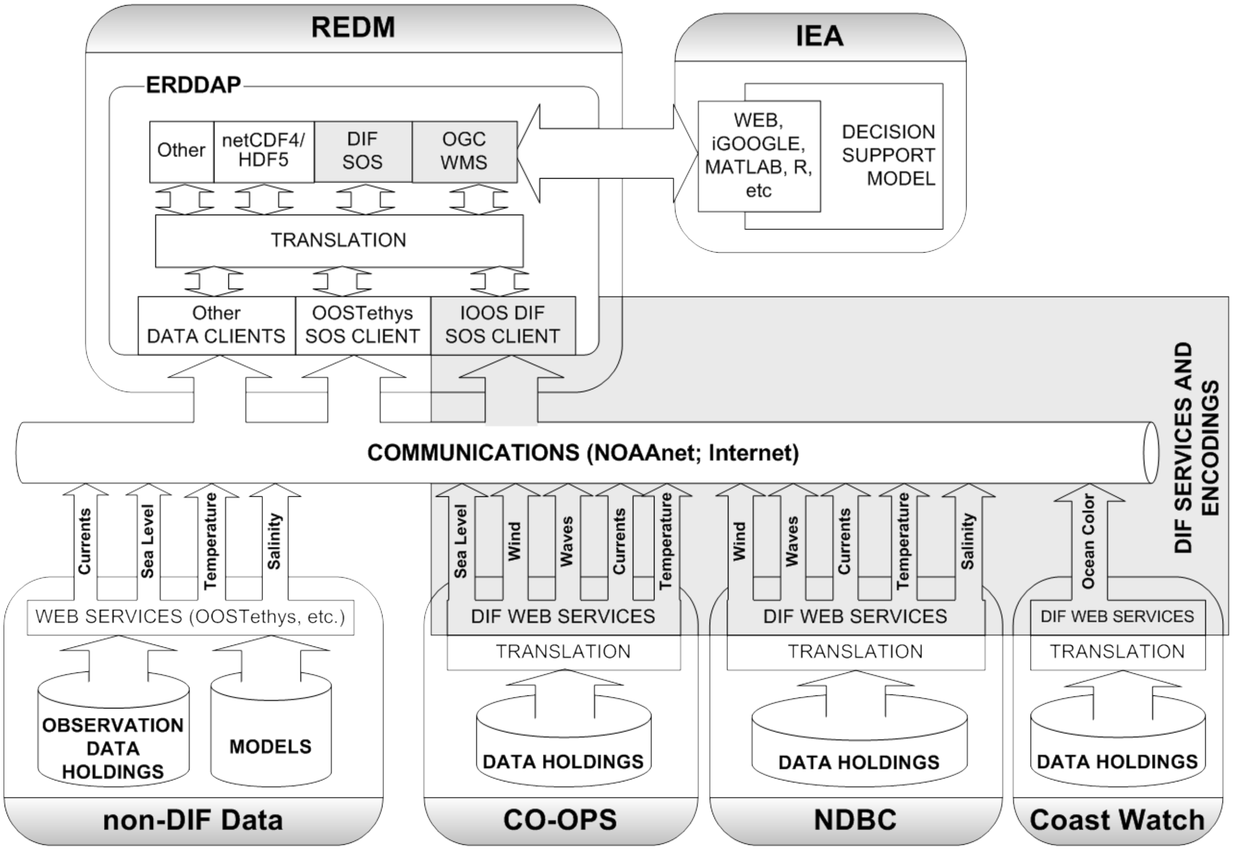


**Figure 3-13 IEA Data Flow Diagram**

### *3.5.3.* Anticipated Benefits

It is expected that this enhancement of ERDDAP will expand the seamless access to the data being served by NOAA providers as well as any provider that serves data via any Web SOS service. That will benefit IEAs, as more data will be available to use in models, calculations and assessments.

The use of ERDDAP data by IEAs will likely result in some social and/or economic gain. However, because the NOAA IEA program is a nascent program, without a specific IEA application yet making use of the data and services, no assessment is possible.

However, ERDDAP development and integration into the REDM architecture will improve the capacity of IEA by adding a number of new capabilities. These new capabilities are direct short-term benefits of the ERDDAP enhancement and implementation.

The list of these short term benefits is shown in Table 3-12.

| # | Benefit | Category[1] |
|---|---|---|
| **1** | Increase in the number of data providers serving DIF data | A |
| **2** | Expansion of the DIF data served by ERDDAP | A |

[1]See Table 3-1

**Table 3-12 IEA Benefits**

### *3.5.4.* Performance Assessment Methodology

#### *3.5.4.1.* Assessment Methods and Benchmarks

The assessment for this project consists primarily of quantifying the level of increase/expansion of the amount of data available to IEAs through ERDDAP. It would seem to be relatively easy to baseline and subsequently quantify the increase in data providers and datasets available after the project compared with before. However, the quantification is problematic because precise dataset and data source definitions can vary based on how and where the data is accessed. It is possible to miscount if, for example, a data center such as NDBC serves some regional data that is now also available in DIF format through ERDDAP. That would show as an increase in the overall number of datasets available in DIF formats, when in fact the one dataset is available from two separate services.

To simplify the problem, project team members were asked to help baseline the metrics, and subsequently to quantify the enhanced performance at the end of the project.

#### *3.5.4.2.* Metrics

The metrics for this project are contained in Table 3-13.

| Benefit | Metric | Measurement Unit |
|---|---|---|
| **1** | Increase in number of unique data sources (providers) serving DIF core variable data in conformity with one or more of DIF standards | Degree of Increase |
| **2** | Increase in number of datasets available through DIF services in forms that can be utilized in IEA development | Degree of Increase |

**Table 3-13 IEA Metrics**

## *3.5.5.* Performance Assessment

### *3.5.5.1.* Results and Anticipated Benefits

Table 3-14 summarizes the results of the project with respect to anticipated benefits and metrics.

| # | Metric | Baseline Performance | Enhanced Performance | DIF IMPACT | |
|---|---|---|---|---|---|
| **1** | **Increase in number of unique data sources (providers) serving DIF core variable data in conformity with one or more of DIF standards**<br>**(degree of increase)** | 10[1] | 23 | ✓ | The number of DIF data sources for core variables more than doubled.[2] |
| **2** | **Increase in number of datasets available through DIF services in forms that can be utilized in IEA development**<br>**(degree of increase)** | NOAA RAs<br>Salinity 2 6<br>Ocean Color 1 5<br>Temperature 2 5<br>Currents 2 5<br>Sea Level 2 4<br>Wind 2 5<br>Waves 1 4<br>TOTAL = 46 | Salinity 10<br>Ocean Color 6<br>Temperature 21<br>Currents 7<br>Sea Level 6<br>Wind 12<br>Waves 2<br>TOTAL = 64[3] | ✓ | Significant increase in the number of datasets available through DIF services for IEAs.[2] |

[1] NOAA and RA providers with one or more of 7 DIF core variables available

[2] Precise dataset and data source definitions can vary based on how and where the data is accessed. It is possible to miscount if, for example, a data center such as NDBC serves some regional data that is now also available in DIF format through ERDDAP. That would show as an increase in the overall number of datasets available in DIF formats, when in fact the one dataset is available from two separate services.

[3] Total number of datasets available via ERDDAP is 307; that number includes all ERDDAP providers serving both core and non-core variables; however, some datasets may not be available in DIF format or via DIF services.

**Table 3-14 IEA Results**

### *3.5.5.2.* Additional Results and Benefits

In addition to the anticipated benefits which were the formal metrics for the project, a number of additional benefits were identified during the course of the project execution:

- Overall, the project expanded the amount of data available via SOS services. ERDDAP can now:
  - Translate non-DIF in-situ data into IOOS GML format
  - Provide that translated data to users via SOS web services
- The IEA activity has laid the infrastructure to:
  - Provide easy access to a broad assortment of environmental biological data
  - Make data easily available to web page developers and to analysts

- The project completed the initial work of providing IEAs with expanded access to a broad suite of data, particularly for the California Current system. The same easy access to this broad suite of data can also be used for marine spatial planning or possibly HAB forecasters.
- The project helped establish an infrastructure that can meet general NOAA needs rather than tailoring it to a particular need.

## 3.6. Summary Assessment of Customer Implementation Projects

Table 3-15 summarizes the results of the customer implementation project metrics, categorized according to the impact of each project's metrics.

| Item | CI | HI | HABs | IEA | TOTAL | % Total |
|---|---|---|---|---|---|---|
| **# of Metrics** | **5** | **2** | **7** | **2** | **16** | **100%** |
| Positive Impact | 1 | 1 | 4 | 2 | **8** | **50%** |
| Negative Impact | 0 | 0 | 0 | 0 | **0** | **0%** |
| Insufficient data for evaluation | 4 | 1 | 3 | 0 | **8** | **50%** |

**Table 3-15 Summary Assessment of Customer Implementation Projects**

# 4.Technical Assessment

In order to execute and evaluate the DIF customer projects, technical infrastructure had to be established to make the DIF core variables available to the data customers. The majority of this infrastructure involved implementation of web services and common data formats at the data providers. For the data to be interoperable, standards and technical specifications has to be selected and adapted under the guidance of the DIF project. Numerous technical decisions had to be made along the way.

This section provides an assessment of the technical decisions that were made during the course of the DIF project. Initially, it was not intended that the DIF would be evaluated according to the technical decisions that were made, but it was determined that such an evaluation would be useful to a) determine if the infrastructure created in the DIF project could be evolved into an initial IOOS DMAC capability, and b) document technical recommendations for IOOS DMAC based on the DIF project experience.

Detailed results of the technical evaluation, along with recommendations for IOOS DMAC, are contained in this section. There are more than 50 individual recommendations contained in Table 4-2, varying from the very specific – such as adopting 'ioos' rather than 'x-noaa' as the namespace – to the more general – such as continuing to use Open Geospatial Consortium (OGC) Web Service metadata.

Subsequent evaluation of the technical decisions made during the DIF execution has yielded clear and concrete recommendations for how to evolve the DIF infrastructure into an initial IOOS DMAC infrastructure.

In addition to the technical recommendations enumerated in this section, key technical achievements that persist beyond the DIF project (in addition to those described in other sections of this document) include:

- Establishment of a core set of initial standards and conventions to enable data interoperability (documented in *Guidance for Implementation of the Integrated Ocean Observing System (IOOS®) Data Management and Communications (DMAC) Subsystem* http://ioos.gov/library/dmac_implementation_2010.pdf)
- Identification of an initial approach to data provider compliance testing
- Significant momentum for moving forward with IOOS DMAC planning and implementation

## 4.1. Assessment Methodology

Evaluation of technical decisions was performed primarily through soliciting input from key stakeholders including IOOS Operations Division personnel, DIF IPT members and members of the customer implementation project teams. Much of the input was obtained informally, through discussions with stakeholders and their review of draft assessments and recommendations. One particular technical decision, the selected SOS encoding format, spurred more controversy than others and therefore a more formal evaluation was conducted to evaluate that decision. An online survey was conducted in January of 2010. The results of the survey are summarized in Section 4.3.3; a more detailed analysis of the results is contained in Appendix 1.

## 4.2. Systems Engineering

It was determined early on that a systems engineering process would be employed to design and develop the DIF. Functional Requirements and Concept of Operations documents were developed with input from the IPT and interviews with members of the communities from the four customer project areas (HI, CI, IEA and HABs). The original DMAC Plan from 2005 was also consulted. A high level design was also documented.

