FINAL REPORT

A Super-Regional Testbed to Improve Models of Environmental Processes on the U.S. Atlantic and Gulf of Mexico Coasts

Principal Investigator:
Richard Luettich, University of North Carolina, Chapel Hill
rick_luettich@unc.edu

Project Director:
L.D. Wright, Director of Coastal and Environmental Research, SURA
wright@sura.org

Project Coordinator:
Elizabeth Smith, SURA
lsmith@sura.org

Award Number:
NA10NOS0120063

Period of Performance:
June 1, 2010 – May 31, 2012

FY10 Budget

$3,526,000 from NOAA/IOOS Office; distributed as follows:
Academia - $2,418,525
SURA Management (incl. IDC and travel) - $627,000
Private Sector/Consultants - $359,919
Non-NOAA Federal Agencies - $120,556

This report will highlight the accomplishments, progress and challenges of the SURA management and information technology activities and each of the Testbed teams for the period of June 1, 2010 – May 31, 2012 and is organized as follows:

Table of Contents

Prepared by Liz Smith (lsmith@sura.org).

1 Executive Summary and Societal Relevance of the Testbed

Coastal ocean and estuarine environments along with adjacent shorelines, wetlands and lowlands are threatened by climate change, sea-level rise, storm-induced flooding, oxygen depleted "dead zones", oil spills and unforeseen disasters. With funding for the reporting period (June 2010-May 2012) from NOAA's Integrated Ocean Observing System (IOOS) Office, the Southeastern Universities Research Association (SURA) has facilitated strong and strategic collaborations among experts from academia, federal operational centers and industry and guided the U.S. IOOS Super-Regional Modeling Testbed through its first highly productive phase.

The <u>mission</u> of the Testbed is *targeted research and development to accelerate the transition of scientific and technical advances from the coastal ocean modeling research community to improved operational ocean products and services.* The long-range <u>vision</u> of the program is *to increase the accuracy, reliability, and scope of the federal suite of operational ocean modeling forecast products to meet the needs of a diverse user community. We consider operational use to cover a wide range of society-critical applications including forecasts (e.g., 4x365 or event based forecasts), hindcasts (e.g., event based forensic studies), risk assessment, and design and system management (e.g., nutrient management regulations).*

The Testbed is serving NOAA's mission to protect, restore and manage coastal resources through an ecosystem approach, as well as to improve NOAA's capability to serve society's needs for weather and water information.

Through this SURA-led, U.S. IOOS-funded project, we have come to better understand and articulate the specific capabilities that we expect a Testbed to provide. These capabilities include:

- Quantitative data on the behavior and implementation requirements of models that are presently in operational use or that are under serious consideration for such use.
- An organized archive of observational data, model inputs and model results that can be used for testing and evaluating both current and future models.
- Tools that leverage or, as necessary, define community standards to enable the efficient access, visualization, skill assessment and other evaluations of multiple model results.
- A research environment, or cyber-infrastructure, where researchers and operational agencies can work together on selected modeling applications as a means of fostering and enhancing the transition of models from research to operations.

Since its initiation in June, 2010, the IOOS Testbed has developed a flexible and extensible community research framework (including a supporting "cyber-infrastructure" as well as an interdisciplinary network of scientists and stakeholders) to advance the testing and evaluation of predictive models of key coastal ocean environmental issues. This framework supports integration, comparison, scientific analyses and archiving of data and model output. The cyber-infrastructure that has been developed includes a repository of data assembled from numerous observations and models as well as tools for comparing and assessing the models and data. Several models, tools and techniques are now being incorporated into NOAA's operational framework. Both the IOOS Office and the National Center for Environmental Prediction (NCEP) have declared the Testbed to be a valuable National resource. Indeed, NCEP is the first operational center to become an "anchor" of the Testbed. The

Testbed addresses major NCEP modeling challenges in coastal predictions by enabling the transition of research improvements into NCEP's operational forecast capability. In addition to numerous academic partners, Federal participants include agency representation from NOAA, Navy, EPA and the U.S. Army Corp of Engineers (USACE).

The first year and a half of the Testbed addressed three research challenges of high socioeconomic relevance: **estuarine hypoxia, shelf hypoxia, and coastal inundation**. The Testbed has yielded new insights for obtaining more accurate model predictions that will help to protect life and property and facilitate economic resiliency. Examples of these include:

- A multi-model ensemble was found to yield more accurate predictions of the "dead zone" in the Chesapeake Bay than any one of the models evaluated in this component of the testbed. The use of ensemble modeling has been recommended to the Chesapeake Bay Program for forecasting future water quality conditions in the Bay.
- Improvements were achieved in the skill of models for predicting the timing and location of hypoxic conditions in the Chesapeake Bay. Such 'dead zones' have a significant impact on living marine resources; predicting their occurrence is critical for ecosystem management within the Bay.
- A simple dissolved oxygen formulation for forecasting the location and timing of seasonal hypoxia was transitioned to NOAA/CSDL's research version of the Chesapeake Bay Operational Forecast System (CBOFS) for evaluation.
- A synoptic scale modeling capability was demonstrated for predicting the timing, duration, spatial extent, and severity of the northern Gulf of Mexico 'dead zone'. This capability has helped reduce uncertainties about the role of nutrient loads in generating the 'dead zone'. Model transition to NOAA/CSDL is underway for evaluation.
- The evaluation of three unstructured grid, coupled wave and surge models demonstrated the efficacy of these models for high resolution predictions of storm surge coastal flooding by nor'easters in the Gulf of Maine. This work demonstrated the efficiency of nesting local area models within a coarse, regional model in areas of relatively steep coastal topography and the importance of wave effects on coastal flooding.
- Storm surge and coastal flooding due to hurricanes in the northwestern Gulf of Mexico were shown to be the result of a forerunner (due to Ekman setup in advance of the storm) as well as direct onshore winds near the time of storm landfall. An evaluation of three unstructured gird, coupled wave and surge models and the widely used SLOSH model demonstrated the sensitivity of model domain size and the parameterizations of surface stress and bottom friction to accurately predicting both the forerunner and the land-falling response.
- A version of SLOSH that is coupled to a wave model was developed in the Testbed and is now being tested by the National Hurricane Center. One of the unstructured grid models will be evaluated during the coming year as part of the NOAA Hurricane Forecast Improvement Program.
- The model archive and tools developed by the cyber-infrastructure component of the Testbed are being used both by Testbed participants and by operational groups that were not part of the initial Testbed (e.g., NRL, National Hurricane Center) to improve model evaluation, skill assessment and visualization.

- A less tangible product of this Testbed, but one that will have a significant lasting impact, is the community-building that resulted from scientists working together on the shared goal of improving model performance. For example, in several cases when one model was found to have less skill than others, the Testbed modelers worked together to figure out the cause of these differences, and to improve the underperforming model. These interactions and feedbacks resulted in significant improvements to several of the models evaluated in the Testbed. For example, the Estuarine Hypoxia team reported that the CHESROMS model realized a 40% overall reduction in RMS difference between predicted and observed bottom dissolved oxygen concentration due to improvements identified during the Testbed.

2 SURA Management

This report covers the entire period of FY 2010 funding from the U.S. IOOS Program (June 2010 through May 2012) which includes the initial one year performance period plus an additional year of no cost extension. Dr. Rick Luettich, University of North Carolina-Chapel Hill, assumed the role of lead principal investigator on the Testbed project on June 1, 2011. Dr. L. Donelson Wright serves as the Project Director for SURA, and Elizabeth Smith is the Project Coordinator.

SURA provided overall project leadership, multi-institution coordination and management of the project and insured that federal funds were expended in a responsible and accountable manner. SURA issued and managed subcontract agreements to 13 Testbed partners including universities (U.S. and Canadian), federal agencies (U.S. and Canadian), industry, and consultants. SURA communicated the expectation for regular monthly invoicing of expenses and maintained close communication with the Principal Investigators throughout the project. SURA monitored, approved, and paid the invoices as they were submitted through SURA's financial and accounting services.

Additional coordination and administrative activities for this reporting period were:

- Conducted two "all-hands" Testbed PI meetings for 70 attendees from academia, federal operational centers and industry including leaders from NCEP and NOS in June of 2010 (Testbed Kick-off meeting) and in June 2011, each for 1.5 days.

- Facilitated 4-5 teleconferences and/or web-based meeting per month among Testbed teams and the TAEG and IOOS Program Office.

A collaborative web site was developed by Cyber-infrastructure team-member, University of Alabama Huntsville: http://Testbed.sura.org. This site is easily accessible to the public and serves as the central archive of all Testbed documents, presentations, publications and meeting announcements. SURA works closely with the website administrators and developers to develop content and to provide feedback on the usability of the site.

2.1 Budget Analysis

SURA requested, and was granted, two six-month No Cost Extensions (NCE) of the Super Regional Modeling Testbed project, extending its performance period from the originally scheduled ending date of May 31, 2011 until May 31, 2012. The NCEs were necessitated by several unforeseen factors, including delays in completing sub-contractual agreements, delays in hiring key support personnel, delays in obtaining necessary computer time and complexities in exchanging model outputs.

Rick Luettich, of UNC at Chapel Hill, assumed the role of lead principal investigator on the Testbed project beginning with the start of the first no-cost extension period.

SURA directly administered 13 sub-awards, of which 3 were to universities that in turn each administered several more Testbed contracts. All-in-all, the Testbed consisted of more than 30 sub-awardees. SURA began to make regular, monthly contact with the PIs for each sub-award in October 2010 to urge them to invoice in a timely manner. This consistent contact subsequently resulted in a more steady and consistent pace of billing by the PIs. The figure below shows the financial summary as of April 9, 2012.

NOAA/IOOS TESTBED GRANT SUMMARY

	BUDGET	INVOICED	REMAINING	% REMAINING	% INVOICED	SOW % COMPLETE
Subawardees:						
DUSA	$ 41,933	$ 41,933	$ -	0%	100%	100%
Galvarino	$ 48,203	$ 48,203	$ -	0%	100%	100%
GMRI	$ 72,304	$ 72,303	$ 1	0%	100%	100%
MSU	$ 603,477	$ 603,477	$ -	0%	100%	100%
UAH	$ 96,404	$ 96,404	$ -	0%	100%	100%
USACE	$ 62,500	$ 62,500	$ -	0%	100%	100%
VIMS	$ 879,716	$ 879,716	$ (0)	0%	100%	100%
NRL	$ 96,556	$ 96,464	$ 92	0%	100%	100%
ASA	$ 238,533	$ 238,426	$ 107	0%	100%	100%
UF	$ 57,843	$ 57,711	$ 132	0%	100%	100%
TAMU	$ 28,921	$ 28,782	$ 139	0%	100%	100%
UNC-RENCI	$ 57,610	$ 57,342	$ 268	0%	100%	100%
UNCCH	$ 615,000	$ 610,196	$ 4,804	1%	99%	100%
Subawardee Total	$ 2,899,000	$ 2,893,458	$ 5,542	0%	100%	
SURA:						
Program Mgmt/OH	$ 627,000	$ 622,157	$ 4,843	1%	99%	
TOTAL	$ 3,526,000	$ 3,515,615	$ 10,385	0%	100%	

Cumulative reimbursements at Apr. 9, 2012:	***100%***
Time elapsed:	***100%***

As a result of billing delays, SURA Financial personnel (Cravens and Bjonerud) and Program Manager (Smith) conferred with IOOS Program office personnel (Baltes and Evans) to develop an action plan for managing the Testbed grant going forward. This was immensely helpful and outlined a clear plan to help keep the invoicing on track and on-time. The following “Best Practices” were agreed upon: SURA will (1) foster relationships with each sub-award contracting officer; (2) tighten up their monitoring/measuring/benchmarking of tasks completed in relation to rate of spending; and (3) ensure PIs interact with account managers at universities.

