Final Report to

Project Number 08121-2801.02.FINAL

Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP)

July 31, 2014

Principal Investigator:
Christopher N. K. Mooers[1], Research Professor
cmooers@cecs.pdx.edu

Patrick J. Hogan[2], Research Scientist

Edward D. Zaron[1], Research Assistant Professor

[1] Department of Civil and Environmental Engineering, College of Engineering and Computer Science, Portland State University, P.O. Box 751, Portland, OR 97207-0751.

[2] Naval Research Laboratory, Stennis Space Center, MS 39529.

LEGAL NOTICE

This report was prepared by Christopher N. K. Mooers as an account of work sponsored by the Research Partnership to Secure Energy for America, RPSEA. Neither RPSEA, members of RPSEA, the National Energy Technology Laboratory, the U.S. Department of Energy, nor any person acting on behalf of any of the entities:

a. MAKES ANY WARRANTY OR REPRESENTATION, EXPRESS OR IMPLIED WITH RESPECT TO ACCURACY, COMPLETENESS, OR USEFULNESS OF THE INFORMATION CONTAINED IN THIS DOCUMENT, OR THAT THE USE OF ANY INFORMATION, APPARATUS, METHOD, OR PROCESS DISCLOSED IN THIS DOCUMENT MAY NOT INFRINGE PRIVATELY OWNED RIGHTS, OR

b. ASSUMES ANY LIABILITY WITH RESPECT TO THE USE OF, OR FOR ANY AND ALL DAMAGES RESULTING FROM THE USE OF, ANY INFORMATION, APPARATUS, METHOD, OR PROCESS DISCLOSED IN THIS DOCUMENT.

THIS IS A FINAL REPORT. THE DATA, CALCULATIONS, INFORMATION, CONCLUSIONS, AND/OR RECOMMENDATIONS REPORTED HEREIN ARE THE PROPERTY OF THE U.S. DEPARTMENT OF ENERGY.

REFERENCE TO TRADE NAMES OR SPECIFIC COMMERCIAL PRODUCTS, COMMODITIES, OR SERVICES IN THIS REPORT DOES NOT REPRESENT OR CONSTIITUTE AN ENDORSEMENT, RECOMMENDATION, OR FAVORING BY RPSEA OR ITS CONTRACTORS OF THE SPECIFIC COMMERCIAL PRODUCT, COMMODITY, OR SERVICE.

THIS PAGE INTENTIONALLY LEFT BLANK

Signature and Date Stamp

Christopher N. K. Mooers
Principal Investigator

31 July 2014
Date

Abstract

The GOMEX-PPP (Gulf of Mexico Pilot Prediction Project) has conducted the applied R&D necessary to implement, evaluate, and establish an operational prediction system for 3D ocean currents in the Gulf of Mexico. Such an operational forecast system is comprised of a numerical ocean circulation modeling subsystem, an ocean observing (satellite and in situ) subsystem with real-time components, a data assimilation subsystem for initializing the forecasts, and an information management subsystem for accessing, displaying, analyzing, and distributing the forecasts. The project was conducted in two phases. For the first phase (MME; multi-model ensemble), several state-of-the-science, mesoscale eddy-admitting baroclinic ocean circulation numerical models participated in a series of forecast experiments for assessment of their skill relative to science-based and applications-based metrics.

During the second phase (SME; single-model ensemble), an extended range (~60 days) ocean forecasting system based on probabilistic methods was implemented for the Gulf of Mexico. A probabilistic prediction has the advantage of providing both a forecast and an uncertainty estimate about that forecast. Both estimates provide information and guidance for a wide range of applications in the public and industrial sectors. Applications for the oil and gas industry are emphasized here. The forecast system generates probabilistic forecasts by perturbing the initial state of the ocean field as well as the surface atmospheric forcing (wind and heat flux fields). A multi-year time series of output from the Global Ocean Forecast System (GOFS 3.0) was used to construct a daily climatology for 3-D lateral boundary conditions associated with the inflow through the Yucatan Channel and outflow through the Straits of Florida. A similar method was used to construct surface boundary conditions from the Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) (Hodur, 1996) atmospheric forecast model implemented for the Intra-Americas Sea domain. Estimates of forecast uncertainty calculated over 32 ensemble members have demonstrated skill in predicting a Loop Current Eddy shedding event that occurred in the spring of 2013. The ensemble forecast system performs a 24-hour forecast every day and 60-day forecasts once per week and has been running quasi-operationally in real-time since January 2013.

Through various model-data analyses, the prediction system is nearly ready to proceed through pre-operational checkout at NAVO.

Table of Contents

Signature and Date Stamp **4**

Abstract **5**

Table of Contents **6**

List of Abbreviations and Acronyms **7**

Executive Summary **10**

Introduction **12**
- Background 12
- Objectives 12
- Scope of Work 13
- Project Organization 15
- Sponsors 15
- Participants 15
  - Principal Investigators 15
  - Chief Analyst 15
  - Phase I Sub-Contractors 15
  - Phase II Sub-Contractor 15
  - Affiliates 16
  - Prime Contractor's Technical Point of Contact 16
  - Prime Contractor's Contractual Point of Contact 16
  - Groups and Committees 16
- Timeline and Schedule 18

Review and Summary of Phase I **19**

Results and Accomplishments of GOMEX-PPP Phase II **22**

Discussion **24**
- Overview of 2013 Conditions 24
- Assessment of GOM-LREFS 25
- Comparison with Phase I Results 26

A Concept of Operations for Sustained Operational Long-Range Forecasts of the Gulf of Mexico **31**

Findings and Recommendations **32**

References **37**

Figures **42**

Appendix I: GOMEX-PPP Technology Transfer **65**

Appendix II: Concept of Operations **68**

Appendix III: Subcontractor's Final Report **71**

List of Abbreviations and Acronyms

ADCP	acoustic Doppler current profiler
AGU	American Geophysical Union
AMS	American Meteorological Society
AMSEAS	NCOM American Seas
ASLO	American Society of Limnology and Oceanography
AXBT	airborne expendable bathy-thermograph
BOEM	Bureau of Energy & Management
CASE-JIP	a consortium of offshore oil & gas companies
CCAR	Colorado Center for Astrodynamics Research
CONOPS	Concept-of-Operations
CPA	Closest-Point-of-Approach metric
CSDL	Coast Survey Development Laboratory/NOS
CSL	Coastal Sea Level
DAS	data assimilation system
DBDB2	NRL Digital Bathymetry Data Base with 2-minute resolution
DwH	Deepwater Horizon
DSRC	Department of Defense Shared Resource Center
EOF	Empirical orthogonal function
FNMOC	Fleet Numerical Meteorology and Oceanography Center
GCOOS RA	Gulf of Mexico Coastal Ocean Observing System Regional Association
GOMEX	Gulf of Mexico
GOMEX-PPP	Gulf of Mexico Pilot Prediction Project
GOM-LREFS	Gulf of Mexico Long Range Ensemble Forecasting System
GTS	Global Telecommunications System
HMI-EW	Horizon Marine Incorporated Eddy Watch
HYCOM	Hybrid Coordinate Ocean Model
IAS	Intra-Americas Sea
IASNFS	Intra-Americas Sea Nowcast/Forecast System
IOOS	Integrated Ocean Observing System program
IOP	Intensive Observing Period
JPL	Jet Propulsion Laboratory
LC	Loop Current
LCE	Loop Current (anticyclonic) Eddy (or, Ring)
LREFS	Long-Range Ensemble Forecast System
MAST	Modeling and Analysis Steering Team (MAST)/IOOS
MITGOM	Massachusetts Institute of Technology Gulf of Mexico ocean model
MME	Multi-Model Ensemble
MMS	Minerals Management Service; relevant portions now in BOEM
MODAS	Modular Oceanographic Data Assimilation System
MODIS	Moderate Resolution Imaging Spectrometer
NAVO	Naval Oceanographic Office
NCEP	National Centers for Environmental Prediction/NWS
NCOM	Navy Coastal Ocean Model
NCSU	North Carolina State University

NGI	Northern Gulf Institute
NIMS	Near-Inertial Motions
NMFS	National Marine Fisheries Service/NOAA
NOAA	National Oceanic and Atmospheric Administration
NOGAPS	Naval Oceanography Global Atmospheric Prediction System
NOMADS	NOAA National Operational Model Archive and Distribution System
NOS	National Ocean Service/NOAA
NRL	Naval Research Laboratory
NWS	National Weather Service/NOAA
OCS	Outer Continental Shelf
PMG	Project Management Group
POM	Princeton Ocean Model
PPP	Pilot Prediction System
PROFS	Princeton Ocean Forecast System
PSU	Portland State University
PU	Princeton University
R & D	Research and Development
RADS	Radar Altimeter Database System
RFP	Request For Proposal
RMSE	Root-Mean-Square Error
ROMS	Regional Ocean Modeling System
RPSEA	Research Partnership to Secure Energy for America, a consortium of oil and gas companies and universities
RTOFS	Real-time Ocean Forecast System
SAC	Scientific Advisory Committee, broad community-based advisory committee provided by the project
SAP	Surface Atmospheric Pressure
SEACOOS	Southeast Atlantic Coastal Ocean Observing System
SECOO-RA	Southeast Coastal Ocean Observing System Regional Association
SIO	Scripps Institution of Oceanography
SF	Straits of Florida
SME	Single Model Ensemble
SOOP	Ships of Opportunity Program
SOW	Statement of Work
SSH	Sea Surface Height
SSS	Sea Surface Salinity
SST	Sea Surface Temperature
SURA	Southeastern Universities Research Association
TAC	Technical Advisory Committee, RPSEA-based project oversight committee
TAMU	Texas A&M University
TOS	The Oceanography Society
UCLA	University of California at Los Angeles
UDW	Ultra Deepwater Program
URL	Uniform Resource Locator

USM	University of Southern Mississippi
USCG	United States Coast Guard
YC	Yucatan Channel

Executive Summary

The Gulf of Mexico Pilot Prediction Project (GOMEX-PPP) was undertaken to evaluate several prototype ocean forecast systems for the Gulf of Mexico, with the view towards implementing an operational forecast system in the future. The project was motivated by the following: (1) the offshore industry's need for forecasts of ocean currents for safety and efficiency of routine operations, (2) the environmental and emergency managers' needs for forecasts of ocean currents which are a prerequisite to physically-based ecological forecasts, and (3) the Navy and NOAA needs for forecasts of ocean currents to support many of their numerous mandated missions. The GOMEX-PPP approach consisted of two phases: (Phase I) the evaluation and skill assessment of several extant models, and (Phase II) a real-time demonstration of a pre-operational, single model ensemble system.

GOMEX-PPP Phase I was conducted in the style of a model test bed in which a suite of models were evaluated in a series of hindcast and forecast experiments designed by consensus of the participants who represented four academic-research models, two U.S. Navy pre-operational models, one NOAA pre-operational model, and one NOAA operational model. The experiments were performed in the following three steps: (Step 1) a year-long hindcast (i.e., a retrospective nowcast or "analysis" for 2010); (Step 2) a series of 3-month retrospective forecasts, conducted monthly, for 2010; and (Step 3) a series of 3-month forecasts, performed in real-time, for three months in 2011 and 2012. The experiments provided the setting for evaluating model nowcast skill and forecast skill, and for evaluating the availability of appropriate environmental data for routine operational initialization of models and systematic verification and validation of their outputs.

Deliverables of GOMEX-PPP Phase I included the following: (1) an evaluation of the individual multiple data-assimilative forecast systems; (2) an assessment of multi-model ensemble forecasts; (3) a demonstration of a suite of prototype mesoscale eddy-admitting ocean prediction systems; (4) the recommendation of a Concept of Operations (CONOPS) for the operational forecast system to be implemented in GOMEX-PPP Phase II; and (5) a recommended demonstration, upgrading, and evaluation program for Phase II.

During GOMEX-PPP Phase II, the Gulf of Mexico – Long Range Ensemble Forecasting System (GOM-LREFS) was implemented and delivered 2-month GOMEX forecasts weekly, in the real-time, demonstration, pre-operational mode. Deliverables of Phase II included the following: (1) setup and configuration of the Relocatable Navy Coastal Ocean Model (R-NCOM) and NCODA assimilation and initialization system with a 3km horizontal resolution and 49 vertical levels that covers the entire Gulf of Mexico; (2) setup of a 32-member ensemble of R-NCOM utilizing the ensemble transform (ET) technique to generate realistic forecast spreads; (3) ensemble calibration through Bayesian averaging was investigated to maximize forecast accuracy by penalizing members with less forecasting skill, benchmarked against recent past observations; (5) GOM-LREFS prototype demonstration was conducted by running the ensemble forecasting system in real-time/pre-operational mode from January 2013 to the present; (6) the GOM-LREFS was run in hindcast mode for the 2010-2011 period for the purpose of validation and comparison with GOMEX-PPP Phase I; (7) a Website was developed for project participants to access model forecasts and view new products derived from the ensemble forecasts; (8) a

Concept of Operations (CONOPS) was proposed for the transition of GOM-LREFS to sustained operational implementation.

Introduction

Background

Though much effort has been invested over the past decade to develop real-time ocean modeling systems in the Gulf of Mexico (GOMEX; see List of Abbreviations and Acronyms), forecasting of Loop Current/eddies (LC/LCE) remains problematic. Meanwhile, the Offshore Oil & Gas Industry's need to forecast the associated currents, which can extend over the full water column, has become even stronger as the Industry ventures further south in the GOMEX and into so-called ultra-deep water. While there are many existing models, none have been demonstrated to perform at the level of accuracy required. Besides the space-time-varying LC/LCE position and strength, other current forecast applications include guiding (1) oil spill response, (2) potential shallow water operations involving hypoxia and produced water, and (3) search-and-rescue.

The Industry loses millions of dollars every year due to drill-rig downtime caused by the presence of LC/LCE. Perhaps more importantly, there have been numerous "near-misses" documented by the Minerals Management Service (MMS) in which rig operators faced potentially serious issues due to strong currents. Accurate forecasts could help operators anticipate and avoid some of these conditions and events through better planning and operations, and avoid potentially dangerous surprises.

Objectives

The Project is focused on evaluating (Phase I; 24 mos) several candidate modeling systems and demonstrating (Phase II; 18 mos) a real-time Pilot Prediction System for the mesoscale (baroclinic) circulation of the Gulf of Mexico.

The general objective is to demonstrate a well-validated operational 3-D modeling system that produces timely and accurate forecasts, nowcasts, and hindcasts of currents across the GOMEX. The aim is to have sufficient accuracy to be useful for a number of applications including LC/LCE forecasts, oil-spill trajectory forecasts, and similar current-dependent forecasts. The numerical products of the modeling system are to be Web-based so that they are, thus, available/accessible to the public and provide substantial benefits for many well-informed users.

More specifically the objectives of this project are to:

- Identify the needs of the user community for various types of prediction outputs.
- Test various methods for quantifying modeling system errors, with special attention on quantifying the errors of relevance to the end users.
- Establish model metrics that will accurately gauge the ability of the modeling system to meet users' requirements.
- Apply model metrics on a routine basis for continuous monitoring of the modeling system performance to help improve the modeling system.
- Better utilize and synthesize on-going observations through data assimilation.

- Develop an archive of hindcasts that can be used for climatological studies (eventually, climate change detection) and diagnostic studies of physical and ecological ocean dynamics in the GOMEX.
- Quantify improvements made by single model ensemble (SME) and multi-model ensemble (MME) forecasts compared to single model forecasts.
- Make the prediction system results and ancillary tools easily accessible (Web-based) and usable by subject matter experts.

After several years of planning GOMEX-PPP, fortuitously and independently, the Project commenced in the same timeframe as the occurrence of the Deepwater Horizon (DwH) incident (20 APR 2010 and beyond) but with no direct connection to it. The Project was conceived a few years earlier as a partnership between RPSEA and the GCOOS-RA. The focus was on forecasting the position of the Loop Current (LC) and the large (ca. 200 km) anticyclonic Loop Current Eddies (LCEs) it sheds every 6-to-24 months, the subsidiary, small (ca. 20km) cyclonic (frontal) eddies (LFEs), the response to tropical cyclones, the continental shelf transient circulation, and their interactions, for example, through entrainment of shelf waters by LCEs and detrainment of LFE waters by shelf flows.

To include the interests of a broad base of potential users, PPP has evaluated Pilot Prediction Systems for their ability to meet the environmental prediction needs of both the offshore oil & gas industry and the GCOOS-RA community; i.e., the large, diverse GOMEX community of "super-users" and "end-users" that consume synoptic environmental information in the spirit of IOOS. The offshore oil & gas industry has several needs for environmental predictions but the primary need is for forecasts of the LC's position and strength and the timing of LCE formation. The prediction horizon desired by Industry is several months. The LC transport and position are known to have, on average, a seasonal cycle comprised of broadband annual and semi-annual components, which suggests the LC & LCE system may have useful predictability at the time scale of a few months. The GCOOS-RA community has several needs, too, including Lagrangian trajectory estimates used in search-and-rescue operations, dispersal studies of fish eggs & larvae, and calculations of transport pathways and dispersal rates for oil & other contaminant spills, especially over the continental margin. It also needs open boundary conditions for down-scaling GOMEX-scale circulation information to relatively high-resolution shelf & estuarine circulation models.

Scope of Work

The RFP from RPSEA states that the primary project aim is to establish an operational prediction (hindcasting, nowcasting, and forecasting) system for strong currents associated with the Gulf of Mexico (GOMEX) Loop Current (LC) and eddy (LCE) system in support of the oil & gas industry's southward extension of its activities into "ultra deepwater". A full water column capability is required, especially for surface and bottom currents, including those along the lower continental slope and continental rise. Secondary applications include guidance for oil spill response and potential shallow water operations involving hypoxia and "produced water". The offshore oil industry is interested in a forecast horizon of a few months for LC & LCE positions, etc. Past industry experience with forecasting LC & LCE positions has indicated a high degree of dependency on the initial conditions, and, by implication, open boundary conditions. Thus,

attention must be given to assessing the quality of data-assimilative model analyses (or nowcasts) used for initial conditions.

PPP proceeded without the support of IOOS funding at the present time while still trying to assist and advance the efforts of GCOOS-RA & SECOORA. Collaborative relationships with BOEM's ongoing GOMEX modeling and observational activities are anticipated. The RFP calls for two phases to the project: Phase I (R&D and Selection Project; initial 24 mos) is comprised of two competitive experiments, and Phase II (Demonstration Project; final 18 mos) is comprised of pilot operational prediction system implementation and assessment. An early step was to firm-up "user requirements" for a deep-water current prediction system that will have a continental shelf capability, and which, in addition to the oil and gas industry, may include, for example, marine transportation, maritime safety, and commercial and recreational fisheries users. The RFP called for a user survey; GCOOS-RA and SECOORA were asked to help in this regard. The "user requirements" helped establish the skill assessment metrics (e.g., Oey et al. (2005)).

The notion of an "operational current prediction system" is not much developed in the RFP, other than that a Website will be established (to be provided by GCOOS-RA) and kept fresh with forecast products that are meaningful to knowledgeable users. It is not made clear whether the aim is to establish, for example, a federally operated prediction system for the GOMEX as a public service or a for-profit private sector system funded by the offshore oil and gas industry (and also operated as a public service) or a hybrid. An example of a hybrid system would be for the federal operational center to provide an operational analysis (nowcast) of the GOMEX on an hourly basis that would be used for initial conditions in a private or academic sector weekly LC & LCE prediction system with a forecast horizon of 3 mos. The RFP allows for the possibility of using multiple models to form ensemble forecasts. In short, a Concept–of-Operations (CONOPS) does not appear in the RFP, so it was intended for the CWG (a subcommittee of the SAC) to be asked to recommend viable alternatives. As the project developed, the CONOPS design was internalized by PSU, NRL, and NAVO.

Another early step was to identify a "historical or target year" in which interesting LC & LCE events occurred, and for which relatively abundant in situ and remote sensing observational datasets exist. The occurrence of a Northern Gulf land-falling hurricane during the target year would be a bonus attribute. For the first experiment, a forecast was to be started at the beginning of each month of the target year and run for a 3-mos period; altogether, there would be nine three-month forecasts to be skill assessed against GOMEX analyses. (In other words, this activity would serve as something of a de facto predictability experiment for GOMEX.) For the second experiment, there would be a 3-mos forecasting test in the blind for a future period, again to be skill assessed against GOMEX analyses. The results of these experiments were to be the main determinants in selecting the model (s) for Phase II.

Operational, quasi-operational, and research modelers agreed to participate with their models. For example, the Navy's Global-NCOM and NOAA's RTOFS-Atlantic – HYCOM operational model output data were made available. Navy and NOAA were welcome to participate in the multi-week R&D forecast experiments of Phase I and the prototype multi-week operational forecasts of Phase II. Leo Oey (PU/MMS contractor and POM modeler), Yi Chao (JPL & UCLA/DAS developer and ROMS modeler) plus Ruoying He (NCSU and SECOORA/quasi-

operational GOMEX ROMS modeler), Dong-Shan Ko (NRL/quasi-operational GOMEX Regional NCOM developer), and Richard Patchen (CSDL/operational GOMEX POM developer) agreed to participate in Phase I (R&D and Selection Project) and compete for Phase II (Demonstration Project). [NOTE: to participate in the proposed project, all of these modelers, of necessity, leveraged their ongoing R&D activities. Also, all of these models already cover the GOMEX Outer Continental Shelf (OCS).] Ann Jochens, Matt Howard, and Steve DiMarco, TAMU were engaged to help with accessing the observational databases, archiving model output, providing a Web portal, and conducting skill assessments, together with Ed Zaron and Chris Mooers, PSU.

To disseminate results and accomplishments of GOMEX-PPP, participants made numerous presentations at scientific meetings and prepared manuscripts for submission to peer-reviewed journals. The activities, which may be described as Technology Transfer, are summarized in Appendix I.

Project Organization

Sponsors

GOMEX-PPP was sponsored by DOE via RPSEA (Research Partnership to Secure Energy for America, a joint offshore oil & gas industry and university consortium), and CASE-JIP (an offshore oil & gas industry consortium, represented by Cort Cooper, Chevron, and Dave Driver, BP America). RPSEA provides the majority of project funding, and CASE-JIP provides the requisite matching funds (namely, 20% of total project costs).

Participants

Principal Investigators
Dr. Christopher N. K. Mooers, Portland State University (PI)
Dr. Cortis Cooper, CHEVRON (CASE-JIP Co-PI)
Mr. David Driver, BP America (CASE-JIP Co-PI)

Chief Analyst
Dr. Edward D. Zaron, Portland State University

Phase I Sub-Contractors
Dr. Yi Chao, Jet Propulsion Laboratory & UCLA
Dr. Ruoying He, North Carolina State University
Dr. Matthew Howard, Texas A&M University
Dr. Dong-Shan Ko, Naval Research Laboratory
Dr. Leo Oey, Princeton University

Phase II Sub-Contractor
Dr. Patrick J. Hogan, Naval research Laboratory

Affiliates
Dr. Frank Bub, U.S. Navy, Naval Oceanographic Office/Ocean Modeling Division, now retired
Dr. Bruce Cornuelle, Scripps Institution of Oceanography
Dr. Avichal Mehra, NOAA/NWS/NCEP/EMC Marine Modeling and Analysis
Mr. Richard C. Patchen, NOAA/NOS/OCS Coast Survey Development Laboratory (CSDL), now retired

Prime Contractor's Technical Point of Contact
Dr. Christopher N. K. Mooers (503) 954-2772 cmooers@cecs.pdx.edu

Prime Contractor's Contractual Point of Contact
Kathleen Choi (503) 725-9652 spa.green@pdx.edu
DUNS Entity Number: 05-222-6800

Groups and Committees
Several groups and standing committees were formed to review GOMEX-PPP progress, and to envision the form and function of a future operational forecast system. Group and committee members represent GOMEX-PPP participants; industry, academic, and governmental stakeholders; and subject matter experts on GOMEX circulation, observations, modeling, and analysis.

Project Management Group (PMG)
Chris Mooers, PSU, PI/SAC Chair
Cort Cooper, Chevron, Co-PI/TAC Co-Chair
Dave Driver, BP America, Co-PI/TAC Co-Chair

Scientific Advisory Committee
A Scientific Advisory Committee (SAC) was formed to (1) review the plans, progress, and prospects of the project; (2) build a broad consensus on the skill of the forecast systems; and (3) recommend a Concept-of-Operations (CONOPS), which defines stakeholders, user requirements, roles and responsibilities, etc. for transitioning the pilot forecasting system to sustained operations. The SAC is responsible for reviewing and advising the PI and Co-PIs on the progress and plans of the GOMEX Forecasting Project. For this purpose, the SAC received progress reports and other documents to review for comment. The SAC was asked to attend the three one-and-one-half-day meetings/workshops planned, as separately described, over the course of the original 30 mo.-project. Several SAC members with operational ocean forecasting roles were asked to serve on a SAC subcommittee called the Concept-of-Operations Development Working Group, as described below.

(Academia/Research)
John Allen, OSU
Bill Schmitz, TAMU-CC
Bob Weisberg, USF & SECOORA
Bob Leben, CU
Nan Walker, LSU
Alan Blumberg, Stevens Institute and consultant to BP

Patrick Hogan, NAVO/USN
Steve Payne, CNMOC/USN
Gregg Jacobs, NRL/USN
Frank Bub, NAVO, now retired
Hendrik Tolman, OMB/NCEP/NWS/NOAA
Frank Aikman, CSDL/NOS/NOAA
Alexis-Lugo Fernandez, BOEM

(Industry)
Michael Vogel, Shell
Robert "Buzz" Martin, Shell
Steve Anderson, Arete Associates
Sergei Frolov, Weather Predict Consulting Inc. (WPC)

TAC CONOPS Development Working Group (CWG)
The CWG, a subcommittee of the SAC, considered the design options for a Concept-of-Operations (CONOPS) for a Gulf of Mexico operational ocean prediction system. The options ranged from a government service providing a "public good," to a non-profit academic service providing a "public good," to a for-profit commercial service providing a "proprietary good," or some hybrid. The design considered user needs; modeling subsystem, observational subsystem, computational, telecommunication, skill assessment, performance assessment, and other technical requirements; and personnel, budgetary, management, administrative, and operational requirements. The management requirements include consideration of the roles and responsibilities of the entities involved as stakeholders, sponsors, operational prediction service providers, etc. Recommendations of the CWG were vetted by the SAC as a whole and forwarded to TAC and others for further consideration.

Cort Cooper, Co-Chair
Dave Driver, Co-Chair
Steve Payne
Frank Bub
Hendrik Tolman
Frank Aikman
Robert "Buzz" Martin
Chris Mooers, ex-officio

Note: CWG was activated with a smaller group of participants than originally planned.