While the documents were useful to gain consensus on the overall requirements and concept of operations of the DIF, they were developed before the specific customer implementation projects were identified and defined. Once the customer implementation projects were defined, the specifics of those projects and the associated statements of work tended to drive the implementation and operations, not the project-wide systems engineering documents. The systems engineering documentation was not heavily used or referred to as the customer projects were developed.

Further, there was little support in the IPT for systems engineering documentation and structure. Many advocated rapid-prototyping or agile-development without documentation, arguing that it allowed the project to be more responsive to user needs and requirements.

Because of the limited scope and somewhat independent nature of the four customer projects, there was no real adverse impact of not using the systems engineering documents to guide implementation efforts. In a larger DMAC context, however, this would not be the case. Due to the large number of diverse users and contributors to IOOS DMAC, implementation will require planning, documentation and management using a systems-of-systems approach to ensure that services that are built are interoperable, changes to services are properly managed and communicated so as not to impact users, and that life cycle support is in place to ensure long-term viability of services.

With respect to systems engineering and documentation, it is recommended that IOOS:

- Leverage the existing DIF systems engineering documentation as inputs to the requirements, concept of operations, and design for DMAC
- Adopt and promote a spiral-development system engineering approach (sometimes referred to as "iterative waterfall" systems engineering), which balances the necessary rigor of the process and documentation with the need to build, test and deploy services as quickly as possible

## 4.3. Design/Architecture Decisions

The sections below describe specific design decisions that were made and provide an assessment of those decisions with respect to contribution to DIF success and lessons-learned for DMAC.

### *4.3.1.* Data Content Standard

The initial work (2007) on standardization for the DIF focused on development of a Data Content Standard (DCS). This document specified the units and formatting for general quantities and attributes such as latitude, longitude, time and station ID, and for observation-specific quantities and metadata. Though a useful first step, the DCS was not sufficient because it addressed only some of the incompatibilities among data providers, and it left data providers free to adopt data access services and data formats of their choice. For example, CO-OPS implemented a SOAP web service with a particular XML encoding, and NDBC implemented a CGI web service with a different XML encoding, yet each complied with the DCS as written. Also, the DCS is a combination of abstract data content ("these fields are required") and concrete encoding conventions that are not necessarily applicable to all formats.

### *4.3.2.* Web Services

In 2008, IOOS DIF adopted particular web services and encoding conventions for various types of geospatial information, as described in the Table 4-1. The selection of web services was made following consultation with subject matter experts within NOAA and the IOOS Regional Associations (RAs) and research regarding optimal methods for representing data types of interest to the IOOS community.

| Data Type | Web Service | Encoding |
|---|---|---|
| Discrete features (e.g., *in-situ* point or profile data, time series, trajectories) | Sensor Observation Service (SOS) | XML - Geography Markup Language (GML) based on Observations & Measurements (O&M) for data; Sensor Model Language (SensorML) for metadata |
| Coverages (e.g., gridded model outputs, Level 3 satellite data, HFR surface currents) | Data Access Protocol (DAP) or Web Coverage Service (WCS) | Binary - Network Common Data Format with Climate & Forecast Conventions (CF/NetCDF) |
| Geo-referenced images of data (features or coverages) | Web Map Service (WMS) | Common image formats (PNG, TIFF, GIF, JPEG) |

**Table 4-1 Web services and data encodings used in the IOOS Data Integration Framework**

#### *4.3.2.1.* OGC Services

SOS, GML, WCS and WMS are standards developed by the Open Geospatial Consortium (OGC). OGC is a standards-developing organization focusing on geo-referenced information and services. Most IOOS federal partners, and over 200 government, commercial and academic entities world-wide are members of OGC. The Global Earth Observing System of Systems (GEOSS), of which IOOS is formally a part, includes many OGC standards in its Standards and Interoperability Registry (http://seabass.ieee.org/groups/geoss/). The European Union mandates the use of OGC standards where applicable through its Infrastructure for Spatial Information in Europe (INSPIRE) directives. OGC standards can therefore enable compatibility with related projects and with the GIS community. For all of these reasons, they were selected for adoption for the DIF project.

There are relevant OGC standards that IOOS has not yet adopted (e.g., Catalog Service for Web [CS/W], Web Processing Service [WPS, Sensor Alert Service [SAS]), primarily because they were not required to meet the limited scope of the DIF project.

OGC standards are often fairly general (i.e., they are applicable to road maps, location-based services, demographic data and environmental observations); additional best practices and conventions are therefore important to ensure interoperability within a particular information community. The OGC-related tools and practices are not always as well developed as the DAP/CF/NetCDF realm.

Relative to the use of OGC standards, it is recommended that IOOS:

- continue to expand the use of OGC Web Services because of broad use by geographic information communities at regional, national and international levels.
- Support the development of best practices, conventions and profiles of OGC standards for oceanographic data.
- Support the development of robust client and server tools that comply with OGC specifications.
- Coordinate between IOOS and other OGC-using organizations within NOAA.

#### *4.3.2.2.* Sensor Observation Service (SOS)

SOS is a service designed to support access to observational data from both *in situ* and remote sensors. SOS is part of the OGC Sensor Web Enablement (SWE) technology area. "Sensors" can be broadly defined to mean any kind of observation procedure, including individual sensors, networks of sensors, sample gathering and analysis, and probing of model outputs. The principal SOS functions enable users to obtain metadata about the service and the observations it offers (GetCapabilities), obtain metadata about a sensor or observation procedure (DescribeSensor), and obtain the result of an observation (GetObservation). Users can issue requests for specific sensors, time ranges, geographic bounding boxes, and observed properties.

SOS was a relatively new standard at the time of adoption by IOOS DIF. Development of the "Sensor Web Enablement" suite of standards began in 2004 or earlier, and SOS version 1.0 was issued in October 2007. Prior to selecting SOS, there was discussion within IOOS about adopting OGC Web Feature Service (WFS) for discrete feature data. WFS is an older standard than SOS and thus has more implementing applications. However, WFS was not developed with observational data in mind and requires considerable specialization. SOS was therefore adopted by IOOS in preference to WFS.

Despite that SOS is focused on observations, some level of specialization was nevertheless required for IOOS purposes. This effort is still in progress, and is also evolving as IOOS continues to assess the success of its implementations to date. Specialized conventions were needed for such elements as: data output format(s); mandatory and optional sensor metadata fields; and identifiers for sensors, stations, phenomenon names and features of interest. Subsequent sections of this document provide discussions of those conventions.

Relative to the use of SOS, it is recommended that IOOS:

- Continue use of SOS for in situ data from sensors and sensor networks.

- Expand use of SOS to trajectory data (VOS, ARGO), biological data, and discrete samples (e.g., water quality).
- Explore additional output formats (including CF/NetCDF) and the use of SOS as a wrapper to a DAP service providing out-of-band data.
- Make available high-performance SOS implementation(s), perhaps as a component contributed to the THREDDS Data Service (TDS) project.
- Document IOOS-specific best practices and conventions for SOS and submit for consideration as a formal OGC Profile of SOS.

#### *4.3.2.3.* Data Access Protocol (DAP)

DAP is in broad use by members of the ocean science community and, along with the CF/NetCDF encoding format, it was already a proposed DMAC standard at the start of the DIF project. DAP and CF/NetCDF has well-developed tools and practices - especially for gridded data - and can handle more complex grids than WCS. Further, there are draft conventions not yet in wide use that allows application of CF/NetCDF encoding to feature data. It is important to note that DAP/CF/NetCDF is not broadly used outside the scientific community. CF/NetCDF has been submitted to OGC for possible adoption as an OGC standard.

Relative to the use of DAP/CF/NetCDF, it is recommended that IOOS:

- Continue use of DAP/CF/NetCDF because of existing use at NOAA
- Support CF/NetCDF adoption by OGC
- Support extension of CF Conventions to feature data
- Explore using CF/NetCDF responses with SOS request protocol (i.e., using CF/NetCDF as binary output format for SOS GetObservation or GetResult operation).

#### *4.3.2.4.* Web Map Service (WMS)

WMS is a mature international standard (ISO 19128) in wide use. Because the primary output of a WMS is a common image format such as PNG, TIFF or GIF, there is no need for client software to decode and process data before presentation to the user. IOOS DIF focused on data access rather than images, and therefore did not make any direct use of WMS, although it was adopted as the web service for geo-referenced images of data. Nevertheless, visualizations are sufficient to solve many of the ocean information needs of prospective users, so WMS could be adopted and implemented more broadly within IOOS. IOOS participated in technical discussions regarding WMS conventions for access to model forecast data and multi-dimensional datasets. IOOS had planned to adopt or develop conventions for symbology and color-coding of data during the DIF, but that did not occur due to resource limitations and lack of immediate DIF customer need.

Relative to the use of WMS, it is recommended that IOOS:

- Continue to expand the use of WMS, such that all data are available both numerically and visually. This could be done either by establishing WMS at each data provider site, or by establishing a standalone WMS capability as a shared component.
- Pursue symbology and color-coding conventions.

### *4.3.3.* In-situ Data Encoding

The SOS specification mandates the use of OGC Observations & Measurements (O&M) schema for providing data in response to an SOS GetObservation request. However, O&M is sufficiently general that a large number of different O&M-compliant encodings are theoretically possible. The DIF project initially attempted to develop encodings based on the Climate Science Modeling Language (CSML), which has enough semantic richness to explicitly encode data from "sampling features" such as Points, Profiles and Trajectories. The encodings were to be compliant with CSML, O&M and OGC Geography Markup Language (GML). In practice, the version of CSML then available was suitable for a homogenous collection of sensors (i.e., a collection of sensors that all report at the same times and have the same vertical profile) but that modifications to CSML would be required to support integration of data from a collection of disparate sensors into a single data response. In addition, some interdependent schemas in the GML/SWE/SensorML family had mutual incompatibilities that needed to be fixed with a private copy of the schema. Due to scheduling constraints and a lack of detailed technical expertise on CSML in the IOOS office, IOOS decided (in May 2008) to defer adoption of CSML for DIF and instead use an IOOS-specific GML application schema that was compliant with O&M (IOOS GML v0.6.1). IOOS collaborators at NOAA NGDC continued experimentation with CSML as part of the Observing System Monitoring Center (OSMC) project.

IOOS GML v0.6.1 was deliberately given a version number less than 1.0 to reflect the fact that it was a draft rather than a complete expression of the scope of IOOS intent and data. The present scope encompasses 6 of the 7 DIF variables (all but gridded ocean color) for time-series data at point, profile and multi-station collections; it does not extend to trajectory data. The IOOS GML v0.6.1 schema, and the associated phenomenon dictionary and data record definitions, are available at http://ioos.gov/dif/schema.html .

The IOOS GML schema is highly structured, hierarchical and semantically rich. Data values are in individual elements, tagged with units of measure, and associated with individual stations, sensors, times and profile bins. A design goal was to make use of this structure such that IOOS GML could serve as a "master format" from which any other simplified format could be derived using industry-standard XML tools such as XSLT (Extensible Stylesheet Language Transformations). In practice, it was found that the desired XSLT templates were difficult to construct, that invoking them as part of the data-request process was somewhat cumbersome, and that applying them to data files required ingesting the entire file into memory first (rather than treating the file in serial fashion, one station at a time). Also, generating the GML, or parsing the output without XSLT, was cumbersome and slow for large data requests. Data providers found it difficult and verbose to generate, and data customers found it difficult to parse.

Validation is technically possible at the XML Schema level, but in practice it requires loading huge schema trees into the validator and provides only a very coarse level of checking (Schematron, or ingestion into a data analysis application, is needed for real syntactic verification). The Data Record Definitions required by O&M were not useful for actual verification. In practice, real data quality control requires checks that are beyond what schema and Schematron can handle.