2.2 Testbed Governance

The key elements of the U.S. IOOS Super-Regional Modeling Testbed governance that have been crucial to success are:

1) The appointment of a non-conflicted Testbed Advisory and Evaluation Group (TAEG) to provide objective, independent guidance, and assessment of progress;

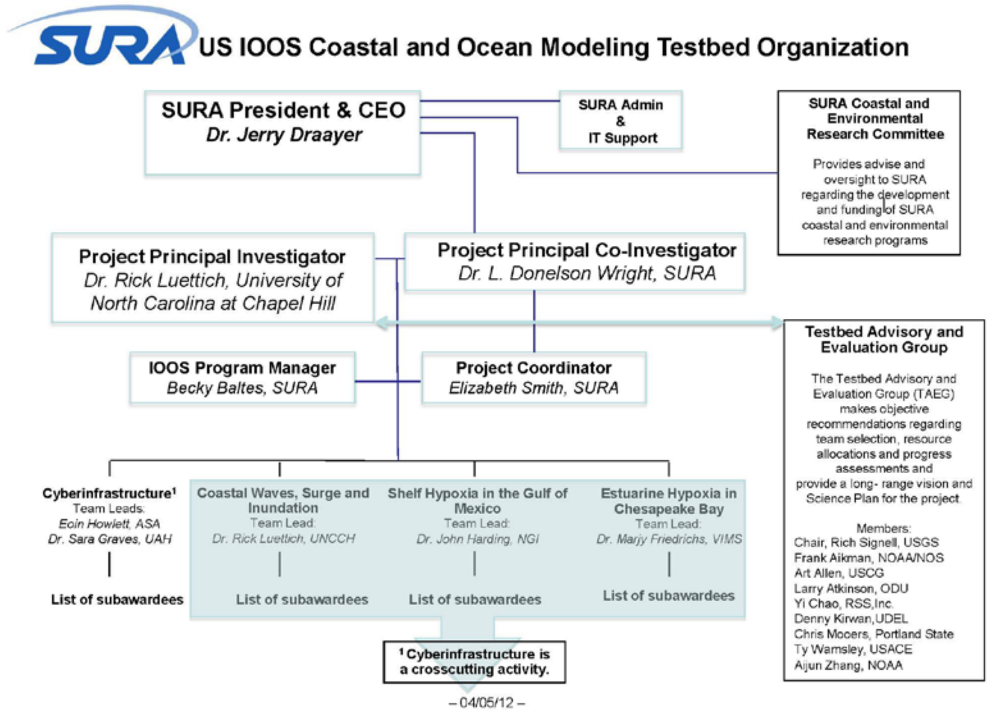
2) Distributed leadership: each team had a team lead and held regular telecons/meetings of all participants;
3) Web site to enable knowledge management and facilitate communication;
4) Centralized management and project coordination by SURA.

SURA's overall management and project coordination for the Testbed not only ensured that the federal funds were expended in a responsible and accountable manner, but also includes the full support by SURA's financial and accounting services, grants and contracts management services, and support services. The Testbed enjoyed the full endorsement of the SURA Coastal and Environmental Research Committee (CERC) which is chaired by Dr. Chris D'Elia of LSU. Although not exercising day-to-day project management, the CERC is the "authorizing" entity for SURA Coastal Research projects. The organization of the Testbed is depicted in the chart below.

The Testbed Advisory and Evaluation Group (TAEG) is an independent, technical and scientific advisory group comprised of scientific and computer experts from academia and the federal government. The TAEG makes objective recommendations regarding team selection, resource allocations and progress assessments and provide a long- range vision and Science Plan for the project.

The Testbed Principal Investigator (PI) works closely with the SURA Project Director, who in turn reports to the SURA CEO. The Testbed PI is responsible for leadership and multi-institutional coordination for the Testbed, interacting with the TAEG as necessary to assess project status and develop a long-range vision. The Project Coordinator at SURA is responsible for the day-to-day activities and communications among Testbed PIs and teams, and also is the liaison among the SURA HQ, the U.S. IOOS Program Office and the Testbed PIs and teams.

The Testbed is comprised of three modeling teams (Inundation, Estuarine Hypoxia and Shelf Hypoxia) each with a team lead, and a Cyber-infrastructure team which is developing the overarching, and sustaining Testbed framework which will be for the benefit of the modeling teams and others wishing to utilize the framework.



2.3 SURA Information Technology Staff Contributions

In 2011, SURA provided great support for the cyber-infrastructure needs of the U.S. IOOS Super-Regional Modeling Testbed through the efforts of SURA and LSU Information Technology (IT) staff. Much of the support described below has been provided at no cost to the project.

2.3.1 Access to HPC Resources

Over the life of the project, SURA IT has acquired 16.45 million service units (SUs), valued at approximately $2,467,500, on large scale computational resources through the development and submission of successful competitive proposals to LONI, TeraGrid, and XSEDE (TeraGrid's NSF funded successor). Access to these systems has been deemed essential to the significant progress achieved by the Inundation Team and is enabling the project team to assess the ability of the identified models to accurately predict the impact of storm related flooding and wave action.

Use of these systems enabled the Inundation Team to run simulations using larger grids, move from 2-D to 3-D models, and explore model coupling such as coupled wind and wave models. In August and September, Hurricane Irene and Tropical Storm Lee hit the U.S. East Coast and provided the opportunity to exercise SELFE and WWM (Wind Wave Model) in a real-time forecast mode. Resources acquired include: Queen Bee at LONI and Ranger at TACC with an additional NSF Petascale Resource Allocation (PRAC) proposal for 46.8 million SUs awaiting an award decision. SURA IT staff also played a key role in coordinating the efforts of model developers, Inundation Team members and HPC systems support staff to port model codes to these large and specialized systems.

2.3.2 Proposal Development

SURA IT contributed to the Testbed team's proposal development efforts by providing strategy for the use and implementation of local and national cyberinfrastructure, writing support, and liaison with NSF

Office of Cyberinfrastructure (OCI) Program Officers and, to date has participated in the development of the BP Oil Spill, NSF Earth Cube, and NSF RCN-SEES proposals.

2.3.3 Testbed Server

SURA provides software and hardware (at no cost to the project) that is used for hosting the Testbed collaboration website (Testbed.sura.org) and the Testbed data and tools repository. LSU provides server hosting, system administration and operational support services and SURA IT provides problem management facilitation and additional technical guidance.

2.3.4 Leverage SURA IT Community

The involvement of SURA IT allows the Testbed to leverage a broader set of SURA IT community projects and resources including:

- Access to interactions with the extended SURA IT Committee (i.e. joint Coastal/IT Committee meeting last Fall that focused on the NSF Earth Cube program)
- The engagement of UAH for the development and maintenance of the Testbed collaboration website and tools.
- Access to convenient in-person and webcast HPC user training. Half a dozen members of the coastal community attended the 2011 SURA-TACC training workshop “Introduction to Scientific Visualization on Longhorn” and additional participants from the coastal community are expected to attend the 2012 training workshops.

3 Modeling and Cyber-infrastructure Team Detailed Reports

3.1 Coastal Waves, Surge and Inundation**

Team Lead: Rick Luettich

3.1.1 Project Summary

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575.

The goal of this component of the Testbed is to provide NOAA and other governmental agencies meaningful guidance on the behavior (e.g., accuracy, robustness, execution speed) and implementation requirements (e.g., resolution, parameterization, computer capacity) of models that are presently in “operational use”, or that are under consideration for such use, for computing waves, storm surge and inundation. The project team is comprised of individuals from nine universities and three groups within NOAA:

R. Luettich – UNC - team lead
B. Beardsley, WHOI – team co-lead, GoM modeling
C. Chen, UmassD – GoM FVCOM/SWAVE modeling
C. Li, LSU - GoMx FVCOM/SWAVE modeling
W. Perrie, BIO – GoM WWIII/SWAN modeling
D. Slinn, UF – SLOSH/SWAN modeling
H. Wang, VIMS – GoM & GoMx SELFE/WWM modeling
B. Weisberg, USF – GoMx FVCOM/SWAN modeling
J. Westerink, UND – GoM & GoMx ADCIRC/SWAN modeling
A. Kramer, A. Haase, A. Myckow NWS MDL – SLOSH modeling

J. Rhome, C. Forbes, NWS NHC – SLOSH modeling

J. Feyen, NOAA CSDL - NOAA Storm Surge Portfolio Manager

A much larger group participates in the effort when associates and students at each university are included. This component of the Testbed operates in two geographical regions: the Gulf of Maine (GoM) with high resolution nesting in Scituate Harbor, MA, and the Gulf of Mexico (GoMx) with enhanced resolution along the Louisiana / Texas coast. Waves, storm surge and inundation in the GoM / Scituate are evaluated in response to two Nor'easters (May 2005, April 2007) and in the GoMx in response to two hurricanes (Rita 2005, Ike 2008).

Early project effort was devoted to assembling the data infrastructure and common methodologies required for this component of the Testbed, e.g., development of common model grids; creation of model forcing data; identification and QA/QC of observational data sets; standardization of file formats and run parameters; and development of skill assessment tools. Once data and methodological issues were resolved, numerous comparative model runs were conducted for tidal forcing and the two storms selected in each domain. Skill assessment and intra- / inter-model comparisons now provide a basis to define model accuracy, implementation requirements, and computational performance to improve the operational use of these models.

3.1.2 Results and Accomplishments

The milestones associated with this component of the project are identified in the following table. A summary of the results and accomplishments for each milestone is provided thereafter. Additional detail is contained in the semi-annual Testbed progress reports and in eight manuscripts (listed in section 3.1.3) that are currently in preparation for a Testbed volume in Journal of Geophysical Research.

		Proposal Milestones						
4.3	***Skill Assessment***	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complete	Task Lead	Finish Date
4.3.1	Identify relevant observational data sets					100	B Beardsley, J Westerink	3/1/2011
4.3.2	IMEDS GUI Enhancements and integration with data standards					100	J Hanson, E Howlett	6/30/2011
4.3.3	Assist with IMEDS utilization by modelers					100	J Hanson, E Howlett	6/30/2011
4.4	***Extratropical Storm Evaluation***	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complete	Task Lead	Finish Date
4.4.1	Identify at least 2 extratropical storms to use for testbed activities.					100	B Beardsley	7/1/2010
4.4.2	Identify observational data for model evaluation					100	B Beardsley	3/1/2011
4.4.3	Place all required data files on project server (grids, forcing input)					100	C Chen	3/15/2011
4.4.4	Execute model runs for selected storms (FVCOM/SWAVE, WW3/SWAN, SLOSH/SWAN, SELFE/WWM, ADCIRC/SWAN)					95	C Chen, J Westerink, H Wang, W Perrie, D. Slinn	12/31/2011
4.4.5	In-depth analysis to determine causes for different model results					95	C Chen, J Westerink, H Wang, W Perrie, D. Slinn	12/31/2011
4.4.6	Evaluation of transition to operations implementation requirements					100	J Feyen, B Beardsley, R Luettich, MDL/NCEP rep?	12/31/2011
4.4.7	Final Analyses/summary writeups					75	group	ongoing
4.5	***Tropical Storm Evaluation***	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complete	Task Lead	Finish Date
4.5.1	Identify at least 2 tropical storms to use for testbed activities					100	J Westerink	7/1/2010
4.5.2	Identify model grid to be used for tropical storms					100	J Westerink	9/15/2010
4.5.3	Identify observational data for model evaluation					100	J Westerink	11/30/2010
4.5.4	Place all required data files on project server (grids, forcing input)					100	J Westerink	7/31/2011
4.5.5	Execute model runs for selected storms					90	J Westerink, H Wang, B Weisberg, C Li, D Slinn	12/31/2011
4.5.6	In-depth analysis to determine causes for different model results					100	J Westerink, H Wang, B Weisberg, C Li, D Slinn	12/31/2011
4.5.7	Evaluation of transition to operations implementation requirements					100	J Rhome, J. Feyen, R Luettich, others?	12/31/2011
4.5.8	Final Analyses/summary writeups					75	group	ongoing

3.1.2.1Summary of Accomplishments for Testbed Milestone 4.3: Skill Assessment

1. Identify relevant observational data sets
    a. Observational datasets were identified for each geographical component. In the GoM these data were comprised of standard NDBC buoy observations. Unfortunately, no data were available within Scituate Harbor during the 2005 and 2007 storms for model skill assessment. Subsequently a water level gauge was installed in Scituate Harbor and its data were used for model skill assessment for tidal forcing. As described in further detail below, substantial data sets were identified and compiled for model skill assessment for tides, Rita and Ike in the GoMx.
2. IMEDS GUI Enhancements and integration with data standards
    a. IMEDS was modified to accept a generic ASCII input data format which was adopted as a data standard within the Testbed. In combination with the existing NetCDF format, IMEDS is now functional for all of the initially anticipated skill assessment applications, which includes waves and water levels. Suggestions for additional IMEDS capabilities

were received during the course of the project and their implementation is being funded from other sources (e.g., programs within the US Army Corps of Engineers).