Timeline and Schedule

December 2010 – February 2012: Phase I Evaluation of Multiple Models.
February 2012: Workshop for subcontractors, SAC, and RPSEA PM to evaluate Phase I and make recommendations for Phase II.
March 2012: Brief RPSEA-TAC on Phase I results and Phase II recommendations to RPSEA.
March 2012 – December 2012: Reconcile Phase I subcontracts and settle the Phase II subcontract.
January 2013 – February 2014: Conduct Phase II Demonstration Project.

Project milestones and progress are summarized on the Gantt chart, below. Note that dates reflect the planned rather than actual project schedule, because the transition from Phase I to Phase II was more complicated than anticipated and necessitated project re-design and a No Cost Extension.

TASK	Owner	Jul '10	Oct '10	Jan '11	Apr '11	Jul '11	Oct '11	Jan '12	Apr '12	Jul '12	Oct '12	Jan '13	Apr '13
1. Project Management Plan & "Kick-Off" Mtg	PSU	x											
2. Technology Status Assessment	PSU	x											
3. Technology Transfer Plan	PSU	x											
4. Monthly Reports	PSU	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx
5. Identify user needs	PSU	xxx	xxx										
5.1 Gather input from users				xx									
5.2 Develop model performance specs and evaluation criteria	PSU	xxx			xx								
5.3 Write Task 5 final report	PSU		x		x								
6. Model selection		xxx	xxx	xxx	xxx	xxx	xxx						
6.1 Conduct modeling work			xxx	xxx	xxx	xxx							
6.1.1 Set up model grid, nowcast/forecast June 2010, & deliver model results to TAMU & PSU	modelers		x	x									
6.1.2 Define historical period, set model metrics, & standardize model delivery & format	PSU/TAMU		x	xxx	x								
6.1.3 Develop & test toolbox for model evaluation	TAMU/PSU			xxx	xxx								
6.1.4 Run CASE statistical model	PSU-Cort					xx							
6.1.5 Generate nowcasts for historical	Modelers				xx								
6.1.6 Generate forecasts for historical period	Modelers				x	x							
6.1.7 Generate forecasts for forecast period	Modelers					xx	x						
6.2 Conduct model inter-comparisons & evaluate results					x	xxx	xxx						
6.2.1 Develop 'true' historical picture of LC/eddy fronts	TAMU/PSU				x	x							
6.2.2 Evaluate individual models	TAMU/PSU					xxx	xxx						
6.2.3 Evaluate multimodel ensembles	PSU/JPL(?)					xxx	xxx						
6.3 Write draft then final report	PSU						xx	x					
6.4 Meet with RPSEA & recommend model								x					
7. Demonstrate Quasi-Operational Forecasts **8. Model Finalization and Technology Transfer**									xxx	xxx	xxx	xxx	x

Review and Summary of Phase I

During Phase I of GOMEX-PPP, a suite of seven dynamical ocean forecasting systems were examined to identify the state-of-the-art for making 3-month forecasts of the GOMEX circulation and LC. A short-range forecast system (AMSEAS) and statistical model for the path of the LC (ForLoop) were also utilized for nowcast and LC frontal position evaluations, respectively. The modeling systems, points of contact, and institutional affiliations of the systems are listed in Table 1.

Table 1: Participating Modeling Systems

	Modeling System Name	Point of Contact	Institution
#1	IASNFS	Dong-Shan Ko	NRL
#2	PROFS	Lie-Yauw Oey	Princeton University
#3	IASROMS-3DVAR	Yi Chao & Ruoying He	UCLA/JPL/NCSU
#4	MITGOM	Bruce Cornuelle	Scripps Inst. Oce.
#5	AMSEAS	Frank Bub	NAVO
#6	NGOM	Rich Patchen	NOAA/NOS/CSDL
#7	RTOFS	Avichel Mehra	NOAA/NWS/NCEP
#8	IASROMS-HYCOM	Ruoying He & Yi Chao	NCSU/UCLA/JPL
#9	ForLoop	Cort Cooper	Chevron

Web URLS:
#1 http://www7320.nrlssc.navy.mil/IASNFS_WWW/GOMEX/GOMEX.html
#2 http://www.aos.princeton.edu/WWWPUBLIC/PROFS/
#3 http://ourocean.jpl.nasa.gov/
#4 http://mitgcm.org/
#5 http://www.ngi.msstate.edu/edac/oceanNomads/AmSeas.php
#6 http://www.nauticalcharts.noaa.gov/csdl/NGOM.html
#7 http://polar.ncep.noaa.gov/ofs/
#8 http://omgsrv1.meas.ncsu.edu:8080/ocean-circulation/
#9 unavailable

The modeling systems were evaluated via intercomparison and model-observation comparisons during a hindcast period (daily analyses produced for 2010), followed by a retrospective forecast period (monthly forecasts of 3-month duration for 2010). A suite of five dynamical models (IASNFS, PROFS, IASROMS-3DVAR, MITGOM, and IASROMS-NHYCOM) continued into a real-time forecasting period (2011-2012).

During Phase I, the systems demonstrated that they can provide forecasts for the GOMEX using realistic, physically-based models. Several systems employ advanced data assimilation strategies to initialize forecasts with a combination of operational data, climatological data, and

background model nowcasts. Ocean forecasts were produced with or without the benefit of long-range (i.e., greater than two weeks) atmospheric forecasts.

Forecast skill was defined in terms of bias, root-mean-square error, and explained variance for sea-surface height fields, sea-surface temperature, subsurface temperature, and subsurface currents. Horizon Marine submitted real-time LC frontal analyses to GOMEX-PPP, which were used as validation data for the 2010 period.

Phase I identified three modeling systems which could be considered for application in a future real-time operational system: IASNFS, IASROMS-3DVAR, and MITGOM. The other models were disqualified for various reasons, as follows: PROFS was withdrawn after Phase I when a coding bug was identified by Oey, NGOM was found to be substantially less skillful than the other systems, RTOFS was occasionally numerically unstable, and IASROMS-HYCOM was superseded by IASROMS-3DVAR once the latter was fully implemented.

The three modeling systems, IASNFS, IASROMS-3DVAR, and MITGOM, demonstrated very similar forecast skill. Summarizing Table 1 of the Phase I Final Report, MITGOM had the best forecast skill for sea-surface height at 2-month lead time (0.3 skill vs. 0.1 skill for IASNFS and IASROMS-3DVAR), and the best agreement with the subsurface temperature (AXBT) data collected during the DwH event. IASNFS appeared most reliable, in some respects, exhibiting the smallest drift and initial transients compared to the other models. IASROMS-3DVAR had the best agreement with initial sea-surface height data and LC frontal location; however, it had the worst agreement with AXBT data in forecasts. Considering the non-definitive nature of the quantitative comparisons, IASNFS was regarded as the most mature candidate for transition to operational mode, primarily because of its reliability, quasi-operational use for several years prior to GOMEX-PPP, and established publication record in applications (both IASROMS-3DVAR and MITGOM were undergoing active development during Phase I).

One metric of LC/LCE predictability is the time for a forecast to lose all skill, when model-data error variance equals observed data variance. Based on satellite observations of sea-surface height during Step 3, the median time to zero skill was about 35 days for persistence and about 50 days for forecast.

Several recommendations were developed based on the findings of Phase I:

1) For improved initialization, validation, and verification of mesoscale forecasts in GOMEX, an enhanced real-time observing system needs to be designed, implemented, evaluated, and managed. Particular attention should be given to the use of satellite imagery to delineate fronts and jets.

2) The statistics of GOMEX fields appear to be non-stationary on the seasonal and annual time scales. Obtaining statistically significant model comparisons will require replicated experiments under nominally identical conditions. Replicate experiments in an operational environment are normally not feasible, unless the full suite of operational data, forcing fields, and other required model inputs can be archived in toto. Hence, it is recommended that pre-operational systems be

configured to archive the necessary inputs for a minimum of three years, so that three nominally independent replicates will be available for future model comparisons.

3) Sustained support for personnel is needed to perform analyses of forecast products and to explore open science questions in the GOMEX, activities which are normally secondary to the operation of a forecast system. Significant questions identified in Phase I are the following:

* Nature and significance of the variability on the GOMEX open boundaries (Yucatan Channel & Straits of Florida)
* Response to summertime tropical cyclone passage and wintertime subtropical frontal passages
* Response to the annual North American Monsoon which is most strongly expressed by a low-level northward wind jet in the western GOMEX during summer
* Passive surface particle trajectory prediction
* Mississippi-Atchafalaya River Plume exchanges with deep water GOMEX
* Nature and significance of shelf and deep-water exchanges throughout GOMEX
* Horizontal and vertical structure of the velocity, temperature, salinity, and potential vorticity fields associated with LC/LCE fronts and jets.

Based on the recommendation of the TAC (Cooper, Driver, and Vogel), Phase II of GOMEX-PPP proceeded with an ocean prediction group led by Pat Hogan at the Naval Research Laboratory (NRL). This recommendation was based on the conclusion that none of the funded Phase I modeling groups had sufficiently improved on the state-of-the-art found in the 2005 DeepStar project, and the forecast error in all GOMEX-PPP models is severely limited by poor model initial conditions. The NRL group has demonstrated experience with multi- and single-model ensemble forecasts, has strong connections with the NAVO operational prediction center, and utilizes the advanced NCODA data assimilation system for model initialization. Hence, the NRL modeling group was invited to take the modeling lead in Phase II.

Results and Accomplishments of GOMEX-PPP Phase II

Numerical Weather Prediction (NWP) has demonstrated that the ensemble-based probabilistic forecasting paradigm is superior to the deterministic forecast approach because it provides flow-dependent uncertainty in addition to an improved single solution (Palmer et al, 2005). Ocean ensemble forecasting and uncertainty estimation has until now been restricted to relatively small domain applications and for short range forecasts (Coelho et al, 2009). The complexity of the deep sea operations involved in supporting operational oil rigs in the northern GOMEX requires a longer range prediction system to allow for planning challenging tasks and to mitigate risk when adverse environmental conditions are likely to occur. Indeed, the shear associated with strong Loop Current eddy activity affects day-to-day operations during offshore oil and gas exploration and production activities, and can play an important role in the intensification of GOMEX hurricanes. Thus, the ability to predict the Loop Current intrusion and eddy separation process can have profound dynamical and socioeconomic implications.

During GOMEX-PPP Phase 2, an ensemble ocean modeling system was implemented to produce routine forecasts of the GOMEX at 60-day lead time. The ensemble forecast system (Figure 1) consists of several component systems, the Navy Coastal Ocean Model (NCOM) (Martin et al., 2009), Navy Coupled Ocean Data Assimilation System (NCODA) (Cummings, 2005) and COAMPS (Hodur, 1996). NCOM is a first principles (primitive equation) finite-difference ocean model that is used operationally at the Naval Oceanographic Office (NAVO). The domain is the entire Gulf of Mexico with 3 km horizontal resolution and 49 hybrid levels (33 sigma levels and 16 z-, or pressure levels, in the boundary layers). The vertical resolution is finest near the surface and the distance between levels progressively increases with depth. The Mellor-Yamada level-2 turbulent closure scheme is used for parameterizing vertical mixing, and the model uses barotropic and baroclinic mode splitting. Tides are included as a lateral boundary condition, and climatological monthly mean transports are used for riverine input. The NCODA system produces an analysis by assimilating all available observations via a 24-hour incremental update cycle.

NCODA is the data assimilation component, also used operationally at NAVO. The version used here is based on a 3-D variational implementation (3-DVar). Observations used by NCODA include all operational data sets available in near-real time. The primary sources are satellite altimetry (SSH) and sea surface temperature (SST), but profile observations from ships, gliders, and floats are also used. The observations pass through a robust quality control algorithm that checks for outliers and other unrealistic anomalies. The Modular Ocean Data Assimilation System (MODAS; Fox et al, 2002) ingests concurrent surface observations of temperature and height, and is used to construct "synthetic" profiles of temperature and salinity that are projected downward through the water column. There is no assimilation of velocity observations in the present system.

The first component of the 60-day forecast is the single (deterministic) forecast designated as the control run. The 24-hour forecast error variances from this control run are then used to generate perturbations to the model initial conditions through the Ensemble Transform (ET, Bishop and Toth, 1999) methodology. These perturbations generated by the forecast error variances are augmented by additional estimates of the model's temporal variability, nowcast/analysis

increments history, and climate variability (C. Rowley, personal communication) so that the ensemble model perturbations will have a spread similar to the best guess of the control run analysis error variance. This procedure is important, so that the distribution of ensemble solutions can encapsulate all detectable and dynamically relevant ocean states.

In addition to perturbing the initial conditions, the surface boundary conditions (the COAMPS wind stress and heat flux) are also perturbed via a space-time deformation technique. This technique relies on the fact that surface boundary condition errors are seldom due to time lags or the displacement of dominant dynamical features already in the forecast. As such, independent random (perturbed) forcing fields for each ensemble member are generated with a random shifting technique every 24 hours, and with a specified de-correlation length. Restated, the forcing is prepared at the same 3-hour interval by linear interpolation of the forcing, but with the values computed at randomly shifted times (Coelho et al, 2009; Wei et al, 2013). This procedure allows the system to use independent forcing for each member by accounting for uncertainty in the surface boundary conditions.

The above-described system was implemented and operated in near-real-time mode during the 2013-2014 period, during which time additional hindcasts of 2010 and 2011 were also performed. Forecast results were archived and disseminated at a purpose-built password-protected web site, which also presented derived risk-analysis products. A complete discussion of the ensemble modeling system, including its development and evaluation at NRL, is presented in the subcontractor's Final Report by P. Hogan, E. Coelho, G. Peggion, and P. Thoppil, submitted to C. N. K. Mooers, which accompanies this report as Appendix III. The Discussion section, below, describes independent analyses and verifications performed at PSU.

Discussion

Overview of 2013 Conditions

Figure 2 illustrates the configuration of the LC during the 13-month period from January 2013 through January 2014. At the start of the year, the LC extended northward into the Gulf, reaching about 27.5N in mid-March, 2013. During this time oscillations of the LC associated with cyclonic vortices amplified, and the LC began to bifurcate, leading to the formation of LCE Kraken. Horizon Marine identified April 2013 as the formation date of LCE Kraken; however, the eddy detached and reattached several times in the April-June time frame before finally translating westward. Subsequent to the detachment event, the LC retreated southward to about 23.5N where it remained from July through December, 2013. Since the start of 2014, the LC has again migrated northward, and it was at approximately 26N during February 2014.

Significant weather events during 2013 included three named tropical storms, as identified by the National Hurricane Center (http://www.nhc.noaa.gov/tafb_latest; Figure 3). Storms affecting the eastern Gulf were Tropical Storm Andrea (June 5-7), Tropical Storm Dorian (July 29-August 3), and Tropical Storm Karen (October 3-6). It does not appear that the storms significantly influenced the LC or formation of the LCE, although this has not been examined in detail.

Recent publications have hypothesized that zonal winds within the GOMEX and in the Caribbean Sea play a role in the formation of LCEs. Chang and Oey (2010) find that steady uniform wind directed to the west delays the onset of the quasi-periodic LCE shedding in a numerical model of the GOMEX. Through considerations of mass and momentum balance, they find that the westward wind induces an eastward near-surface flow in the deep GOMEX that opposes the westward translation of baroclinic perturbations in the LC intrusion. The delayed westward translation allows more time for buoyancy to accumulate in the LC intrusion, so that, when an LCE is finally shed, a rapid exchange of mass between the eastern and western Gulf occurs. Another role for wind forcing is hypothesized in the observational study of Chang and Oey (2013), who observed a correlation between the westward Trade Winds in the Caribbean Sea, which are linked to the oceanic meridional temperature gradient in the Caribbean, and baroclinic instability of the LC near the YC. The latter is hypothesized to provide the generation mechanism for eddies which form on the cyclonic edge of the LC, leading to a dynamical instability of the LC intrusion and separation of LCE.

To explore dynamical links to regional winds, time series of wind stress in several potentially-significant regions are examined using the NCEP daily surface wind analyses for 2013. Figure 4 (top) displays the surface wind vectors on March 1, 2013 for context. The westward Trade Winds are prominent throughout the Caribbean. The region from 80-70W and 12.5-17.5 is the location of the Caribbean Low-Level Jet (CLLJ; Amador, 1998), and it is the average zonal wind in this region that was identified as significant in Chang and Oey (2013). Figure 4 (middle) shows a time series of zonal wind stress, averaged over the CLLJ region, for 2013. The wind stress amplitude was maximum in January and in the June/July time periods, as is typical of the Trade Winds; and minima occurred in May and, again, in September through November. The large pulse of westward wind stress in January 2013 may have played a role in preconditioning

the Caribbean and creating conditions for baroclinic instability of the LC, consistent with the mechanisms hypothesized by Chang and Oey (2013).

The zonal wind stress averaged along 85W, shown in Figure 4 (bottom), may also be significant to the timing of LCE Kraken. If increased westward wind stress delays LCE formation (Chang and Oey, 2010), then one would expect decreased easterlies to accelerate LCE formation. And, indeed, in March 2013 there was a weakening and even reversal of the easterlies which persisted for several weeks. The weakened easterlies coincided with the first major cutoff of LCE Kraken from the LC. Further dynamical analysis would be necessary to determine causation, but the coincidence of favorable winds and rapid LC/LCE development is suggestive of a link.

Assessment of GOM-LREFS

During Phase II, the primary metric for quantitative assessment of GOM-LREFS was the comparison of model forecasts with subsequent verifying analyses. Since the analyses are made using the same modeling system as the forecast, this assessment technique cannot detect certain systematic errors or biases. Nonetheless, at the scale of the LC, this assessment is thought to be a fair indicator of model forecast skill.

Figure 5 summarizes the performance of the model through the entire 2013-2014 time period of Phase II, where comparisons are based on sea-surface height (SSH) within the key region of LC variability, from (89W,22N) to (83W,28N) and depth greater than 150m. The solid black line shows the root-mean-square (RMS) spatial variability of the SSH, relative to the time mean; variability was largest at the start of 2013, decreased to a minimum in September, and slightly increased or held steady until 2014. The blue line indicates the same RMS quantity for the forecast at 4-week lead time. The two time series are within a few centimeters of each other during the entire year, and demonstrate that the forecasting system is well-calibrated and yields realistic SSH variability, even at 4-week lead time.

The solid red curve in Figure 5 illustrates the RMS error (forecast vs. verifying analysis) at the same 4-week lead time. The RMS error was variable from March to mid-July, and peaked in early July, when it was as large as the RMS SSH signal itself. In the early part of the year the LC shed LCE Kraken, which separated during June-July, and the LC evidently followed an unpredictable path in late July. The SSH was much more predictable after mid-July.

The dashed red curve in Figure 5 illustrates the standard deviation (STD) of the ensemble forecasts over the same region and 4-week lead time. During much of the year, particularly after mid-July, the ensemble STD agrees well with the RMS error. However, in May and July, the ensemble STD greatly under-predicts the actual RMS error. Thus, during these time periods, the oceanic state was unpredictable, and the uncertainty of the forecasts was also not predicted. The reasons for this outcome are not yet known, but it suggests that the ensemble modeling system may be missing some sources of variability that are sometimes present in the natural system.

Figure 6a shows the SSH spatial anomaly correlation vs. forecast lead time for all 2013-2014 forecasts. Forecasts are much more skillful than persistence, with the period of useful skill (correlation > 0.6) being about 45 days. Model skill varied substantially over the year, and this is

illustrated in Figures 6b and 6c which show the correlation during the April-July time period and during the rest of the year, respectively. The useful forecast time is only 30 days during the April-July period, while it extends to 50 days during the remainder of the year. Taken together, Figures 5 and 6 illustrate some of the non-stationarity of dynamics and processes in the GOMEX.

Table 2. Forecasts/Hindcasts Completed and Analysis

Year	@NRL	@PSU
2013	Pre-operational real-time forecasts and development of forecast products.	Forecast skill assessment and ensemble error assessment.
2014	Real-time forecasts continue through May, 2014	
2010	Retrospective forecasts and daily re-analysis.	Skill assessment and comparison with GOMEX-PPP Phase I (IASNFS).
2011	Retrospective forecasts and daily re-analysis.	

Comparison with Phase I Results

In early 2014, NRL completed a series of retrospective forecasts for the year 2010, using the same methodology as the real-time forecasting system which was implemented in 2013 (and continued to run through 2014). The forecast cycle of the GOM-LREFS system differed slightly from the forecast cycle used in GOMEX-PPP Phase I, as detailed in Table 3. The main difference concerns the number of forecasts within the year. The Phase II forecast cycle was timed to produce 56-day forecasts every week, while the Phase I forecast cycle was timed to produce 90-day forecasts every month. Thus, there were many more independent forecasts produced during Phase II, each of shorter duration, than during Phase I.

As a preliminary exercise, the GOM-LREFS forecasts for 2013 were compared to the IASNFS forecasts for 2010 (Figure 7). Note that the time periods were similar with respect to LCE formation: Eddy Franklin formed in May 2010, and Eddy Kraken formed around the same time in 2013. The comparison between years is not conclusive, because the models are compared over different time periods and used very different forecast cycles (Table 3). It is interesting to note that persistence appeared to lose skill much more rapidly in in Phase II versus Phase I. Whether this difference is related to differences in the modeling systems or oceanic conditions is not yet known.

Table 3. Forecast Cycle in GOMEX-PPP Phase I and II.

Attribute	GOMEX-PPP Phase II (GOM-LREFS)	GOMEX-PPP Phase 1 (IASNFS)
Forecast initialization	Once per week	Once per month
Forecast duration	56 days (8 weeks)	90 days (3 months)
Output products	Once per week	Once per day

Figure 8 summarizes the main findings from the comparison of Phase II with Phase I. GOM-LREFS was found to have significantly higher skill, as measured by spatial anomaly correlation over the LC/LCE region, than IASNFS during the 2010 hindcasts. Both models were significantly more skillful than persistence, which lost all useful skill after about 35 days. However, the GOM-LREFS model lost skill at a rate of 1/(200 day), while IASNFS lost skill at a rate of 1/(150 day). GOM-LREFS displayed useful skill over the entire 56 day forecast period, while IASNFS lost all useful skill over the same period.

Because of differences in the forecast cycle (Table 3), there are many more degrees of freedom in the skill estimate for GOM-LREFS, as compared to IASNFS, and there is some question as to whether the skill difference is statistically significant. The statistical significance of the observed correlation difference may be evaluated by considering the number of degrees of freedom for the two cases. For GOM-LREFS, the 95% confidence interval for the 8-week anomaly correlation is estimated to be 0.7 to 0.8, based on 52 weekly forecasts. The confidence interval for IASNFS is estimated to be 0.4 to 0.7, based on 12 monthly forecasts. Although there are many assumptions involved in the significance estimate, the confidence intervals overlap only slightly, and the difference in skill is thought to be a robust feature of the comparison. This conclusion is supported by other quantitative and qualitative differences between the models, shown below.

The differences in anomaly correlation described above are measured by comparison of the forecast with the verifying analysis produced by each model separately. Thus, the forecast skill is evaluated by comparing each model with a different version of the “true” fields, the verifying analyses. To evaluate whether the difference in skill is an artifact of the self-verifying methodology, the forecast skill was also computed with respect to a common reference SSH analysis, the AVISO delayed-time multi-satellite altimeter product (Figure 9). The results of this comparison produced essentially the same results as the self-verifying analysis. Both models are strongly correlated with the AVISO analysis initially. The rate of loss of correlation with AVISO is similar to the rate of loss of correlation with respect to the self-verifying analysis. Thus, the forecast correlation displayed in Figure 8 appears to be a realistic and robust feature of the GOM-LREFS retrospective forecasts for 2010. Note that the GOM-LREFS correlation with AVISO is initially somewhat smaller than the IASNFS correlation. It is hypothesized that the GOM-LREFS initial condition is more accurate that the AVISO analysis, due to the dynamical content of the GOM-LREFS analysis, and the lack of initial correlation represents the additional information content of the GOM-LREFS system.

Similar to findings for 2013, there is a significant time-dependence of the model skill. This is illustrated in Figure 10, which shows statistics of SSH variability and forecast error at 4-week

lead time. During the first half of 2010 the RMSE of GOM-LREFS is low, about 7 cm, and significantly better than IASNFS, which ranges from 10 cm to 15 cm. Later, in July, the RMSE of GOM-LREFS rises sharply to the 15 cm level, before dropping again in September.

An interesting and un-explained difference between the 2010 and 2013 results is illustrated in Figure 11, which overlays the GOM-LREFS forecast skill during the two years. The SSH forecast loses skill at a much lower rate in 2010 as compared to 2013. At the end of the forecast window, on average, the 2010 forecasts still had useful skill, while the 2013 forecasts did not. The reason for this difference is unclear, but there are several possible explanations, as follows:

1. Changes in ad hoc observing system: The Jason-1 satellite altimeter reached its end-of-life in 2011, while the new CryoSat and Altika satellite altimeters were in operational use during 2013, beginning in July 2010 and March 2013, respectively. Also, the DwH event prompted acquisition of large in situ data sets during 2010, which were not present in 2013.
2. Changes in forecast system: During late 2013 GOM-LREFS changed the open boundary condition climatology used for forecasts, and the new climatology (based on Global HYCOM, as compared to previous Global NCOM results) was used during the hindcasts of 2010.
3. Changes in LC dynamics: The conditions upstream in the Caribbean Sea, in the LC, and the LCE shedding region may have differed significantly between 2010 and 2013 in a manner which influenced the predictability of the system.

The differences in model skill between GOMEX-PPP Phase I (IASNFS) and Phase II (GOM-LREFS) are described above in terms of SSH anomaly correlation. Similar differences in skill are evident when other field variables are considered. Figure 12 illustrates the anomaly correlation of forecasts for temperature at 100-m depth and 250-m depth. Once again, GOM-LREFS displays useful skill out to the 56-day limit of the forecasts.

The time series of statistics of the 250-m temperature 4-week forecast is summarized in Figure 13. Similar to the SSH statistics, an increase in RMSE is noted in the July-September time period. Comparison of GOM-LREFS and IASNFS finds lower RMSE in GOM-LREFS, except during the July-September period.

For completeness, Figures 14 and 15 illustrate the comparison of GOM-LREFS and IASNFS forecast metrics for the current speed at 100-m depth. The vector velocity field is much less predictable than the other (scalar/0 fields so far considered, and this is reflected in the reduced skill and predictability of the speed field. GOM-LREFS skill drops to the 0.6 level after approximately 35 days, about 10 days after IASNFS skill drops to this level. Comparison with 2013 results shows, as above, that 2010 was much more predictable than 2013.