These issues were revealed in discussions with collaborators and documented more formally in a survey (Dec 2009, see 5.1).

With regard to the IOOS GML data encoding, it is recommended that IOOS:

- Offer additional encoding format(s) besides IOOS GML
- Retain the GML option for servers that already have it, at least until any existing users have switched to other formats.
- Not extend IOOS GML to broaden its scope to other data types not already handled, unless there is explicit demand from a particular customer(s).
- Consider whether CSML has evolved sufficiently that it could be adopted without further schema development as a GML-compliant schema.

The DIF did not adopt an alternative SOS encoding known as SWE Common (SWEC). SWEC is also compliant with O&M but is not a true GML application schema. SWEC is a mixture of rows of plain-text comma-separated values (CSV) wrapped in XML elements. Unlike CSML, SWEC does not have formal semantics for sampling features (points, profiles, trajectories). SWEC is used by some, but not all, OGC SOS implementations. The failure to adopt SWEC caused disagreement with its proponents. Conversely, adopting SWEC would have caused disagreement with the CSML proponents. The resulting middle ground (IOOS GML) was not satisfactory to any of the parties. In short, the question of SOS encodings was the most contentious issue in the DIF project.

Other O&M encodings for SOS include Weather Information Exchange Schema (WXXS), Groundwater Markup Language (GWML), and Water Markup Language (WaterML) v2.0 now under development. These were learned of after the DIF project was underway.

With regard to SOS encodings, it is recommended that IOOS:

- Offer several encodings to satisfy different user communities (not all SOS servers would need to offer all possible encodings, but a translation capability should be established to handle conversions), candidates include:
  - pure CSV for direct ingest into spreadsheet applications
  - SWE Common for existing SWEC clients
  - KML for direct ingest into popular virtual earth applications
  - CF/NetCDF, perhaps via an out-of-band SOS/DAP mechanism, to combine DAP binary data access with SOS metadata access
  - Shapefile for existing GIS clients.
- Continue the IOOS collaboration with NOAA/NMFS ERDDAP and NFS OOI-CI to explore hosting of format conversion services on cloud-computing resources.

### *4.3.4.* Service Registry/Data Catalog/Data Viewer

Three important components in IOOS DMAC are:

- Service Registry – to allow customers to discover a data providers' data access services and summary metadata
- Data Catalog - to allow customers to discover detailed spatial and temporal extents of available data
- Data Viewer – to allow customers a mechanism to view the actual data, without use of their own client software.

The DIF was conceived as a project involving three (3) data providers and four (4) customers, all of whom had knowledge of each other. As such, it was not necessary to establish a formal Service Registry of data access services or a Data Catalog of available data. Further, each of the four customer projects already had its own domain-specific application for viewing data, so it was not necessary to establish a generic Data Viewer capability. The DIF did not attempt to establish a Service Registry, Data Catalog or Data Viewer so no assessment of these components is necessary.

In expanding the DIF prototype towards more complete IOOS DMAC capability, however, it will be necessary to have a Service Registry, Data Catalog and a web-based Data Viewer that serves as a general-purpose user interface to the Catalog. Such an effort is now underway. The benefits are expected to be numerous, including: the ability for new and existing users to find data from a single entry point; the capability for the Program Office to demonstrate IOOS status and progress; the establishment of system monitoring tools; the discovery and remediation of insufficient metadata or of incompatibilities between servers; and the establishment of an embodiment of IOOS into which prospective data providers can register their offerings and receive acknowledgement thereof.

Relative to Service Registry, Data Catalog, and Data Viewer, it is recommended that IOOS:

- Establish a prototype of these capabilities in FY2010 to increase the in-house familiarization/expertise with the technical and operational aspects
- Augment the prototype capabilities and ensure operationally reliable hosting in FY2011 and beyond
- Enable GEOSS and Geodata.gov to harvest IOOS Registry and Catalog, and thereby enable IOOS data providers to meet the requirement to register servers at GEOSS and Geodata.gov simply by registering with IOOS.
- Explore methods to enable commercial search engines to harvest IOOS Data Catalog such that users can find data simply by searching the web.

### *4.3.5.* Metadata

Several classes of metadata were of possible relevance to the DIF effort; not all were addressed and/or implemented. The classes and the work done are described below.

- **Service-level metadata:** the "table of contents" describing data server holdings. Such metadata is provided by all OGC Web Services--WMS provides a list of the map layers on any given server and their spatial and temporal coverage; SOS has a list of "observation offerings" (e.g., buoys or networks thereof); WCS has a list of gridded coverages. In DIF these were populated as required by the relevant OGC specification. The TDS implementation of DAP/WCS has a limitation in this regard, however: there is no single table of contents for a given TDS. Instead, there is a directory tree-like which can be navigated until you reach an actual dataset, and the only way to build up a table of contents is to traverse the tree. Service-level metadata in DIF is not currently expressed in ISO 19119/19139 XML.

Relative to Service-level metadata, it is recommended that IOOS:

  - Continue using OGC Web Service metadata.
  - Partner with NOAA's Global Earth Observation-Integrated Data Environment(GEO-IDE) Unified Access Framework project being implemented by the DMIT that is developing a crawler to generate summary metadata from a THREDDS Data Service (TDS) directory tree.
  - Use the future Service Registry to harvest and verify service metadata, and to make service metadata available in ISO 19139 format using the OGC Catalog Service for Web (CS/W).

- **Dataset-level metadata:** the overall description of a dataset (including spatial and temporal coverage, data quality, access points, etc). Dataset metadata are usually formatted as US Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) or ISO 19115/19139, which are related standards that are slated to merge. The DIF project did not develop additional dataset metadata, largely because the datasets involved were well-known to the DIF customers. However, evolution towards IOOS DMAC capability for finding and archiving all data will require investment in basic metadata development for the purposes of the Data Catalog and more complete metadata including QA/QC information and other details thereafter.

  Relative to dataset-level metadata, it is recommended that IOOS:

  - Promote ISO 19115/19139 metadata development by offering complete examples of good metadata, suggesting tools to help write metadata (e.g., MEST, MERMAID), requiring IOOS-funded data providers to provide good metadata, and rating all IOOS data providers (even volunteer ones) according to the quality of their metadata (as part of a suite of Capability Maturity Level rankings).

- **Sensor and station metadata:** detailed descriptions of a particular buoy, station, observing platform or sensor. Such metadata can be made available using the SOS DescribeSensor operation in OGC Sensor Model Language (SensorML). As part of the DIF, CO-OPS developed detailed SensorML descriptions of their stations. NDBC developed rudimentary descriptions, and will enhance them in FY2010 to finalize the DIF work. An overall metadata model showing linked sets of metadata records to describe types of platforms, types of sensors, and specific instances of each of those, was drafted during the DIF but not fully established. Work done to date did not cover moving platforms such AUVs, gliders or vessels.

  Relative to sensor and station metadata, it is recommended that IOOS:

  - Define minimum SensorML descriptions for IOOS sensors and platforms.
  - Ensure data providers make such metadata available, and display corresponding CM ranking.
  - Work with sensor manufacturers and Smart Sensor Consortium to promote the use of SensorML by vendors to describe each of their products.
  - Ensure linkages between metadata records to support overall metadata model.

- **Time-dependent metadata:** information associated with particular measurements that changes with time and must be included with the actual data. For example, CO-OPS current meters include heading, pitch, and roll information about the sensor along with the speed and direction of the current. Fields for these values were included in the IOOS GML schema for the DIF and data providers populated those slots in their GML response.

  Relative to time-dependent metadata, it is recommended that IOOS:

  - Continue providing time-dependent metadata about observations as appropriate based on customer need and metadata availability.

## *4.3.6.* Other standards and conventions

### *4.3.6.1.* Phenomenon identifiers

For the SOS work, an IOOS Phenomenon Dictionary was defined. This dictionary defined names (e.g., 'Winds') and formal URLs (e.g., 'http://www.csc.noaa.gov/ioos/schema/IOOS-DIF/IOOS/0.6.1/dictionaries/phenomenaDictionary.xml#Winds') for each measured variable, and also stated what scalar (i.e., single-valued) quantities make up each composite phenomenon (for example, Winds is a combination of wind speed, direction, guest and vertical speed). Use of a phenomenon dictionary is required by the OGC Sensor Web Enablement (SWE) specifications.

For the DAP/CF/NetCDF work, the CF Standard Names mandated by the CF conventions was used. CF provides names, but not formal URLs, for scalar quantities. CF does not provide names for composite quantities like winds. CF/NetCDF does not have a formal concept of composite quantities; instead, each CF/NetCDF file states in the header what arrays it contains. The Marine Metadata Interoperability (MMI) ontology registry associates URLs for each CF name.

The IOOS phenomenon names used in the DIF SOS were not the same as the CF standard names (even for scalar quantities), and the IOOS phenomenon URLs were not those defined by MMI (even for scalars). This lack of semantic interoperability between the SOS and DAP/CF work is not necessary. Therefore, the following recommendations have begun (2010) to be implemented.

Relative to phenomenon identifiers, it is recommended that IOOS:

- Adopt CF names, and corresponding MMI URLs, for scalar quantities.
- Revise the IOOS phenomenon dictionary to define composite phenomena in terms of CF scalars.
- Add names and URLs of composite phenomena to MMI ontology.
- Work with the CF Metadata group to define new CF names for scalars as IOOS expands to other variables[2].
- Work with MMI to add new names and URLs to ontology. Include GCMD keywords in metadata about services, datasets and observing systems to aid in catalog searching and classification.
- Work with GCMD to add additional keywords as necessary.

### *4.3.6.2.* Station and sensor identifiers

The IOOS SOS servers provide data from moored buoys and other stations. Each station has one or more sensors associated with it. The stations may have a WMO number assigned by NDBC, or an identifier assigned by CO-OPS, or another identifier. There may be ambiguities--for example, it is not possible to tell without further information whether '42001' is a WMO number or some other number. For IOOS, each station and sensor has been assigned a Uniform Resource Name (URN) that provides more semantic information than simply a number. The structure of these URNs is based upon the pattern defined in the OGC Best Practice paper "Definition identifier URNs in OGC namespace."

Example station IDs:

urn:x-noaa:def:station:noaa.nws.ndbc::23401

urn:x-noaa:def:station:NOAA.NOS.CO-OPS::cb0102

Example sensor IDs:

urn:x-noaa:def:sensor:noaa.nws.ndbc::23401:tsunameter0

urn:x-noaa:def:sensor:NOAA.NOS.CO-OPS::cb0102:Nortek-ADP-555

[2] NOTE: The following new CF names have been established at IOOS request: water_surface_height_above_reference_datum for tide gauge water levels, sea_floor_depth_below_sea_surface for tsunameter water levels, and sea_surface_wave_mean_period as a generic term for wave period (existing terms specified particular calculation methods).

The 'urn:' prefix is mandatory for URNs. 'x-noaa' is the namespace identifier; it begins with 'x-' because NOAA has not formally registered 'noaa' as a namespace with the Internet Assigned Numbers Authority (IANA). 'def' is used in the OGC pattern to signify a URN that defines something. 'station' or 'sensor' is used as appropriate. The next field is defined by the operator of the SOS server (NDBC or CO-OPS); this was not the best choice, as described below. The next field is null (two colons in a row) and is reserved for a version number in case IDs are reassigned. The next field provides the WMO number or other number of the station. Sensors have an additional field to identify the specific sensor.