3. Assist with IMEDS utilization by modelers
    a. In addition to the modification to IMEDS to accept generic ASCII input, output translators were developed to convert individual model output into an IMEDS compatible format, IMEDS documentation was developed, and several training webcasts were held for the modelers.

3.1.2.2 Summary of Accomplishments for Testbed Milestone 4.4: Extratropical Storm Evaluation (Gulf of Maine)

1. Identify at least 2 extratropical storms to use for Testbed activities.
    a. Storms from 2005, 2007 that caused significant flooding in Scituate area were selected for evaluation.
2. Identify observational data for model evaluation
    a. All known data were identified for the 2005 & 2007 storms. These data are located outside of Scituate Harbor. A water level gauge was recently installed in Scituate Harbor; data from this gauge for the month of May 2010 were obtained for model evaluation to tidal forcing.
3. Place all required data files on project server (grids, forcing input)
    a. Files have been on server since the fall of 2010, although subsequent adjustments were made to the project grid based on initial model results.
4. Execute model runs for selected storms (FVCOM/SWAVE, WW3/SWAVE/SWAN, SLOSH/SWAN, ELCIRC/SELFE/WWM, ADCIRC/SWAN)
    a. FVCOM/SWAVE and WW3/SWAN runs were completed for the 2005 and 2007 storms for the greater GoM domain. FVCOM/SWAVE, SELFE/WWM and ADCIRC/SWAN runs were completed for the 2005 and 2007 storms, as well as for 2010 tides, for the Scituate domain. Runs were also completed using the coupled SLOSH/SWAN models on several of NOAA's operational grids.
5. In-depth analysis to determine causes for different model results
    a. Water levels generated using FVCOM, SELFE and ADCIRC are virtually identical at the Scituate water level gauge for the 2010 tides simulation. Harmonic decomposition of the model elevation output showed that all three models had high skill in both amplitude and phase, for the major diurnal and semi-diurnal tidal constituents.
    b. In the case of the 2005 storm simulation without the inclusion of wave-current interaction, FVCOM, SELFE and ADCIRC predict nearly identical water surface responses throughout most of the Scituate model domain and specifically at eight of nine designated model comparison locations. At the ninth comparison location, SELFE and FVCOM predict inundation while ADCIRC does not. Depth-averaged velocities at the eight stations are consistent between the models although they are more variable than surface elevations. The primary reason for this variability is because each model computes velocity in a different location (ADCIRC – element corners, FVCOM – element centers, SELFE – element faces).
    c. In the case of the 2005 storm, when wave-current interaction is included in the model runs, ADCIRC/SWAN, FVCOM/SWAVE, SELFE/WWM predict similar water levels and inundation at all nine designated comparison stations and generally throughout the Scituate domain. Larger differences are observed between the model velocity fields, due to the development of relatively small scale velocity features such as eddies,

recirculation zones and areas of significant lateral shear. Based on the similarity in water surface elevations, it does not appear that this small scale velocity variability significant affects water flux through the majority of the domain.

   d. In the case of the 2007 storm without wave-current interaction, all three models predict nearly identical water levels throughout the Scituate domain and consistent depth-averaged velocities, similar to the 2005 storm. In this case, no significant differences occur in the models' predictions of inundation at the inland comparison station.
   e. In the case of the 2007 storm when wave-current interaction is included in the model runs, all three models predict nearly identical water levels throughout most of the Scituate domain. However, in this case ADCIRC/SWAN predicts no inundation at the on land comparison station while the other two models do inundation this area. While FVCOM/SWAVE and SELFE/WWM both inundate this area, they do so at significantly different times and SELFE/WWM is approximately 5 cm greater than FVCOM/SWAVE. While observations indicate that this area did inundate, unfortunately there is no data to provide a time history or actual water levels in the area.
   f. Including wave-current interaction in the models substantially strengthens and in some cases changes the direction of the velocity field in coastal areas outside of Scituate Harbor, at the entrance to Scituate Harbor and within Scituate Harbor. Wave-current interaction generates additional flux into Scituate Harbor and enhances small scale velocity features within the harbor. While these do not translate to large differences in water level within the harbor, they do influence the inundation of areas having elevations that are close to the astronomical high tide level.
   g. Evaluation of SLOSH and SLOSH/SWAN model results in this area was inconclusive. These models are not configured to use the meteorological forcing used by the other models or to include tidal forcing (which in this area contributes substantially more to the total water level than the storm surge) and therefore results are difficult to directly compare to the other models. The resolution of the existing SLOSH grids is sufficiently coarse that it is impossible to evaluate the model response inside of Scituate Harbor (which is smaller than a single SLOSH grid cell).
   h. At the larger Gulf of Maine scale, comparisons indicate that wave model significant wave height (Hs) and peak period (Tp) are comparable between SWAVE and the nested WW3/SWAN at buoy locations. Hs is fairly sensitive to the spectral and directional resolution with combinations of 35 frequencies, 72 directions giving lower Hs near the peak of the 2007 storm than the initially proposed combinations of 21 frequencies, 36 directions. Both models appeared to be biased low in Hs, although at times this is at least partially due to the contribution of low frequency swell that neither model captured and may have propagated in through the boundary of the domain rather than being internally generated.

6. Evaluation of transition to operations implementation requirements
   a. FVCOM/SWAVE is currently being used as part of a quasi-operational forecast system in the Gulf of Maine and Scituate Harbor running at UMassD. Insights developed during the Testbed, including revisions to the model grid, required spectral and directional resolution in the wave model and the importance of wave-current interaction in the greater Scituate Harbor area have all impacted the development of this system. Additional perspectives on the transition of this technology to operations may be forthcoming as the final analysis is completed.
7. Final Analyses/summary write-ups

a. There are few differences in the predictions between FVCOM/SWAVE, ADCIRC/SWAN and SELFE/WWM in the Scituate Harbor study. Of the differences the most significant was the small scale velocity structure induced by the wave-current interaction. None of these coupled wave/current models has been rigorously evaluated at the scales present in this problem and therefore it is not clear if any provide a more accurate solution than the others.
b. It is noteworthy that in these comparisons ADCIRC was run in 2-D, depth-averaged form, while SELFE and FVCOM were run in 3-D.
c. More detailed results from the Extratropical Storm Evaluation are currently being finalized and prepared in two manuscripts to be submitted for publication in a Testbed volume in the Journal of Geophysical Research.

3.1.2.3Summary of Accomplishments for Testbed Milestone 4.5: Tropical Storm Evaluation (Gulf of Mexico)

1. Identify at least 2 tropical storms to use for Testbed activities
    a. Hurricanes Ike (2008) and Rita (2005) were selected because they represented two significant, recent hurricanes that impacted the western Louisiana and northeastern Texas coasts and substantial observational data sets exist for both.
2. Identify model grid to be used for tropical storms
    a. Originally, the Computational Hydraulics Lab out of the University of Notre Dame provided the group with SL+TX-GoM+Atl-Lite-01, a ~2 million node unstructured mesh that included the Gulf of Mexico, Eastern Seaboard and Atlantic Ocean out to the 60°W, but this proved to be unworkable for the group, so the domain was reduced to the Gulf of Mexico and the mesh resolution was decreased to create SL+TX-GoM-UL-01, a 424,485 node mesh. Internal boundary features, such as levees were also found to be unusable by FVCOM and SELFE, so they were removed, and a group approved SL+TX-GoM-UL-01-LR (417,642 nodes) was created. The bathymetry and nodal attributes for this unstructured mesh were linearly interpolated from SL18+TXv33 [Hope et al. 2012]. This unstructured mesh was used by the unstructured models (ADCIRC, FVCOM, and SELFE), whereas SLOSH, which uses a curvilinear quadrilateral grid system, often particular to the landfall target of the approaching hurricane, employed several of its model specific meshes. SLOSH was simulated on a local mesh and gulf mesh for each of the hurricanes. For hurricane Ike, egl3 (45790 nodes) was the local "Galveston" mesh and egm3 (185409 nodes) was the "Gulf of Mexico" mesh, while for hurricane Rita, ebp3 (77827 nodes) was the local "Sabine Pass" mesh and egm3 was the "Gulf of Mexico" mesh.
3. Identify observational data for model evaluation
    a. The group identified 59 NOAA stations that contained tidal constituent data that were used for the tidal analysis. Observational hydrographs at over 200 locations, plus high water marks, have been identified for Ike. A smaller, but substantial number of data sets have been identified for Rita. Datasets were converted from original their original source format to the IMEDS ascii format.
4. Place all required data files on project server (grids, forcing input)
    a. Final grid, tidal forcing & observational data (59 locations), Ike & Rita wind forcing & observational hydrograph data were posted during the course of the first year of the Testbed.

5. Execute model runs for selected storms
    a. A variety of 2D and 3D runs were executed. The following breaks down the different runs.
        i. Tides
            1. 2D (ADCIRC,FVCOM, SELFE)*
            2. 3D (ADCIRC,FVCOM, SELFE)*
        ii. Hurricane Ike:
            1. 2D (ADCIRC, FVCOM, SELFE)* (SLOSH)^
            2. 2D w/ waves (ADCIRC+SWAN, FVCOM/SWAN, SELFE+WWMII)*
            3. 3D (ADCIRC, FVCOM, SELFE)*
            4. 3D w/ waves (ADCIRC+SWAN, FVCOM/SWAN, SELFE+WWMII)*
            5. waves only (SWAN, WWMII)*
        iii. Hurricane Rita:
            1. 2D (ADCIRC, FVCOM, SELFE)* (SLOSH)^
            2. 2D w/ waves (ADCIRC+SWAN, FVCOM/SWAN, SELFE+WWMII)*
            3. 3D (ADCIRC, FVCOM, SELFE)*
            4. 2D w/ waves (ADCIRC+SWAN, FVCOM/SWAN, SELFE+WWMII)*
            5. waves only (SWAN, WWMII)*
    b. Planned model runs with the coupled FVCOM/SWAVE by C. Li have not been completed.