Figures 16, 17, and 18 provide a visual comparison of the GOM-LREFS and IASNFS forecasts in relation to the HMI Eddy Watch Frontal Analyses (HMI-EW) provided for the 2010 period. Figure 16 shows the subsurface (250-m) temperature in GOM-LREFS (left) vs. IASNFS (right) starting from a forecast initialized at the start of April, when both models had relatively low RMSE (cf., Figure 13). There is initially good agreement in the location of the LC front between

HMI-EW and GOM-LREFS. Although the frontal contours diverge over the course of the 3-weeks shown, the location of the western and northwestern boundaries of the LC is forecast accurately. In contrast, both the initial and subsequent location of the IASNFS LC boundary diverge from HMI-EW.

At the start of June (Figure 17) the initial agreement of the models and HMI-EW is poorer. The evolution of the LC/LCE in GOM-LREFS displays changes in shape and orientation similar to those in the HMI-EW analysis; however, the frontal locations differ substantially after 3 weeks. Evolution of the LCE (Franklin) in IASNFS diverges sharply from the HMI-EW analysis, with the model LCE apparently more symmetric and rapidly translating westward.

The forecast initialized on August 1 displays a rapid divergence from the HWI-EW analysis (Figure 18). The LCE evolves rather similarly in the two models, but it translates westward too rapidly compared to the HMI-EW analysis. The main path of the LC evolves considerably, also, and GOM-LREFS appears to better represent its initial location and evolution than IASNFS.

Note that the above comparisons with HMI are based on the thermal structure at depth (250 m). Other comparisons are possible; however, it is not known what attribute of the models corresponds best to the HMI front. The HMI frontal estimates utilize remotely sensed imagery and surface Lagrangian drifter tracks, where the drifters are drogued to follow the current at a nominal 50 m depth. There is considerable vertical shear in the upper ocean, between the surface mixed layer and 200 m depth. Hence, comparisons between the HMI LC front estimate and horizontal current isotachs, below, are intended to compare qualitatively the large-scale features of both fields, rather than providing a definitive quantitative comparison.

Figures 19 and 20 present a view of the May 1 and August 1 forecasts 4 weeks after being initialized, corresponding to the forecasts in Figure 15 and Figure 17, respectively. The sub-domain shown has been selected to focus on the LC and LCE, with water current speed at 100-m depth (SPD100) illustrated together with the HMI-EW analysis overlaid. Figure 19 compares the GOM-LREFS verifying analysis, 4-wk. forecast, and 4-wk. persistence SPD100 fields (top row). At that time of year, the GOM-LREFS forecast of the LC front agrees well with HMI-EW analysis. The anomaly fields, and corresponding anomaly correlation values (bottom row), display the extent of agreement of spatial pattern.

Figure 20 illustrates the lack of agreement which occurs subsequent to the August 1 forecast. The verifying analysis (top left) agrees well with the HMI-EW analysis, and shows a separated LCE, Eddy Franklin. The 4-wk. forecast (top center) for this date contains a weak LC intrusion, and the LCE is about 100 km west from its location in the HMI-EW analysis. The persistence forecast (top right), which is also the initial field for the forecast, illustrates the complexity of the flow field at the start of the forecast. The anomaly fields for this date (bottom row) are poorly correlated between the forecast and analysis.

In summary, the comparison of GOMEX-PPP Phase II (GOM-LREFS) and Phase I (IASNFS) models over a common 2010 reference period finds that the GOM-LREFS system has significantly more forecast skill than IASNFS. Based on RMSE metrics of SSH and upper-ocean fields, the forecast skill of GOM-LREFS is reduced during time periods of the LCE shedding

event when the LC and LCE system are rapidly evolving, separating, and re-attaching. The reduced accuracy of GOM-LREFS associated with the LCE shedding is similar to what was observed in 2013. A difference in forecast skill between 2010 and 2013 is documented, with 2010 being substantially more predictable, but the reasons for this difference are unclear.

A Concept of Operations for Sustained Operational Long-Range Forecasts of the Gulf of Mexico

A Concept of Operations (CONOPS) is a Systems Engineering construct for efficiently managing a complex system with multiple sponsors, providers, and users. A CONOPS identifies roles and responsibilities, and performance metrics and standards. An operational environmental prediction system (e.g., GOMEX-PPP), comprised of observing, modeling, and data analysis and management subsystems, is an example of a complex system that requires a CONOPS. Only the real-time modeling subsystem is considered here.

The Navy has developed and formalized the transition of modeling subsystems from R&D to operational use (Appendix II). With input from the academic community, NRL develops the Navy's prediction models, and NAVO runs the models in an operational (i.e., routine, continuing) fashion. Both NRL and NAVO are involved in the various transition processes; e.g., model skill assessment. They have a documented CONOPS for their interactions. GOMEX-PPP has proceeded in the standard fashion of conducting a formal transition to meet Navy needs. After a modeling subsystem completes the transition process, the Navy (viz., NAVO) also distributes the basic forecast fields to the civilian community via NOAA (viz., NCEP).

While the Navy is meeting its needs thanks to its rigorous transition process and CONOPS, the civil sector (e.g., industry, NOAA, and academia) has additional requirements; for example, to conduct retrospective re-analyses, and observing system simulation experiments (Appendix II).

Findings and Recommendations

FINDING 1: A GOMEX circulation forecast system, GOM-LREFS, has been implemented.

An extended range (~60 days) real time ocean forecasting system, the "Gulf of Mexico Long Range Ensemble Forecast System" (GOM-LREFS), based on probabilistic (single-model-ensemble) methods has been developed for GOMEX. The probabilistic prediction system provides both a forecast of time-evolving ocean circulation, including currents, water temperature, salinity, and sea-surface height, and an estimate for the uncertainty of that forecast.

The GOM-LREFS system has run in real-time pre-operational mode since January 27, 2013, to present. The forecast cycle produces a 60-day forecast once-per-week, using analyses which are produced daily by the system. In addition to the real-time forecast, hindcasts for the year 2010 and other historical periods have been produced for retrospective skill assessments.

The system provides the foundational data for forecast products relevant to a range of applications in the public/industrial sector, such as the probability of strong surface currents. The forecast products are derived from the forecast ensemble and can be used for planning and risk mitigation.

RECOMMENDATION 1: Continue to operate GOM-LREFS.

The GOM-LREFS system should continue to operate in order to disseminate, validate, and verify the forecast products. In the manner of a model testbed, the forecast system should undergo continuous quality assessment and improvement and be used to facilitate development of the GOMEX ocean observing system. Continued operation would also provide much needed data for cost-benefit analysis relevant to extended range ocean forecast systems.

FINDING 2: The forecast skill of GOM-LREFS was documented.

Skill assessment of GOM-LREFS was performed by computing an anomaly correlation of forecasts with verifying analyses as a function of forecast lead time. Additional qualitative assessments were performed by examining forecasts during periods of LCE genesis and separation. The forecast duration of useful model skill, defined as anomaly correlation greater than 0.6, extends from two to eight weeks, depending on the season and the field variables considered. Forecasts of the dynamically-significant SSH field contain useful skill at 30 days during periods of rapid LC evolution; during more quiescent periods the useful skill extends past 50 days.

An assessment of the GOM-LREFS forecast uncertainty was performed by comparing actual RMSE with the ensemble spread of SSH forecasts during 2013. Agreement between forecast and actual uncertainty was found during some periods; however, forecasts with large RMSE were not predicted by the ensemble spread. In other words, the ensemble spread was too small; it provided an overly-optimistic indicator of forecast error.

RECOMMENDATION 2: Validate forecast uncertainty estimates.

Because a large number of the GOM-LREFS forecast products depend on the forecast uncertainty, which is computed from the ensemble spread, more validation of the ensemble error metrics should be performed. Validation of forecast uncertainty, and other probabilistic quantities, requires many independent realizations or case studies in order to reduce sampling error, thus sustained operation of the GOM-LREFS system is required.

FINDING 3: Model skill was improved in GOMEX-PPP Phase II versus Phase I.

GOM-LREFS produced a hindcast of 2010 for evaluation with respect to the same reference period (i.e., control runs) as used in GOMEX-PPP Phase I. Comparisons with the 2010 hindcast produced by IASNFS during Phase I illustrate the positive impact of newer methodologies employed in GOM-LREFS during Phase II.

Evaluation of GOM-LREFS and IASNFS via the same metrics demonstrated significant improvement in GOM-LREFS, as measured by both smaller RMSE and longer period of useful forecast skill. For example, forecasts of upper ocean current speed were skillful for 40 days in GOM-LREFS, as compared with 25 days in IASNFS.

RECOMMENDATION 3: Establish a regional modeling testbed.

A testbed for modeling subsystems (and observing, data assimilation, and data management subsystems) should be established to foster the rapid development of quantitative regional prediction capabilities for R&D and pre-operational and operational groups.

FINDING 4: Seasonal and interannual variability is present.

Virtually all model skill metrics examined showed trends or seasonal-scale variability during the course of the year, in both 2010 and 2013. Forecast RMSE, as well as spatial variance of the physical fields, was larger during the first half of the year than in the second half of the year. Changes in RMSE and spatial variance were also pronounced during periods of LCE development and shedding.

A large change in GOM-LREFS forecast skill was noted between 2010 and 2013, with 2010 being, on average, more predictable than 2013. There are several possible explanations for this difference in predictability, but the actual cause is presently unknown.

RECOMMENDATION 4: Study the factors controlling LC predictability.

The cause of the change in predictability between 2010 and 2013 should be identified in order to determine whether it is controllable, i.e., related to changes in the ocean observing system or overall forecasting system, or uncontrollable, i.e., related to intrinsic variability of the LC system.

FINDING 5: Collaboration has benefits for the R&D and operational prediction communities.

Multi-use prediction systems can advance more rapidly and rigorously when R&D and pre-operational activities coordinate infrastructure, experiments, etc., in a sustained and collaborative fashion.

RECOMMENDATION 5: Form a consortium of sponsors, users, and observationalists concerned with Intra-Americas Sea (IAS) environmental prediction.

Adopt a Systems Engineering approach by forming a consortium of sponsors, users, and observationalists for the IAS to coordinate model requirements (e.g., prediction grids and resolution, ecosystems, advanced air-sea coupling) and optimum observing system design.

References

Amador, J. 1998. A climatic feature of the tropical Americas: the trade wind easterly jet. Top. Meteor. Oceanogr., 5: 91-102.

Barker, C.H. 2011. A statistical outlook for the Deepwater Horizon oil spill. In: Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophysical Monograph Series, vol. 195. Washington, D.C.: American Geophysical Union. Pp. 237–244.

Bishop, C.H. and D. Hodyss. 2011. Adaptive Ensemble Covariance Localization in Ensemble 4D-VAR State Estimation. Mon. Wea. Rev., 139:1241–1255.

Bishop, C.H. and Z. Toth. 1999. Ensemble transformation and adaptive observations. J. Atmos. Sci., 56: 1748-1765.

Candela, J., J. Sheinbaum, J. Ochoa, A. Badan, and R. Leben. 2002. The potential vorticity flux through the Yucatan Channel and the Loop Current in the Gulf of Mexico. Geophys. Res. Lett. 29:2059–2062.

Candela, J., S. Tanahara, M. Crepon, B. Barnier, and J. Sheinbaum. 2003. Yucatan Channel flow: Observations versus CLIPPER ATL6 and MERCATOR PAM models. J. Geophys. Res. 108(C12):3385.

Chang,Y.-L. and L.-Y. Oey. 2010. Why can wind delay the shedding of Loop Current eddies? J. Phys. Oceanogr 40:2481-2495.

Chang, Y.-L. and L.-Y. Oey. In press. Why does the Loop Current tend to shed more eddies in summer and winter? Geophys. Res. Lett., doi:10.1029/2011GL050773.

Chang, Y.-L., L. Oey, F.-H. Hu, H.-F. Lu, and A. Fujisaki. 2011. 2010 oil spill: trajectory projections based on ensemble drifter analyses. Ocean Dynamics 61(6):829-839.

Chang,Y.-L. and L.-Y. Oey. In press. Loop Current Cycle: coupled response of Loop Current and deep flows. J. Phys. Oceanogr.

Chassignet, E.P., H.E. Hurlburt, E.J. Metzger, O.M. Smedstad, J. Cummings, G.R. Halliwell, R. Bleck, R. Baraille, A.J. Wallcraft, C. Lozano, H.L. Tolman, A. Srinivasan, S. Hankin, P. Cornillon, R. Weisberg, A. Barth, R. He, F. Werner, and J. Wilkin. 2009. U.S. GODAE: Global Ocean Prediction with the Hybrid Coordinate Ocean Model (HYCOM). Oceanography 22(2):64-75.

Chen, K. and R. He. 2010. Numerical Investigation of the Middle Atlantic Bight Frontal Circulation using a High Resolution Ocean Hindcast Model. Journal of Physical Oceanography DOI: 10.1175/2009JPO4262.1

Chin, T.M., M.J. Turmon, J.B. Jewell, and M. Ghil. 2007. An ensemble-based smoother with retrospectively updated weights for highly nonlinear systems. Monthly Weather Review 135:186-202.

Coelho, E. G. Peggion, C. Rowley, G. Jacobs, R. Allen, E. Rodriguez. 2009. A note on NCOM temperature forecast error calibration using the ensemble transform. J. Mar. Sys., 78 (supplement): S272-S281.

Counillon, F. and L. Bertino. 2009a. High-resolution ensemble forecasting for Gulf of Mexico eddies and fronts. Ocean Dynamics 59:83-95.

Counillon, F. and L. Bertino. 2009b. Ensemble optimal interpolation: multivariate properties in the Gulf of Mexico. Tellus 61A:296-308.

Crosby, D.S., L.C. Breaker, and W.H. Gemmill. 1990. A definition for vector correlation and its application to marine surface winds. National Meteorological Center Office Note No. 365, NOAA/National Weather Service, Washington, D.C., 50 pp.

Crosby, D.S., L.C. Breaker, and W.H. Gemmill. 1993. A Proposed Definition for Vector Correlation in Geophysics: Theory and Application. J. Atmos. Sci. 10:355-367.

Cummings, J.A. 2005. Operational multivariate ocean data assimilation. Quarterly Journal of the Royal Meteorological Society 131:3583-3604.

Cummings, J.A. 2006. The NRL real-time ocean data quality control system. NRL Technical Note (available from cummings@nrlmry.navy.mil).

Cummings, J.A. 2011. Ocean Data Quality Control. In: Schiller, A. and G.B. Brassington, eds. Operational Oceanography in the 21st Century. New York: Springer-Verlag. Pp. 91-121.

Ezer, T., L.-Y. Oey, W. Sturges, and H.-C. Lee. 2003. The variability of currents in the Yucatan Channel: Analysis of results from a numerical ocean model. J. Geophys. Res. 108(C1):3012.

Fox, D.N., C.N. Barron, M.R. Carnes, M. Booda, G. Peggion, and J.V. Gurley. 2002. The modular ocean data assimilation system. Oceanography 15:22-28.

Fu, L.-L. and A. Cazenave. 2001. Satellite Altimetry and Earth Sciences. International Geophysics Series, Vol. 69. San Francisco: Academic Press. 463 pp.

Hamilton, P., J.C. Larsen, K.D. Leaman, T.N. Lee, and E. Waddell. 2005. Transports through the Straits of Florida. J. Phys. Oceanogr. 35:308-322.

Hodur, R.M., 1996: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS), *Mon. Weather Rev*., 125, 1414-1430.

Hyun, K.H. and R. He. 2010. Coastal Upwelling in the South Atlantic Bight: A Revisit of the 2003 Cold Event Using Long Term Observations and Model Hindcast Solutions. Journal of Marine Systems 83:1-13.

Jenkins, G.M. and D.G. Watts. 1968. Spectral Analysis and Its Applications. San Francisco: Holden-Day. 525 pp.

Jolliff, J.K., J.C. Kindle, I. Shulman, B. Penta, M.A.M. Friedrichs, R. Helber, and R.A. Arnone. 2009. Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment, Journal of Marine Systems 76:64-82.

Kantha, L., J.-K. Choi, K.J. Schaudt, and C.K. Cooper. 2005. A Regional Data-Assimilative Model for Operational Use in the Gulf of Mexico. In: Circulation in the Gulf of Mexico: Observations and Models. Geophysical Monograph Series 161. Pp. 165-180.

Kenyon, S., J. Factor, N. Pavlis, and S. Holmes. 2007. Towards the Next Earth Gravitational Model. Society of Exploration Geophysicists 77th Annual Meeting, San Antonio, Texas. http://earth-info.nga.mil/GandG/wgs84/gravitymod/new_egm/papers.html.

Ko, D.S., R.H. Preller, and P.J. Martin. 2003. An Experimental Real-Time Intra Americas Sea Ocean Nowcast/Forecast System for Coastal Prediction. Proceedings, AMS 5th Conference on Coastal Atmospheric & Oceanic Prediction & Processes.

Ko, D.S., P.J. Martin, C.D. Rowley, and R.H. Preller. 2008. A real-time coastal ocean prediction experiment for MREA04. J. Marine Systems 69:17-28.

Larsen, J.C. and T.B. Sanford. 1985. Florida Current volume transports from voltage measurements. Science 227:302–304.

Leben, R.R. 2005. Altimeter-derived Loop Current metrics. In: Sturges, W. and A. Lugo-Fernandez, eds. Circulation in the Gulf of Mexico: Observations and Models. Washington, D.C.: American Geophysical Union. Pp. 181–201.

Leben, R.R. and G.H. Born. 1993. Tracking Loop Current eddies with satellite altimetry. Adv. Space Res. 13(11):325-333.

Leben, R.R., G.H. Born, B.R. Engebrath. 2002. Operational altimeter data processing for mesoscale monitoring. Marine Geodesy 25:3-18.

Leben, R.R. and K. Corcoran. 2005. Horizon Marine Inc. Numerical Ocean Model Nowcast and Forecast Skill Assessment in the Gulf of Mexico. Horizon Marine, Inc. http://www.horizonmarine.com/library/Horizon%20Marine%20Inc.%20Numerical%20Ocean%20Model%20Nowcast%20and%20Forecast%20Skill%20Assessment%20in%20the%20Gulf%20of%20Mexico.pdf. 31 pp.

Lee, T.N., and E. Williams. 1988. Wind-forced transport fluctuations of the Florida Current. Journal of Physical Oceanography 18:937-946.

Lee, T.N., W.E. Johns, R.J. Zantopp and E.R. Fillenbaum. 1996. Moored observations of western boundary current variability and thermohaline circulation at 26.5N in the subtropical North Atlantic. Journal of Physical Oceanography, 26:962-983.

Li, Z., Y. Chao, J.C. McWilliams, and K. Ide. 2009. A Three-Dimensional Variational Data Assimilation Scheme for the Regional Ocean Modeling System. Journal of Atmospheric and Oceanic Technology 25:2074-2090.

Lin, X.-H., L.-Y. Oey, and D.-P. Wang. 2007. Altimetry and drifter assimilations of Loop Current and eddies. Journal of Geophysical Research 112:C05046.

Liu, Y. and R.H. Weisberg. 2011. Evaluation of trajectory modeling in different dynamic regions using normalized cumulative Lagrangian separation. J. Geophys. Res. 116:C09013.

Liu, Y., R.H. Weisberg, C. Hu, and L. Zhang. 2011. Trajectory forecast as a rapid response to the Deepwater Horizon oil spill. In: Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophysical Monograph Series, vol. 195. Washington, D.C.: American Geophysical Union. Pp. 153-165.

MacFadyen, A., G.Y. Watabayashi, C.H. Barker, and C.J. Beegle-Krause. 2011. Tactical modeling of surface oil transport during the Deepwater Horizon spill response. In: Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophysical Monograph Series, vol. 195. Washington, D.C.: American Geophysical Union. Pp. 167-178.

P. J. Martin, C. N. Barron, L. F. Smedstad, T. J. Campbell, A. J. Wallcraft, R. C. Rhodes, C. Rowley, T. L. Townsend and S. N. Carroll. 2009. User's Manual for the Navy Coastal Ocean Model (NCOM) version 4.0. *NRL Report* NRL/MR/7320--09-9151.

Mehra, A. and I. Rivin. 2010. A Real Time Ocean Forecast System for the North Atlantic Ocean. Terr. Atmos. Ocean. Sci. 21(1):211-228.

Mooers, C.N.K., C.S. Meinen, M.O. Baringer, I. Bang, R. Rhodes, C.N. Barron, and F. Bub. 2005. Cross validating ocean prediction and monitoring systems. Eos, 86(29): 269-273.

Murphy, A. H., and E. S. Epstein. 1989. Skill scores and correlation coefficients in model verification. Mon. Wea. Rev., 117: 572-581.

Naeije, M., E. Doornbos, L. Mathers, R. Scharroo, E. Schrama, and P. Visser. 2002. Radar Altimeter Database System: exploitation and extension (RADSxx). NUSP-2 report 02-06, NUSP-2 project 6.3/IS-66. Delft Institute for Earth-Oriented Space Research (DEOS), Delft, Netherlands.

Oey, L.-Y. 1996. Simulation of mesoscale variability in the Gulf of Mexico: Sensitivity studies, comparison with observations, and trapped wave propagation. J. Phys. Ocean. 26(2):145-175.

Oey, L.-Y., T. Ezer, and W. Sturges. 2004. Modeled and observed Empirical Orthogonal Functions of currents in the Yucatan Channel, Gulf of Mexico. J. Geophys. Res. 109:C08011.

Oey, L.-Y., H.-C. Lee, and W.J. Schmitz. 2003. Effects of winds and Caribbean eddies on the Loop Current eddy shedding: A numerical model study. J. Geophys. Res. 108:3324.

Oey, L.-Y., T. Ezer, G. Forristall, C. Cooper, S. DiMarco, and S. Fan. 2005. An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico. Geophys. Res. Lett. 32:L12611.

Pavlis, N.K., S.A. Holmes, S.C. Kenyon, D. Schmidt, and R. Trimmer. 2004. A Preliminary Gravitational Model to Degree 2160. Paper presented at the IAG International Symposium on Gravity, Geoid and Space Missions 2004 (GGSM2004) Porto, Portugal. http://earth-info.nga.mil/GandG/wgs84/gravitymod/new_egm/papers.html.

Palmer, T.N., F.J. Doblas-Reyes, R. Hagedorn, and A. Weisheimer. 2005 Probabalistic prediction of climate using multi-model ensembles: from basics to applications. Phil. Trans. R. Soc. B, 360: 1750.

Powell, B.S. and R.R. Leben. 2004. An Optimal Filter for Geostrophic Mesoscale Currents from Along-Track Satellite Altimetry. J. Atm. and Ocean. Tech. 21:1633-1642.

Rio, M.-H. and F. Hernandez. 2003. A Mean Dynamic Topography computed over the world ocean from altimetry, in-situ measurements and a geoid model. J. Geophys. Res. 109:C12032.

Rio, M.H., S. Guinehut, and G. Larnicol. In press. The New CNES-CLS09 global Mean Dynamic Topography computed from the combination of GRACE data, altimetry and in-situ measurements. J. Geophys. Res., doi:10.1029/2010JC006505.

Rousset, C. and L.M. Beal. 2010. Observations of the Florida and Yucatan Currents from a Caribbean Cruise Ship. J. Phys. Oceanogr. 40:1575-1581.

Shay, L.K., B. Jaimes, J.K. Brewster, P. Meyers, E.C. McCaskill, E. Uhlhorn, F. Marks, G.R. Halliwell Jr., O.M. Smedstad, and P. Hogan. 2011. Airborne ocean surveys of the Loop Current complex from NOAA WP-3D in support of the Deepwater Horizon oil spill. In: Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise, Geophysical Monograph Series, vol. 195. Washington, D.C.: American Geophysical Union. Pp. 131-151.

Shoosmith D.R., M.O. Baringer, and W.E. Johns. 2005. A continuous record of Florida Current temperature transport at 27°N. Geophys. Res. Lett. 32:L23603.

SSALTO/DUACS. 2011. SSALTO/DUACS User Handbook: (M)SLA and (M)ADT Near-Real Time and Delayed Time Products. Issue 2.6, CLS-DOS-NT-06-034, Nomenclature : SALP-MU-P-EA-21065-CLS. 57pp.

Sturges, W., J.C. Evans, S. Welsch, and W. Holland. 1993. Separation of warm-core rings in the Gulf of Mexico. J. Phys. Oceanogr. 23:250-268.

Sturges, W., N. Hoffmann, and R. Leben. 2010. A Trigger Mechanism for Loop Current Ring Separations. J. Phys. Oceanogr. 40:900-913.

Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res. 106: 7183-7192.

van Leeuwen, P.J. and G. Evensen. 1996. Data assimilation and inverse methods in terms of a probabilistic formulation. Mon. Wea. Rev. 124: 2898-2913.

Wei, M., C. Rowley, P. Martin, C. N. Barron, and G. Jacobs. 2013. The US Navy's RELO ensemble prediction system and its performance in the Gulf of Mexico. Q. J. R. M. S., doi:10.1002/qj/2199.

Zhao, Y. and R. He. In Press. Cloud-free Sea Surface Temperature and Color Reconstructions for the Gulf of Mexico: 2003-2009. Remote Sensing Letters.

Figures

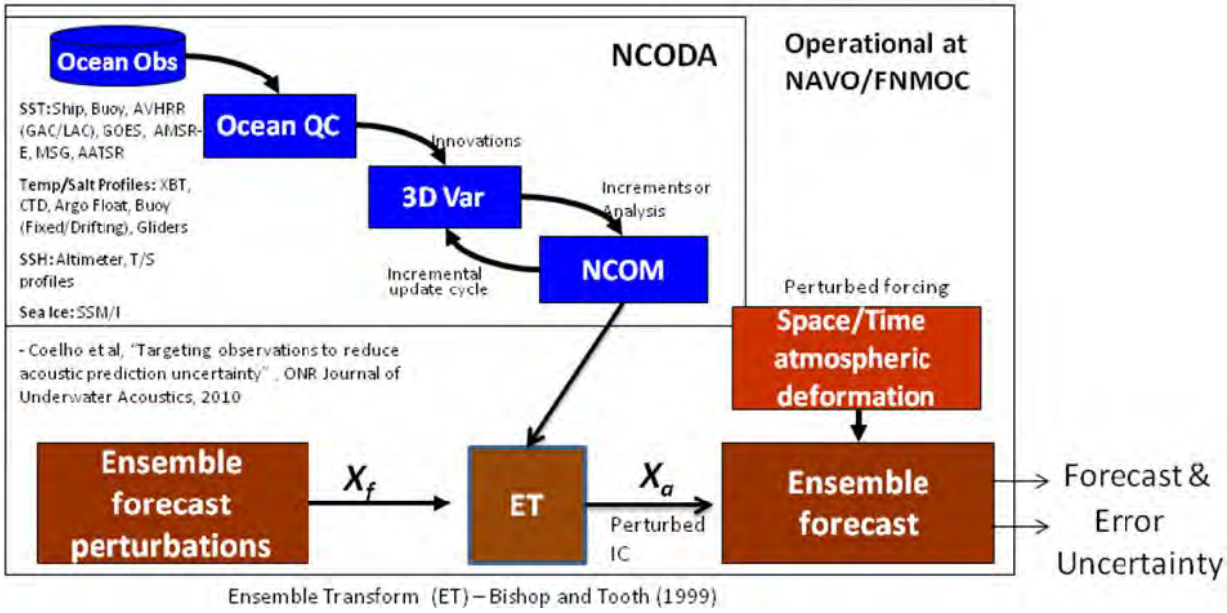


Figure 1. Schematic diagram of the ensemble forecast system. The components of the NCODA system are the blue boxes in the inset. Both the surface forcing and initial conditions are perturbed to generate the individual ensemble members.