Using the operator of the SOS in the URN has caused a problem when the station itself is owned or operated by a different group. NDBC retransmits some observations gathered by the IOOS RAs or from US Minerals and Management Service (MMS) offshore oil platforms. Each of the IOOS RA platforms in question has been assigned a WMO number by NDBC (which has authority to do so) for use when the data are sent out over the WMO GTS. Each RA and MMS station has also been assigned a URN as described here, with 'noaa.nws.ndbc' in the SOS operator field. That URN does not identify the station owner. The station owner information will be available in metadata (SOS DescribeSensor operation) as part of enhancements in 2010, but the semantics of the URN itself could be improved to make station ownership obvious.

It is important for IOOS partners to agree on identifier schemes in order to detect or avoid duplicate stations in a data record when data are gathered and retransmitted from multiple sources. For example, the IOOS Observation Registry map shows duplicate stations because it harvests station lists from IOOS RAs and from NDBC and CO-OPS. Funded partners could perhaps be forced to adopt identifiers chosen by IOOS, or IOOS and its partners could agree on a mutually acceptable URN pattern that allows partners to assign identifiers which are then used throughout the system.

Relative to station and sensor identifiers, it is recommended that IOOS:

- Revise the IOOS URN pattern for station and sensor identifiers. Include the owner of the station, the number assigned to the station by WMO or another authority, and the authority in question.[3]
- Adopt 'ioos' rather than 'x-noaa' as the namespace, and register the 'ioos' namespace with IANA.
- Establish a URN resolver in order to link station and sensor IDs to the actual web-accessible resource describing that station or sensor (e.g., inputting the station ID in the resolver takes you to the SensorML metadata for that station). A URN resolver is also an IANA requirement for namespace registration.
- Work with MMI or OGC to use or learn from their URN resolver.

## 4.4. Development and Testing

The sections below provide an assessment of the DIF development and testing processes, procedures and tools.

### *4.4.1.* Development

As discussed in Section 4.2, development of systems to support the four customer implementation projects was driven primarily by a series of Statements of Work (SOW) for the data providers and customer project teams. These SOWs described the tasks to be performed, the functionality to be provided in the developed systems, project deliverables, responsibilities and milestones schedules. Overall this approach to development proved effective and efficient – the teams were relatively small and met frequently to discuss progress and issues. IOOS Operations Division personnel led the coordination to ensure cost, schedule and technical specifications were met. The teams generally took an iterative approach to development – interim releases of software/capability were provided to allow for early integration and risk reduction. Subsequent releases provided bug fixes and additional functionality.

[3] NOTE: A possible scheme for revised IOOS station URNs is urn:ioos:station:OWNER:AUTHORITY:DATE:ID, where OWNER = station or platform owner; AUTHORITY=WMO or other authority, ID=number assigned by that authority, and DATE=date the number was assigned (to account for later reassignments, especially of WMO numbers). The "owner" could be the actual owner or a project name such as 'TAO' or 'ARGO'. This requires considerable discussion.

The approach used was extremely effective and could be a model for future development. In the case of the DIF, the project teams could operate somewhat independently from one another. This would not be the case in the larger DMAC – where all the systems would need to interoperate more seamlessly.

The initial DIF data providers were NDBC, CO-OPS and CoastWatch.

CoastWatch implemented DAP and WCS access to a collection of satellite ocean color (chlorophyll) observations. CoastWatch used the THREDDS Data Server (TDS) software from Unidata (an NSF-funded program). TDS offers DAP, WCS and WMS access services.

NDBC and CO-OPS implemented SOS. Each group already had in-house code to read from their internal database and emit data upon request, but that software did not support SOS. At the start of the DIF project, both groups investigated existing open-source implementations of SOS. Both concluded that it would be easier to modify their in-house code to support the SOS interface specification, and more difficult to modify the open-source code to support their internal database. Each group therefore implemented their own SOS software. This is likely due to the level of sophistication and in-house resources of these data providers. Future, less sophisticated data providers may require server-side reference implementations, or other support in order to implement the recommended services.

NDBC implemented an SOS interface in PHP, connecting to an RDBMS on a Linux server. This implementation is available on the SOS web site (http://sdf.ndbc.noaa.gov/sos/). The software has also been installed at CENCOOS, SCCOOS and WHOI. NDBC provided a Lessons Learned report at the end of FY2009. NDBC also implemented TDS to serve HFR data, although that data was not used by the DIF customers in their customer implementation projects.

CO-OPS implemented an SOS interface in JSP, connecting to an RDBMS on a Linux server. The SOS functionality is embedded directly within the existing CO-OPS software and is not available independently.

Relative to approaches for development, it is recommended that IOOS:

- Play an active role in monitoring each project throughout the development stage, to ensure adherence to technical, budget and schedule requirements.
- Maintain small, agile teams that can “build a little, test a little” to minimize risk and more quickly deploy services
- Provide a central point of coordination across all development teams to ensure interoperability

### *4.4.2.* Testing and Validation

Testing and validation in the DIF context can be described in three categories:

- Functional – verification that the delivered systems perform all the required functions
- Standards Compliance – verification that the delivered systems are compliant with applicable standards, schemas, etc.
- Performance – assessment of how well the delivered systems will perform under various loads; high data volume, network traffic, volume of user requests, etc.

Testing in these three areas varied based on the customer project. Functional testing was performed as a matter of course during integration and acceptance by the customers. Standards compliance was not explicitly tested, with the exception of some SOS service compliance testing (described in more detail below) that was conducted on DIF data providers and Regional Associations’ servers after completion of the DIF customer projects. In general, performance testing was not conducted, except in the case of the Coastal Inundation project where some informal user load tests were conducted to gauge robustness of the CO-OPS server in responding to numerous simultaneous requests during a hurricane event.

Because of the DIF’s role as a proof-of-concept, risk-reduction project, rigorous testing was not of prime concern. This will not be the case if IOOS data providers are expected to support operational systems with high availability and reliability requirements.

Following implementation of SOS services at NDBC, CO-OPS and a number of the IOOS Regional Associations, IOOS conducted SOS service compliance testing using an OGC test suite. The test results revealed that many of the implementations had some level of compliance issue, varying from trivial to serious. The issues with the test can be grouped into three categories based on ease of resolution:

a. Errors that can be easily addressed in the near-term (typos, incorrect references, etc.).

b. Errors that do not need to be corrected in the near-term (wrong or inconsistence exception code messages, etc.).

c. Errors due to the unnecessary strictness of the OGC test (truncated seconds in timeDate value, non-OGC URNs, HTTP instead of HTTPS in URLs, etc.)

IOOS is working with the data providers to highlight the compliance issues and get them corrected; NDBC is considered a primary target as the majority of the regions have NDBC SOS software installed. IOOS is also working with the OGC Compliance Test development group to update the test scripts. There is also an ongoing work to align SOS Web service testing to the IOOS Data Catalog development. Full OGC specification compliance is a reasonable goal for FY 2011.

For testing IOOS RA THREDDS Data Services, staff at MACOORA are running the NetCheck Web application (http://netcheck.mine.nu/resources). NetCheck is a program developed by CoastWatch, which periodically performs functional tests of network resources and sends emails to various subscribers whenever there is a change in the status of a test (when a test initially fails, or when the nature of the failure changes, or when a test that had been failing now passes). Currently, NetCheck can run three types of tests:

- HttpTest tests if the content retrieved from a URL includes various desired text strings and doesn't include various undesired text strings.
- SftpTest tests if a file can be uploaded to another computer, renamed, and downloaded via sftp.
- OpendapTest tests if information can be obtained from a gridded data file served by OPeNDAP.

Relative to testing and validation, it is recommended that IOOS:

- Define a methodology/approach for testing each of these three areas that can be communicated to IOOS participants
- Investigate available tests tools, harnesses or test services that can be leveraged. Determine if investment is required to develop test tools that can be used across data providers.
- Define a certification and validation process for IOOS data providers.

## 4.5. Summary Recommendations of Technical Assessment

Table 4-2 summarizes the recommendations from the technical assessments.

| # | Category | Recommendation |
|---|---|---|
| 1. | Systems Engineering | Leverage the existing DIF systems engineering documentation as inputs to the requirements, concept of operations, and design for DMAC |
| 2. | Systems Engineering | Adopt and promote a spiral-development system engineering approach (sometimes referred to as “iterative waterfall” systems engineering), which balances the necessary rigor of the process and documentation with the need to build, test and deploy services as quickly as possible |
| 3. | OGC Standards | Continue to expand the use of OGC Web Services because of broad use by geographic information communities at regional, national and international levels. |
| 4. | OGC Standards | Support the development of best practices, conventions and profiles of OGC standards for oceanographic data. |

| # | Category | Recommendation |
|---|---|---|
| 5. | OGC Standards | Support the development of robust client and server tools that comply with OGC specifications. |
| 6. | OGC Standards | Coordinate between IOOS and other OGC-using organizations within NOAA. |
| 7. | SOS | Continue use of SOS for in situ data from sensors and sensor networks. |
| 8. | SOS | Expand use of SOS to trajectory data (VOS, ARGO, AUVs, and gliders), biological data, and discrete samples (e.g., water quality). |
| 9. | SOS | Explore additional output formats (including CF/NetCDF) and the use of SOS as a wrapper to a DAP service providing out-of-band data. |
| 10. | SOS | Make available high-performance SOS implementation(s), perhaps as a component contributed to the THREDDS Data Service (TDS) project. |
| 11. | SOS | Document IOOS-specific best practices and conventions for SOS and submit for consideration as a formal OGC Profile of SOS. |
| 12. | DAP/CF/NetCDF | Continue use of DAP/CF/NetCDF because of existing use at NOAA |
| 13. | DAP/CF/NetCDF | Support CF/NetCDF adoption by OGC |
| 14. | DAP/CF/NetCDF | Support extension of CF Conventions to feature data |
| 15. | DAP/CF/NetCDF | Explore using CF/NetCDF responses with SOS request protocol (i.e., using CF/NetCDF as binary output format for SOS GetObservation or GetResult operation). |
| 16. | WMS | Continue to expand the use of WMS, such that all data are available both numerically and visually. This could be done either by establishing WMS at each data provider site, or by establishing a standalone WMS capability as a shared component. |
| 17. | WMS | Pursue symbology and color-coding conventions. |
| 18. | IOOS GML data encoding | Offer additional encoding format(s) besides IOOS GML |
| 19. | IOOS GML data encoding | Retain the GML option for servers that already have it, at least until any existing users have switched to other formats. |
| 20. | IOOS GML data encoding | Not extend IOOS GML to broaden its scope to other data types not already handled, unless there is explicit demand from a particular customer(s). |
| 21. | IOOS GML data encoding | Consider whether CSML has evolved sufficiently that it could be adopted without further schema development as a GML-compliant schema. |
| 22. | SOS encodings | Offer several encodings to satisfy different user communities (not all SOS servers would need to offer all possible encodings, but a translation capability should be established to handle conversions), candidates include:<br>• pure CSV for direct ingest into spreadsheet applications<br>• SWE Common for existing SWEC clients<br>• KML for direct ingest into popular virtual earth applications<br>• CF/NetCDF, perhaps via an out-of-band SOS/DAP mechanism, to combine DAP binary data access with SOS metadata access<br>• Shapefile for existing GIS clients |
| 23. | SOS encodings | Continue the IOOS collaboration with NOAA/NMFS ERDDAP and NFS OOI-CI to explore hosting of format conversion services on cloud-computing resources. |