    *SL+TX-GoM-UL-01-LR with OWI wind stress (if applicable), ^SLOSH Basins with SLOSH internal wind forcing

6. In-depth analysis to determine causes for different model results
    a. Model skill assessment (using six statistical skill metrics contained in IMEDS) for tides was conducted for ADCIRC, FVCOM and SELFE. Each model was run for 105 days; the first 30 days served to ramp up the forcing and allow the model response to settle while half-hourly time series were recorded during the final 75 days for analyses at 59 NOAA stations covering the 5 Gulf coast states. The Matlab-based program T_Tide was used to decompose the time series into 38 harmonic constituents including four diurnal (O1, K1, Q1 and P1) constituents and four semi-diurnal (N2, M2, S2 and K2) constituents. Overall, the three models showed very similar skill, with R-squared values in the range of 0.35 for the amplitudes of the four diurnal constituents and 0.75 for the four semi-diurnal constituents. Phase results were better across the board with an average R-squared of 0.8. Model skill was best in relatively simple coastal settings; ADCIRC runs using a higher resolution grid (~9.2 million horizontal nodes) yielded substantially higher skill.
    b. Tidal results from ADCIRC, FVCOM and SELFE were fairly insensitive to the model's frictional formulation provided that reasonable parameter values were used by each. However, the models' responses to hurricanes Rita and Ike were sensitive to both the surface stress and the bottom friction. Considerable effort was expended to ensure these were identical between the models, although this required insight into the models' formulations and parameterizations. An important lesson learned was that these models cannot be treated as black boxes with default parameter settings and expected to yield robust predictions. Rather, operational use of any of these models should be overseen by individuals who are well versed with the particular model.
    c. Surge and inundation during Hurricane Ike were characterized by two distinct phases; during the 24 hours prior to landfall, shore parallel winds along the western Louisiana

and northeastern Texas coasts created over 2 m of Ekman setup (also called forerunner). This response is quite sensitive to a model's bottom friction formulation and parameter values. An additional 3 m of surge occurred as the storm made landfall due to the onshore blowing winds. ADCIRC(/SWAN), FVCOM(/SWAN) and SELFE(/WWM) all captured the forerunner and the main surge for Ike at comparable levels in western Louisiana and Texas. Model skill was good in relatively open (e.g., coastal) settings. However, in many of the more complex and inland areas, the resolution of the Testbed grid was inadequate to accurately propagate surge and inundation. Comparisons with a much higher resolution ADCIRC run (~9.2 million node grid) indicated that the higher resolution run was much better at resolving the more geometrically complex areas (similar to the tidal simulations). SLOSH (with and without SWAN) on both the local Galveston Basin (egl3) and the Gulf of Mexico wide basin (egm3) did not reproduce any appreciable forerunner. SLOSH on the local Galveston Basin was generally 1-2 m below the other models in representing the total surge, while on the Gulf of Mexico wide basin it more closely reproduced peak water levels in open coastal areas. While the qualitative aspects of this finding appears to be robust (e.g., completely missing the forerunner), a rigorous quantitative comparison between SLOSH and the other models is hampered because SLOSH does not use the same wind / surface stress fields as the other models, but rather uses a parametric wind field with a much simpler surface drag relationship. UND Testbed participants are in the process of updating SLOSH to remove this limitation with wind / wind stress.

d. Hurricane Rita made landfall approximately 150 km to the east of Ike. Shore parallel winds created much less Ekman setup which was limited to the area immediately west of the Mississippi River delta. The contribution of this setup was minor in the areas near landfall and the greatest surge and inundation were largely due to the onshore blowing winds in the northeast quadrant of the storm at landfall. As was the case with Ike, ADCIRC, FVCOM and SELFE give comparable and reasonably accurate results in open locations. For this storm, SLOSH generated surge and inundation that were substantially greater than the other models. We believe the varying performance of SLOSH in both Ike and Rita are due to the surface and bottom stress parameterizations that are built into the model, although final understanding awaits the ability to force SLOSH with the same wind / wind stress fields as used in the other models.
e. As mentioned above, UND compared ADCIRC results from the standard Testbed grid, SL+TX-GoM-UL-01_LR, with a high resolution grid, SL18+TX33 (~9.2 million node). In more open settings, model response including both water surface elevations and wave parameters were relatively insensitive to the increased grid resolution (which extended over much of the continental shelf areas). Rather, the primary response area affected by resolution was near the inundation front. The coarse mesh was unable to resolve many barrier islands or waterways and the resulting smooth coastline allowed for a relatively uniform inundation. Inundation patterns in the high-resolution mesh reflected the better resolved topography, including ridges, valleys and existing waterways. In areas of large storm surge, the water overwhelmed much of the topography and was often similar between the model runs. Significant differences mostly occurred at modest surge / inundation levels.
f. Both 2D and 3D model formulations appear able to capture the surge / inundation response due to Rita and Ike. Due to the differences in the way that bottom friction is represented in 2D vs 3D, it is difficult to translate model parameters from one model

formulation to the other. Thus, parameters must be selected individually for each and will strongly influence conclusions about model behavior.

   g. Timings were collected for each simulation and uploaded to the Testbed with the run's metadata, although due to the use of different numbers of processors, loose coupling between FVCOM/SWAN (involving data file exchange on two different computer platforms) vs tight coupling between ADCIRC/SWAN and SELFE/WWM (involving direct memory passing on a single computer platform) and conservative time step selections by the modelers, a rigorous comparisons between the model performance was difficult to extract from this information. Consequently, a dedicated performance run was conducted for a 10-day Ike simulation (with and without waves) on 512 cores on the Ranger computer at the Texas Applied Computational Center. The results without waves were 3:24:13 (SELFE dt=60s), 4:26:27 (FVCOM dt=1.2s), and (6:18:55 (ADCIRC dt=1s), and with waves were 10:51:46 (SELFE+WWMMII dt=60s), and 11:03:39 (ADCIRC+SWAN dt=1s). However, even these results are equivocal because 1) model time step sizes were not evaluated to give optimal model performance, 2) models scale differently depending on the number of processors available, 3) the models are coupled in different ways (i.e. FVCOM/SWAN not tightly coupled), and 4) no model output was written (which can significantly impact the relative timings between the models).

7. Evaluation of transition to operations implementation requirements
   a. The tropical component of the Testbed has begun to illustrate short comings in the operational storm surge model (SLOSH) used to predict hurricane storm surge within NOAA / NWS. In the present case this is a combination of basin size, resolution and model parameterization (specifically the surface and bottom stress). This identification of need is important to guide future improvements to SLOSH or to motivate the transition to new modeling technology.
   b. The version of SLOSH coupled to the SWAN wave model that was used in the Testbed was provided to the storm surge group at the National Hurricane Center along with initial setup and training on computers available to the NHC surge group.
   c. Also, NOAA will evaluate an ensemble of ADCIRC+SWAN model runs in the Gulf of Mexico as part of the Hurricane Forecast Improvement Program during the 2012 hurricane season. Distribution and visualization of model results will utilize the cyber-infrastructure methodology and tools developed as a part of the Testbed.
   d. These Testbed activities have focused on model hindcasting for the explicit purpose of evaluating model skill and comparative performance. These are critical quantities for understanding the viability of any model under consideration for operational use. However, there are other issues that must be considered (e.g., difficulty of setting up and executing model runs, frequency of model instability, run time on available computing resources) before a model can be transitioned into operational use, particularly into a real time forecast environment.
   e. The Testbed has highlighted the critical need for well-trained modelers and sufficient computational resources to enable the use of the current state-of-the-art in wave, storm surge and inundation models. Presently these impediments are limiting the transition of more sophisticated models to operational use more so than the availability of the models themselves.
8. Final Analyses/summary write-ups
   a. The implementations of FVCOM/SWAN, ADCIRC/SWAN and SELFE/WWM in the Gulf of Mexico component of the Testbed provided similar results for tidal to wave/surge/

inundation calculations provided the same forcing and model parameterizations were used. Based on the model results and on the available run time information, all three model sets are capable of performing high fidelity wave, surge and inundation predictions. The selection of a specific model for operational use from among these three may depend on the wave/surge/inundation specific features that are built into each model, the specific HPC architecture available to run the model and the knowledge base available to oversee the use of the model.

b. Due to substantial differences in its implementation (including its grids and frictional parameterization), SLOSH gave systematically different results from the other models.
c. Details of the differences between the models and their parameterization dependencies are being more fully presented in manuscripts that are currently in development for publication in a Testbed volume in the Journal of Geophysical Research.

3.1.3 Communication and Dissemination

Inundation Testbed results were presented and discussed in depth at the Testbed all hands meeting in June 2011. The meeting was well attended by representatives from multiple operational federal agencies.

Rick Luettich and Bob Beardsley presented posters covering results from the inundation Testbed at the Coastal Modeling Gordon Conference at the end of June 2011.

Linda Akli presented a poster covering results from the inundation Testbed at the TeraGrid11 Conference in July.

John Harding presented material from the inundation Testbed as part of a Testbed overview presentation at the Coastal Zone Conference in July 2011.

Jamie Rhome, who leads the storm surge group at the National Hurricane Center, invited inundation lead, Rick Luettich, to visit the NHC on Sept 6-8, 2011, to get a first hand appreciation of how hurricane forecasting is done and of the operations of the storm surge group. It was a great time to visit since there were several active storms in the Atlantic basin. While there Luettich presented a seminar covering recent advancements in storm surge modeling with emphasis on the inundation Testbed. The seminar was attended by the NHC Director, Bill Reed, forecasters, members of the storm surge group and colleagues from Florida International University (the NHC is on the FIU campus).

Rick Luettich gave a talk on the inundation Testbed at the 12th International Workshop on Wave Hindcasting and Forecasting / 3rd Coastal Hazard Symposium, Kohala, HI in October 2011.

Rick Luettich gave a talk on the inundation Testbed at the 12th International Conference on Estuarine and Coastal Modeling, St. Augustine, FL in October 2011.

In December 2011, inundation co-PI Don Slinn visited the National Hurricane Center and provided the coupled SLOSH - SWAN model code he had developed and utilized as part of the Testbed as well as initial training on its use. Slinn is continuing to assist the NHC implement the coupled models for use around the Caribbean Islands including Puerto Rico.

Additional talks on this component of the Testbed were given at the 2012 AMS Annual Meeting and the 2012 Ocean Sciences Meeting. These will be reported on as communication activities in Testbed #2.

The following manuscripts are currently in preparation that report on the wave/surge/ inundation Testbed activities:

1. BIO group lead – all participants as additional authors, “Inter-model comparison experiments of surface waves during 2005 and 2007 extratropical storm events in the Gulf of Maine”

2. UMassD/WHOI group lead – all participants as additional authors, “Inter-model comparison experiments for 2005 and 2007 extratropical storm events in Scituate, Massachusetts”

3. Donahue, A.S., R.C. Martyr, P.C. Kerr, J.J. Westerink, R.A. Luettich, Jr., M.E. Hope, H.J. Westerink. “SURA-IOOS Coastal Inundation Testbed, Sensitivity of Tidal Response to Domain Size, Mesh Resolution and Frictional Parameterization."

4. Kerr, P.C., A. S. Donahue, J. J. Westerink, R. A. Luettich Jr., L. Y. Zheng, R. H. Weisberg, H. V. Wang, D. N. Slinn, J. R. Davis, Y. Huang,Y. Teng, D. Forrest, A. T. Haase, A. Kramer, J. R. Rhome, J. Feyen, R. P.Signell, J. Hanson, A. Taylor, M. E. Hope, R. Estes, R.Dunbar, L. Semeraro, H. J. Westerink, A. Kennedy, J. M. Smith, M. D. Powell, V.J. Cardone, A. T. Cox. “SURA-IOOS Coastal Inundation Testbed, Gulf of Mexico Inter-Model Comparison: Tidal, Wind, Wave and Surge simulation for Hurricanes Rita and Ike.”

5. Kerr, P.C., M.E. Hope, R.C. Martyr, J.J. Westerink, R.A. Luettich, Jr., A.S. Donahue, J.C. Dietrich, C. Dawson, H.J. Westerink, L.G. Westerink. “SURA-IOOS Coastal Inundation Testbed, Sensitivity of Hurricane Wave and Surge Response to Domain Size, Mesh Resolution and Frictional Parameterization."

6. Wang, H and Yi-cheng Teng, “The Effect of Bottom Boundary Layer Dynamics on the Forerunner Simulation during Hurricane Ike in the Gulf of Mexico”.

7. Zheng, Weisberg, Luettich, Westerink, Kerr, Donahue, Crane, and Akli, “Implications of 2D vs 3D model formulation on Hurricane Ike storm surge”

8. Huang Y., R.H. Weisberg and L. Zheng, Gulf of Mexico hurricane wave simulations using SWAN: Bulk formula based drag coefficient sensitivity for Hurricane Ike.

3.2 Estuarine Hypoxia Modeling in Chesapeake Bay

Team Lead: Marjy Friedrichs

3.2.1 Project Summary

The overall goal of the Estuarine Hypoxia (EH) component of the Super-Regional Modeling Testbed was to evaluate existing hydrodynamic and water quality models used or likely to be used for operations and/or for regulation (e.g., establishing nutrient TMDLs – total maximum daily loads) in the Chesapeake Bay.