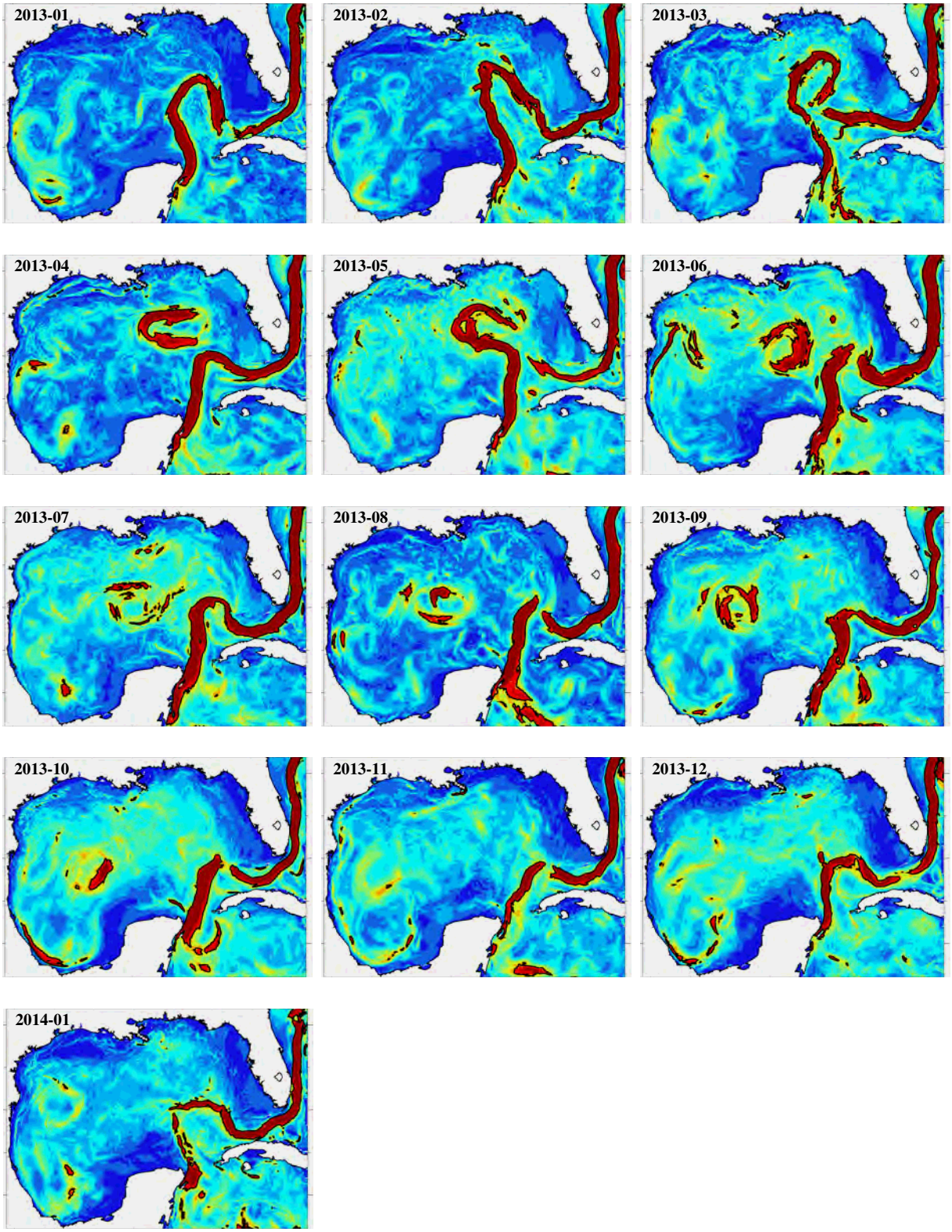


Figure 2. Evolution of the LC during 2013. Panels show analysis of near-surface speed at monthly intervals from January 2013 through January 2014. The color scale ranges from 0m/s (dark blue) to 1m/s (dark red).

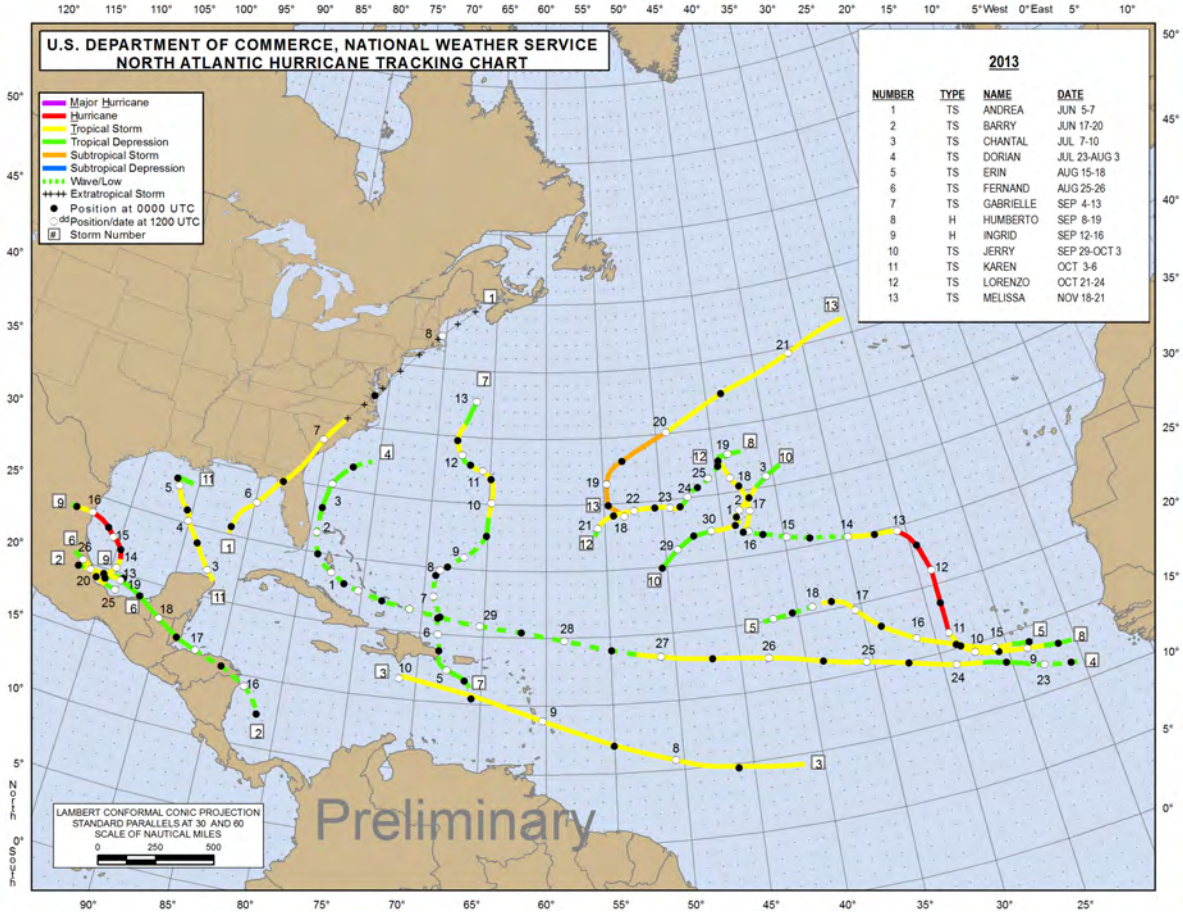


Figure 3. Atlantic tropical storms, 2013.

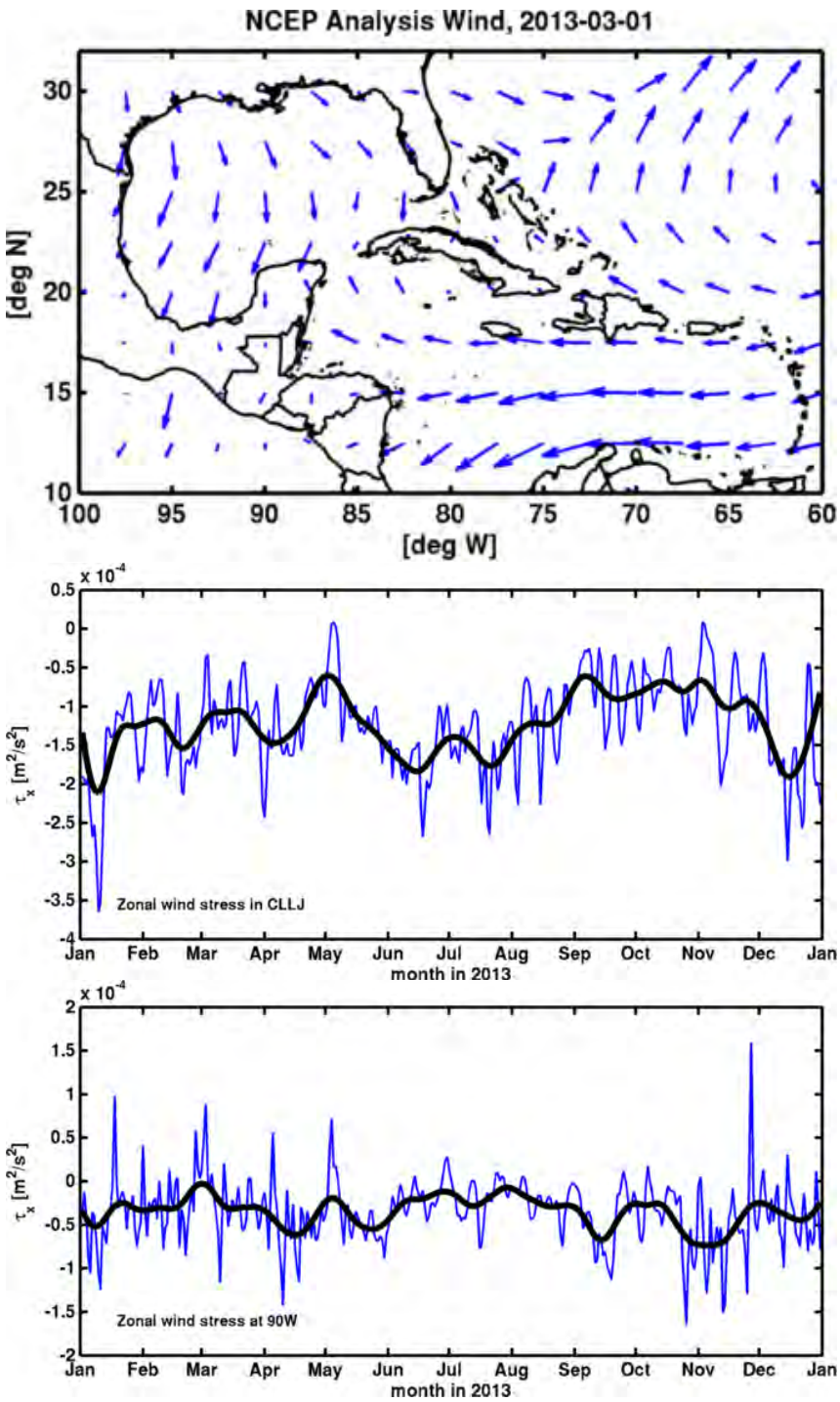


Figure 4. GOMEX and Caribbean winds. Easterly trade winds are prominent throughout the Caribbean in March, 2013 (top); maximum vector magnitude shown is 12 m/s. The zonal component of winds stress, averaged over the region of the Caribbean Low-Level Jet (CLLJ), shows that the annual cycle of the trade winds was typical in 2013 (middle), with largest westward stress occurring twice per year in June/July and December/January. Zonal wind stress averaged along 90W within the GOMEX (bottom) was near zero in early March, coincident with the initial formation of LCE Kraken.

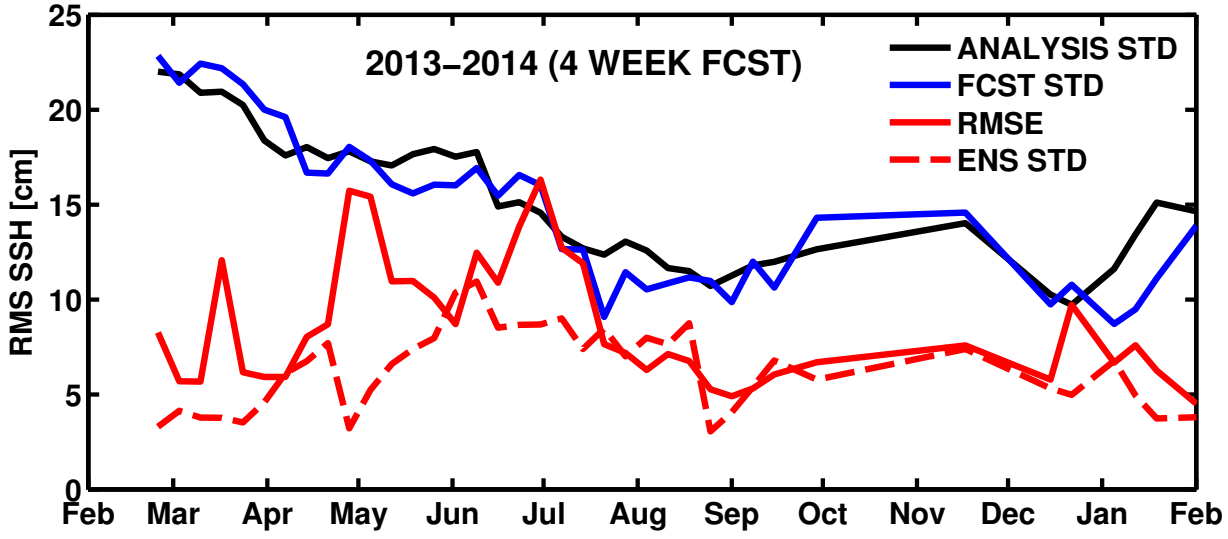


Figure 5. Summary of GOM-LREFS sea-surface height (SSH) anomaly statistics during Phase II. The verifying analysis (black) and 4-week forecast (blue) exhibit nearly equal root-mean-square (RMS) spatial variability (i.e., spatial standard deviation; STD). The RMS forecast error (solid red; RMSE) and ensemble standard deviation (dashed red; ENS STD) diverge during periods of poor predictability (May and July).

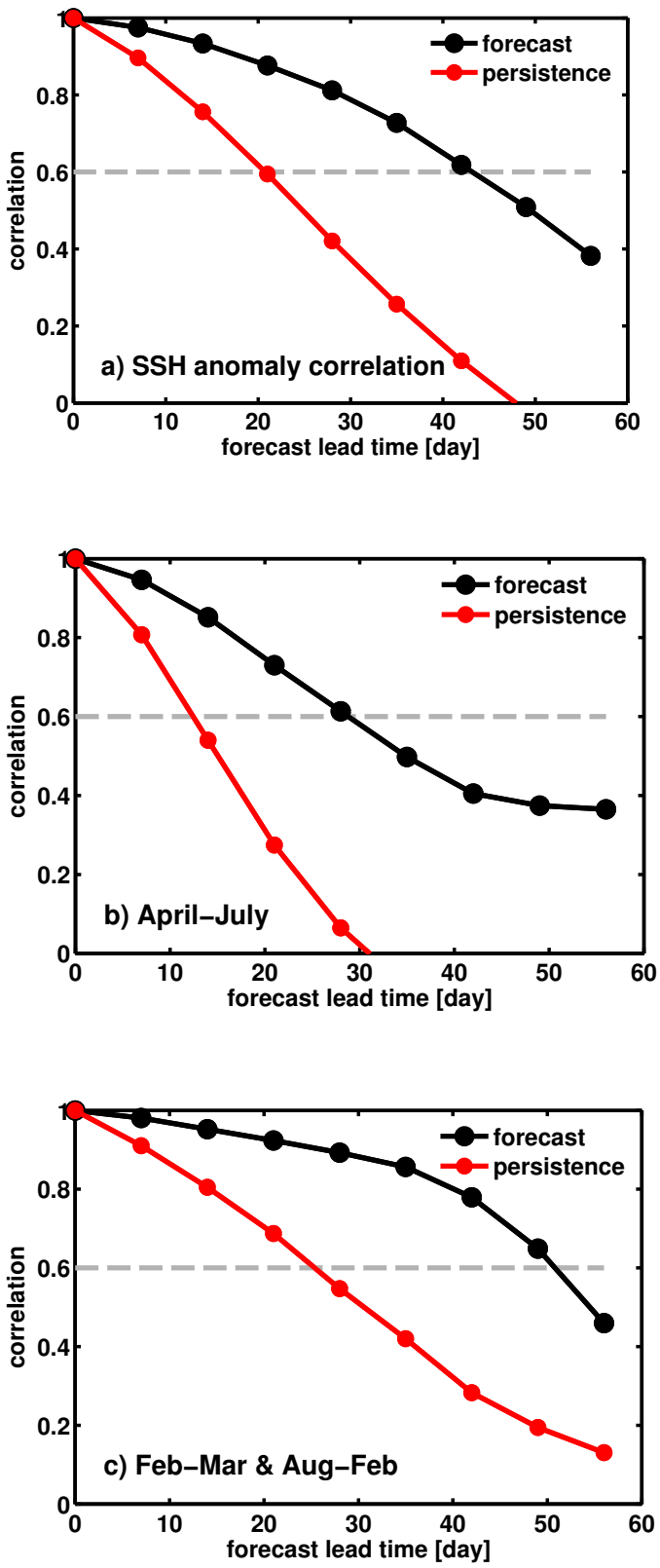


Figure 6. Anomaly correlation of SSH vs. forecast, Phase II; a) average over 2013; b) average over summer months 2013; c) average over winter months 2013.

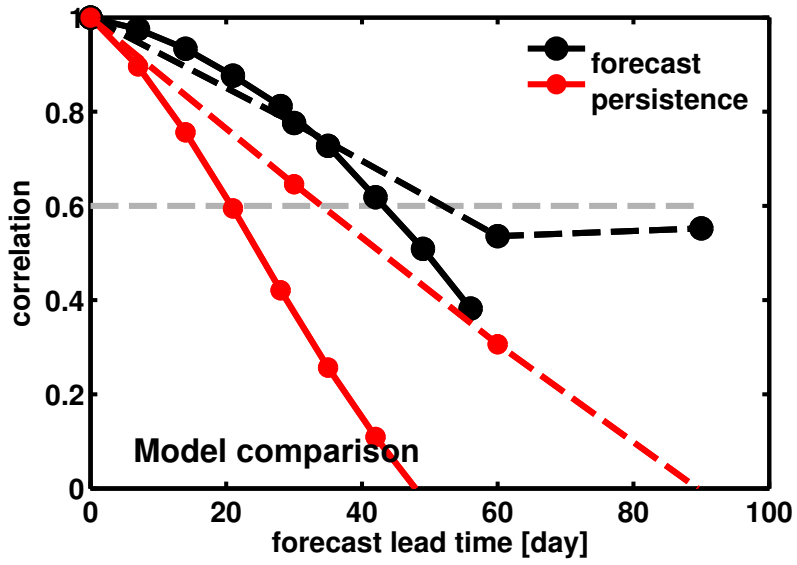


Figure 7. Comparison of Phase I and Phase II results. SSH anomaly correlation for GOM-LREFS (solid black, forecast; solid red, persistence) and IASNFS (dashed black, forecast; dashed red, persistence) within the LC/LCE region (89W-22N to 83W-28N and water depth greater than 150m). Note that the comparison involves different years, 2013 (Phase II; solid curves) vs. 2010 (Phase I; dashed curves).

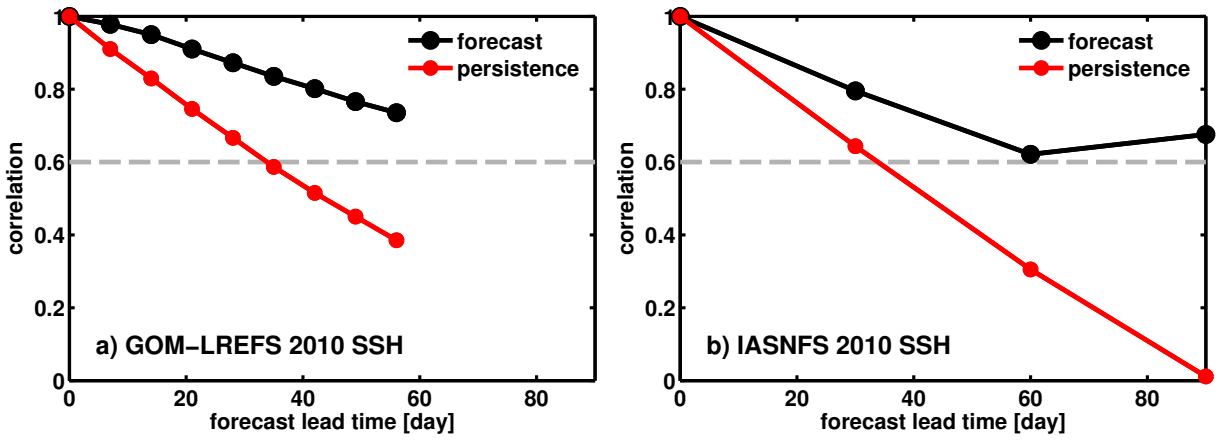


Figure 8. Correlation with Self-Verifying Analyses, 2010. Correlation between anomaly SSH of forecast and analysis as a function of forecast time is shown for (a) GOM-LREFS and (b) IASNFS for the 2010 time period. Anomaly SSH is defined with respect to both the time-average of SSH within the subdomain, 22-28N 89-83E with depth greater than 250 m. Based on correlation with the self-verifying analysis, GOM-LREFS is more skillful than IASNFS. The rate of loss of correlation is about 1/(200 day) for GOM-LREFS vs. 1/(150 day) for IASNFS. Alternately, GOM-LREFS forecasts are skillful for about 3 weeks longer than IASNFS.

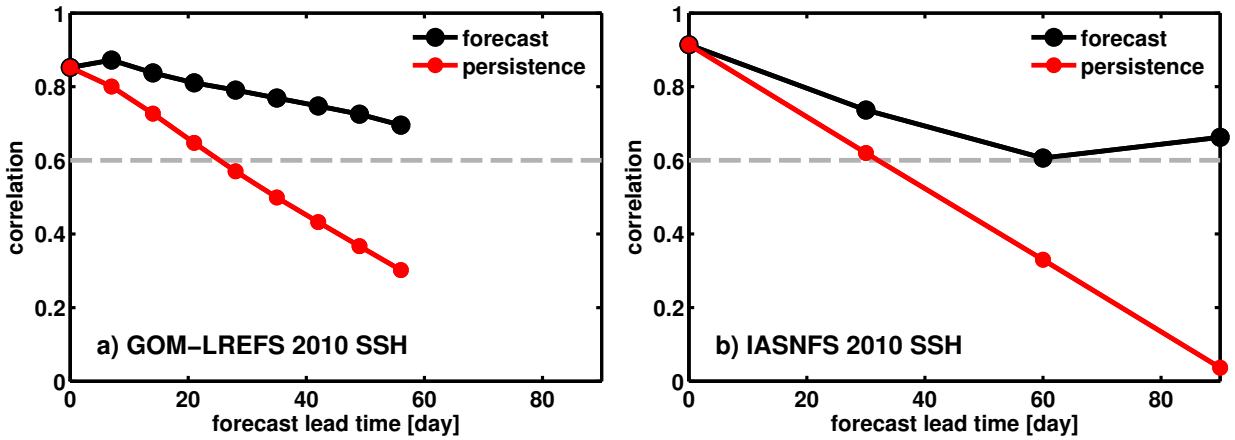


Figure 9. Correlation with AVISO Analyses, 2010. Correlation between anomaly SSH of forecast and analysis as a function of forecast time is shown for (a) GOM-LREFS and (b) IASNFS for the 2010 time period.