| # | Category | Recommendation |
|---|---|---|
| 24. | Service Registry/Data Catalog/Data Viewer | Establish a prototype of these capabilities in FY2010 to increase the in-house familiarization/expertise with the technical and operational aspects |
| 25. | Service Registry/Data Catalog/Data Viewer | Augment the prototype capabilities and ensure operationally reliable hosting in FY2011 and beyond |
| 26. | Service Registry/Data Catalog/Data Viewer | Enable GEOSS and Geodata.gov to harvest IOOS Registry and Catalog, and thereby enable IOOS data providers to meet the requirement to register servers at GEOSS and Geodata.gov simply by registering with IOOS. |
| 27. | Service Registry/Data Catalog/Data Viewer | Explore methods to enable commercial search engines to harvest IOOS Data Catalog such that users can find data simply by searching the web. |
| 28. | Service-level metadata | Continue using OGC Web Service metadata. |
| 29. | Service-level metadata | Partner with NOAA DMIT Unified Access Framework project that is developing a crawler to generate summary metadata from a THREDDS Data Service (TDS) directory tree. |
| 30. | Service-level metadata | Use the future Service Registry to harvest and verify service metadata, and to make service metadata available in ISO 19139 format using the OGC Catalog Service for Web (CS/W). |
| 31. | Dataset-level metadata | Promote ISO 19115/19139 metadata development by offering complete examples of good metadata, suggesting tools to help write metadata (e.g., MEST, MERMAID), requiring IOOS-funded data providers to provide good metadata, and rating all IOOS data providers (even volunteer ones) according to the quality of their metadata (as part of a suite of Capability Maturity Level rankings). |
| 32. | Sensor and station metadata | Define minimum SensorML descriptions for IOOS sensors and platforms. |
| 33. | Sensor and station metadata | Ensure data providers make such metadata available, and display corresponding CM ranking. |
| 34. | Sensor and station metadata | Work with sensor manufacturers and Smart Sensor Consortium to promote the use of SensorML by vendors to describe each of their products. |
| 35. | Sensor and station metadata | Ensure linkages between metadata records to support overall metadata model. |
| 36. | Time-dependent metadata | Continue providing time-dependent metadata about observations as appropriate, based on customer need and metadata availability. |
| 37. | Phenomenon identifiers | Adopt CF names, and corresponding MMI URLs, for scalar quantities. |
| 38. | Phenomenon identifiers | Revise the IOOS phenomenon dictionary to define composite phenomena in terms of CF scalars. |
| 39. | Phenomenon identifiers | Add names and URLs of composite phenomena to MMI ontology. |

| # | Category | Recommendation |
|---|---|---|
| 40. | Phenomenon identifiers | Work with the CF Metadata group to define new CF names for scalars as IOOS expands to other variables. |
| 41. | Phenomenon identifiers | Work with MMI to add new names and URLs to ontology. |
| 42. | Phenomenon identifiers | Work with MMI to add new names and URLs to ontology. Include GCMD keywords in metadata about services, datasets and observing systems to aid in catalog searching and classification. |
| 43. | Phenomenon identifiers | Work with GCMD to add additional keywords as necessary. |
| 44. | Station and sensor identifiers | Revise the IOOS URN pattern for station and sensor identifiers. Include the owner of the station, the number assigned to the station by WMO or another authority, and the authority in question. |
| 45. | Station and sensor identifiers | Adopt 'ioos' rather than 'x-noaa' as the namespace, and register the 'ioos' namespace with IANA. |
| 46. | Station and sensor identifiers | Establish a URN resolver in order to link station and sensor IDs to the actual web-accessible resource describing that station or sensor (e.g., inputting the station ID in the resolver takes you to the SensorML metadata for that station). A URN resolver is also an IANA requirement for namespace registration. |
| 47. | Station and sensor identifiers | Work with MMI or OGC to use or learn from their URN resolver. |
| 48. | System Development | Play an active role in monitoring each project throughout the development stage, to ensure adherence to technical, budget and schedule requirements. |
| 49. | System Development | Maintain small, agile teams that can “build a little, test a little” to minimize risk and more quickly deploy services |
| 50. | System Development | Provide a central point of coordination across all development teams to ensure interoperability |
| 51. | Testing and Validation | Define a methodology/approach for testing each of these three areas that can be communicated to IOOS participants |
| 52. | Testing and Validation | Investigate available tests tools, harnesses or test services that can be leveraged. Determine if investment is required to develop test tools that can be used across data providers. |
| 53. | Testing and Validation | Define a certification and validation process for IOOS data providers. |

**Table 4-2 Summary of Technical Recommendations**

# 5.IOOS Regional Associations

IOOS RA data was not intended to be included in the DIF project (other than regional data that is served by NDBC). However, several factors contributed to the RAs implementation of DIF standards and services and general advancement in the homogeneity of data from the RAs. Factors that contributed to this progress include:

- Establishment of regular and formal coordination/collaboration between RA data managers and IOOS DIF team members
- DIF adoption of specific web services and encoding schemes that could be implemented
- Availability of an SOS code base, due to NDBC implementation of DIF SOS services, that formed the basis of RA implementations

As a result of the RAs implementing standard methods of serving their data, they transitioned from a group of nascent, fairly independent data providers to a more advanced data management group. Across the RAs there is now more uniformity of services of data formats.

Specific RA infrastructure and applications that remain in place due to this collaboration include:

- DIF-compliant SOS services operating at 9 of the 11 RAs (see Table 5-1)
- OGC WCS or OPeNDAP services operating at all 11 RAs
- Interoperability of model data output across all 11 RAs
- The GCOOS Data Portal that consolidates data from 10 local data nodes into a single aggregated SOS according to DIF standards. All GCOOS regional data nodes offer their data via a standardized IOOS interoperable interface; this is a direct result of the IOOS DIF activity. They nodes provide an up-to-the-minute accurate representation of what parameters are currently available via the continually updated XML files they host for the IOOS Regional Observation Registry. The local data nodes have adopted a common vocabulary (that was recently submitted to MMI - See The "IOOS Parameter Vocabulary V1.0 http://mmisw.org/orr/#http://mmisw.org/ont/ioos/parameter). This common vocabulary was developed with input from the IOOS DIF effort.

Examples of how this infrastructure is being used include:

- An OPeNDAP service (THREDDS server) at GLOS that is used by scientists at the NOAA Great Lakes Environmental Research Laboratory (GLERL) to compare the Great Lakes Coastal Forecasting System (GLCFS) modeled wave heights with the observations from the offshore Great lakes buoys. The comparisons are generated 2x per day via automated scripts and the results pushed to the GLERL Web site (http://www.glerl.noaa.gov/res/glcfs/thredds/opendap-wvh.php?buoy=45005). The tool proved its value due to power of THREDDS to serve virtual year-long data sets (appended on the fly from 6-hr data sets), and the speed of data acquisition over the network.
- Scientists in the Mid-Atlantic Coastal Ocean Observing System Regional Association (MACOORA) use THREDDS to serve High Frequency Radar sea surface currents observations that are used to forecast currents for the next 24 hours. The observations and forecast are now delivered operationally to the U.S. Coast Guard and integrated with satellite and model data to better predict location of missing persons or drifting vessels in search & rescue operations. Improvements in the quality of understanding of the dynamics of the ocean allows search and rescue controllers to optimize their search plans.
- NERACOOS has developed a tool which compares Northeast model forecasts with actual wave height and water level observations within the model domains (http://www.neracoos.org/products/modeldata). The product is driven by various SOS' for both the model and observation time series plots, and no local database cache is involved. Observations are obtained directly from the NOS, NDBC and NERACOOS SOS'. The model time series are obtained from an OOSTethys SOS which interfaces with remote THREDDS Data servers at USGS in Woods Hole and University of Massachusetts Dartmouth.

This section includes detailed information on RA implementation of the DIF standards and the model data interoperability project.

## 5.1. IOOS DIF Standards Implementation in IOOS Regional Associations (RAs)

Figure 5-1 illustrates the key milestones in the progression of RA implementation of DIF services. In late 2008, the Regional DIF Team was established, comprised of data managers from the RAs and IOOS DIF personnel. The purpose of this team was to move toward establishment of common methods for formatting and serving RA data. Over the course of the following two years, a series of meetings and workshops resulted in extensive implementation of DIF standards throughout the RAs.

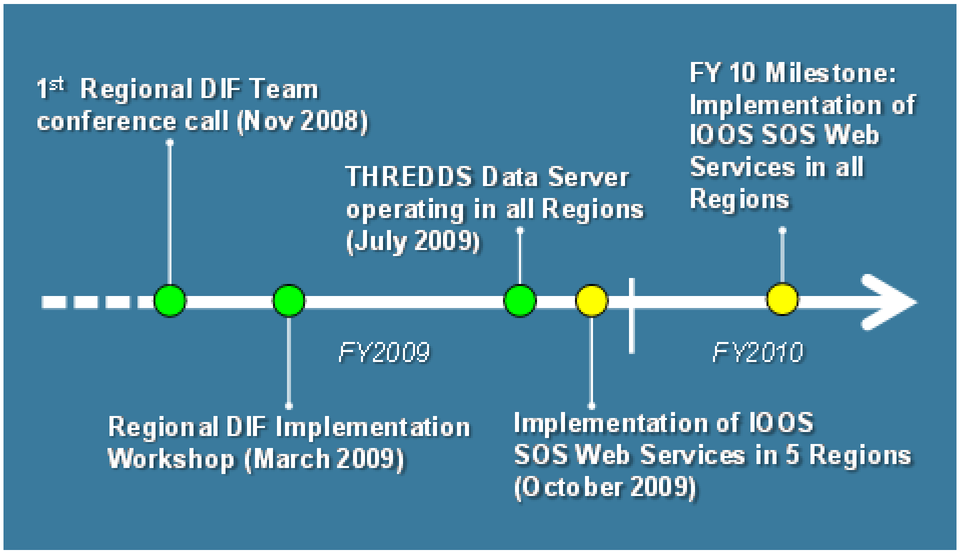


**Figure 5-1 Regional DIF Implementation: Key Milestones (by C. Alexander)**

**The knowledge and experience acquired during development of the proposed standards for data encoding and access services, and the implementation of those services at NDBC and CO-OPS, allowed IOOS to help the Regional Associations (RAs) to deploy the DIF Standards. As a result of that activity, all 11 RAs are now capable of serving data in a standardized format via OGC Web Services or OPeNDAP. Although the RAs are at disparate level of deployment (from initial testing to a fully operational status), a considerable volume of data has already been made available to users. The status of the data standards implementation as of March of 2010 is presented in**

Figure 5-2. Service implementation is an ongoing and evolving process, so this graphic is merely a snapshot of the status of service implementation throughout the RAs. Table 5-1 provides information about RA Sensor Observation Service implementations, including valid entry-point URLs. As previously noted, there is a wide disparity in the level of SOS implementation; some regions are operating an SOS server but do not yet use it to serve data. Others are actively serving data and supporting customer applications with their SOS services.

| Region | NDBC codebase | DIF output schema | SOS URL |
|---|---|---|---|
| **AOOS** | Yes | Yes | http://ak.aoos.org/ows/sos.php |
| **CaRA** | Yes | No | http://dm1.caricoos.org/sos/ |
| **CeNCOOS** | Yes | Yes | http://204.115.180.244/server.php |

| Region | NDBC codebase | DIF output schema | SOS URL |
|---|---|---|---|
| **GCOOS** | Yes | Yes | http://gcoos.rsmas.miami.edu/dp/sos_server.php |
| **GLOS** | Yes | Yes | http://michigan.glin.net/glos/sos/sos.php |
| **MACOORA** | No | Yes | http://www.weatherflow.com/sos/sos.pl |
| | | No | http://155.206.18.164/cgi-bin/sos/oostethys_sos.cgi |
| **NANOOS** | No | Yes | http://data.stccmop.org/ws/sos.py |
| **NERACOOS** | No | Yes | http://www.neracoos.org/cgi-bin/sos/V1.0/oostethys_sos_dif.cgi |
| | No | No | http://www.neracoos.org/cgi-bin/sos/V1.0/oostethys_sos.cgi |
| **PacIOOS** | No | No | http://oos.soest.hawaii.edu/oostethys/sos |
| **SCCOOS** | Yes | Yes | http://sccoos-obs0.ucsd.edu/sos/server.php |
| **SECOORA** | No | Yes | http://neptune.baruch.sc.edu/cgi-bin/difSOS.cgi |
| | | No | http://neptune.baruch.sc.edu/cgi-bin/oostethys_sos.cgi |

**Table 5-1 Regional SOS Implementation (July 2010)**

Nearly all of the implementations use the IOOS DIF GML schema for output data encoding, and more than half use the codebase developed by NDBC. Typically, RAs focus their efforts on development of products rather than data dissemination. The DIF SOS server software developed by NDBC, along with either IOOS DIF GML or OOSTethys SWE Common data encoding, was leveraged by the regions and allowed them an efficient way to establish their SOS services. In that way, the NDBC SOS server has become reference implementation software for IOOS. Going forward, NDBC will likely enhance and maintain the software, allowing extension to output formats other than DIF GML (e.g. CSV/TSV, KML, etc.). RAs will again be able to take advantage of that develop to enhance their own services.