By engaging experts from NOAA, EPA and the Army Corps, we leveraged existing community resources to evaluate a diverse suite of models for Chesapeake Bay (CB) hydrodynamics and oxygen dynamics. The hydrodynamic models compared included CBOFS2, CH3D, UMCES ROMS, ChesROMS, and EFDC. Comparisons were made to observed temperature, salinity, and dissolved oxygen over multiple spatial and temporal scales. A range of standardized hind-cast boundary conditions, including those downscaled from other agency models, were used to force and test the CB models, during both wet and dry years, and under diverse wind conditions. A wealth of long-term observational data from the EPA Chesapeake Bay Monitoring Program was utilized for model-data comparison.

Model skill assessment was performed based on metrics such as bias and root mean squared differences and assessments were illustrated via Target diagrams. We were also able to leverage Testbed resources at CCMP and at the NSF-funded Community Surface Dynamics Modeling Program through its existing Chesapeake Focus Research Group (http://csdms.colorado.edu/wiki/CCMP).

The results of the EH Testbed provided the CB modeling community with a better understanding of the uncertainty inherent in predictions by agency and community models of salinity, temperature and dissolved oxygen concentrations. These results furthermore clearly demonstrated the utility of agencies simultaneously applying several models to obtain ensemble estimates of hydrodynamics and oxygen for use in programmatic decision-making.

ESTUARINE HYPOXIA UPDATE		Proposal Milestones						
5.3	*MAB hydrodynamic model comparison*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complet	Task Lead	Finish Date
5.3.1	Define hindcast/nowcast/forecast periods					100%	EH Team	1-Oct-10
5.3.2	Assemble common forcing data sources; identify data sets for model testing					100%	J. Wilkin / J. Levi	1-Jun-11
5.3.3	Establish and refine metrics for model-model-data intercomparison					100%	J. Wilkin / J. Levi	
5.3.4	Downscale from basin-scale to MAB to CB forcing					100%	J. Wilkin / J. Levi	1-Apr-11
5.3.5	Execute hydrodynamic model comparison for selected hindcast/nowcast/forecast periods (run ROMS, HOPS, COAWST, RTOFS, HyCOM-NCODA, NCOM, Mercator)					100%	J. Wilkin / J. Levi	
5.3.6	Perform Model Skill Assessment					100%	J. Wilkin	
5.3.7	Provide agencies with model error bound estimates						J. Wilkin	
5.3.8	Provide agencies recommendations for ensemble model transition to operations						J. Wilkin	
5.4	*Assess CB hydrodynamic model intercomparison*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complet	Task Lead	Finish Date
5.4.1	Define hindcast/nowcast/ periods					100%	Distributed acro	1-Oct-10
5.4.2	Assemble common forcing data sources; including those down-scaled from basin-scale to MAB to CB forcing; identify data sets					100%	Distributed acro	1-Feb-11
5.4.3	Establish and refine metrics for model-model-data intercomparison					100%	Distributed acro	1-Mar-11
5.4.4	Execute hydrodynamic model comparison for selected hindcast/nowcast/forecast periods (run ROMS, ChesROMS, CBOFS2, CH3D, EFDC, SELFE,FVCOM)					100%	Distributed acro	
5.4.5	Perform formal model skill assessment					100%	M. Friedrichs / A	
5.4.6	Provide agencies with model error bound estimates					100%	M. Friedrichs/C.	
5.4.7	Provide agencies recommendations for ensemble model transition to operations					100%	M. Friedrichs/C.	
5.5	*Assess CB hypoxia model intercomparison*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	% Complet	Task Lead	Finish Date
5.5.1	Define hindcast periods					100%	Distributed acro	1-Oct-10
5.5.2	Assemble common forcing conditions, including output from above CB hydrodynamics models; identify data sets for model					100%	Distributed acro	1-Feb-11
5.5.3	Establish and refine metrics for model-model-data intercomparison					100%	Distributed acro	1-Mar-11
5.5.4	Execute hypoxia model comparison for selected hindcast periods (run ChesROMS-NPZD, ChesROMS-ODU, EFDC, CE-QUAL-					100%	Distributed acro	
5.5.5	Perform Model Skill Assessment					100%	M. Friedrichs / A	
5.5.6	Provide agencies with model error bound estimates					100%	M. Friedrichs/C.	
5.5.7	Provide agencies with recommendations for ensemble model transition to operations					100%	M. Friedrichs/C.	

Prepared by Liz Smith (lsmith@sura.org).

3.2.2 Results and Accomplishments

Middle Atlantic Bight Hydrodynamic Model Comparison (5.3 above)
The Rutgers group was tasked by the IOOS Modeling Testbed to provide a skill assessment of the seven circulation models that run operationally for the MAB and have output available via web services. These models are: global HyCOM/NCODA (NRL), global NCOM (NAVO), COAWST (USGS), Rutgers ESPreSSO, NYHOPS (Stevens Tech.), UMassDHOPS (U. Mass.) and Mercator (EU). Model-data comparisons were computed with respect to (1) CTD data from gliders (8 deployments in 2010) , (2) NOAA ECOMON cruises, and (3) CODAR surface currents from the Rutgers regional node of the national HF-Radar network. Extensive assessments were made and are being assembled for a peer-reviewed journal article. Examples of the character of these quantitative skill assessments are presented below.

Figure 1 shows assessments of model temperature using glider data from March 2010 and ECOMOM dataset from Oct 2010. The Taylor diagrams depict centered RMS error as distance from OBS to MODEL (green contours), correlation as azimuth angle (blue lines) and standard deviation error (radius). Mean bias is tabulated. Values are normalized by data variance. No single model emerges as consistently performing best across all metrics (RMSE, correlation, bias, or glider vs. ship data) or across all CTD deployments. The analysis has been refined to ensembles of observations grouped according to (i) ocean depth (divided at approximately the 40 m isobath), (ii) along-shelf location (to distinguish the environs of the mouth of the 3 major estuaries of the MAB), and (iii) season. The analysis is being extended to include glider datasets through January 2012 as final figures are produced for a paper in preparation. Again, no single “best” model emerges, though we can recommend some models with respect to certain metrics.

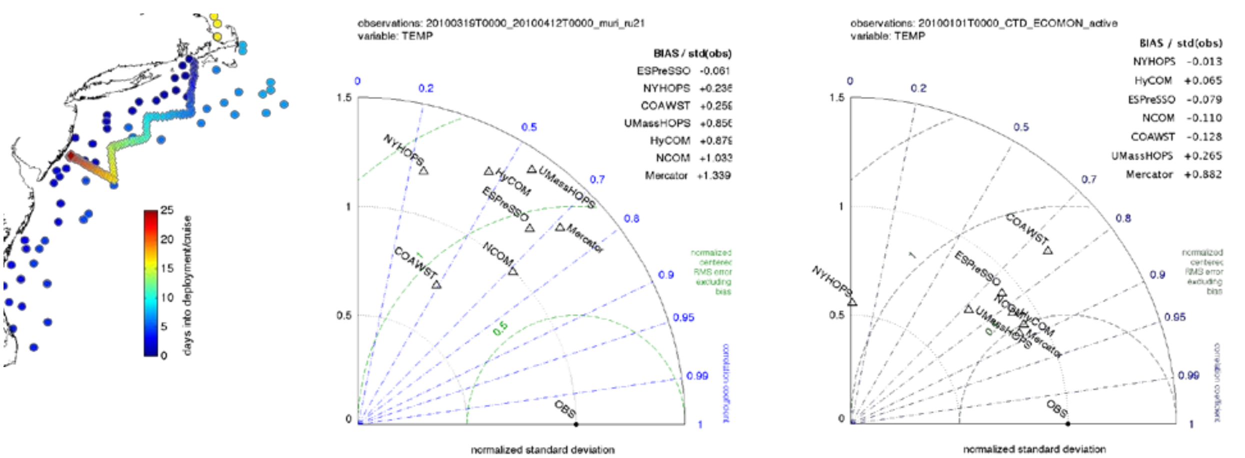


Figure 1: Left: Data locations for glider and ECOMON data. Taylor diagrams assess 7 mesoscale models covering the MAB. Center: glider transect temperature. Right: ECOMON cruise temperature.

Models of estuarine circulation need information not only on temperature and salinity at their inner shelf open boundaries, but also velocity and sea level. Tidal harmonic variability and local wind forcing dominate high-frequency dynamics at these open boundaries for the Chesapeake and Delaware, but sub-inertial frequency flow at the estuary may be also be significant to estuary-shelf exchange processes and therefore the shelf-wide operational models offer information on this aspect of the open ocean state. Accordingly, we have also assessed the seven operational models with respect to how well they reproduce daily variability, and annual mean surface currents in the MAB as observed by the HF-Radar network.

As an example, Figure 2 shows results for the 3 regional models supported by IOOS through the MARACOOS regional observatory effort. Comparisons for the other 4 models have been completed but are not shown.

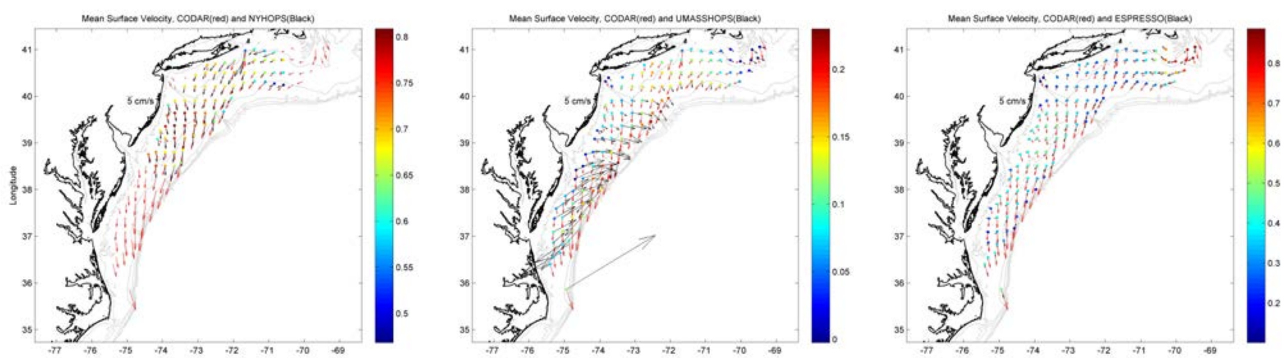


Figure 2: Comparison of annual mean CODAR velocity (red vectors) to 3 MARACOOS models (black vectors). Left: NYHOPS. Center: UMassHOPS. Right: RU-ESPreSSO. Colored dots show vector correlation of daily velocity variance (note change in color scale between panels for the correlations).

At the mouth of the Chesapeake estuary, the 2010 CODAR data set is relatively sparse compared to elsewhere in the MARACOOS radar network, and the radar data have not passed our quality control for the inner shelf. However, we have extended our analysis to include all CODAR data in 2011 and are reporting on these comparisons in the skill assessment paper in preparation.

Provide agencies with model error bound estimates (5.3.7 above)

The skill of RU-ROMS(Espresso), Stevens-NYHOPS, UMassD-HOPS, USGS-COAWST, HyCOM, NCOM and Mercator has been quantified with respect to CODAR surface currents and CTD data from gliders and ECOMON cruises in 2010-2011 in terms mean bias, correlation, and standard deviation error. These quantities can be depicted graphically on Taylor and target diagrams, and form the basis of a paper on the quantitative skill assessment in preparation for the Testbed special issue of JGR.

Considering glider and Ecomon data sets separately, temperature and salinity separately, and performing the comparison separately for 4 subregions that divide the inner from outer shelf, and north/south of Atlantic City, no single model emerges as consistently more skillful.
Efforts to derive statistical significance tests for the skill metrics is the final research task being undertaken to complete the draft publication. Certain models do come out as achieving the best skill score more frequently than others. For the inner shelf sub-regions that would provide boundary conditions to local models of the Chesapeake, Delaware, Hudson/Raritan and Narragansett estuaries, these models are NCOM, Espresso, NYHOPS, in no particular order.