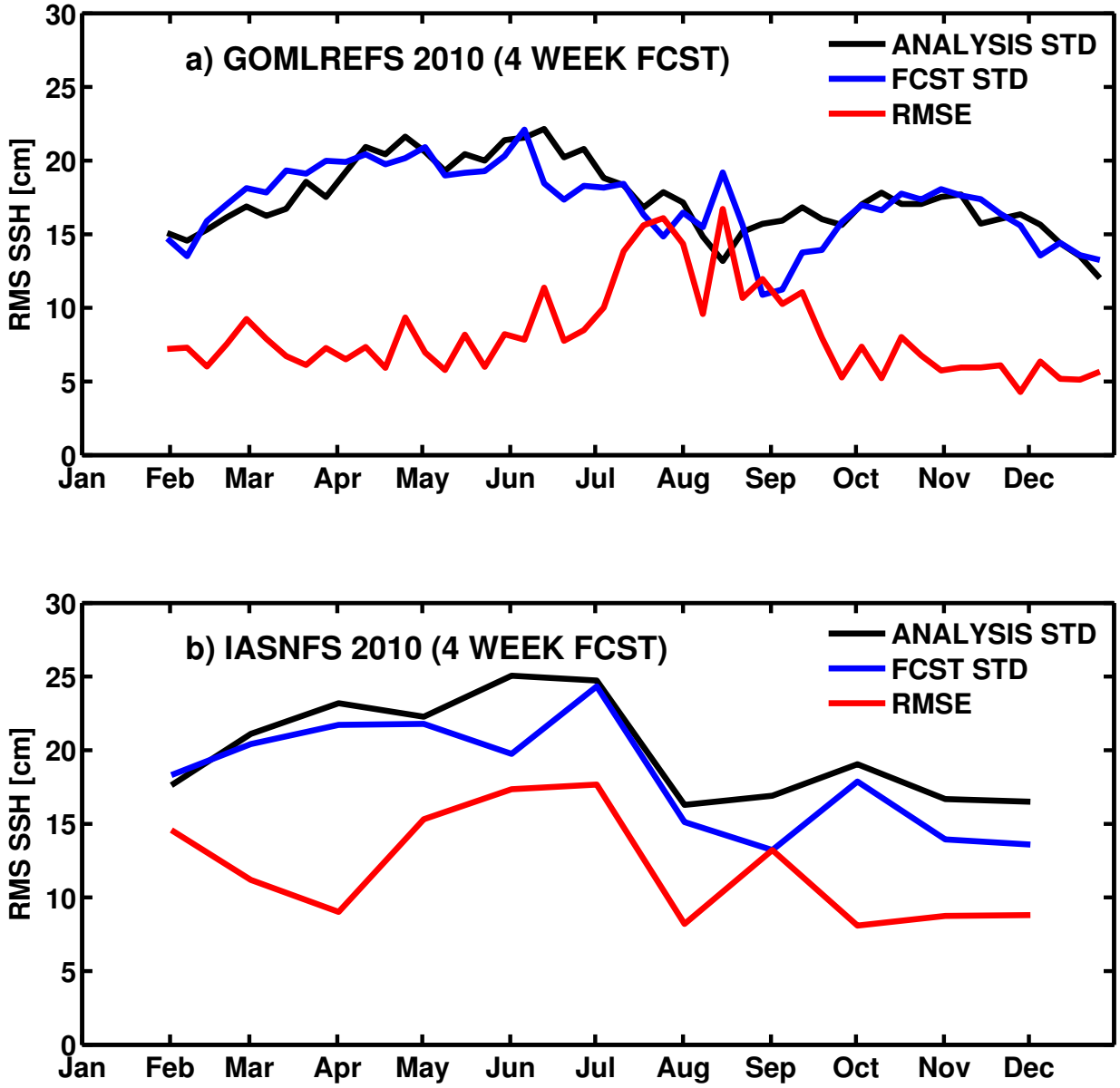


Figure 10. Performance at 4-week Lead Time, 2010. The spatial variance of anomaly SSH is used to compare the self-verifying analysis (black), the forecast at 4-week lead time (blue), and the forecast error with respect to the self-verifying analysis (red). The RMSE of IASNFS (12.5 cm) is larger than the RMSE of GOM-LREFS (8.8 cm), and the growth of error in IASNFS begins earlier in the year than in GOM-LREFS.

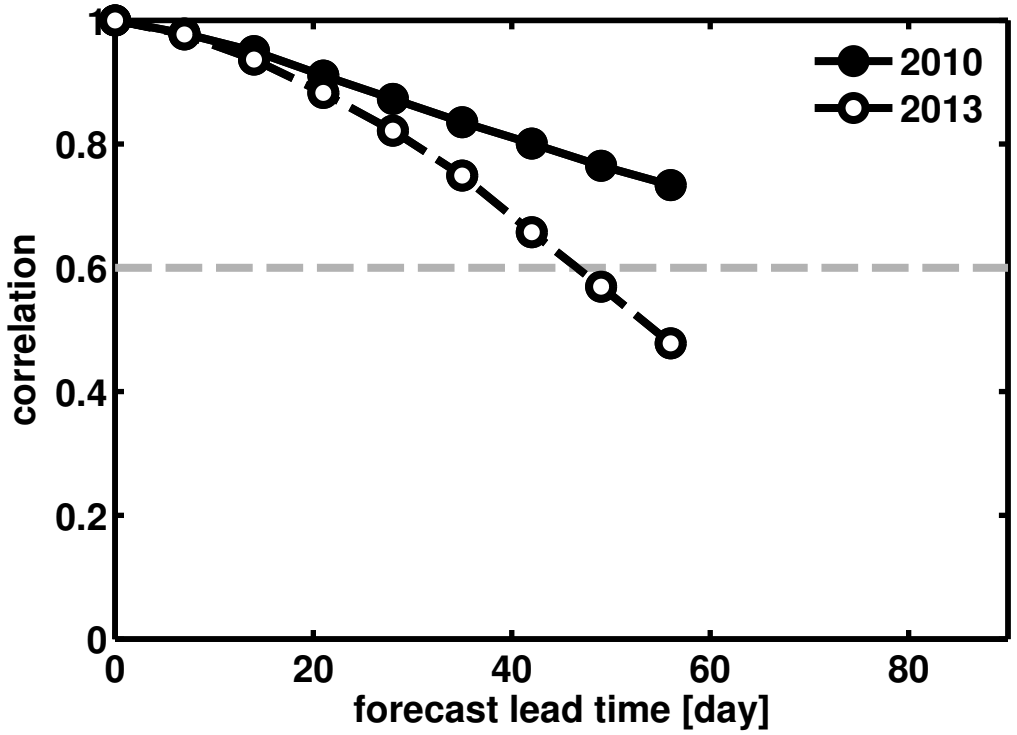


Figure 11. GOM-LREFS Forecast Skill, 2010 versus 2013. Forecast skill, represented by SSH anomaly correlation over the same subdomain used above, is shown for the 2010 retrospective forecasts (solid) and the 2013 near-real-time forecasts (dashed). Model forecast skill differed greatly in the 2 years, with nearly constant skill loss rate in 2010 and accelerating skill loss rate in 2013.

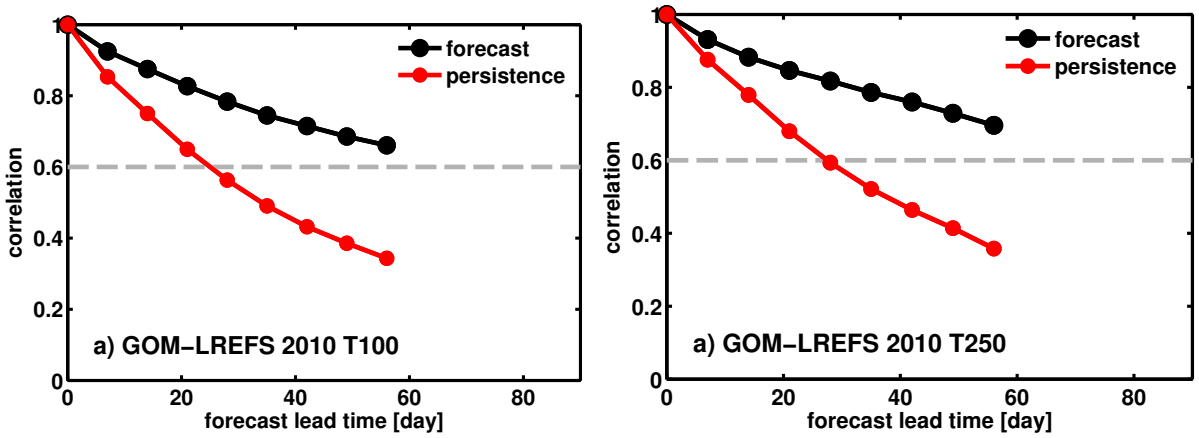


Figure 12. GOM-LREFS Forecast Skill, 2010. Anomaly correlation for temperature at 100 m (T100, left) and 250 m (T250, right) is predictable out to 60 days in 2010. Comparison with Figure 16, which shows anomaly correlation during 2013, demonstrates the significant positive increment for model skill in 2010 that occurred for temperature as well as SSH.

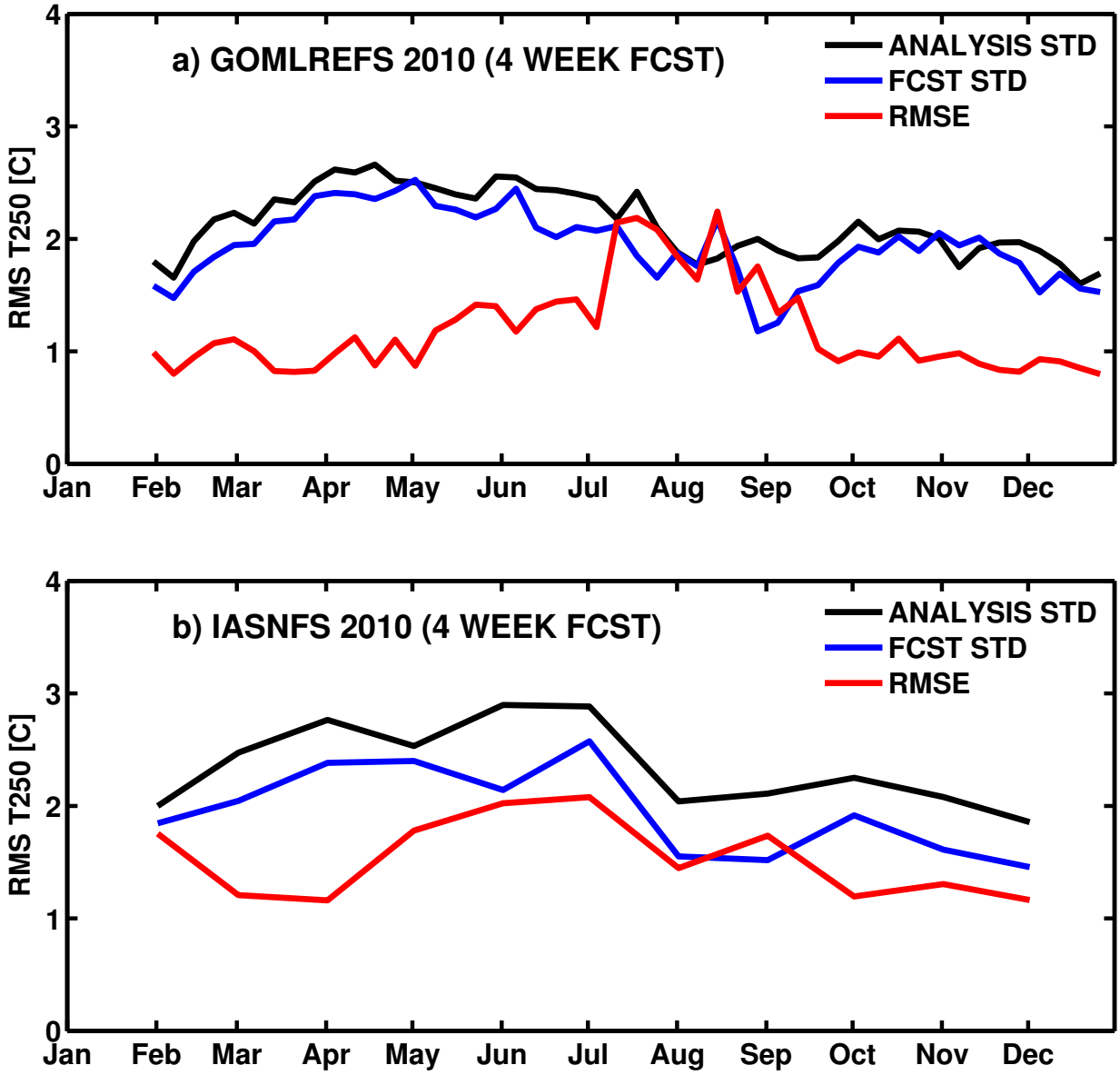


Figure 13. Performance at 4-week Lead Time, 2010, Temperature at 250 m (T250). Compare with Figure 30, which shows 4-wk forecast statistics for SSH. During most of 2010 (from February through July, and October through December), the RMSE of GOM-LREFS is approximately 1 C, whereas the RMSE of IASNFS is approximately 1.3 C, or 30% worse. However, during the July to September period, the GOM-LREFS RMSE exceeds the IASNFS RMSE.

An additional difference is noted in the standard deviation (STD) of the analysis. The discrepancy between the forecast and analysis STD is noticeably larger in IASNFS compared with GOM-LREFS.

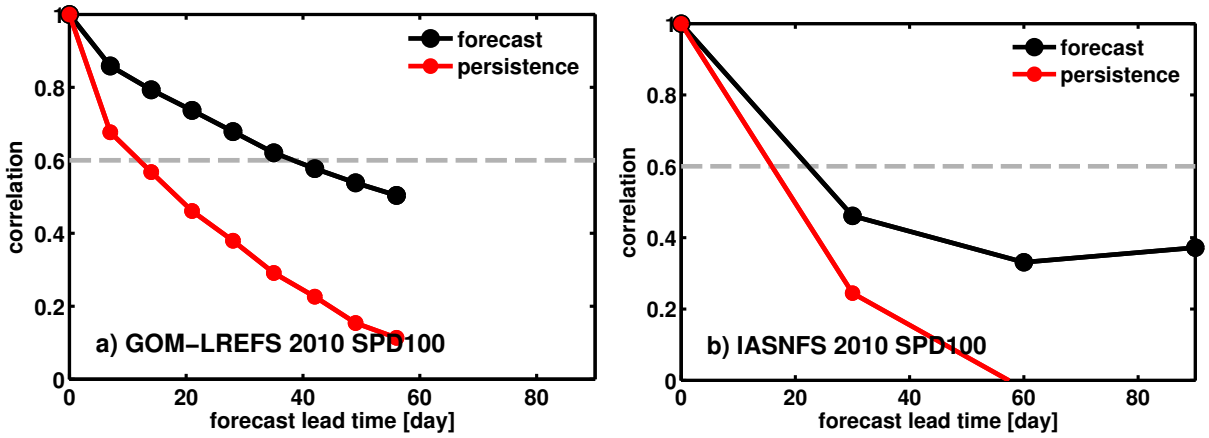


Figure 14. Anomaly correlation of current speed at 100-m depth (SPD100) during 2010. Correlation between anomaly SPD100 of forecast and analysis as a function of forecast time is shown for (a) GOM-LREFS and (b) IASNFS for the 2010 time period. Anomaly SPD is defined with respect to both the time-average of SPD100 within the subdomain, 22-28N 89-83E with depth greater than 100 m.

The enhanced forecast skill of GOM-LREFS is evident. Current speed is predicted with useful skill out to 40 days in GOM-LREFS, and there is a large excess of skill compared to persistence. In contrast, IASNFS loses all useful skill before day 25, and IASNFS shows a smaller skill increment compared to persistence.

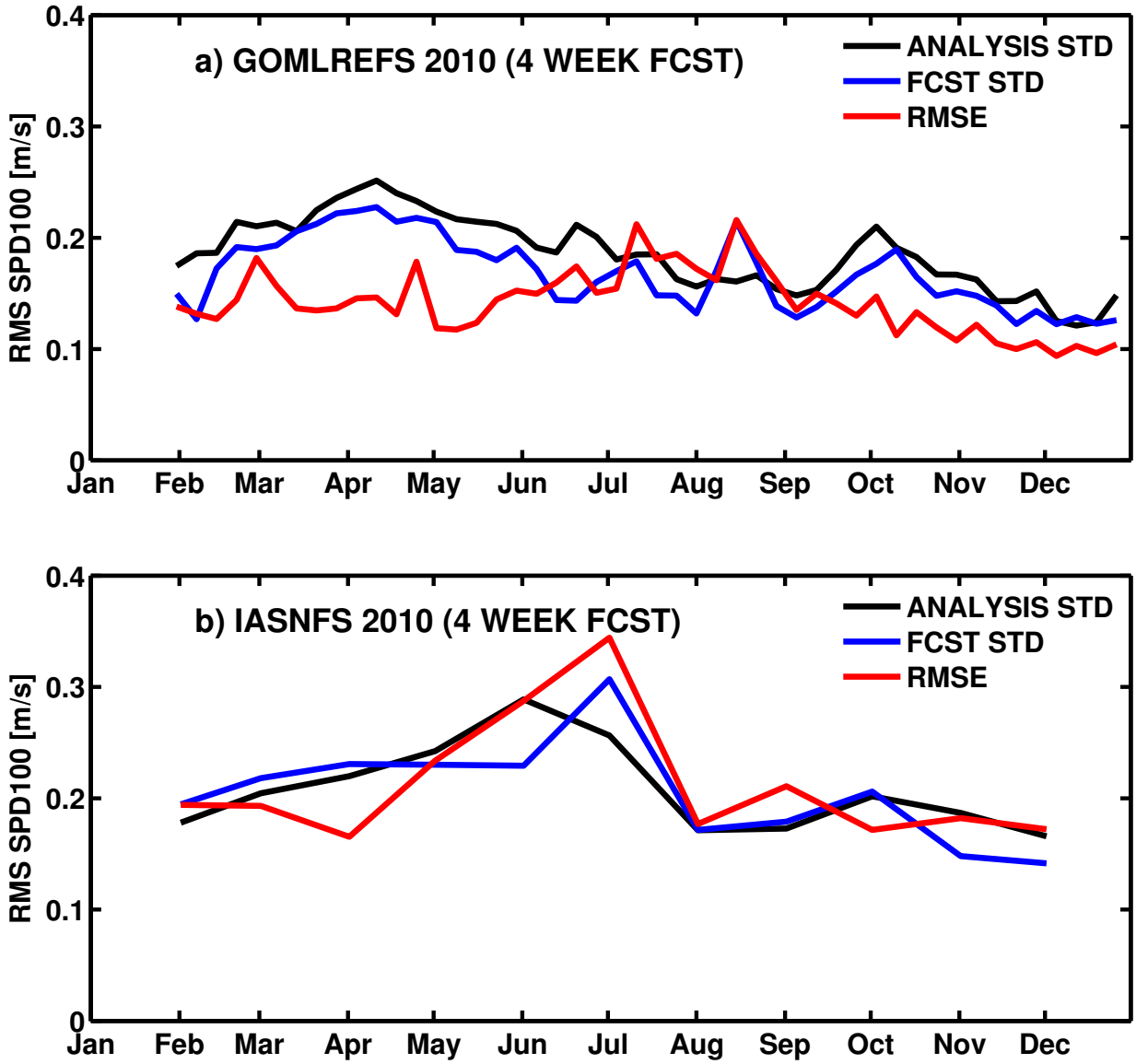


Figure 15. Performance for SPD100 at 4-week Lead Time, 2010. The spatial variance of anomaly 100-m current speed (SPD100) is computed for the self-verifying analysis (black), the forecast at 4-week lead time (blue), and the forecast error with respect to the self-verifying analysis (red). The RMSE of IASNFS (0.2 m/s) is larger than the RMSE of GOM-LREFS (0.15 m/s), and the growth of error in IASNFS begins earlier in the year than in GOM-LREFS.

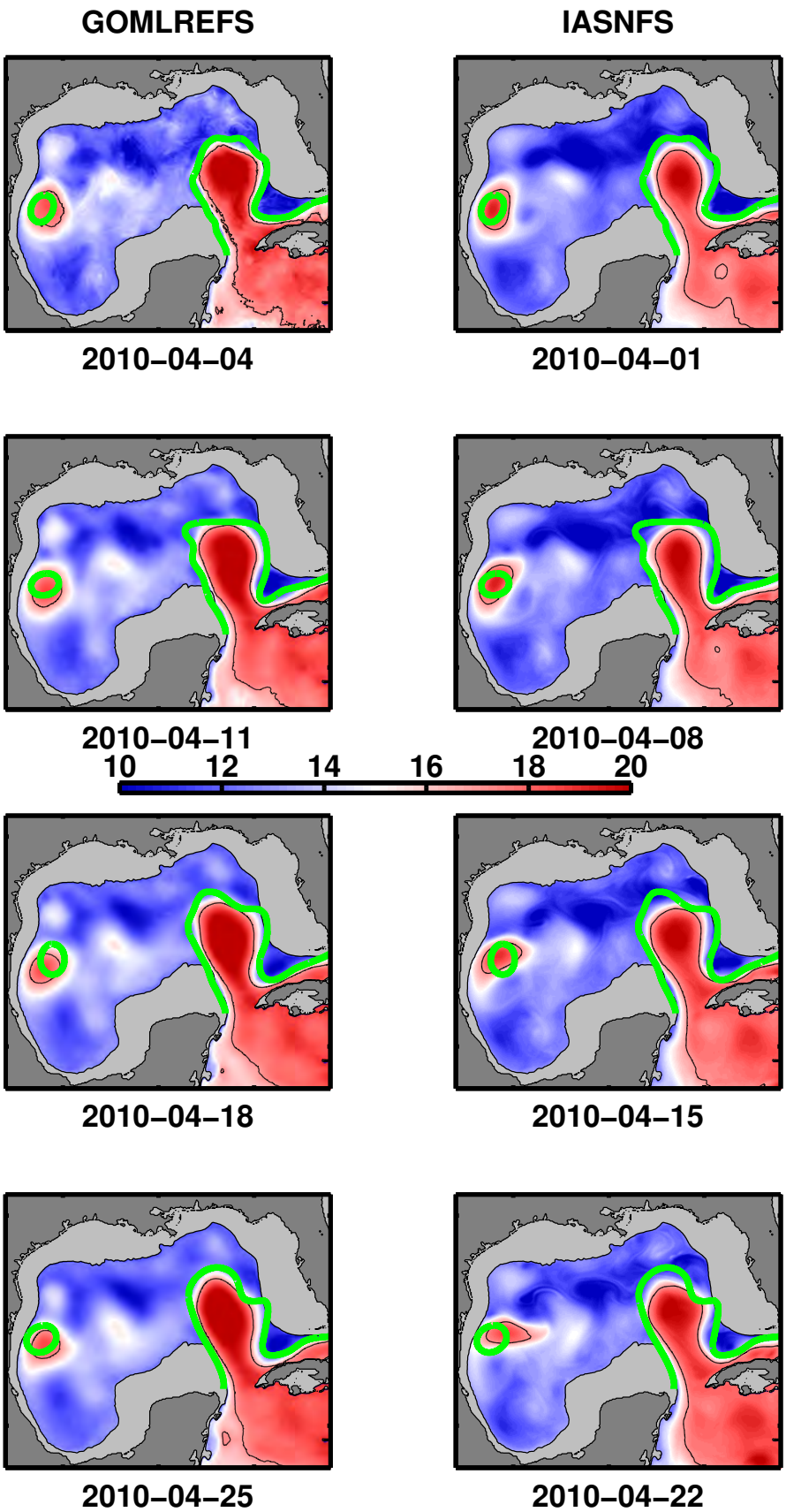

GOMLREFS
IASNFS
2010–04–04
2010–04–01
2010–04–11
2010–04–08
10
12
14
16
18
20
2010–04–18
2010–04–15
2010–04–25
2010–04–22

Figure 16 (previous page). GOM-LREFS and IASNFS comparison to HMI-EW frontal analyses during April 2010. Panels show the temperature at 250-m depth (black contour shows 18C isotherm) in relation to HMI-EW (solid green contour), during April 2010, a time period when both models had lower-than-average RMSE for the 4-week forecast (cf., Figure 33). The top panel (dated 2010-04-04 for GOMLREFS and 2010-04-01 for IASNFS) shows the initial analysis, and subsequent rows show the forecast at 1-week intervals. During this time period both models track the HMI-EW front with good accuracy, and GOM-LREFS has visibly better agreement with the HMI-EW analysis.

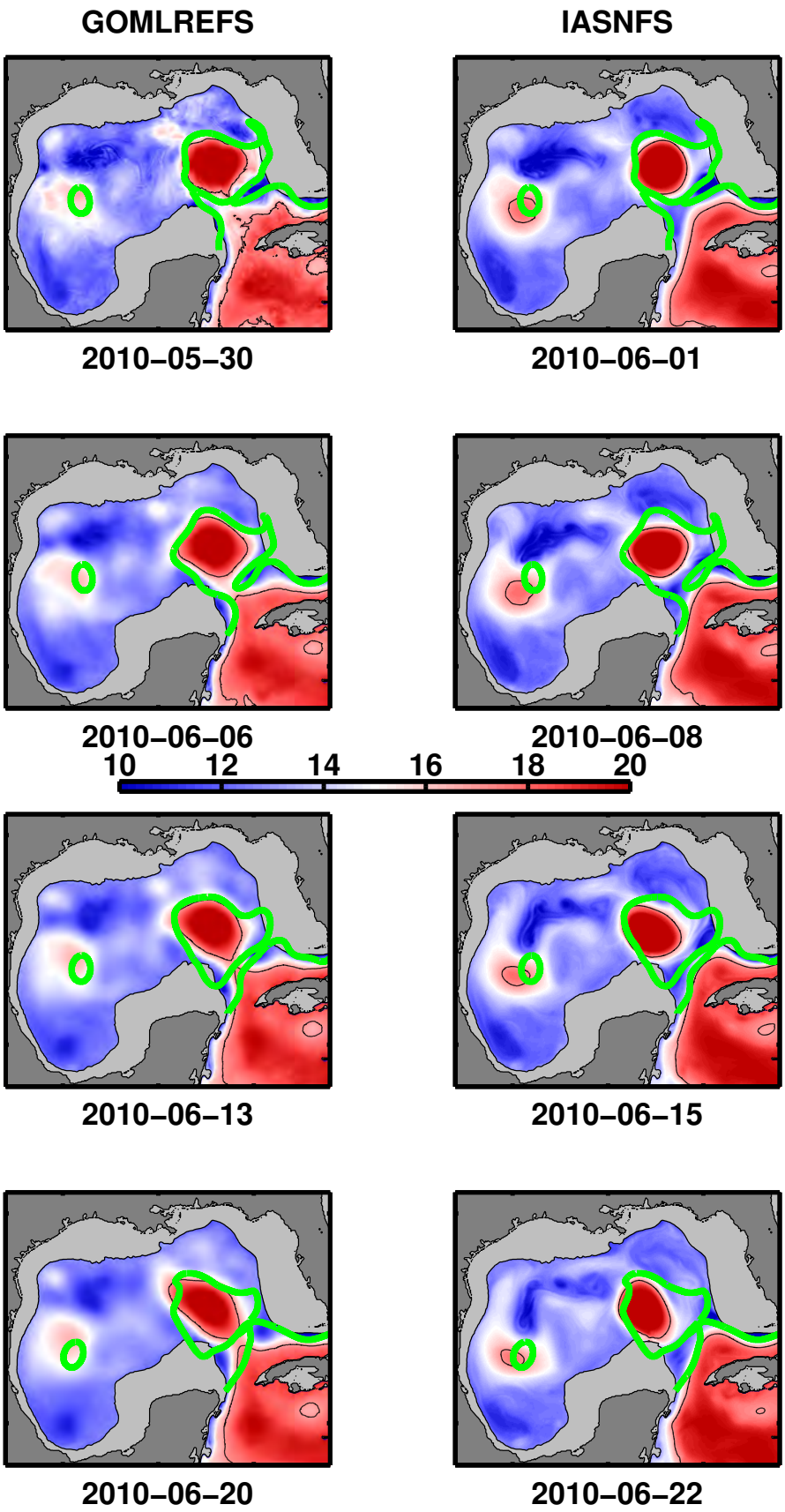

GOMLREFS
IASNFS
2010–05–30
2010–06–01
2010–06–06
2010–06–08
10
12
14
16
18
20
2010–06–13
2010–06–15
2010–06–20
2010–06–22

Figure 17 (previous page). GOM-LREFS and IASNFS comparison to HMI-EW frontal analyses during June 2010. Panels show the temperature at 250-m depth (black contour shows 18C isotherm) in relation to HMI-EW (solid green contour), during June 2010, a time period when RMSE for the 4-week forecast was increasing (cf., Figure 33). The top panel (dated 2010-05-30 for GOM-LREFS and 2010-06-01 for IASNFS) shows the initial analysis, and subsequent rows show the forecast at 1-week intervals. During this time period GOM-LREFS initially agreed well with the HMI-Ew analysis but subsequently diverged from it. Conversely, the frontal location was not well-represented in IASNFS, but subsequent forecasts did a good job of predicting the location of its northwest boundary.