All 11 RAs have also implemented OGC Web Coverage Service (WCS) and/or OPeNDAP. Nearly all implementations are based on Unidata THREDDS Data Server software, which supports both WCS and OPeNDAP, and has a many other advantages, e.g. more efficient serving of NetCDF, HDF and GRIB data. The complete catalog of IOOS THREDDS servers and their content can be found at the URL http://coast-enviro.er.usgs.gov/thredds/ioos_catalog_top.html . The servers provide both observations (e.g. High Frequency Radar currents data) and model data. The model data is "conditioned" in a way that maximizes interoperability between model outputs, using a technique developed and implemented through the IOOS DIF Model Data Interoperability project (see Section 5.2).

Although all RAs have implemented SOS and WCS/OPeNDAP servers, the usability of these services varies from region to region depending on the region's data architecture, i.e. data is holding as flat files on a disk or in a database. For example, NANOOS is using OPeNDAP because their database is too large to be effectively handled by an SOS Web Service, which is directed more toward light-weight database data; for a similar reason SCCOOS prefers THREDDS as it was easier to implement because their database structure was already compliant with the THREDDS file structure.

Now that standardized access services are in place, the focus can shift to the ability of users to discover the services and datasets. The visibility of services and datasets remains relatively low (e.g. GCOOS has noticed that practically no one in the region is pulling data via WCS). Potential data users should be able to discover the data as well as the access service entry point; this can be done via a Service Registry and Data Catalog such as the IOOS Data Catalog. The Catalog allows users to find the data they need, for the location and time period of interest, from all available IOOS partners without having to know in advance what partners operate the actual observing systems and data servers. The first version of the IOOS Data Catalog is now available on the IOOS Web site (http://ioos.gov/catalog). The functionality is still evolving, and not every IOOS observing platform is included yet.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **COLOR KEY: Implementation Status** | **Eligible Service (Not Operating)** | **Eligible Service (Operating) =** | | **Non-Eligible Service** | "Eligible" is being used only with respect to the IOOS Regional DIF Implemention effort. Two main elements: 1) is data being collected; and 2) is the web service supported. | | |
| **VALUE KEY: Operational Status** | **1 = Planned** | **2 = In Progress** | **3 = Min. Capacity** | **4 = In Testing** | **5 = Approved** | **6 = Certified** | **7 = Operational** |
| **SERVICE KEY:** | **SOS = Sensor Observation Service** | | **WCS = Web Coverage Service** | | **WMS = Web Mapping Service** | | **DAP = Data Access Protocol** |

| **DIF CORE VARIABLE** | **CV: Water Temperature** | | | | **CV: Water Level** | | | | **CV: Salinity** | | | | **CV: Surface Waves** | | | | **CV: Surface Winds** | | | | **CV: Surface Currents** | | | | **CV: Ocean Color** | | | | **CV: Model Output (Bonus)** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SERVICE TYPE** | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP | SOS | WCS | WMS | DAP |
| AOOS | 4 | 2 | 2 | 2 | 4 | | | | 4 | 2 | 2 | 2 | 1 | | | | 4 | 2 | 2 | 2 | 1 | | 2 | 2 | 4 | | | | | 4 | 4 | 4 |
| CARA | 1 | | 1 | 1 | 1 | | 1 | 1 | 1 | | 1 | 1 | 1 | | 2 | 1 | 3 | | 3 | 3 | 1 | | 1 | 1 | | | 1 | 1 | | | | |
| CENCOOS | 4 | 7 | 7 | 4 | 4 | 7 | 7 | 4 | 4 | 7 | 7 | 4 | | 7 | 7 | | 4 | 7 | 7 | 4 | 4 | | | 4 | | 7 | 7 | 3 | | 3 | 3 | 3 |
| GCOOS | 5 | | | 2 | 5 | | | 2 | 5 | | | 2 | 2 | | | 1 | 5 | | | 2 | 5 | | | 2 | | 2 | | 3 | | 3 | | 1 |
| GLOS | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | | | | | | | | | | | | | 1 | 3 | 4 | 3 | | | | | | | | |
| MACOORA | | | 4 | 4 | | | 4 | 4 | | | 4 | 4 | | | | 4 | 4 | | | | | | 4 | 4 | | | | 4 | | | 4 | 4 |
| NANOOS | 4 | | | 4 | 4 | 2 | | 4 | 4 | 4 | | 4 | 4 | 2 | | 4 | 4 | 2 | | 4 | 4 | 4 | | 4 | 4 | 2 | | 4 | | | 3 | 3 |
| NERACOOS | 4 | | 4 | 4 | | | | | 4 | | 1 | 4 | 4 | | 1 | 4 | 4 | | 4 | 4 | 4 | | 4 | 4 | 4 | | | | | | 4 | 4 |
| PACIOOS | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| SECOORA | 4 | | 4 | 4 | 4 | | 4 | 4 | 4 | | 4 | 4 | 4 | | 4 | 4 | 4 | | 4 | 4 | 4 | | 4 | 4 | | | | 3 | | | | |
| SCCOOS | | | | | | | | | 3 | | | | 3 | | | | | | | | | | | 4 | 1 | | | | | | | |

**Figure 5-2 IOOS Regional DIF Implementation Matrix (March 2010)**

## 5.2. IOOS DIF Model Data Interoperability project

The IOOS DIF Model Data Interoperability project (see Reference Document 9) was undertaken to enable interoperability between varieties of heterogeneous structured grid model outputs generated by IOOS partner organizations. The project showed how existing model output can be brought into CF-Compliance via NetCDF Markup Language (NcML), and made accessible to the scientist's desktop via standards-based tools.

To achieve the project goal, technology had to be implemented that would improve the RAs capability to provide information from differing models in a way that can be widely usable by various applications and platforms. The ideal solution would be to enable interoperability between models while placing minimal burden on the providers and users: the providers serve the data in native output formats, and the users are able to integrate the data using their preferred tools.

The project goal was achieved because it employed the DIF approach of overlaying standard services on the existing data output infrastructure rather than changing each underlying system. Interoperability of existing model outputs was done by implementing transformation tools in conjunction with the standardized Web Services (WCS or OPeNDAP).

Each model output provider started out with non-standard output files and their own web sites with various graphical displays of model results. This limited the ability to obtain model data and compare results of models across multiple providers. Many structured grid coastal models (ROMS, POM, ECOM, HYCOM, Delft3D) have curvilinear orthogonal coordinates in the horizontal and stretched terrain and free-surface following vertical coordinates (e.g. sigma, s-coordinate, double sigma, sigma over z).

The CF Conventions allow these to be specified in a standard way via metadata, allowing the possibility of standard-access to data on the native model grid. Again, the approach taken was to implement the missing or incorrect metadata via NetCDF Markup Language (NcML) as a middleware layer rather than force providers to rewrite their output files with proper conventions. Figure 5-3 illustrates the overall approach. In general, model interoperability was achieved through the following basic steps:

- Each modeling center places NetCDF files on their web site, or installs a THREDDS Data Server (more efficient and a unified way of serving NetCDF, HDF and GRIB)
- Use NcML to create virtual datasets in a standard form (CF)
- Each RA creates a central THREDDS catalog
- Create central catalog for all 11 RAs

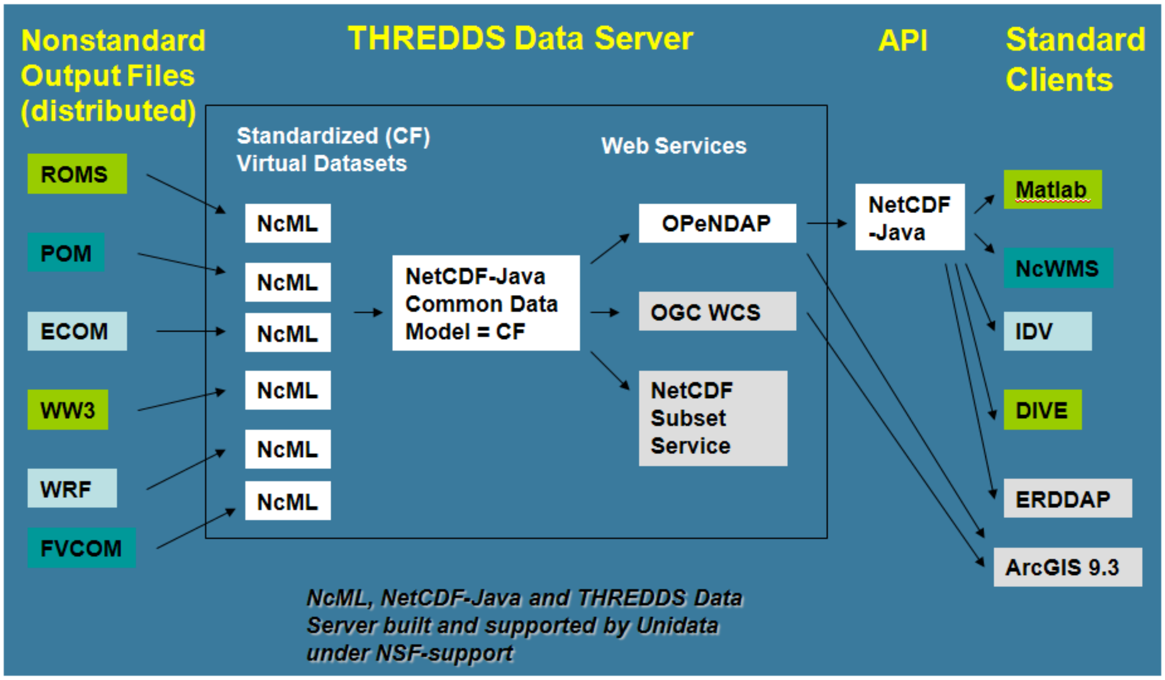


**Figure 5-3 Model Data Interoperability (R. Signell)**

This approach has proven to be very effective for structured grids, but for unstructured grids new standards conventions are needed as well as implementation of standards in an API (e.g. NetCDF-Java); in addition, the clients need to be modified to use the new methods (e.g. IDV, Matlab, ncWMS).

The first implementation of this system was in the Gulf of Maine, but the system has now been implemented in all 11 IOOS regions. This approach is also being used by the USGS National Hydrologic Modeling Structure Project, and the GEO-IDE Unified Access Framework project to provide common access to gridded data throughout NOAA. The OOI-CI Data Exchange project seems to further expand upon this successful model, implementing the middleware NcML and Web Services on the cloud, and adding registration, scalability, and handling of unstructured grid.