Provide agencies recommendations for ensemble model transition to operations (5.3.8 above)
Until it can be shown that there are sufficient in situ data for statistical significance tests to distinguish the skill of individual models from each other, we cannot envision an objective process by which the weights of a multi-model ensemble could be derived. As this time we do not recommend the compilation of an ensemble model as an operational strategy, though multiple models could be used to derive an estimate of ensemble model spread to place error bounds on an ensemble mean. But we cannot rule out that the ensemble mean may still be inferior to the most skillful individual model. This requires further work to quantify.

Assess Chesapeake Bay Hydrodynamic Model Comparisons (5.4 above)
1 Define hindcast periods (100% accomplished)
Hindcast period is a two-year period including 2004 and 2005.

2 Assemble forcing and data sets for skill assessment (100% accomplished)
Common forcing data were assembled and are available for future modelers at:
ftp://ocsftp.ncd.noaa.gov/Lyon/CBayMEE/

Chesapeake Bay Program station data (temperature, salinity and dissolved oxygen) are used for skill assessment.

3 Establish metrics for model-comparison (100% accomplished) The metrics used for model comparison include bias, root mean squared difference (RMSD) and unbiased RMSD. In addition, a novel Frechet Distance metric was also introduced.

4 Execute comparison (100% accomplished)
Throughout the duration of this project we have assessed 27 model runs in terms of their hydrodynamic skill and 29 models runs in terms of their dissolved oxygen skill.

For each model run, skill is computed for ten variables: bottom/surface oxygen, bottom/surface temperature, bottom/surface salinity, hypoxic volume (< 0.2 and < 2.0 mg/L), maximum vertical salinity gradient, and depth of this maximum vertical gradient.

5 Perform skill assessment (100% accomplished)
On seasonal time-scales, the models all do well in capturing fundamental aspects of the hydrodynamic and oxygen fields, although the intensity of the pycnocline is underestimated. Models with constant net respiration independent of nutrient supply reproduce hypoxia nearly as well as much more complex, nutrient-dependent ecological models. Seasonal variation in DO was insensitive to seasonal cycles in the respiration rate, freshwater input, and density stratification. Rather, seasonal variation in DO was found to be very sensitive to seasonal variations in wind, likely due to wind-induced lateral upwelling of hypoxic areas. The overall intensity of stratification and resulting hypoxic volume was also found to be sensitive to numerical formulations of turbulence closure and advection. Another significant finding with regards to future modeling strategies is the result that the ensemble hindcast for dissolved oxygen using multiple models was more accurate than the hindcast from any one model. All modelers, with the exception of those running the Chesapeake Bay Program model, which as a result of its regulatory status remains frozen, are continually adjusting parameters and formulations within their models and improving the model-data fit. The modelers have been in close communication throughout the project through weekly emails, phone conferences, workshops and national meetings.

As a result, the modelers have continued to learn from each other in regards to what types of model modifications help improve the skill of the models. As a result, we have documented a significant improvement in skill for the models participating in this Testbed project. For example, the ChesROMS-NPZD model has undergone > 40% reduction in RMSD during the time frame of this project.

6 Provide agencies with model error bound estimates (100% accomplished)

Throughout the project we have worked closely with managers and model experts at NCEP, NOAA and EPA and they have been continually informed of our model error bound estimates. The Chesapeake Bay Program Management Board has been particularly interested in our model comparison and skill assessment efforts, and as a result they are currently assessing the feasibility of implementing multiple management models in the Chesapeake Bay Program for regulation of nutrient TMDLs within the Chesapeake Bay. These feasibility assessments would not be occurring if it were not for the successes demonstrated by the EH Testbed team.

7 Provide agencies recommendations for model transition to operations (100% accomplished)

The EH Team has made very important progress in transitioning its model formulations and scientific insights for use by Federal Agencies. For example, Lyon Lanerolle, a modeler at NOAA-CSDL and a member of the Estuarine Hypoxia Team, has already incorporated within the research version of NOAA- CSDL's Chesapeake Bay Operational Forecast System (CBOFS) a promising hypoxia formulation developed by the EH Team. A meeting between lead PIs on the EH Team and NOAA operational forecasters (described more fully below) hammered out transition steps for moving a fully operational version of the CBOFS model, including hypoxia, to NCEP.

3.2.3 Communication and Dissemination

Results obtained from this project have formed the basis for one published paper[1], one submitted manuscript, and an additional 7 papers in preparation. In addition, EH team members have given more than 35 presentations on COMT results at numerous national and international conferences and workshops.

In addition, multiple correspondences to the Chesapeake Bay Program (CBP) managers urging them to consider using multiple models have directly resulted from our COMT activities[3,4], and a recent positive response from the CBP director[5] indicates that these are having an effect and significantly changing the direction of future CBP activities: "We expect to begin this Bay Program-wide consideration of the role of multiple models in April 2012… The EPA is now examining the potential to fund a few prototype shallow water models this year."[5]

Finally, four workshops have directly been inspired by EH model comparison results, and EH PIs have been essential to the planning and organization of all four workshops by participating in their respective steering committees:

"Chesapeake Bay Hydrodynamic Modeling", June 9-10, 2011, Edgewater, MD, 55 participants[2]

Multiple members of the EH team (M. Friedrichs, C. Friedrichs, K. Sellner, R. Hood, W. Long) served as steering committee members for this workshop, which included overview talks on hydrodynamic model comparisons, as well as presentations on specific hydrodynamic models that the CBP could consider as alternatives or supplements to the existing CBP model. Presentations on the sensitivities of these models to coastal boundary conditions and turbulence closure schemes were included as well. At the conclusion of the workshop, the steering committee identified five recommendations for future CBP hydrodynamic modeling efforts. These included the use of multiple open-source community models and formal skill assessment.

"Chesapeake Bay Eco-Forecasting Workshop II: Research to Applications", July 3-4, 2011, MD and DC, ~50 participants.

One of the objectives of this workshop, led by David Green (NWS), Doug Wilson (NOS/NCBO) and Chris Brown (NESDIS/STAR), was to provide an update on the transition to operation of prototype forecast products and services. An additional goal was to identify possible tactical approaches to operate and maintain a Chesapeake Bay ecological forecasting system, and workshop discussions focused largely around recent modeling results produced by the EH Testbed team.

"Multiple Management Models Pilot Project", Gloucester Point, VA, April 2012, ~25 participants

This workshop will be convened upon request of the Chesapeake Bay Program's Modeling Team to discuss: (1) How multiple estuarine models can be used in a regulatory environment, and (2) What model qualifications, model outcomes, and skill assessment would be necessary for a successful multiple hydrodynamic shallow water modeling pilot project within the Chesapeake Bay. This workshop is the result of face-to-face communications with Chesapeake Bay Program managers regarding recommendations to incorporate multiple models in future modeling efforts. Bay Program managers have requested that this workshop is conducted prior to May 31, 2012. EH Team lead Marjorie Friedrichs was asked to act as Chair of this upcoming workshop.

"Multiple Management Models in Regulatory Decisions", Washington, DC, tbd, July-October 2012, ~50-75 participants.

This workshop is the result of multiple conversations and correspondences with CBP management on the topic of using multiple management models. During these discussions, CBP managers made it clear that they would need to identify how multiple models could be used throughout the CBP's modeling suite before committing to the implementation of various models. By carefully considering case studies of other multiple modeling projects, and identifying the benefits and drawbacks related to each, the CBP sees this workshop as an essential first step in exploring the use of multiple models throughout the Bay Program's modeling suite.

[1] http://www.springerlink.com/content/rmhp6v712856133p/
[2] Workshop report available at:
http://www.chesapeake.org/pubs/257_Friedrichs2011.pdf
[3] http://www.chesapeake.org/pubs/279_ChrisPyke2011.pdf
[4] http://www.chesapeake.org/pubs/283_ChrisPyke2012.pdf
[5] http://www.chesapeake.org/pubs/267_NickDiPasquale2012.pdf

3.3 Shelf Hypoxia Modeling in the Northern Gulf of Mexico

Team Leads: John Harding, Katja Fennel, Rob Hetland, Jerry Wiggert

3.3.1 Project Summary

The long-term goal of the shelf hypoxia project is the evaluation, and transition to operations, of a coupled, biogeochemical/physical model capable of forecasting the real-time evolution of shelf ecosystem processes in the northern Gulf of Mexico.

The objectives of year 1 efforts were to:

(1) Expose the CI team to the cyberinfrastructure challenges of a case study to aid in their design and development of the super-regional Testbed intended to enhance academic/operational collaboration and transitions,
(2) Address the hypothesis that regional boundary conditions impact the initiation and evolution of synoptic scale shelf hypoxia events in the northern Gulf of Mexico,
(3) Provide a preliminary comparison of NOAA and EPA research approaches to synoptic scale shelf hypoxia prediction in the northern Gulf of Mexico,
(4) U.S. Navy implemented a *regional* circulation prediction system for the Gulf of Mexico and Caribbean as a baseline operational capability applicable to future planned shelf hypoxia prediction capabilities as well as relevant to real-time Coast Guard search and rescue operations, harmful algal bloom tracking, oil spill response applications, and other marine-related needs in the region.
(5) Transition an initial distribution capability for retrospective results of real-time operational ocean predictions for the Gulf of Mexico and elsewhere for future science and operational applications.

3.3.2 Results and Accomplishments

SHELF HYPOXIA APPROACH & MILESTONES		Proposal Milestones								
3.5	*Assess Hydrodynamic Skill*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	Jul to Sep	Oct to Dec	% Complete	Task Lead - Hetland	Finish Date
3.5.1	Compile/organize/store all physical data							100%	Hetland	31-May-11
3.5.2	Provide hindcast outputs (NOAA GOM POM, IASNFS, GoM HYCOM)							100%	Hetland/ Harding/ Ko/ Patchen	31-Mar-11
3.5.3	Couple H/ROMS to hindcast outputs							100%	Hetland/ Morey	31-Mar-11
3.5.4	Run Hindcast models (H/ROMS uncoupled, H/ROMS coupled to Hind cast outputs, nGOM FVCOM coupled to NGOM POM.)							100%	Hetland	30-Jun-11
3.5.5	Evaluate/Develop skill assessment tools (physical)							100%	Hetland/ Lipphardt	30-Jun-11
3.5.6	Evaluate/compare coupled runs (physical)							100%	Hetland	30-Sep-11
3.5.7	Final Analyses/ Summary Report							100%	Hetland	31-Dec-11
3.6	*Assess Ecosystem Model Skill*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	Jul to Sep	Oct to Dec	% Complete	Task Lead - Fennel	Finish Date
3.6.1	Compile/organize/store all biogeochemical data							100%	Fennel	30-Apr-11
3.6.2	Evaluate/Develop skill assessment tools (biogeochemical)							100%	Fennel/ Lipphardt	30-Jun-11
3.6.3	Run uncoupled and coupled hindcast physical models with hypoxia code included							100%	Fennel	31-Jul-11
3.6.4	Evaluate/compare coupled runs (biogeochemical)							100%	Fennel	30-Sep-11
3.6.5	Run EPA hypoxia code coupled to IASNFS NCOM							100%	Ko	30-Jun-11
3.6.6	Evaluate/compare Hindcast H/ROMS coupled to IASNFS NCOM (biochemical)							100%	Fennel/ Ko/CI Team	31-Oct-11
3.6.7	Final Analyses/ Summary Report							100%	Fennel	31-Dec-11
3.7	*Evaluate transition NAVO AMSEAS NCOM*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	Jul to Sep	Oct to Dec	% Complete	Task Lead - Wiggert	Finish Date
3.7.1	Identify real-time evaluation data (physical ocean & atmos.)							100%	Wiggert	31 Apr 11
3.7.2	Accumulate real-time evaluation data (physical ocean & atmos)							100%	Wiggert/ Bub	30-Jun-11
3.7.3	Provide 3 hrly, 3 km, Real-Time Forecast NAVO AMSEAS NCOM output on NGI/NCDDC EDAC OceanNOMADS							100%	Harding/ Bub/ Cross	Continuing
3.7.4	Provide NAVO OPTEST Evaluation Tools							100%	Wiggert/ Bub	31-Mar-11
3.7.5	Evaluate/develop additional real time & post-time skill assessment tools							100%	Wiggert/ Lipphardt/ Fitzpatrick	31-Aug-11
3.7.6	Evaluate NAVO AMSEAS							100%	Wiggert	31-Dec-11
3.7.7	Evaluate NAVO AMSEAS atmospheric forcing							100%	Fitzpatrick	31-Jul-11
3.7.8	Transition NAVO/AMSEAS from pre-operational to operational							100%	Wiggert/ Bub	30-Sep-11
3.7.9	Transition NGI Developmental OceanNOMADS to Operational NCDDC Ocean NOMADS							100%	Harding/Cross	30-Jun-11
3.7.10	Final Analyses/ Summary Report							100%	Wiggert/Harding	31-Dec-11
3.8	*Overall SURA Shelf Hypoxia Project*	July to Sep	Oct to Dec	Jan to Mar	Apr to Jun	Jul to Sep	Oct to Dec	% Complete	Task Lead - Harding	Finish Date
3.8.1	Consolidated Summary Report							100%	Harding/ Hetland/ Fennel/ Wiggert	31-Dec-11