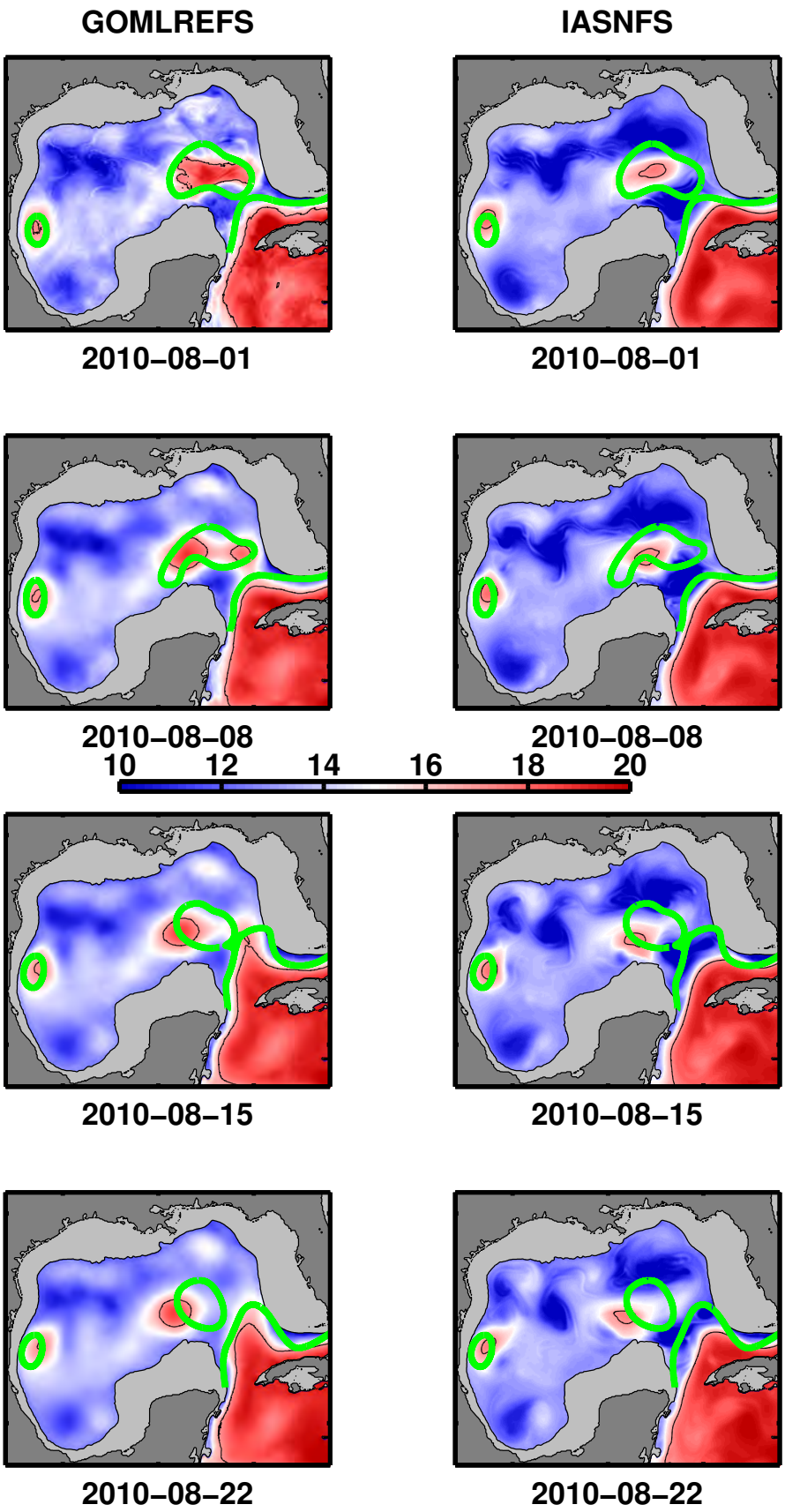

GOMLREFS
IASNFS
2010–08–01
2010–08–01
2010–08–08
2010–08–08
10
12
14
16
18
20
2010–08–15
2010–08–15
2010–08–22
2010–08–22

Figure 18 (previous page). GOM-LREFS and IASNFS comparison to HMI-EW frontal analyses during August 2010. Panels show the temperature at 250-m depth (black contour shows 18C isotherm) in relation to HMI-EW (solid green contour), during August 2010, a time period when RMSE for the 4-week forecast was large in both models (cf., Figure 33). The top panel (dated 2010-08-01 for GOM-LREFS and 2010-08-01 for IASNFS) shows the initial conditions, and subsequent rows show the forecast at 1-week intervals. During this time period GOM-LREFS appeared to have a much better initial analysis than IASNFS; however, the frontal evolution occurred differently than predicted by both models.

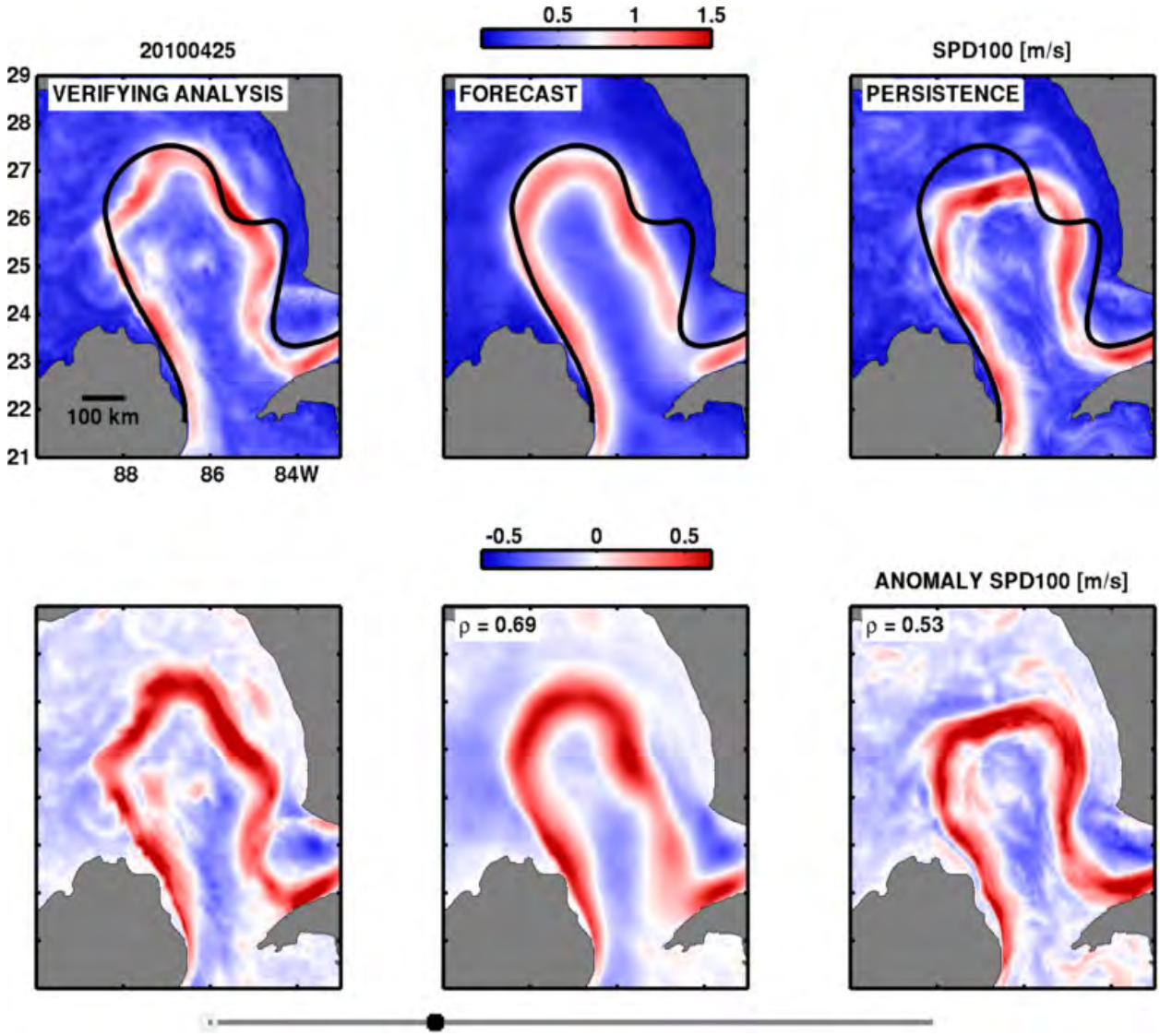


Figure 19. Current speed at 100-m depth (SPD100). The figure (taken from an animation of GOM-LREFS) shows the 4-week forecast currents in the LC region in relation to the HMI-EW analysis (black contour). The strongest currents in the verifying analysis for 2010-04-25 (top left) are similar to the HMI analysis. The 4-week forecast (top center) is similar; although, the detailed locations of the undulations on the LC are not represented. The northwestern frontal location is well represented in the forecast (top center) compared to persistence (top right). Bottom panels show the corresponding anomaly fields, where the time-mean of the SPD100 has been removed, and show the good correlation of the forecast and the analysis (rho = 0.69). Note that this snapshot is during the time period when GOM-LREFS forecasts are highly accurate (cf., Figure 30a).

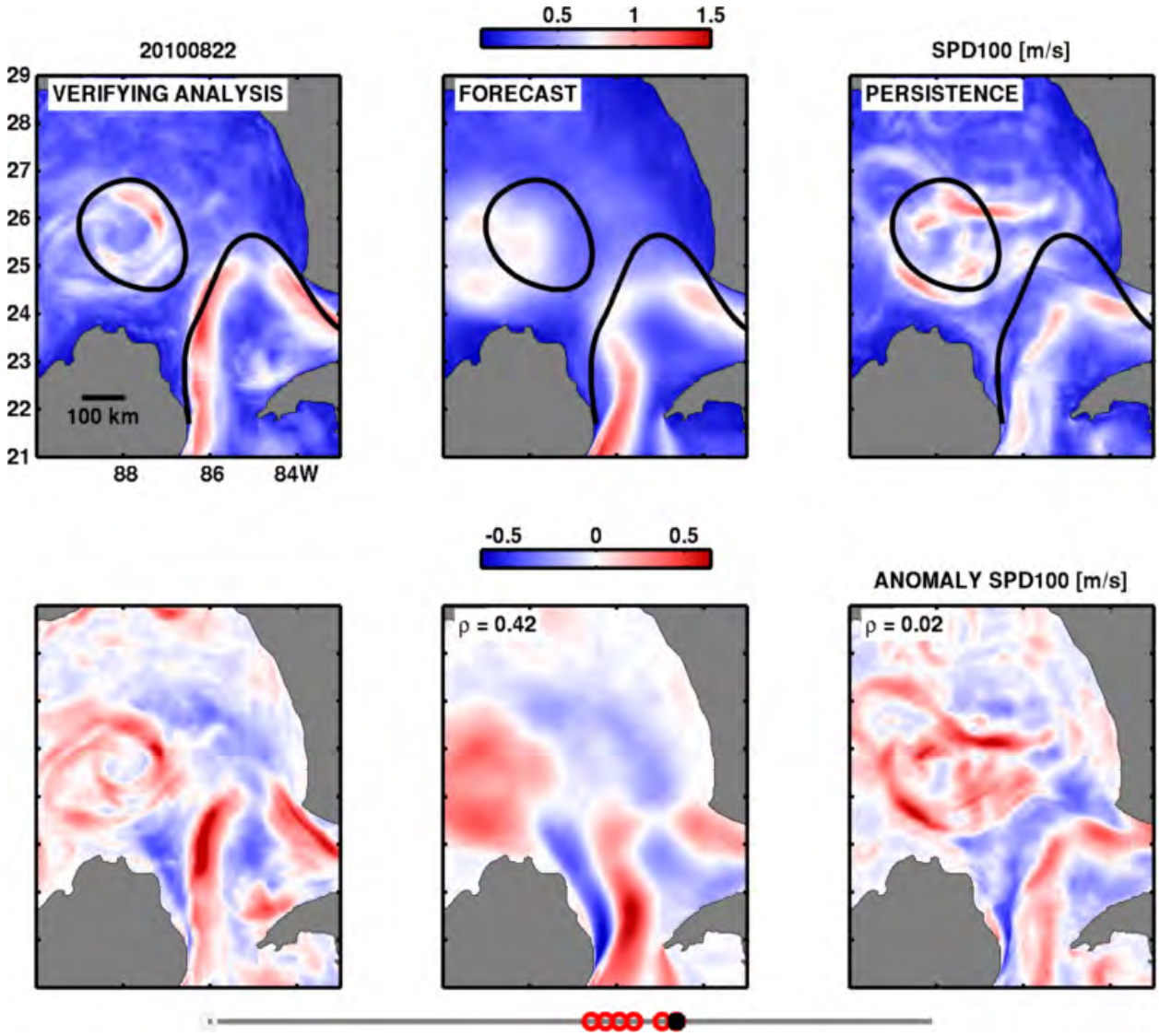


Figure 20. Current speed at 100-m depth (SPD100). As in Figure 39, except the snapshot is for the date 2010-08-22 when GOM-LREFS skill is low (cf., Figure 30a).

Appendix I: GOMEX-PPP Technology Transfer

Background

RPSEA requires that at least 2.5% of the total funding (i.e., ca. $39K, which is rounded up to $40K for convenience below) be assigned to technology transfer.

Activities

The technology transfer activities of GOMEX-PPP are comprised of several elements:

(NOTE: * designates a RPSEA mandate.)

*(1) attend the briefings of RPSEA TAC thru UDW Meetings every four months (YR1/Q2, YR1/Q3, YR1/Q4, YR2/Q2, YR2/Q3, YR2/Q4, YR3/Q2, and YR3/Q3);

(2) make presentations at annual GCOOS-RA and/or SECOORA meetings in YR1/Q2, YR2/Q2, & YR3/Q2;

*(3) present peer-reviewed papers at annual MTS/IEEE OCEANS Meetings in YR2/Q2 and YR3/Q2 and OTC Meetings in YR2/Q1 and YR3/Q1;

(4) submit peer-reviewed papers to the *Journal of Geophysical Research*, *Journal of Operational Oceanography*, and/or *Marine Technology Society Journal* at the end of Phase I (YR2/Q2) and again at the end of Phase II (YR3/Q2);

(5) conduct a GCOOS-RA & SECOORA Stakeholder & Super-User Workshop in YR3/Q1;

(6) maintain and enhance the GOMEX-PPP Website via the GCOOS Website, continuing from YR1/Q1; and

(7) provide a final technical report with recommendations for an operational ocean prediction system for the Gulf of Mexico circulation in YR3/Q2.

Summary

The status of technology transfer activities through Phase I is summarized below:

(1) Accomplished.
(2) Accomplished for GCOOS-RA but not SECOORA because there is enough overlap and cross-talk between GCOOS-RA and SECOORA to make it unnecessary to brief SECOORA separately.
(3) Was not accomplished for OCEANS and OTC Meetings; however, instead (for stronger peer review), nine oral & poster presentations (see Appendix II) were made at the American Meteorological Society (AMS) and The Oceanography Society (TOS), American Geophysical Union (AGU), and Association for the Sciences of Limnology and Oceanography (ASLO), Annual & Biennial, resp., Meetings in 2012/Q1.

(4) An overview manuscript covering Phase I is being planned for the Bulletin of the American Meteorological Society (BAMS) now that the Phase I results are in hand; a follow-on manuscript is planned for the end of Phase II.
(5) The Stakeholder & Super User Workshop is still planned for late in Phase II.
(6) The GOMEX-PPP Website is maintained at TAMU and is synergistically linked to the GCOOS-RA Website. It is steadily evolving into a real-time Web-Portal, a Phase II objective.
(7) The final technical report with recommendations regarding an operational ocean prediction system for GOMEX is a Phase II objective.

Peer-Reviewed Publications:

(1) E. D. Zaron, P. J. Fitzpatrick, S. L. Cross, Harding, J. M., F. L. Bub, J. D. Wiggert, D. S. Ko, Y. Lau, K. Woodard, and C. N. K. Mooers, 2014: Initial evaluations of a U.S. Navy rapidly relocatable Gulf of Mexico/Caribbean ocean forecast system in the context of the Deepwater Horizon oil spill disaster, *Ocean Modelling*, in review.

(2) Chang, Y.-L., and L.-Y. Oey, 2012: Why does the Loop Current tend to shed more eddies in summer and winter? *Geophys. Res. Lett.*, 39 (5), L05605, doi:10.1029/2011GL050773.

(3) Zhao, Y. and He, R., 2012: Cloud-free Sea Surface Temperature and Color Reconstructions for the Gulf of Mexico: 2003-2009, *Remote Sensing Letters*, 3 (8), 697-706.

Presentations at Scientific Meetings:

Mooers, C. N. K., E. D. Zaron, and P. J. Hogan. 2014. A Pre-Operational Prediction System for the Gulf of Mexico. Presented at the AMS Annual Meeting in Atlanta, GA, February 3, 2013.

Zaron, E. D., P. J. Hogan, and C. N. K. Mooers. 2013. The Gulf of Mexico Pilot Prediction Project: Evaluation of 60-day Ensemble Forecasts of the Loop Current System. Presented at the GODAE OceavView Symposium, College Park, MD, November 4-6, 2013.

Mooers, C.N.K. and E. Zaron. 2012. Multi-Model Comparisons of the Mesoscale Circulation in the Gulf of Mexico in 2010. Presented at the AMS Annual Meeting, 22-26 January 2012, New Orleans, LA. (a contribution to the embedded AMS 10th Symposium on the Coastal Environment).

AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting 2012 held in Salt Lake City, Utah, 20 to 24 FEB 12:

The presentations below comprised one-half of those made in the Special Session on Gulf of Mexico Circulation and Ecosystem Numerical Modeling.

Chang, Y. L. and L.-Y. Oey. 2012. Why Does the Loop Current Have Seasonal Preferences for Shedding Eddies? Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Farrara, J.D., Y. Chao, Z. Li, X. Wang, H. Zhang, P. Li, R. He, and H. Qian. 2012. A ROMS-based Data Assimilating Ocean Forecast System for the Gulf of Mexico. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Gopalakrishnan, G., B. Cornuelle, I. Hoteit, D. Rudnick, and W. Owens. 2012. State Estimates and Forecasts in the Gulf of Mexico. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Howard, M.K., E. Zaron, C. Mooers, Y. Chao, B. Cornuelle, R. He, D. Ko, L. Oey, A. Mehra, and R. Patchen. 2012. Gulf of Mexico Pilot Prediction Project (GOMEX-PPP): Model-Data Comparisons. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Ko, D.S. 2012. A Long-Term Ocean Forecast Experiment for Gulf of Mexico Applying IASNFS. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Mooers, C.N., E.D. Zaron, M. Howard, Y. Chao, B. Cornuelle, R. He, D.S. Ko, L. Oey, A. Mehra, and R. Patchen. 2012. The Gulf of Mexico Pilot Prediction Project (GOMEX-PPP). Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Wiggert, J.D., J.M. Harding, F.L. Bub, P.J. Fitzpatick, and K.C. Woodward. 2012. Evaluation of the AMSEAS Gulf of Mexico/Caribbean Regional Forecast System: A SURA Super-Regional Modeling Testbed Activity. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Xue, Z., R. He, K. Fennel, W. Cai, and S. Lohrenz. 2012. Modeling Seasonal and Interannual Variability of Circulation and Biogeochemical Processes in the Gulf of Mexico. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Zaron, E.D., C.N. Mooers, M.K. Howard, Y. Chao, B. Cornuelle, R. He, D.S. Ko, A. Mehra, L.Y. Oey, and R. Patchen. 2012. Gulf of Mexico Pilot Prediction Project (GOMEX-PPP): Forecast Skill and Model Intercomparisons. Presented at the AGU, TOS, & ASLO Sponsored Ocean Sciences Meeting, 20-24 February 2012, Salt Lake City, UT.

Appendix II: Concept of Operations

A Perspective on CONOPS, provided by Dr. Frank Bub, formerly of NAVO

To be answered: who, what, where, when, how, and why?

Note: while the GOMEX-PPP work concentrates on the Gulf of Mexico, NAVO must be able to apply this capability both there and in other parts of the world that are of Navy interest.

1. Introduction (what & why)
    a. Background
    b. Description (resolution, physical, ecological, forecast horizon, etc.)
    c. Justification
    d. Transition: Research-to-Operations
    e. Agreements with OGA (other Government agencies) such as DOE, NOAA, NASA
    f. Responsibilities

2. Running the LREFS at NAVO (what, how, where, when, who)
    a. Software
        i. Description
        ii. Required libraries, etc.
        iii. Required scripts
        iv. Interface with NAVOCEANO Model Operations (ROAMER)
    b. Cost estimates
        i. Computational requirements – FLOPS
        ii. Personnel time to set up, monitor, repair, consult
    c. Data management – volumes, file sizes, etc.
        i. Inputs – (1) observations, (2) atmospheric forcing, (3) river data, (4) bathymetry, (5) what else?
        ii. Internal storage (intermediate data, scratch)
        iii. Outputs – (1) data files (NetCDF KML, other formats), (2) graphics (JPEG, GIF, etc.), (3) derived products (ex: text file of the 1.5 knot isotach)
    d. Computational requirements (above will be a cost estimate, this will be the DSRC footprints)
        i. Processors
        ii. Run wallclock time (pre-, main, post-)
        iii. Run start/stop/delivery times
        iv. Storage
        v. Monitoring and interface to ROAMER (our model monitoring tool)
    e. Personnel requirements (our goal is automated operations with personnel involved only if there is a problem)

        i. NAVOCEANO TPOC with NRL
        ii. Navy Administrative Modeling Oversight Panel (AMOP) transition steps
        iii. Movement of system from NRL R&D to NAVOCEANO operations
        iv. Monitoring of daily runs
        v. Expert maintenance and troubleshooting
        vi. External interactions and consultations
            1. Difficulties with deliveries
            2. Responses to questions, problems
    f. Documentation
        i. Usual requirements: Software design document (SDD), software users manual (SUM), validation test report (VTR), Operational Evaluation (done by NAVOCEANO)
        ii. WIKI-based operations manual, including troubleshooting instructions

3. Delivery of products to external (non-Navy) users (what, how)

    a. There are no direct Web services from NAVOCEANO
        i. Could use NRL or NOAA
        ii. Long-term data storage at OceanNOMADS
        iii. This will have to be arranged
    b. What products?
        i. What fields?
            1. Currents, temperature, salinity, density, etc.
            2. Multi-level (daily for some, hourly for "zooming in")
            3. Full and/or decimated grids
            4. Depths (only the surface?)
            5. Presentation of uncertainty
        ii. Time steps
        iii. Data, graphics, etc.
        iv. Derived products: need to relate to applications
    c. Legal issues
        i. Navy will assume no liability
        ii. Describe the proper versus improper use of the output

4. Timeline (steps from R&D to OPS)

A Perspective on Operational Ocean Models, provided by Dr. C. N. K. Mooers, Portland State University

Attributes of operational ocean models (in contrast to R&D ocean models):

- Certified to meet (evolving) user (e.g., human forecaster) requirements within known error bars
- Sustained sponsorship for real-time (relative to timescale of applications) forecast delivery
- Robust, resilient, and redundant operations

Ocean forecasting systems:

- Four component subsystems: modeling, observing, data assimilation, and data management
- Observing subsystem must also serve needs of verification and validation schemes
- Operational observing sub-systems need subsurface (e.g., CTD and velocity profilers) as well as surface components
- Forecasting system may be optimized w/a continually operating ("strategic") system supplemented with a rapidly-deployable ("tactical") system for more precise & focused state estimations

The testbed concept:

- Use of multiple models is proving desirable in forecasting environmental systems due to uncertainties about the physics, initial and boundary conditions, air-sea forcing, etc.
- Next steps in the joint community of R&D and operational forecasters may be sustained, multi-model regional testbeds for providing on-call prediction systems, conducting prediction experiments, evaluating observing system designs, cultivating human resources, etc.
- Cyberinfrastructure may play critical roles in facilitating forecast systems and their utilization

The broader ocean observing and modeling community:

- GCOOS-RA does not plan to support an OGCM for GOMEX in the next few years
- However, a GCOOS-RA modeling task team is being formed to represent a broad array of multi-disciplinary application model types (which will generally need OBCs from one or more OGCMS)
- GCOOS-RA will also increase awareness among civil sector users of Navy & NOAA operational models, and R&D GOMEX models through its web pages

Appendix III: Subcontractor's Final Report

The Final Report submitted by P. Hogan and collaborators from the Naval Research Laboratory is attached on the following pages.

NRL Final Report

Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP) Phase II

Extended Range Ensemble Forecasting in the Gulf of Mexico

July 22, 2014

Patrick Hogan (Naval Research Laboratory,
Patrick.hogan@nrlssc.navy.mil)
Emanuel Coelho (University of New Orleans,
emanuel.coelho.ctr.po@nrlssc.navy.mil)
Germana Peggion (University of New Orleans)
germana.peggion.ctr.po@nrlssc.navy.mil)
Prasad Thoppil (Naval Research Laboratory)
Prasad.Thoppil@nrlssc.navy.mil)

Submitted to C.N.K. Mooers, Research Professor (cmooers@cecs.pdx.edu)
Principal Investigator of the GOMEX-PPP

Abstract

An extended range (~60 days) ocean forecasting system based on probabilistic (ensemble) methods has been developed for the Gulf of Mexico under funding from the Research Partnership to Secure Energy for America (RPSEA) through a subcontract with Portland State University. A probabilistic prediction has the advantage of providing a forecast as well as the uncertainty about that forecast. Both provide information and guidance for a wide range of applications in the public/industrial sector. The emphases here are on applications for the oil and gas industry. The forecast system uses a set of possible forecast states (ensemble) generated by perturbing the initial state of the ocean field as well as the surface atmospheric forcing (winds and heat fluxes). A multi-year time series of output from the Global Ocean Forecast System (GOFS 3.0) is used to construct a daily climatology for 3-D lateral boundary conditions. A similar method is used to construct surface boundary conditions from the Central America Coupled Ocean Atmosphere Mesoscale Prediction System (COAMPS) (Hodur, 1996) atmospheric forecast model. The premises are to accept that the analyses are to be considered as "truth", since they ingest all available observations. Estimates of forecast uncertainty calculated over this 32 member ensemble have demonstrated skill in predicting Loop Current Eddy shedding events that occurred in the spring of 2010, 2011, 2013 and perhaps 2014 (at the time of this writing). The ensemble forecast system delivers a short range 24-hour forecast every day and a 60-day long range forecast once per week and has been running in real-time since January 2013. Anomaly correlations have been calculated and used to quantify the forecast skill relative to the best corresponding analyses for the prognostic model variables at several depths. Output from the system is used to derive risk assessment products such as risk of strong currents or risk of combined error in the forecast. These are calculated from probability distribution functions over the length of the forecast and can be used for planning, risk mitigation, search and rescue, etc.