# APPENDIX A SOS Encoding Survey

Early on in the DIF project, the DIF Integrated Products Team (IPT) recommended adoption of the OGC GML specification as a foundation for encoding of in-situ data served via SOS. At present, the IOOS GML v0.6.1 is a primary encoding standard for in-situ data served via SOS by NOAA data providers, and some Regional Associations (RAs). However, as more data providers have been implementing IOOS GML for serving their data, and in view of the DIF project transition to more broad DMAC framework, IOOS has recognized potential benefits of adopting other standards, either powered by XML like Google KML, or hybrid like SWE Common.

In order to evaluate the operational utility of IOOS GML specification, IOOS determined it would be valuable to formally solicit feedback on the IOOS GML specification from data providers and data consumers. It was determined that the most expedient method for doing this would be an online survey. A survey was developed and distributed to the primary audience involved with the IOOS GML, either as a data provider, a data consumer, or a participant in the development of the schema. The survey focused only on the output format used for in-situ observations, i.e. on the XML format used for encoding in-situ observations served by the IOOS SOS at NDBC, CO-OPS and some RAs. The survey did not solicit input on SOS itself. The paragraphs below provide selected questions asked in the survey along with a summary of the responses. See Reference Documents for the source of the complete compilation of the survey questions and responses.

**Question #1: What is your involvement with SOS? (Check all that apply)**

Twenty specialists from NOAA and the RAs took part in the survey. Figure 5-4 illustrates the role of the various survey participants in the specification development, client/server code development, and/or data consumption.

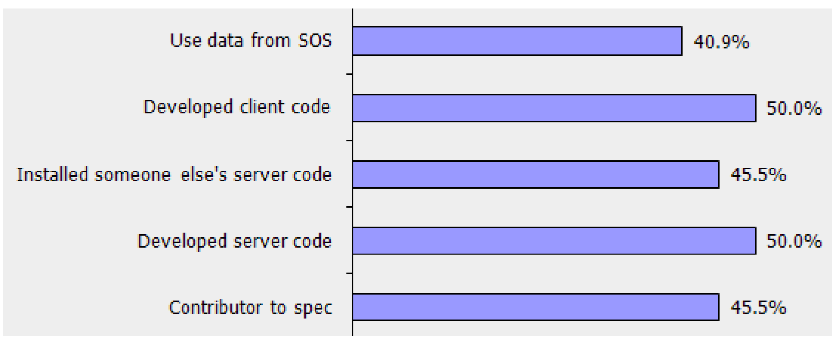


**Figure 5-4 IOOS GML Survey Participants**

**Question #2: Which of the following encoding formats do you have experience with? (Check all that apply)**

The survey participants appear to be very well-versed in the technologies and issues at hand: about 70% have worked with the IOOS GML encoding itself, and the overwhelming majority has been exposed to two or more data encoding formats (Figure 5-5).

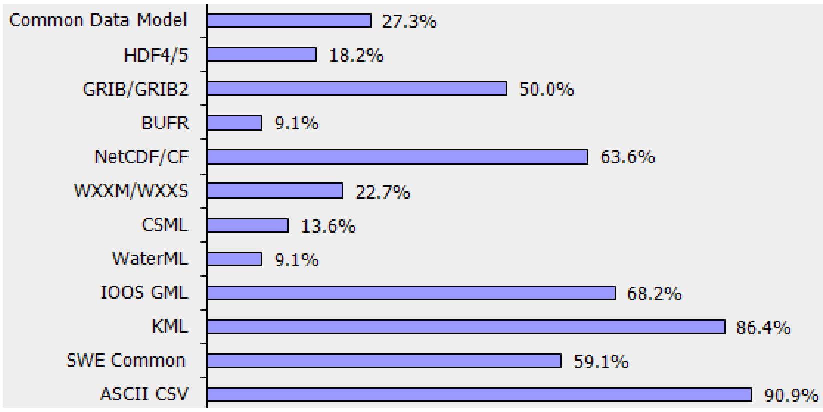


**Figure 5-5 Survey Participants' Experience with Data Formats**

**Question #3: In your opinion, how desirable are structure and hierarchy when representing observational data as XML?**

Next, the survey attempted to gauge what level of structure and hierarchy is desirable for in situ data encoding. All XML-based formats are characterized by a certain level of structure (i.e., each value has its own XML element or attribute), and hierarchy (i.e., elements are nested in a hierarchical relationship among observation stations, measurement times, and observed property). The survey queried the participants as to whether or not these parameters (and, in a sense, XML in general) are really important for in-situ data encoding.

The results showed that the GML data model itself, the chosen XML schema, and the level of implementation are the main reasons why the participants do not regard IOOS GML as an adequate single format for in-situ data encoding.

The majority of the participants generally agreed that both structure and hierarchy are desirable, although they disagreed in the degree of desirability (Figure 5-6). In total, three quarters of the audience ranked structured between "somewhat desirable" and "highly desirable"; only one quarter considers structure to be either "undesirable" or of "no use at all". In the case of hierarchy, that ratio increases to 95% to 5%. It is interesting that all 5% seem to believe that hierarchy is not needed at all.

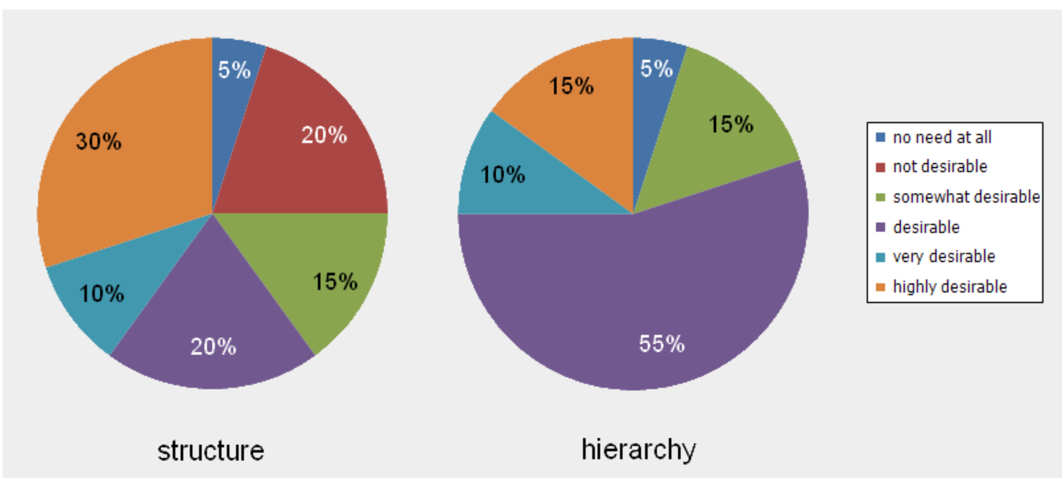


**Figure 5-6 Participant Opinion of Desirability of Structure and Hierarchy for In-situ Data Encoding**

**Question #4: In your opinion, how structured and verbose is IOOS GML v0.6.1?**

With just 25% of the participants regarding structure as generally undesirable, it is no surprise that the majority, 66.7%, of participants consider the level of structure of the IOOS GML to be high, moderate, and even insufficient

(Figure 5-7). However, one third of the audience evaluates the level of structure of IOOS GML as "excessive", a clearly negative assessment. In fact, an even higher percentage of the participants polled are dissatisfied with the excessive verbosity of the IOOS GML, which turns out to be one of the major arguments against expanding its implementation.

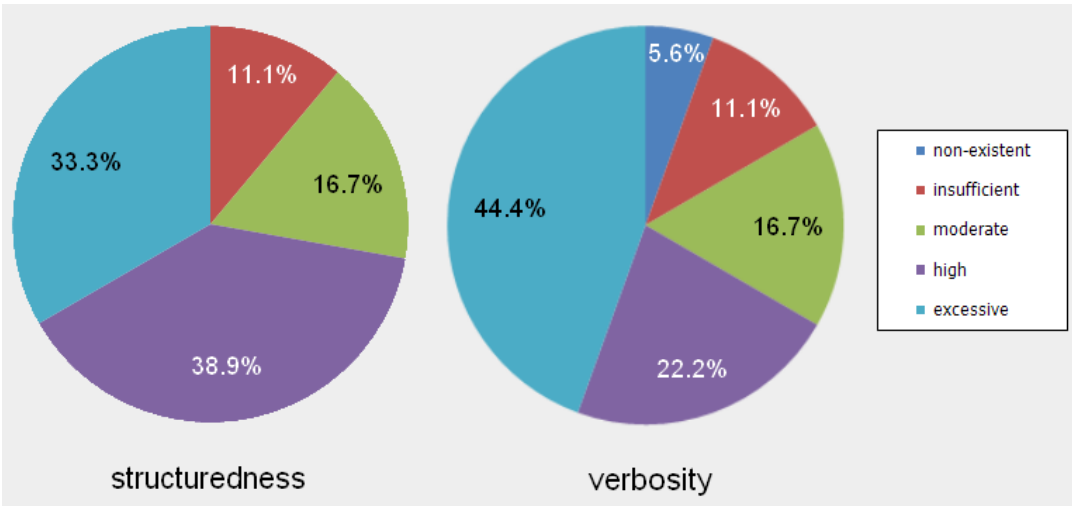


**Figure 5-7 The audience perception of the structuredness and verbosity of the IOOS GML**

**Question #5: Should IOOS SOS implementations offer GML as an output format?**

Survey participants were asked for their opinion on the IOOS GML format with respect to other formats. The survey results clearly demonstrated that, despite general acceptance of the IOOS GML encoding, the audience is not particularly favorable to it, and is willing to try an alternative. Only 5% of the audience strongly believes that there is no need for any encoding format other than IOOS GML. The overwhelming majority (about 80%) would like a variety of encoding formats, and wants IOOS GML to be just one of them. At the same time less than 16% believes the IOOS GML should be abandoned (Figure 5-8).

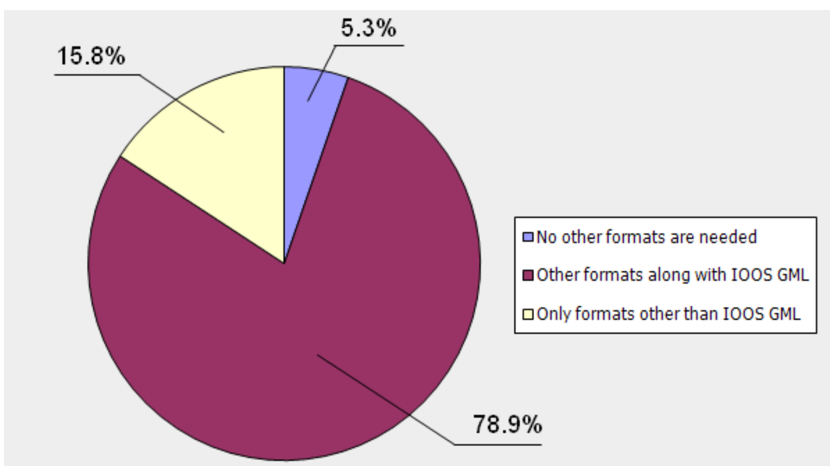


**Figure 5-8 Participant Attitude Toward IOOS GML Usage**

**Question #6: What should IOOS do with GML if it is offered?**

There is also no unified view on the future of the IOOS GML. Whereas 40% of the participants posited that IOOS GML either should be left alone or just have small changes made, another 40% would like to see more radical changes – either switch to a different GML schema (6.7%), or adopt an entirely non-GML schema (33.3%). Half that many (20.0%) selected the "in-between" approach; keep the IOOS GML on the condition that major changes shall be implemented (Figure 5-9).