3.3.2.1 Description of significant research results, protocols developed, and research transitions

1. Established direct collaboration between academic hypoxia researchers and NOAA CSDL operational Gulf of Mexico hypoxia model developers. **[Tasks 3…]**

2. Nested vs unnested physical simulations of the northern Gulf shelf show impact on horizontal salinity distributions resulting from the nesting. A strong spike in signal-to-noise ratio in summer likely results from strong, small-scale eddies formed on the edge of the Mississippi/ Atchafalaya river plume front. **[Tasks 3.5…]**

3. Created consolidated, error-checked, multi-year hypoxia data set and provided to SURA CI team and to NODC for future availability via NOAA Hypoxia Watch Data Portal. **[Task 3.6.1]**

4. While the nested physical simulations show improvements in salinity distributions compared to the unnested simulations, initial analyses of the shelf biogeochemical model nested within different physical Gulf models does not show definitive improvement in response to the physical boundary condition treatment. Given that the hypoxic layer is confined to the bottom few meters, the results from these comparisons do suggest that subsequent research needs to focus on the model representation of the bottom boundary layer (e.g. vertical resolution, diffusivity etc.) and the biogeochemical interaction between the bottom waters and the underlying sediment (especially sediment oxygen consumption). **[Tasks 3.6…]**

5. Specific Improvements to Models: Linking Sediment Transport and Biogeochemical Models within ROMS. **[Tasks 3.6…]**

6. Analysis of ROMS near bottom trajectories during June-July 2007 showed residence times > 90 days in some hypoxic areas and hypoxic water masses originating offshore, near the shelf break. **[Tasks 3.6…]**

7. ROMS simulations with realistic boundary conditions and instantaneous remineralization provide a more realistic inshore position of hypoxic area relative to EPA GEMS with comparable representation of hypoxic area size. The EPA GEMS hypoxic area and phytoplankton biomass are consistently too far offshore for each of the four years. ROMS with climatological boundary conditions and instantaneous remineralization gets the inshore location correct but appears to under-represent the size of hypoxic area relative to EPA GEMS. Future hypoxic zone area and phytoplankton biomass comparisons of ROMS, using the Hetland and DiMarco sediment oxygen consumption formulation, with EPA GEMS would be of interest as these ROMS simulations better represented the size of the hypoxic area even with climatological boundary conditions but, as with the EPA GEMS, tended to be too far offshore relative to the instantaneous remineralization cases. **[Tasks 3.6…]**

8. Supported the transition of output of the U.S. Navy operational Gulf of Mexico regional ocean nowcast/forecast system to NCDDC. **[Tasks 3.7…]**

9. Provided insight relevant to NOAA CSDL operational Gulf of Mexico coastal nowcast/forecast system developers. **[Tasks 3.7…]**

10. Restructured NCDDC/NGI developmental EDAC for retrospective OceanNOMADS capability (as Navy "White Front Door" for operational ocean nowcast/forecast products). **[Task 3.7.9]**

3.3.3 Communication and Dissemination

3.3.3.1 Publications

Brown, C. W., R. R. Hood, W. Long, D. L. Ramers, C. Wazniak, **J. Wiggert**, J. Xu, R. Murtugudde, M. D. Wilson, J. Jacobs and R. Wood, Ecological Forecasting In Chesapeake Bay Using A Hybrid Mechanisic-Empirical Modeling Approach, J. Mar. Sys., ***in prep***.

Fennel, K., Hetland, R., Feng, Y., DiMarco, S., 2011. A coupled physical-biological model of the Northern Gulf of Mexico shelf: Model description, validation and analysis of phytoplankton variability, Biogeosciences 8, 1881-1899, doi:10.5194/bg-8-1881-2011

Fennel, K., Hu, J., Laurent, A., Hetland, R., Marta-Almeida, M., Lehrter, J., Model predictions of hypoxic area on the Texas-Louisiana Shelf are highly sensitive to sediment oxygen consumption parameterizations and stratification. In preparation.

Harding, J., S. Cross, F.Bub, and M. Ji. OceanNOMADS: Real-time and retrospective access to operational U.S. ocean prediction products. Draft in preparation for *EOS*.

Marta-Almeida, M., R. D. Hetland, X. Zhang. Evaluation of model nesting performance on the Texas-Louisiana continental shelf. Draft in preparation for *Ocean Modeling*.

Wiggert, J. D., R. R. Hood, W. Long, J. Xu, L. W. J. Lanerolle, M. B. Prasad, and C. W. Brown, Assessment of Biophysical Variability in a Coupled Physical-Biogeochemical Model Developed to Enable Water Quality and Ecological Forecasts of Chesapeake Bay, J. Mar. Sys., *in prep*.

Xu, J., W. Long, **J. D. Wiggert**, L. W. J. Lannerolle, C. W. Brown, R. Murtugudde, and R. R. Hood, Climate Forcing and Salinity Variability in the Chesapeake Bay, USA. Estuaries and Coasts: DOI 10.1007/s12237-011-9423-5, 2011.

Zhang, X., **M. Marta-Almeida, R. D. Hetland**. An operational nowcast/forecast system for the Texas-Louisiana continental shelf. Draft in preparation for *Journal of Operational Oceanography*

3.3.3.2Presentations

Brown, C. W., R. R. Hood, W. Long, D. L. Ramers, C. Wazniak, **J. Wiggert**, J. Xu, R. Murtugudde, M. D. Wilson, J. Jacobs and R. Wood, Ecological Forecasting In Chesapeake Bay Using A Hybrid Mechanisic-Empirical Modeling Approach, Meeting: 2011 Advances in Marine Ecosystem Modelling Research, Plymouth, UK, 27-30 June 2011.

Brown, C. W., R. R. Hood, W. Long, D. L. Ramers, C. Wazniak, **J. Wiggert**, R. Murtugudde, M. B. Decker, D. Wilson, and J. Xu, Development of a Chesapeake Bay Ecological Prediction System, Meeting: 2011 Aquatic Sciences Meeting, ASLO, San Juan, Puerto Rico, 13-18 February 2011.

Brown, C. W., R. R. Hood, W. Long, D. L. Ramers, C. Wazniak, **J. Wiggert**, R. Murtugudde, Mary Beth Decker, G. C. de Magny, D. Wilson, and J. Xu, Development of a Chesapeake Bay Ecological Prediction System, Meeting: 2010 ICES Annual Science Conference, Nantes, France, 20-24 September 2010.

Cross, S., J. Harding, and A.R. Parsons, Retrospective access to operational U.S. ocean model products through the OceanNOMADS system. 2012 Ocean Sciences Meeting, Salt Lake City, Utah, February 20-24, 2012

Feng, Y., Jackson, G., DiMarco, S., **Hetland, R., Fennel, K.,** Friedrichs, M.A., Understanding hypoxic area variability in the Northern Gulf of Mexico from a three-dimensional coupled physical-biogeochemical model, 2012 Ocean Sciences Meeting, Salt Lake City, Utah, February 20-24, 2012

Fennel, K., Hu, J., Hetland, R., DiMarco, S., Simulating hypoxia on the Texas-Louisiana shelf in the northern Gulf of Mexico, CERF 21th Biennial Conference, Daytona Beach, Florida, November 6-10, 2011 (oral presentation)

Harding, J., OceanNOMADS: Transition of a NOAA ocean prediction archival capability**,** Presented at the Marine Technology Society TechSurge 2011: Oceans in Action, Biloxi, MS, 22-23 August 2011. (Invited)

Harding, J., Super-Regional Testbed for Improving Forecasts of Environmental Processes for the U.S. Atlantic and Gulf of Mexico Coasts: Shelf Hypoxia Progress/ Plans, Presented at the 2nd Annual Gulf of Mexico Research Coordination Workshop, Bay St. Louis, MS, 31 March – 1 April 2011.

Harding, J., C. Friedrichs, and R. Luettich, The Role of the SURA Testbed in the Improvement of U.S. Coastal and Estuarine Prediction, Presented at Coastal Zone 2011, Chicago, ILL, 17-21 July 2011.

Harding, J., K. Fennel, R. Hetland, and **J. Wiggert**, An overview of shelf hypoxia efforts in the SURA Super-Regional Modeling Testbed, To be presented at 10th Symposium on the Coastal Environment at the the 92nd Annual Meeting of the American Meteorology Society, New Orleans, LA, 22-26 January 2012.

Harding, J., K. Fennel, R. Hetland, and **J. Wiggert**, The SURA Super-Regional Modeling Testbed: an overview of the shelf hypoxia team's activities, Presented at the 2011 Northern Gulf Institute Annual Conference: An Earth Systems Approach to Northern Gulf Science and Management, Mobile, AL, 17-19 May 2011.

Harding, J., S. Cross, F.Bub, and M. Ji, OceanNOMADS: Real-time and retrospective access to operational U.S. ocean prediction products, Presented at the American Geophysical Union 2011 Fall Meeting, San Francisco, CA, 5-9 December 2011. (Invited)

Harris, C.K., J.J. Birchler, J.P. Rinehimer, C.R. Sherwood, and L. Sanford. Sediment transport models: putting sediment into biology, and biology into sediment. Chesapeake Bay Modeling Forum, VIMS, Gloucester Point, VA. May, 2011.

Harris, C., Coupling Sediment Transport and Biological Processes within a Numerical Ocean Model, Gordon Research Conference on Coastal Ocean Modeling, 26 June 2011, South Hadley, MA

Harris, C.K., J.J. Birchler, J.P. Rinehimer, C.R. Sherwood, and L. Sanford. Sediment transport models: putting sediment into biology, and biology into sediment. Poster presented at the Chesapeake Hydrodynamic Workshop, Edgewater, MD, June 9 – 10, 2011.

Harris, C. K., Fennel, K., Hetland, R., Wilson, R., Coupling Sediment Transport to Biogeochemical Processes: Effects of Resuspension on Oxygen Consumption, 2012 Ocean Sciences Meeting, February 20-24, Salt Lake City, Utah.

Hetland, R. D. Physical mechanisms that control seasonal hypoxia on the Texas-Louisiana continental shelf. Poster presented at the Gordon Research Conference on Coastal Ocean Modeling, Mt. Holyoke, NH, June 26 – July 1, 2011.

Hood, R. R., C. W. Brown, **J. D. Wiggert**, W. Long, J. Xu, R. Wood, J. Jacobs, M. B. K. Prasad, and L. W. J. Lanerolle, CBEFS: The Chesapeake Bay Ecological Forecasting System, Meeting: 2011 Aquatic Sciences Meeting, ASLO, San Juan, Puerto Rico, 13-18 February 2011.

Long, W., R. Hood, C. Friedrichs, M. Friedrichs, K. Sellner, **J. Wiggert**, J. Xu, and L. Lanerolle, The Community Operational Chesapeake Bay Regional Ocean Modeling System (ChesROMS) and Its Skill Assessment, Meeting: 92nd Annual Meeting of the American Meteorology Society, New Orleans, LA, 22-26 January 2012.

Marta-Almeida, M., R. D. Hetland, X. Zhang. Evaluation of model nesting performance on the Texas-Louisiana continental shelf. Poster presented at the Gordon Research Conference on Coastal Ocean Modeling, Mt. Holyoke, NH, June 26 – July 1, 2011.