Introduction

Ensemble-based probabilistic forecasts are superior to single run deterministic forecasts because they provide flow-dependent uncertainty along with improved best guess solutions (e.g. Dobla-Reyes et al, 2005). Because they are resource intensive, ocean ensemble forecasting and uncertainty estimation has until now been restricted to relatively small area applications and short range forecasts (e.g. Coelho et.al, 2008). The complexity of the deep sea operations involved in supporting operational oil rigs in the northern Gulf of Mexico requires a longer range prediction system to allow planning for challenging tasks and to mitigate risk of adverse environmental conditions. As an example, the shear associated with strong Loop Current eddy activity affects day-to-day operations during offshore oil and gas exploration and production activities, and can play an important role in the intensification of Gulf of Mexico Hurricanes. Thus, the ability to predict the Loop Current intrusion and eddy separation process can have profound socioeconomic implications.

Ensemble Modeling System

The Navy Coastal Ocean Model (NCOM) (Martin et al., 2009), Navy Coupled Ocean Data Assimilation System (NCODA) (Cummings, 2005) and COAMPS form the primary components of the ensemble forecast system (Figure 1). NCOM is a first principles (primitive equation) finite difference ocean model that is used operationally at the Naval Oceanographic Office (NAVO). The domain used in the present work is the entire Gulf of Mexico with 3 km horizontal resolution and 49 hybrid levels (33 sigma levels and 16 z-, or pressure levels). The vertical resolution is finest near the surface and the distance between levels progressively increases (becomes more coarse) with depth. The Mellor-Yamada level 2 turbulent closure scheme is used for vertical mixing. Tides are included as a lateral boundary condition and monthly mean transports are used for riverine input. The model is initialized from an analysis produced by the NCODA system, thus allowing the assimilation of all available observations via a 24-hour incremental update cycle.

NCODA is the data assimilation component, also used operationally at the NAVO. The version used here is based on a 3-D variational implementation (3-DVar). Observations used by NCODA include all operational observations available in near-real time. The primary sources are satellite altimetry and sea surface temperature (SST), but profile observations from ships, gliders, and floats are also used. The observations pass through a robust quality control algorithm that checks for outliers and other unrealistic anomalies. The Modular Ocean Data Assimilation System (MODAS; Fox et al., 1990) ingests concurrent surface observations of temperature and height, and is used to construct “synthetic profiles” that are projected downward through the water column and used to constrain the ocean interior. At the present time there is no assimilation of velocity observations by the system.

The first component of the 60-day forecast is the single (deterministic) forecast designated as the control run. The 24-hour forecast error variances from this control run are then used to generate perturbations to the model initial conditions through the Ensemble Transform (ET, Bishop and Toth, 1999) methodology. The method is used to run Monte-Carlo simulations in operational frameworks (e.g. Coelho et al., 2008). Here, the Ensemble Transform uses the NCODA derived analysis error variance to transform forecast perturbations from a prior Monte-Carlo ensemble into equally likely posterior analysis perturbations. The method finds distinct linear combinations of the prior forecast perturbations that preserve the best guess of the error covariance and impose that the posterior analysis error variance will be consistent with the new set of observations used for the model analysis (Coelho, UCUP). In addition, the initial condition perturbations generated by the forecast error variances are augmented by additional estimates of the model temporal variability, nowcast/analysis increments history, and climate variability (C. Rowley, personal communication) so that the ensemble model perturbations will have a

spread similar to the best guess of the control run analysis error variance. This approach is important, so that the distribution of ensemble solutions can encapsulate all detectable and dynamically relevant ocean states.

In addition to perturbing the initial conditions, the surface boundary conditions (the COAMPS wind and heat forcing described below) are also perturbed for the duration of the forecast via a space-time deformation technique. This technique relies on the fact that surface boundary condition errors are seldom due to time lags and/or displacement of dominant dynamical features already in the forecast. As such, independent random (perturbed) forcing fields for each ensemble member are generated with a random shifting technique every 24 hours with a specified de-correlation length. Restated, the forcing is prepared at the same 3 hour interval by linear interpolation of the forcing, but with the values computed at randomly shifted times (Coelho et al., 2008, Wei et al., 2013). This approach allows the system to use independent forcing for each member by accounting for uncertainty in the surface boundary conditions.

Climatological boundary forcing

A short range (< 1 week) ocean forecast typically uses forecast output from an atmospheric prediction system to drive the surface momentum and heat fluxes. The atmospheric model being used here (COAMPS) has forecast lengths of 5 days. Thus, for an extended range forecast (> 5 days), it was necessary to construct "climatological" forcing functions for the lateral and surface boundary conditions. For this purpose, 7 years of 3-hourly output (2005-2012) from the Central Americas COAMPS (Hodur, 1997) was used to generate an annual climatology at 3-hour intervals by averaging the several years analysis for each calendar day with 3-hour resolution (thus we refer to it as an annual climatology of high frequency forcing). This allows the climatological forcing to retain more synoptic character than monthly forcing (for example). At the start of 2014, we added 2013 to the time series used to construct the climatological atmospheric forcing. Subsequently, for each 60-day forecast, the forecast from Central Americas COAMPS was used for the first 5-days of the forecast, and then smoothly transitioned to the climatological forcing afterwards. A similar method was used for the lateral boundary forcing. For each 60-day forecast, the first 5-days were from global NCOM, and transitioned to a daily climatology that was constructed from 12-years of output from a global HYCOM reanalysis (1994-2005). [NCOM climatology was used prior to November 2013] In November 2013, the initial 5-days from NCOM were switched to the HYCOM-based GOFS 3.1, as the operational global model was changed from NCOM to HYCOM in November 2013 at NAVO.

Routine ensemble forecast products

The first real-time "production" 60-day, 32-member forecast was performed on January 27, 2013. Daily 24-hour and weekly 60-day forecasts have continued to run in real-time with few interruptions since then. Considerable effort went into writing the scripts and programs used to post-process the ensemble output, with emphasis on what quantities were of most value to the oil and gas industry. That post-processing is now all completely automated, and once a forecast is finished all of the plots and animations ("products") are automatically posted to a password protected Website. The most common calculations are the ensemble mean (if designed properly, the ensemble mean should produce a better forecast than a single deterministic model run), and the ensemble spread (or forecast variance representing the uncertainty, but formally the standard deviation calculated across the ensemble members). Derived products and plots can typically be added, removed, or modified with moderate effort. One caveat of the plots and animations is that although each weekly forecast is (exactly) 60-days in duration, the plots are generated from weekly output files, thus the "last plot" in an analysis/forecast sequence may be the last Sunday of the 60-day forecast. The current products linked to the Website include (both plots and animations) of the following:

Speed (surface, 500m, 1000m, 2000m) mean and variance
Salinity (surface, 200m, 500m, 1000m) mean and variance
Temperature (surface, 100m, 200m, 500m, 1000m) mean and variance
Sea Surface Height (10cm/17 cm contours) mean and variance
De-tided Sea Surface Height (10cm/17cm contour) mean and variance

The following paragraphs describe various examples of 2-month predictions of the Loop Current eddy separation obtained with the forecast system. The latest results from 2014, which cannot yet be confirmed, are described first, followed by the Eddy Kraken event in 2013. Both of these were run in real-time. A reanalysis covering the years 2010-2011 was also performed. Examples from both of those years are subsequently shown. In the following sections, ensemble mean and standard deviations are calculated across the 32 ensemble members.

Forecasts for the April-June, 2014 timeframe

For the latest 60-day forecast (April 20 (analysis) to June 15), weekly snapshots of de-tided ensemble mean SSH (color background) with the mean 17 cm SSH contour superimposed are shown in Figure 2. The corresponding ensemble variance (or spread), including the 17 cm de-tided SSH contour from each member is shown in Figure 3. In the snapshot of the SSH analysis (Figure 2a), the Loop Current penetrates into the northern Gulf to about 27.5N and strong cyclonic eddy activity is evident on the east side of the Loop Current extension, a state that often indicates the extension may become unstable and a LCE shedding event may occur soon. The weekly forecasts (Figures 2(a)-2(i)) show the evolution of the Loop Current eddy shedding process, although the actual forecast can not be confirmed at the time of this writing (late April, 2014).

The ensemble variance supports the strong possibility that a LCE shedding event is likely. At the analysis time (April 20) all ensemble solutions are very similar to the control run (by design). As the forecast progresses, the ensemble spread increases (the color backgrounds in Figure 3), and the individual ensemble members begin to evolve into dynamically realistic but (towards the end of the 60-day forecast) distinct states (the black contour lines). In this case, the ensemble mean (denoted by the green contour) exhibits a closed 17 cm contour on May 25 (Figure 3f), but the control run (denoted by the red contour) lags that by one week (Figure 3g). By the end of the 60-day forecast (Figure 3i), two distinct clusters of solutions associated with the multi-modal probability distribution of ocean states are evident. The more frequent cluster depicts LCE separation, while the less frequent cluster does not.

Figures 4 and 5 are similar (ensemble mean and spread for the 60-day forecast April 20 to June 15), but for surface speed (calculated from the individual velocity components). In Figure 4, the .77 m/s (1.5 kt) isotachs are superimposed (in these cases the isotachs encapsulate areas where speeds > 1.5 kt are colored in red). This threshold was chosen because current speeds higher than this, only associated with LCEs or extreme wind events, are detrimental to oil industry drilling operations. Similar to the SSH spread, Figure 5 depicts the ensemble surface speed spread. These plots look noisy because of the encapsulating nature of the 1.5 kt isotachs, but the same general character of Loop Current evolution that was seen in the SSH is evident here as well.

To neglect the impact of surface and near surface effects (winds, inertial and tidal oscillations, diurnal heating and cooling, etc.) subsurface eddy structure is often examined. One means to do this is to map the structure of 18C water at 200m. The ensemble mean temperature with the 18C isotherm superimposed is shown in Figure 6, and the corresponding ensemble spread in Figure 7. As is expected, the general features of the Loop Current are evident, as is the remnant of a LCE in the western Gulf

(near 95W, 25N). Similar to the SSH and surface speed, the water temperature at 200m shows the northward penetration of the Loop Current extension and strengthening of the cyclonic eddy activity on the east side of the LC, suggesting that conditions are favorable for a shedding event sometime in the May-June timeframe of 2014.

Ensemble likely mode estimation

In ensemble ocean forecasting, the ensemble mean is often used as a proxy for the "best" forecast, and although typically more representative than a single deterministic forecast (the control run for example) this is not always the case, especially in dynamically complex regions. In such cases, a maximum likelihood estimate can be used to extract the most likely state for a given forecast, which can give a more accurate forecast than any of the ensemble members, the ensemble mean, or the control run (deterministic) forecast. Methods to do this are relatively straightforward for probability distribution functions with relatively few dominant modes (for example) and quite complicated for joint probability distribution functions of high dimensionality. The Gulf of Mexico is closer to the former description, and we have implemented a method for extracting the likely mode from the ensemble population.

The methodology is as follows. An ensemble set is defined with N (32+1 in this case) members that includes all possible realizations. The Ensemble Mode Estimate is then defined as the mean of the K (adjustable, but 5 in the present case) members that are closest to the dominant mode peak. To identify the subset of these K elements, an iterative method is used that starts by computing the mean of the population N and then selects the N-I members that are closest to this mean. The process is then repeated n times until N-n*I (where n is the number of iterations) is smaller than K. I is presently taken as 4, so that each iteration discards the four members that are further away from the iteration mean. The "distance" of each run from the mean is computed using a Root Mean Square difference between each member and the iteration mean for each variable (T-S-U-V and Elev.). These are then normalized by the variable standard deviation computed through the full ensemble mean. The final "distance" is then computed by a weighted average over the individual variables distance. The present scheme uses the same relative weight of 0.25 for T-S and Elevation and 0.125 for U and V.

An example of the application of this method is shown in Figure 8. The mean de-tided SSH for the April 20/June 15 analysis/forecast pair is shown in Figure 8a and 8c, and the corresponding SSH variance or ensemble spread in Figure 8c and 8d, respectively. For the analysis, the observations constrain the initial ensemble spread, so the ensemble mean and likely mode estimate practically overlay each other, but these diverge as the forecast error grows with the forecast period. The difference between the location of the ensemble mean and most likely mode estimate is most easily evident in Figure 8c, where the most likely mode (shown in dashed grey) has more of a NE-SW orientation than does the ensemble mean (black contour).

Example of predicting the 2013 LCE shedding event (Kraken)

While the validity of the June 2014 LCE shedding event described in the preceding paragraph will not be known for several weeks, Figure 9 depicts a sequence of maps during March-May 2013 that are designed and chosen to demonstrate the predictability of a LCE shedding event in a succinct fashion. As can now be confirmed by observations, but which were not available at the time the forecast was being performed (because it was done in real-time), during the spring of 2013 the Loop Current extension became unstable and experienced a shedding event, giving birth to Eddy Kraken.

From Figure 9, the top left and right panels depict the ensemble mean SSH for two analysis times, March 10 and April 21, 2013, respectively. The black contour outlines the location of the 17 cm SSH

ensemble mean. The corresponding 60-day ensemble mean SSH forecasts with respect to those analyses are shown in the left (May 05) and right (June 16) middle top panels. As the March 10 analysis shows, the Loop Current extension was penetrating deep into the northern Gulf with the signature of a cyclonic eddy causing undulation. That LC extension subsequently became unstable and shed Eddy Kraken as can be seen in the April 21 analysis. The forecasts of May 05 and June 16 show LCE separation, westward translation, and weakening of Eddy Kraken. These forecasts are confirmed by an independent analysis of altimeter data downloaded from the Colorado Center for Astrodynamics Research (CCAR) SSH Data Viewer (http://eddy.colorado.edu/ccar/ssh/nrt_gom_grid_viewer). An important point to note is that while the analyses did assimilate altimeter data (but track by track data through NCODA, not the gridded analyses from the CCAR Website), the forecasts did not. Finally, the bottom left (right) shows the ensemble spread, or forecast uncertainty or forecast variance, for May 05 (June 16). These also show the location of the 17 cm SSH contour as black isolines. That more 17 cm SSH isolines depict a separated LCE is consistent with the mean SSH forecast for the same day. The forecast variance for June 16 is of particular interest, because it shows a tri-modal distribution for the height field, i.e. some show the LC entering through Yucatan and exiting through the FS (no shedding), some show moderate penetration into the interior of the Gulf, and some show a separated LCE (actually a majority show this mode, as is consistent with the mean forecast on the same day). This tri-modal distribution of dynamically possible/realistic forecast states is an important attribute of the ensemble forecast system.

2010 ensemble forecast system reanalysis (Franklin)

A re-reanalysis of the ensemble forecasting system covering 2010 and 2011 was performed. A re-analysis is different from a real-time forecast in that it covers periods of time that have already passed, thus analysis-quality (lateral and surface) boundary conditions are available. However, in order to keep the skill of the forecasts consistent with those run in real-time, those analysis quality boundary conditions were not used, i.e., the re-analysis used the same climatological boundary conditions as the real-time runs. The purpose of the reanalysis was to analyze additional LCE shedding events and to take advantage of additional observations that were taken during the Deep Water Horizon (DWH) spill for additional system validation. As with any reanalysis (now including Eddy Kraken in 2013), the character and timing of past LCE shedding events is now know. Examination of the LREFS output through the first half of 2010, CCAR altimetry analyses, and HMI Eddy Watch Maps (not shown) all indicate that the shedding event that occurred in late spring/early summer (Eddy Franklin) was a complicated one, in that there were periods of detachment and re-attachment before most observations were in agreement that actual shedding and westward translation of Franklin was well underway in late summer.

Figure 10 depicts the weekly time evolution of de-tided SSH ensemble spread for analysis-8-week forecast pairs (analysis/forecast) from April 11-June 06, 2010 (Figure 10a) through June 13/August 08, 2010 (Figure 10j). As with the SSH ensemble spread analyses for 2014 and 2013 discussed previously, the SSH analyses here are constrained well by the available observations, and thus show little spread by design. The analyses from April 11 through May 23 (Figure 10a-10g) all show the northward penetration of the LC, but starting around May 30 (Figure 10h) the contours in the neck region start to converge, consistent with stronger cyclonic circulation on the southeast side of the extension near 25N, 85W. On June 06 (Figure 10i), the CCAR altimeter analysis (lavender contour), the control run (red contour), and ensemble mean (green contour) all show a closed contour of the 17cm SSH isoline. Thus, LCE separation has occurred with respect to this particular metric.

As expected, there are larger differences between the CCAR altimetry analysis, control run, and ensemble mean in the 8-week forecasts (right sides of the pairs in Figure 10). The altimetry shows a closed contour in the July 04 forecast, but the control run and ensemble mean do not (Figure 10e). In

the July 25 forecast, the ensemble mean and altimetry show a closed contour, but the control run does not. Although some of these differences are likely due to the fact that we are keying on a particular SSH contour (17 cm here), they do demonstrate the complicated nature of the shedding event during this time. Indeed, the altimetry analyses show a closed contour (separated eddy) in the July 04 through July 25 forecasts (Figures 10e-10h), but a re-attachment in the August 01-08 forecasts (Figures 10i-10j).

In Figures 10a and 10i, the June 06 forecast is compared with the (corresponding) June 06 analysis. In this comparison the analysis is treated as truth for comparison to the forecast. In the forecast, the ensemble mean, control run, and altimetry analysis do not show a closed contour (shed eddy), although they are close to doing so. The ensemble mean and control run in the corresponding analysis on the other hand do, but again they are close. Thus, in this case, there are differences in the details of exactly when the LCE shedding event occurred, but qualitatively, both support that the shedding event definitely took place in early to mid June, 2010.

2011 ensemble forecast reanalysis results (Galileo, Hadal, and Icarus)

The ensemble calculations were ported to the DoD Shared Resource Center (DSRC) to take advantage of additional computational resources. Thus, the 2010 reanalysis was extended into 2011, but that year had just finished at the time of the writing of this report, so, limited analyses and verifications were performed. However, 2011 was a time of significant instability and energetic activity with respect to the LC extension and eddy shedding.

For instance, in January-February, the Loop Current extended to the northwest as a relatively thin feature and showed a short-lived detachment (~2 weeks) in mid-February with respect to the ensemble mean 17 cm height contour, although the altimetry analysis did not (not shown). There was another brief (~1 week) detachment in mid-March. From that time through the end of April the LC penetrated northwestward showing many meanders and undulations ('necking' episodes where the LC appears to be on the verge of a shedding event accompanied with strong cyclonic circulation on the east side). On May 1, 2011, the LC and altimetry both showed closed 17 cm height contours, and the size of the eddy suggested that this may be a robust separation event (Figure 11a), but the analyses depict that the eddy reattached ~3 weeks later (not shown). This aspect may be the model depiction of Eddy Galileo, reported by some based on proprietary observations (HMI Eddy Watch maps), but less evident in most operational observing systems. The 8-week prediction for this shedding event is shown in figure 11b. In that depiction, Eddy Galileo had re-attached, although the 2, 4, and 6-week predictions did show a detached eddy more consistent with observations (not shown). Given that this was a short-lived detachment, the fact that the 8-week prediction depicted a re-attachment is not completely unexpected, although the detached eddy is only present is a few of the ensemble members as shown in the standard deviation of the SSH (figure 11c).

Throughout May-July, 2011, the model LC and altimetry depicted an unstable elongated LC extension that typically penetrated to ~28N. On July 31, the model LC separated (closed 17 cm height contour) as Eddy Hadal (figure 11d). The altimetry did not show a closed contour until August 21, but it was nearly closed and consistent with the mesoscale circulation features in the Gulf. The 8 week prediction for this event, shown in figure 11e, is quite similar to the analysis in both the shape, size, and strength of both the detached eddy and the resultant LC extension. The individual ensemble members accurately reflect the bi-modal distribution of the detached Eddy Hadal and the resultant LC extension (figure 11f).

Although Eddy Hadal was a true separation event (no reattachment) and translated westward for the next several months, the LC extension did not retreat as far southward to the Yucatan as is typically seen in previous separation events. Indeed, during August-October 2011, the LC continues its penetration

towards the northwest into the interior of the Gulf, and undergoes another separation event on November 6 (figure 11g). This is the Eddy Icarus shedding event and the closed SSH contour is also seen in the altimetry analyses (lavender colored contour in figure 11g). Icarus remains a separate eddy as indicated by the closed 17 cm contour criterion, but a relative high SSH remains between the extension and the eddy for the next several weeks (in fact several of the altimetry analyses alternate between closed and open 17 cm contour depictions). The 8-week forecast for this event is shown in figure 11fh, and is fairly consistent with the November 6 analysis, although the remnants of Eddy Hadal are much weaker in the prediction than the analysis, but mostly in amplitude, as there are several ensemble members that depict a closed 17cm SSH contour in that area.

Verification Metrics

Verification and validation is a critical part of any ocean prediction system, but depend on a sufficient number of independent observations (i.e., not assimilated into the analysis). When and where those observations are available, we have examined the forecast skill. Figure 12 is an example of an assessment of forecast skill as a function of depth and forecast length. It shows profiles of temperature at a location (denoted by red triangle on inset map) in the western Gulf (left set of panels) and the eastern Gulf (right set of panels). For each location, the left panels are the temperature vs. depth profiles for each member (black), the ensemble mean (green), and the observation (red). The middle panel shows the temperature observation minus the temperature of each ensemble member for a 3(top), 10 (middle), and 17 (bottom) day forecast. The right panels show the bias (mean error) (red) and RMS error (black). The bias and error at the location in the western Gulf are relatively small (between 0-1C) as this is a comparatively quiescent area of the Gulf. The errors at the eastern location in the Gulf are larger, as this is the area dominated by LCE dynamics. At both locations, the bias and error are largest between ~100-200m, the approximate depth of the thermocline.

Another commonly used verification metric of forecast skill is the anomaly correlation. The anomaly correlation provides a measure of the spatial similarity between forecasts and corresponding best analyses and is calculated as

$$AC(f,a) = \frac{\sum (f-\bar{c})(a-\bar{c})}{\sqrt{\sum (f-\bar{c})^2 \sum (a-\bar{c})^2}} \tag{1}$$

Where f represents the forecast, a represents the analysis, and $\bar{c}$ is the climatological (model) mean that spans the 2013 analysis period. The anomaly correlation can be calculated at any time during the forecast horizon and for any model variable at any model depth at all model grid points. Figure 13 shows the spatial anomaly correlation of the speed at 1, 3, 5, and 8 weeks into the horizon (each quad of panels) for the surface (left), and 250m (right). Red indicates a correlation of 1.0, i.e. a perfect correlation between the forecast and corresponding analysis, while dark blue indicates no correlation. That the maps become less correlated (more blue) as the forecast length increases is to be expected. As any forecast will eventually revert to a free-running model solution as the forecast length increases. Put differently, the model forecast eventually looses any memory of the initial condition. All the depths show the same general pattern as a function of forecast length, and also similar patterns of higher correlations in the north central Gulf. This is because this region is dominated by the large energetic Loop Current Extension and Eddie(s) which are more constrained by mesoscale designed observing systems, which in turn allows for higher forecast skill (higher correlation) in these regions.

Figure 14 shows the spatial anomaly correlations for salinity. The basic patterns are similar to that of speed, except that overall the correlations are notably higher at 250m depth (right quad). This is likely due to the relatively thick subsurface salinity maximum that is known to episodically exist in the part of

the Gulf at this depth. The same series of plots for temperature are shown in Figure 15. The striking feature about the correlations of temperature is that the surface temperature correlations remain very high throughout the 8-week forecast period. This may be due to the fact that during the spring through fall months, the surface temperatures in the Gulf become quite homogeneous and show few thermal gradients, at least in the interior.

In Figures 16-18, the spatial anomaly correlations have been averaged over the central Gulf in the Loop Current region (88-82W, 22-26N). The black curves show the anomaly correlation for the 8-week forecast, the red curves for a forecast of persistence (no change). An anomaly correlation of greater than 0.6 indicates forecast skill (Murphy and Epstein, 1989). Thus, the surface speeds show forecast skill at all three depths out to about 28 days (indicated by black arrows that points to where the anomaly correlation intersects the value of 0.6 on the y-axis). Additionally, the speed anomaly correlations are all relatively barotropic in nature. The salinity at the surface and 100m show forecast skill out to about 14 days (likely impacted the scarcity of salinity values as well as by fresh water inflow from the Mississippi and other rivers), but at 250m the forecast is skillful out to about 35 days, consistent with the subsurface salinity maximum mentioned earlier. The surface temperature forecast shows high skill throughout the 8-week forecast period, likely due to the large number of SST observations, but that skill drops off to ~35 days in the subsurface. Figure 19 shows the anomaly correlations of the SSH. Except in 2013, the forecast shows high skills extending up to 8 weeks. The large difference between the forecast and persistence in 2011 suggest that forecast is superior to the persistence. Interannual differences in the anomaly correlations for the different variables may be due to the different types and amounts of observations available. For example, 2010 included the interleaved Jason altimetry while 2013 did not. It may be cautioned that the forecast skill comparisons from 2011 are in part due to the difference in the number of forecasts being used, which will be confirmed as more forecasts become available.