The non-GML schema advocates form the largest group among those polled, reinforcing a general discontent with the IOOS GML. When added to the participants that want no or only small changes to IOOS GML enhancement, the group becomes a majority of about 73%, leaving only about a quarter of the participants advocating for continuation of the IOOS development trend, even at the cost of adopting a different GML schema.

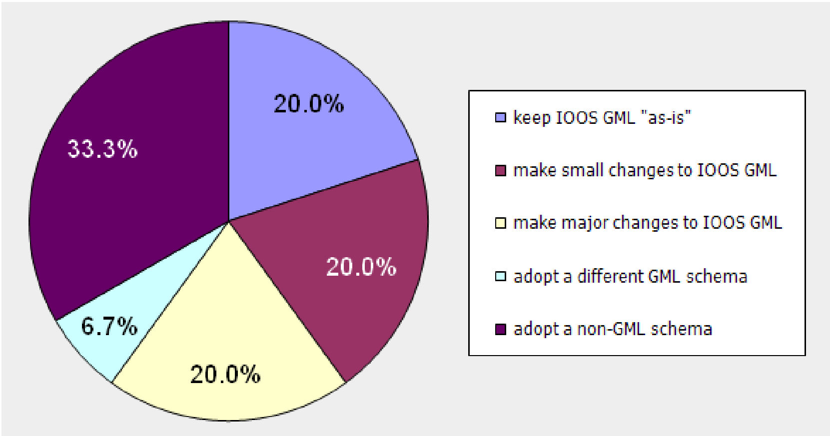


**Figure 5-9 The audience attitude toward further IOOS GML enhancement**

**Question #8: If you recommended the adoption of a different XM or GML schema in question 6, please specify [which]?**

The result of this questions shows that among the data encoding format candidates considered for addition to the IOOS SOS implementation, the OGC Sensor Web Enablement Common Data Model (SWE Common) received the vast majority of votes. Surprisingly, the Google KML format received only half as many votes as SWE Common. CSML and WXXM/WXXS XML incarnations together barely produced half as many votes as KML (Figure 5-10).

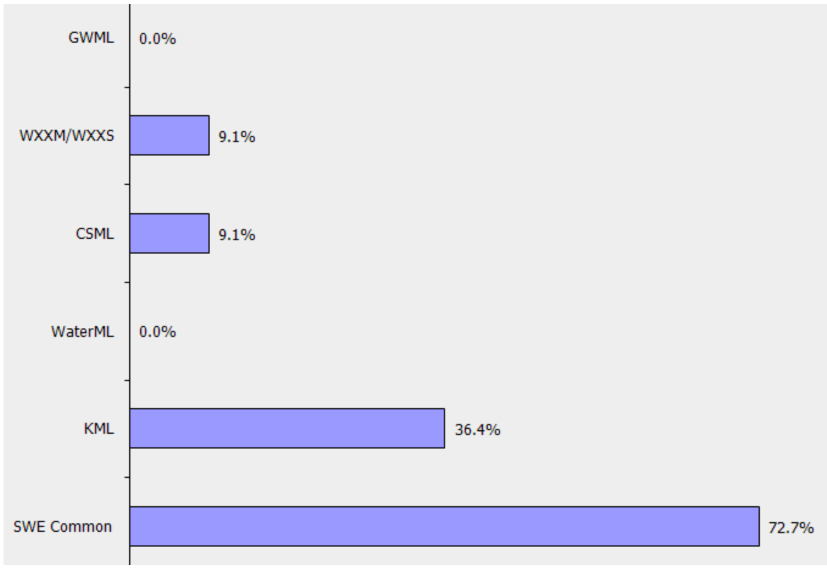


**Figure 5-10 The audience preference of the alternative encodings**

The result clearly shows that the only two encoding formats that should be implemented in addition to the IOOS GML are SWE Common and KML.

The popularity of SWE Common is likely due to the fact that it meets the community needs:

1. it is an open OGC standard (initially, SWE Common data structures were defined as part of the OGC SensorML v1.0 specification but will be defined as a separate standard in future versions);
2. it allows equally effective delivery of simple data types and aggregate data types;
3. it allows a high level of hierarchy without excessive structure and verbosity;
4. by supporting formats other than XML-based encoding formats, i.e. text blocks, binary blocks, and reference to well known standard encodings like JPEG or GIF, it allows source data in some binary format (e.g. netCDF, BUFR, GRIB2) to be just wrapped in the XML envelope, potentially saving a lot of computational resources, bandwidth and time;
5. numerous reference software tools have been developed by projects such as OOSTethys

The KML format is required because it facilitates providing data to a Google Earth client, which is a de-facto standard viewer for many data consumers.

In addition, there was an assumption, prior to the survey, that implementation of the CSV comma-delimited ASCII format would be also beneficial, and appreciated by the participants as it is a well known and widely-used format that allows users to parse and feed data directly into Excel spreadsheets for analysis without any intermediate conversion. The survey confirmed that assumption: more than 90% of the respondents have used the ASCII CSV encoding in their operations (Figure 5-5). To support the community's need, IOOS had sponsored the successful CSV output encoding format implementation at NDBC; it is expected that other data providers will also offer in-situ observations in CSV format in the near future.

# APPENDIX B Acronyms

| | |
|---|---|
| ADCIRC | Advanced Circulation Hydrodynamic Model |
| ADCP | Acoustic Doppler Current Profiler |
| AOML | Atlantic Oceanographic and Meteorological Lab |
| AOOS | Alaska Ocean Observing System |
| AUV | Autonomous Underwater Vehicles |
| AWIPS | Advanced Weather Interactive Processing System |
| CalCOFI | California Cooperative Oceanic Fisheries Investigations |
| CaRA | Caribbean Regional Association |
| CDMO | Centralized Data Management Office |
| CeNCOOS | Central and Northern California Ocean Observing System |
| CF | Climate Forecast (netCDF) |
| CI | Coastal Inundation |
| CIO | Chief Information Officer |
| CIP | Customer Implementation Project (DIF) |
| C-MAN | Coastal-Marine Automated Network |
| CODAR | Coastal Ocean Dynamics Applications Radar |
| COMPS | Coastal Ocean Monitoring and Prediction System |
| CO-OPS | Center for Operational Oceanographic Products and Services |
| CSC | Coastal Services Center |
| CSDL | Coast Survey Development Lab |
| CSML | Climate Science Modeling Language |
| CSV | Comma-separated values |
| CTD | Conductivity, Temperature, Depth |
| DAP | Data Access Protocol |
| DCS | DIF Data Content Standard |
| DIF | Data Integration Framework |
| DMAC | Data Management and Communications |
| DMIT | Data Management Integration Team |
| EAO | Enterprise Architect's Office |
| EPA | Environmental Protection Agency |
| ERD-DAP | Environmental Research Division Data Access Program |
| ETSS | Extratropical Storm Surge |
| FGDC | Federal Geographic Data Committee |
| FIPS | Federal Information Processing Standards |
| FTP | File Transfer Protocol |
| GCOOS | Gulf of Mexico Ocean Observing System |
| GEOSS | Global Earth Observing System of Systems |
| GLOS | Great Lakes Observing System |
| GML | Geography Markup Language |
| GOM | Gulf of Mexico |
| GSFC | Goddard Space Flight Center |

| | |
|---|---|
| GTS | WMO Global Telecommunications System |
| GTSPP | Global Temperature and Salinity Profile Program |
| HAB | Harmful Algal Bloom |
| HAB-FS | Harmful Algal Bloom Forecasting System |
| HDF | Hierarchical Data Format |
| HF | High frequency (radar) |
| HFR | High Frequency Radar |
| HI | Hurricane Intensity |
| HLS | Hurricane Local Statement |
| HTTP | Hyper Text Transfer Protocol |
| HWRF | Hurricane Weather Research and Forecasting |
| HYCOM | Hybrid Coordinate Ocean Model |
| IDS | Required Data Sets |
| IEA | Integrated Environmental Ecosystem Assessments |
| IGOSS | Integrated Global Ocean Services System |
| IMS | Information Management System |
| IOC | Initial Operating Capability |
| IOC | Intergovernmental Oceanographic Commission |
| IODE | International Oceanographic Data and Information Exchange committee |
| IOOS | Integrated Ocean Observing System |
| IPT | Integrated Project Team |
| IT | Information Technology |
| JCOMM | Joint Commission on Oceanography and Marine Meteorology |
| KML | Keyhole Markup Language |
| MACOORA | Mid-Atlantic Coastal Ocean Observing System |
| MDL | Meteorological Development Lab (NWS) |
| MEDS | Canada's Marine Environmental Data Service |
| MMI | Marine Metadata Interoperability |
| MMS | Minerals Management Service |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| NAM | North American Mesoscale |
| NANOOS | Northwest Association of Networked Ocean Observing Systems |
| NCCOS | National Centers for Coastal Ocean Science |
| NCDDC | National Coastal Data Development Center |
| NCEP | National Center for Environmental Prediction |
| NDBC | National Data Buoy Center |
| NEP | National Estuary Program |
| NERACOOS | Northeastern Regional Association of Coastal and Ocean Observing Systems |
| NERRS | National Estuarine Research Reserve System |
| NESDIS | National Environmental Satellite, Data, and Information Service |
| netCDF | Network Common Data Form |
| netCDF/CF | Network Common Data Form/Climate Forecast |
| NGDC | National Geophysical Data Center |
| NGOM | Northern Gulf of Mexico |
| NIST | National Institute of Standards and Technology |

| | |
|---|---|
| NMFS | National Marine Fisheries Service |
| NOAA | National Oceanic and Atmospheric Administration |
| NODC | US National Oceanographic Data Center |
| NOS | National Ocean Service |
| NWLON | National Water Level Observation Network |
| NWS | National Weather Service |
| NWSTG | NWS Telecommunications Gateway |
| OAR | Oceanic and Atmospheric Research |
| OGC | Open Geospatial Consortium |
| OPeNDAP | Open-source Project for a Network Data Access Protocol |
| PacIOOS | Pacific Regional Integrated Ocean Observing Systems |
| PMEL | Pacific Marine Environmental Laboratory |
| POM | Princeton Ocean Model |
| PORTS® | Physical Oceanographic Real-Time System® |
| QA | Quality Assurance |
| QC | Quality Control |
| RA | IOOS Regional Association |
| RD | Reference Document |
| RDBMS | Relational Database Management System |
| SCCOOS | Southern California Coastal Ocean Observing System |
| SDP | SLOSH Display Program |
| SeaWiFS | Sea-viewing Wide Field-of-view Sensor |
| SECOORA | Southeastern Coastal Ocean Observing Regional Association |
| SLOSH | Lake, and Overland Surges from Hurricanes |
| SOS | OGC Sensor Observation Service |
| SURA | Southeastern Universities Research Association |
| SWE | Sensor Web Enablement |
| SWEC | Sensor Web Enablement - Common |
| TABS | Texas Automated Buoy System |
| TGLO | Texas General Land Office |
| TDS | THREDDS Data Server |
| TPC | Tropical Prediction Center |
| TSV | Tab-separated values |
| URN | Uniform Resource Name |
| USF | University of South Florida |
| USGS | United States Geological Survey |
| WCS | OGC Web Coverage Service |
| WFS | OGC Web Feature Service |
| WMO | World Meteorological Organization |
| WMS | OGC Web Map Service |
| WSDE | IOOS Working Group on Standards and Data Encoding |
| XML | Extensible Markup Language |
| XSLT | Extensible Stylesheet Language Transformation |