Wiggert, J.D., J.M. Harding, F.L.Bub, P.J. Fitzpatrick, and K.C. Woodard, Evaluation of the AMSEAS Gulf of Mexico/Caribbean regional forecast system: A SURA Super-Regional Modeling Testbed Activity. 2012 Ocean Sciences Meeting, Salt Lake City, Utah, February 20-24, 2012.

Wiggert, J. D., R. R. Hood, W. Long, J. Xu, C. W. Brown, B. K. Mathukumalli, and L. W. J. Lanerolle, Assessment of the Biogeochemical Fields from a Coupled Physical-Biogeochemical Model used in Providing Water Quality and Ecological Forecasts of the Chesapeake Bay, Meeting: 2011 Advances in Marine Ecosystem Modelling Research, Plymouth, UK, 27-30 June 2011.

Wiggert, J. D., W. Long, J. Xu, R. R. Hood, B. K. Mathukumalli, L. W. J. Lanerolle, and C. W. Brown, Assessment of a Coupled Physical-Biogeochemical Model Developed for Water Quality and Ecological Forecast use in Chesapeake Bay, Meeting: 2011 Aquatic Sciences Meeting, ASLO, San

Juan, Puerto Rico, 13-18 February 2011.

3.4 Cyberinfrastructure

Team Leads: Eoin Howlett, ASA and Sara Graves, UAH

3.4.1 Background and Project Summary

The primary goal of the Testbed Cyberinfrastructure is to develop a unified search, access, analysis and visualization environment that allows scientists to work efficiently and powerfully with all IOOS model data. This involves developing standards, web services and standards-based tools that work across a variety of different model types and conventions, enabling interoperability and software reuse by following IOOS DMAC principles. Because the toolset needed by scientists to create complex and customized model/data assessment applications is largely the same toolset needed by developers to create complex and customized applications for decision makers and the public, all IOOS stakeholders (and the entire geosciences community) benefit from this foundational development. A secondary goal of the Testbed Cyberinfrastructure is to serve the specific needs of the scientists participating in the Testbed. This involves developing and maintaining the Testbed web sites, computing resources, and custom code that performs tasks such as format conversion and executing specific workflows. The Testbed made significant progress on both of the goals in year 1. All tools are available via the Testbed website at <http://Testbed.ioos.us>.

Foundational Infrastructure
Search: The Testbed CI team worked with the ncISO team led by Ted Habermann to establish a standard protocol that maps information from NetCDF model files to ISO metadata. The CI team also worked with the GI-CAT and ESRI Geoportal Server teams to establish a standard protocol that harvests the ISO metadata and delivers the results via OpenSearch and CSW.

Access: The Testbed CI team worked with an international group to establish the first standard for unstructured grid (triangular grid) model data. The team added an unstructured grid class to the Unidata NetCDF-Java (a.k.a. Common Data Model) library, allowing interoperable access to all four types of unstructured grid models currently used in IOOS. This interoperability was enabled in NCTOOLBOX, a standards-based toolbox for Matlab, developed jointly with the NSF Ocean Observatories Initiative (OOI), and the effectiveness proven by investigators on the Inundation team, who were able to perform model-model comparisons with just a few lines of non-model specific code. Observational data was encoded into NetCDF files with CF Conventions, and an OGC SOS service was developed to plug into Unidata's THREDDS Data Server, to allow access of time series data via SOS or OPeNDAP.

Analysis and Visualization: The CI team developed a prototype web tool to connect to the back end services for graphically browsing and comparing model results using OGC Web Map Services (WMS). The team also added functionality to the NCTOOLBOX for enabling extraction and subsampling based on lon/lat and date constraints, interpolation onto arbitrary vertical and horizontal sections, extraction of time series data and simulated glider tracks from the 4D model output. The IMEDS toolbox used by the inundation community to analyze and compare time series of water levels was enhanced to read standards-based (CF compliant) data.

Testbed Websites: The project home website http://Testbed.ioos.us was developed using Drupal, allowing flexible and easy customization. The site hosts discussion, documents, links to data and project-developed tools, and tracks the progress of the projects. A sister web site was also developed, Testbedapps.sura.org, which hosts much of the project data as well as the project services, such as the THREDDS Data Server and GI-CAT. This site will be maintained and can be easily expanded and further developed to support future Testbed activities and provide a central repository of code, tools, data access, and relevant information.

Specific Tools: Tools were developed to convert the NHC SLOSH model custom output files into CF-Compliant NetCDF files. This enables the SLOSH output to be analyzed with a wide array of interoperable tools. Tools were also developed to convert the Chesapeake Bay 20 year hindcast (CH3D) from custom data formats to compliant NetCDF, the initial work on this has been shared with the community and it is on-going.

3.4.2 Results and Accomplishments

As discussed, there were very specific solutions required to meet the needs of the Testbed as data and model results were generated, and longer term developments to serve the larger community.
Longer term solutions that can be, and are being, used by the community included implementation and testing of a number of community Opensource technologies and standards including :

- Thredds
- F-TDS
- ncWMS
- ERDDAP
- Ramadda
- GI-CAT
- ncISO
- ESRI GeoPortal
- OpenSearch
- NetCDF Java
- SOS
- IMEDS Skill Assessment
- Matlab
- Python

While the value of the implementation and experimentation of these tools and standards may not have been immediately clear to the modeling teams, there has been a wealth of information learnt about what works in a practical sense when it comes to use of these technologies for managing large and diverse data sets. A “hidden gem” of this Testbed activity that is not explicitly defined in the deliverables is an understanding of how the suite of community tools work, and where we face challenges. The Testbed project demonstrated that combining a dedicated team of CI experts with modelers allowed us to evaluate and build upon these technologies in a Testbed environment. These lessons can then be shared with the operational community.

Some examples of the findings include:

- Ramadda had potential as a solution for a catalog but had a small footprint in the community, needed some development, and had did not have full time developer support
- ERDDAP is an excellent tool for data transformation and delivery, but lacked support for SOS and unstructured grids in its current version
- ncWMS provides good support for visualization of scalar values in regular gridded data, but did not offer visualization of vector data – the Testbed did add this to ncWMS and is now being used by the community. ncWMS is valuable software but its performance may not be sufficient for very large datasets and it currently does not support unstructured grids
- F-TDS provided “regridding“ support, but was very difficult to deploy and configure
- ISO (and ncISO) provides a metadata standard and the project enabled adaptation so it could support complex model and observation data
- Catalog implementation was an incredible challenge to allow for management of model and observation data with a useful programmatic interface that supports faceted and unfaceted search

- We made modifications to NetCDF Java libraries to support unstructured grids, but we discovered that it is a complex development environment
- We started the project with a focus on the use of Matlab, but it became apparent that there were many Python users and we started a transition so we have a parallel Matlab/Python development path
- Very valuable experimentation running Matlab "as a service"
- TDS is a valuable technology but needs more developer support based on the "mission critical" status that it now holds for many data providers and agencies.

The team focused on the following tasks and made significant progress in Year 1 :

NCToolbox Development

- Refactoring/simplification of code
- Introduction of point object for methods based on points/collections of points/profiles
- Flexible grid subsetting code
- New methods to aid comparison of data, including unit conversions
- Delivered to the community and has been used for data access by scientists
- Coordination with the OOI-CI team where the toolbox is being used
- Implementing support for ugrids based on ugrid java
- Returns 4-d depth fields when model depths change with time as well as lat/lon
- Subsetting on orthodox CF-compliant files
- Successful implementation for Adcirc/Selfe comparisons
- Documentation updates
- Download the latest repository as a zip file instead of accessing source control repository.

Running Matlab on the server as a service

- Large breakthrough to configure Matlab as a server application
- Making use of nctoolbox for data interoperability
- Wrote a web service interface so clients, such as a web client) can run Matlab processes on the server – no desktop license required
- Exploring numerous applications where this concept can be leveraged
- Being utilized as a processing server for other projects alongside SURA
- Found way to legally create multiple instances of matlab within the matlab com server to assist with the thread lock performance issues, not sure if this will help or hurt performance yet
- Experimented with running similar server on Linux using pipes instead of com

NetCDF Java Library – Unstructured

- Implementation plan and test cases identified
- Connected to Matlab toolbox
- Time series extraction from UGRIDs by Lat/Lon point was completed for:
  - FVCOM
  - SELFE
  - ELCIRC
  - ADCIRC
- Subsetting by geographic extents

Web based Model Browser connected to Collaboration Site

- Draft Catalog interface implementation complete
- ncWMS configuration
- Ability to extract time series from model grids served by TDS
- Side by side model comparison visualization

IMEDS Support

- IMEDS posted to server
- IMEDS updated for user-selected output formats
- Added Target plots to output
- Added multiple Model/station data selection and preprocessor to IMEDS menu
- Added multiple Model/station comparison tool
    - Different data views incorporated to GUI (Table and Plot View)
    - Data filter options using model run/station name
    - Model/Obs Data Peak analysis plot

ncWMS Enhancements

- Coordination with Ethan Davis (Unidata) and Jon Blower (U. Reading)
- Added WMS 1.1.1 and 1.3.0 compliant time-rounding integrated into NcWMS. Returns rounded time in HTTP response header
- Integrated new code for vector display

Testbed Archive Data Management

- Installed monitoring and reporting tools
- Log notification services
- General software maintenance and support
- NCML generation and coordination with modelers to ensure data is compliant

Completion Status

Task	**Completion**
NCToolbox Development	100%
Server side processing/NetCDF Java Library - Unstructured	75%* Subsetting works, and integration with nctoolbox is 100% complete. Integration with other Unidata components needs to be addressed. Discussions that future work developments may be Python-based
Python Server side processing Tools	100%
Testbed Archive Server Support and Data Management	100%
Web based Model Browser connected to Collaboration Site	100%
IMEDS Support	100%
ncWMS Enhancements	100%
Testbed Archive Data Management	100%
Support Testbed Collaborative Portal	100%

Conclusions
The model Testbed CI team successfully completed tasks so that model and observation data generated by the modeling teams are now available in compliant standards on the Testbed server and accessible to the entire community using common technologies. Lessons learned about implementation challenges have been shared with the community and tools such as NCtoolbox and IMEDS are being used by the modeling teams to analyse and compare model results.

3.4.3 Testbed Collaborative Portal

Team Lead: Sara Graves, UAH

The main objective of the SURA IOOS Coastal Modeling Testbed is to develop and maintain a cyber-infrastructure to support the collaboration of coastal modelers and federal agency decision makers through the hosting, testing and analysis of coastal models for possible promotion from research to production. To reach this objective the Testbed's cyber-infrastructure is providing a **web-based environment supporting the collaboration** of participants, as well as designing and developing an underlying data management infrastructure to support analysis and skill assessment.

The Coastal Modeling Testbed portal [http://Testbed.sura.org] has been designed and developed to be the focal point for both the project participants and the modeling community to obtain project information and gain access to the Testbed archive datasets and tools. A public portion of the portal provides information about each of the Testbed's modeling efforts, information about the participants and a general overview of the project's goals. The collaboration portion of the portal is access-controlled to allow only Testbed team members to contribute content and utilize project tools. It currently contains all the publications, presentations, reports and related materials dealing with the Coastal Modeling Testbed.

The data management team has continued to develop and support portal functionality that directly supports the collaboration of Testbed participants and activities, as well as providing a public view of the Testbed to the modeling community. This activity has included the integration of Testbed tools in the portal for easy access by all participants. The team has worked closely with the overall CI team and the Testbed management team to insure the portal reliably provides the necessary functionality.

A data catalog search application has been developed as an example application that is integrated with the portal environment. Using the open search protocol/standard, the Data Search tool provides users with a way to graphically define their spatial area of interest and then search based on keywords, data names and temporal coverage. By utilizing the open search standards, the Data Search tool is able to interface with any data catalog that supports that search protocol. The results of the data searches provide download links to the identified datasets. Eventually, users will be able to save the search results in their own user-space in Testbed portal.