Derived products and decision-making aides

The uncertainty estimates provided by the ensembles allows for the construction of probability distribution functions describing the ocean states that can be used as inputs to Operational Risk Management tools. As part of the transition of the Underwater Common Uncertainty Picture (UCUP) system to the U.S. Navy, Risk Assessment Codes (RACs) have been developed that closely follow the terminology detailed in the OPNAVINST 3500-39C. The RACs weigh the probability of an occurrence of specific error levels or environmental conditions versus the severity of possible impacts affecting identified missions, platforms and sensors. Although originally designed for Navy applications, the RACs have great utility and generality for planning and risk mitigation purposes. The RACs allow the complicated and voluminous output from an ensemble forecast system to be presented in compressed, easy to comprehend visual aides. In that context, a high probability of a severe event is depicted as red and a low probability of a minimal event as green (Figure 20). This approach has inherent flexibility because the "user" is the one who defines the probabilities as well as the events, including the thresholds of what constitutes a severe or minimal event. The probabilities are computed from the joint probability distribution functions for the user defined criteria estimated from the ensemble runs using non-parametric methods. These functions are then used to estimate the likelihoods less than 0.1, between 0.1 and 0.25, between 0.25 and 0.5 or above 0.75, of the occurrence of the selected criteria within the prescribed threshold ranges, representing the levels of very high critical impact, high or severe impact, low impact, or very low impact. Risk Assessment Codes are then assigned such that a code 1 will correspond to likelihood above 0.75 of impacts above the "high impact" or above 0.5 for "very high impacts", and a code 5 to likelihoods below 0.1 for "low impacts" or below 0.25 for "very low impacts".

Examples of RACs generated by the LREFS are shown in Figure 21. Numbers 1 to 5 are assigned to each geographical location on a grid describing the trade-off between probabilities of occurrence vs. severity of impacts consistent with those shown in Figure 20. Figure 20a and 20c show the risk of

occurrence of currents for the April 20 and June 19 analysis and forecast, respectively. The color-coded risk depends on the current thresholds according to: currents with speed > 0.77 m/s – very high impact; currents with speed > 0.5 < 0.77 m/s – high impact; currents with speed > 0.25 < 0.5 m/s - low impact, currents with speed >0.15 < 0.25 m/s – very low impact, and currents with speed <0.15 m/s no impact. To reiterate, these thresholds are chosen by the user for specific applications, thus providing flexibility and generality.

As stated, RACs can depict the risk for a single variable or be computed from joint probability distribution functions. Whereas the risk assessment shown in Figures 21a and 21b are for surface currents only, Figures 21b and 20c take into account the combined risk of sea state (elevation, which also takes into account the surface wind speed) and surface currents. The thresholds for the surface currents are the same as in Figure 20a and 20c, and the thresholds for the sea state are: elevation > 2.0m – very high impact, elevation >1.5 < 2.0m – high impact, elevation >1.25 < .5 m – low impact, elevation > 0.75 < 1.25 m – very low impact, elevation < 1.25 m – no impact. Like the surface currents, the thresholds for each chosen variable are chosen by the user for specific applications. As is evident, the risk of strong currents and/or combined error of the forecast are greatest where ocean features are most energetic and variable, i.e. the area of the Loop Current.

The RAC maps are to be used by planners to assess the risk of occurrence of damaging currents that may have an impact over the missions under consideration, along the forecast times and along the several possible regions of operations. Once the risks are known and accepted and areas and times selected the planners can use the control run and ensemble mean to estimate the most likely conditions that will be encountered during the mission execution. They can also look into the ensemble spread to have an idea of the most extreme conditions that may be locally encountered in order to prepare their contingency plans and brief/train the teams involved in the execution. More complex cases can quantify the joint occurrence of variables or derived functions (e.g., plumes, tracks, Lagrangian trajectories, etc.).

Computational requirements

The real-time system (2013-2014) ran on a Linux cluster at NRL. The daily (control run) took ~2 hours of wallclock time running on 96 processors. The once per week 60-day forecast was configured to run five members on 120 processors, which took ~3.5 wallclock hours. Thus the 32 member ensemble took approximately 23 hours. Each member generated ~6 GB in netCDF format (~200 GB for a 32 members of a 60-day forecast saving model output once per day). At the beginning of 2014 the system was ported to an IBM iDataPlex6 at the DSRC, located at NAVO-Stennis. Similar numbers of processors are used there as well, but because the IBM has many more processors, there is inherently more system flexibility.

Summary

An extended-range forecast system for the Gulf of Mexico has been developed. The system provides a 60-day forecast and the uncertainty of that forecast on a weekly basis. The uncertainty of the forecast is computed using ensemble methods. The ensemble populations are generated by perturbing the initial conditions and surface forcing. A multi-year time series of output from the Global Ocean forecast System (GOFS 3.0) is used to construct a daily climatology for 3-D lateral boundary conditions. A similar method is used to construct surface boundary conditions from the COAMPS atmospheric forecast model. Estimates of forecast uncertainty calculated over 32 ensemble members have demonstrated skill in predicting LCE shedding events that occurred in the spring of 2010, 2011 (a particularly energetic year), the summer of 2013, and perhaps the spring of 2014 (the latest forecast for which observations are not yet available for verification). The forecasts are qualitatively verified by

comparing to independent altimetry and sea surface temperature analyses and all other available observations. The ensemble forecast system delivers a short range 24-hour forecast every day and a 60-day long range forecast once per week and has been running in real-time since January 2013. Output from the system is used to derive risk assessment products such as risk of strong currents or risk of combined error in the forecast. These are calculated from probability distribution functions over the length of the forecast and can be used for planning, risk mitigation, search and rescue, etc.

List of Figures

Figure 1: Schematic diagram of the ensemble forecast system. The components of the NCODA system are the blue boxes in the inset. Both the surface forcing and initial conditions are perturbed to generate the individual ensemble members.

Figure 2: Weekly ensemble mean SSH. (a) is analysis for April 20, 2014, (b) is 1-week forecast relative to the analysis for April 27, etc. (i) is ~60 day forecast (June 15, 2014).

Figure 3: Ensemble spread of SSH with 17 cm SSH contour from each ensemble member (black contour) and control run (red contour) superimposed. Shaded region is the standard deviation of SSH across the 32 ensemble members. Panels cover same dates as Fig. 2.

Figure 4: Ensemble mean of the surface speed with the 0.77 m/s (1.5 kt) isotach superimposed. Panels cover same dates as Fig. 2.

Figure 5: Ensemble spread of the surface speed with the 1.5 kt isotach from the individual members superimposed. Shaded region is the standard deviation of speed across the 32 ensemble members. Panels cover same dates as Fig. 2.

Figure 6: Ensemble mean temperature at 200m with the 18C isotherm superimposed. Panels cover same dates as Fig. 2.

Figure 7: Ensemble spread of temperature at 200m with the 18C isotherm from individual members superimposed. Standard deviation of temperature at 200 m is shaded. Panels cover same dates as Fig. 2.

Figure 8: Example of likely mode estimation in the April 20/June 15, 2014 analysis/forecast cycle. The mean de-tided SSH for the analysis and forecast (color background) are shown in (a) and (c), and the corresponding SSH variance or ensemble spread in (b) and (d), respectively. Similar to Figures 2 and 3, the 17 cm contours are highlighted. Here, the likely mode estimate for the 17 cm SSH is depicted by the dashed grey contour. For clarity, it has more of an NE-SW orientation than the mean SSH, and is easiest to see in the mean SSH forecast (c).

Figure 9: Demonstration of 60-day forecast skill of the Loop Current eddy separation. Left panel from top to bottom: Sea Surface Height (SSH) (17 cm contour) for the analysis on March 10, 2013, the corresponding 2-month forecast of the 32-member ensemble mean SSH (i.e. for May 05), SSH (anomaly) verification from real-time mesoscale altimetry from Colorado Center for Astrodynamics Research (CCAR), forecast variance of SSH calculated from 32 perturbed ensemble members (black contours are 17 cm SSH from each ensemble member). Right panels are for April 21, 2013 analysis, 2-month forecast (June 16), corresponding SSH anomaly verification from CCAR and forecast variance with 17 cm contours superimposed. Figure 10: Example of forecast skill during June-July 2010.

Figure 10. Weekly pairs of de-tided SSH analysis (left) and 8-week forecast (right) variance or ensemble spread (color) with the 17 cm contours from the individual members (black), the CCAR altimetry analysis (lavender), the control run (red), and the ensemble mean (green contour) overlaid. Figure 10a is the April 11/June 06 (analysis/forecast) pair, Figure 10b is the April 18/June 13 (analysis/forecast) pair, etc. The June 06 forecast in (a) corresponds with the analysis in (i), and the differences are due to forecast error. (k) is the CCAR altimetry analysis, (l) is the CCAR altimetry analysis (contours) with MODIS chlorophyll underlain in color, (m) is the same as (l) with GHR SST underlain in color, and (n) is the same as (l) with the MODIS SST underlain in color.

Figure 11: Ensemble mean SSH (de-tided) (color background) for (a) the May 05 2011 analysis (Eddy Galileo), (b) the corresponding 8-week forecast (relative to the March 06 analysis), (c) the standard deviation of the SSH for the May 05 forecast. Panel (d) is the July 31 ensemble mean SSH analysis (Eddy Hadal), (e) the corresponding forecast relative to the June 05 analysis, and (f) the standard deviation of the SSH. Panel (g) is the SSH analysis for November 06 (Eddy Icarus), (h) the corresponding 8-week forecast relative to the Sept. 11 analysis, and (i) the standard deviation of the SSH analysis.

Figure 12: Validation of temperature profiles for two locations in the GoM on Feb. 13, 2013. The panels on the left (right) are for a location in western (eastern) Gulf. The left panels are temperature vs. depth (black are individual ensemble members, green is ensemble mean, red is observation). Middle panel is observation minus each ensemble member for 3 (top), 10 (middle), and 17 (bottom) day forecast. The right panels are the bias (mean error) (red) and RMS error (black).

Figure 13: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) speed for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the

predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

Figure 14: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) salinity for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

Figure 15: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) temperature for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

Figure 16: Anomaly correlation of surface, 100 m, and 250 m forecast of current speed with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Figure 17: Anomaly correlation of surface, 100 m, and 250 m salinity with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Figure 18: Anomaly correlation of surface, 100 m, and 250 m temperature with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Figure 19: Anomaly correlation of de-tided SSH with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47, 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Figure 20: Risk Assessment Code. Red indicate a high probability of a severe event, but the severity of the event is application dependent and chosen by the user.

Figure 21: Examples of Risk Assessment Codes (RACs) in the Gulf of Mexico. (a) and (c) are the risk of occurrence of surface currents for specified thresholds (see text) for the analysis on April 20, 2014 and the forecast on June 19, 2014, respectively. (b) and (d) are the risk of combined error in the forecast for the same analysis/forecast. The risk of combined error also takes into account sea state (and thus surface winds) in addition to surface currents. Red areas are high probability of severe impact for specific thresholds.

References:

Bishop, C. and Z. Toth, 1999. Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, 56, 1748-1765.

Coelho, E.F., G. Peggion, C. Rowley, G. Jacobs, R. Allard, and E. Rodriguez, 2008. A Note on NCOM Temperature Error Calibration Using The Ensemble Transform. *Journal Marine Systems Rapid Environmental Assessment special issue*, 78, S272-S281.

Cummings, J.A., 2005. Operational multivariate ocean data assimilation. *Quart. J. Royal Met. Soc*., 131, 3583-3604.

Dobla-Reyes, F., R. Hagerdorn, and T.N. Palmer, 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination, *Tellus*, 57A, 234-252.

Fox, D.N., W.J. Teague, C.N. Barron, M.R. Carnes, and C.M. Lee, 2002. The Modular Ocean Data Assimilation System (MODAS). *Journal of Atmospheric and Oceanic Technology* 19:240–252, http://dx.doi.org/10.1175/1520-0426(2002) 019<0240:TMODAS>2.0.CO;2.

Hodur, R.M., 1996. The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS), *Mon. Weather Rev*., 125, 1414-1430.

Martin, P.J., C. N. Barron, L. F. Smedstad, T. J. Campbell, A. J. Wallcraft, R. C. Rhodes, C. Rowley, T. L. Townsend and S. N. Carroll, 2009.
User's Manual for the Navy Coastal Ocean Model (NCOM) version 4.0
NRL Report NRL/MR/7320--09-9151

Metzger, E.J., O.M. Smedstad, P.G. Thoppil, H.E. Hurlburt, J.A. Cummings, A.J. Wallcraft, L. Zamudio, D.S. Franklin, P.G. Posey, M.W. Phelps, P.J. Hogan, F.L. Bub, and C.J. Dehaan, 2014. US Navy Operational Global Ocean and Arctic Ice Prediction Systems, *Oceanography* (in press).

Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. Bishop, and X. Wang, 2006. Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus*, 58A, 28-44.

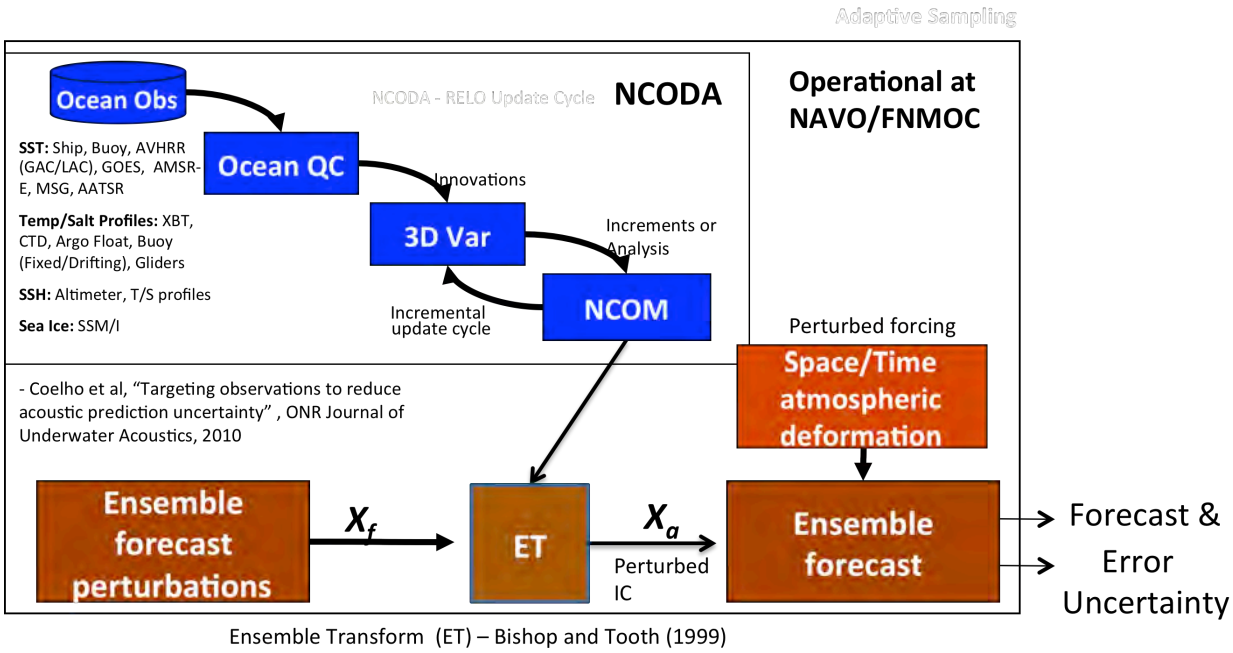


Figure 1. Schematic diagram of the ensemble forecast system. The components of the NCODA system are the blue boxes in the inset. Both the surface forcing and initial conditions are perturbed to generate the individual ensemble members.

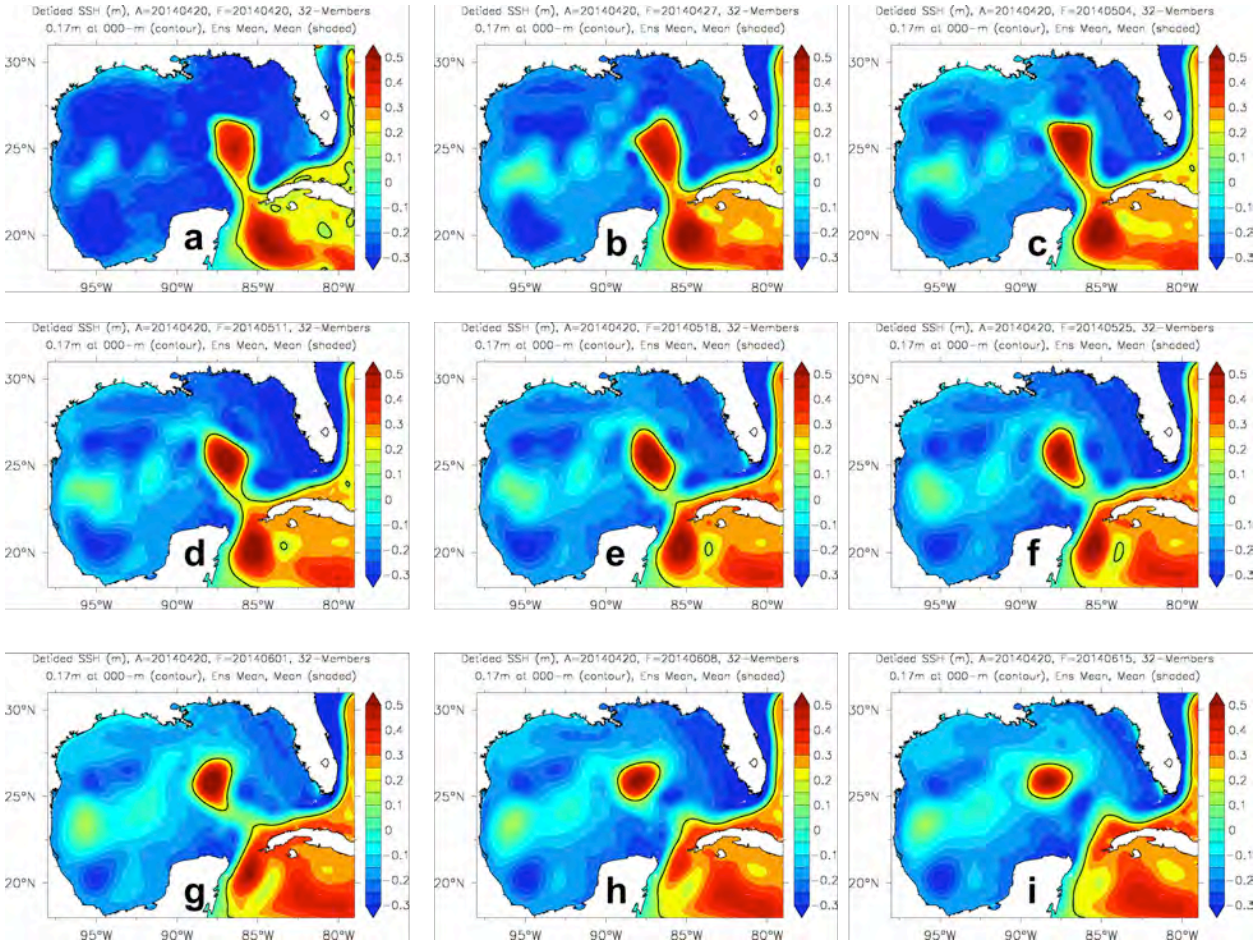


Figure 2. Ensemble mean SSH with the 17 cm contour superimposed for (a) April 20, 2014 analysis (b-i) 60-day forecast relative to the analysis on April 20, 2014 showing every Sunday.

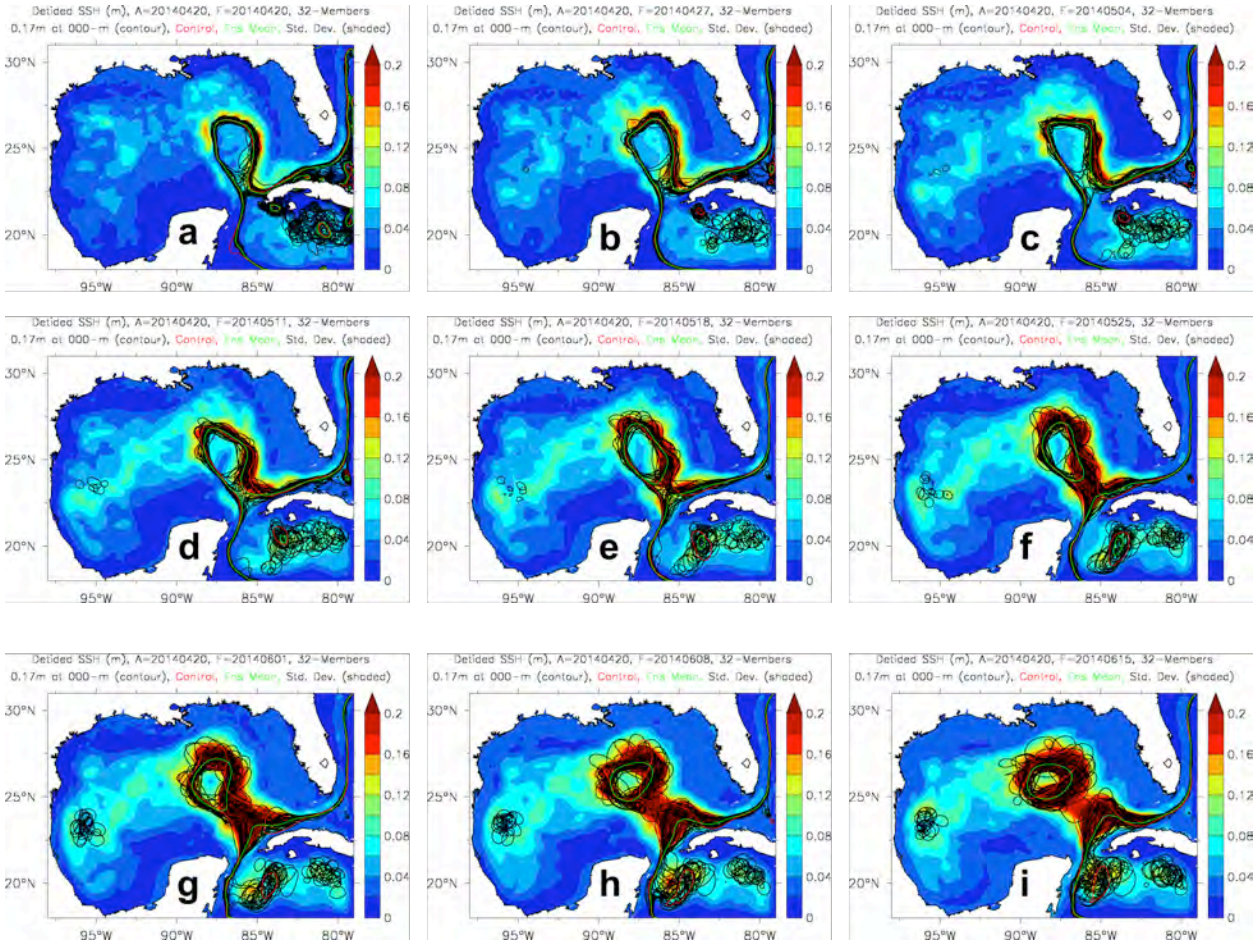


Figure 3. Ensemble spread of SSH with 17 cm SSH contour from each ensemble member (black contour) and control run (red contour) superimposed. Shaded region is the standard deviation of SSH across the 32 ensemble members. Panels cover same dates as Fig. 2.

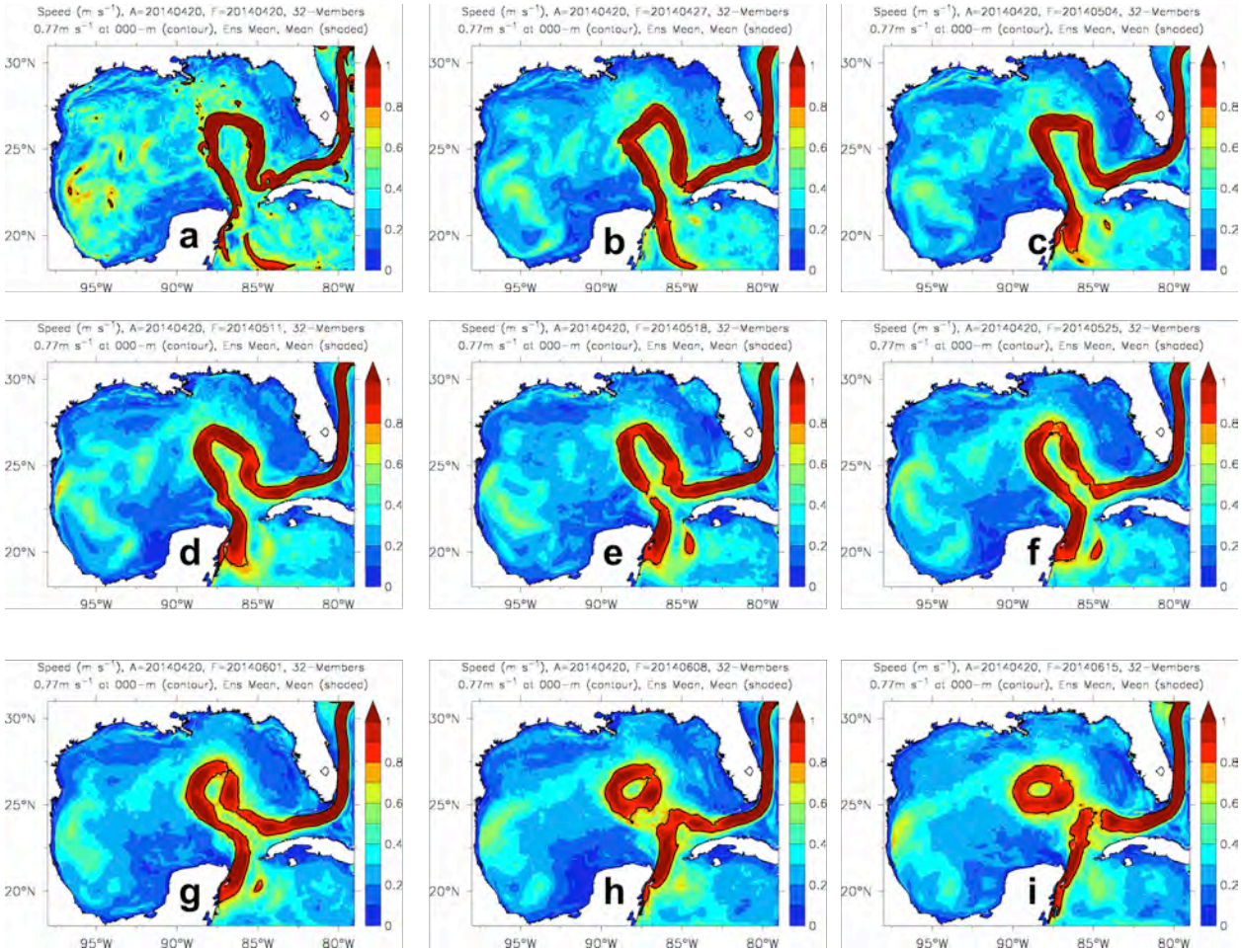


Figure 4. Ensemble mean surface speed with the 0.77 m/s (1.5 kt) isotach superimposed. Panels cover same dates as Fig. 2.

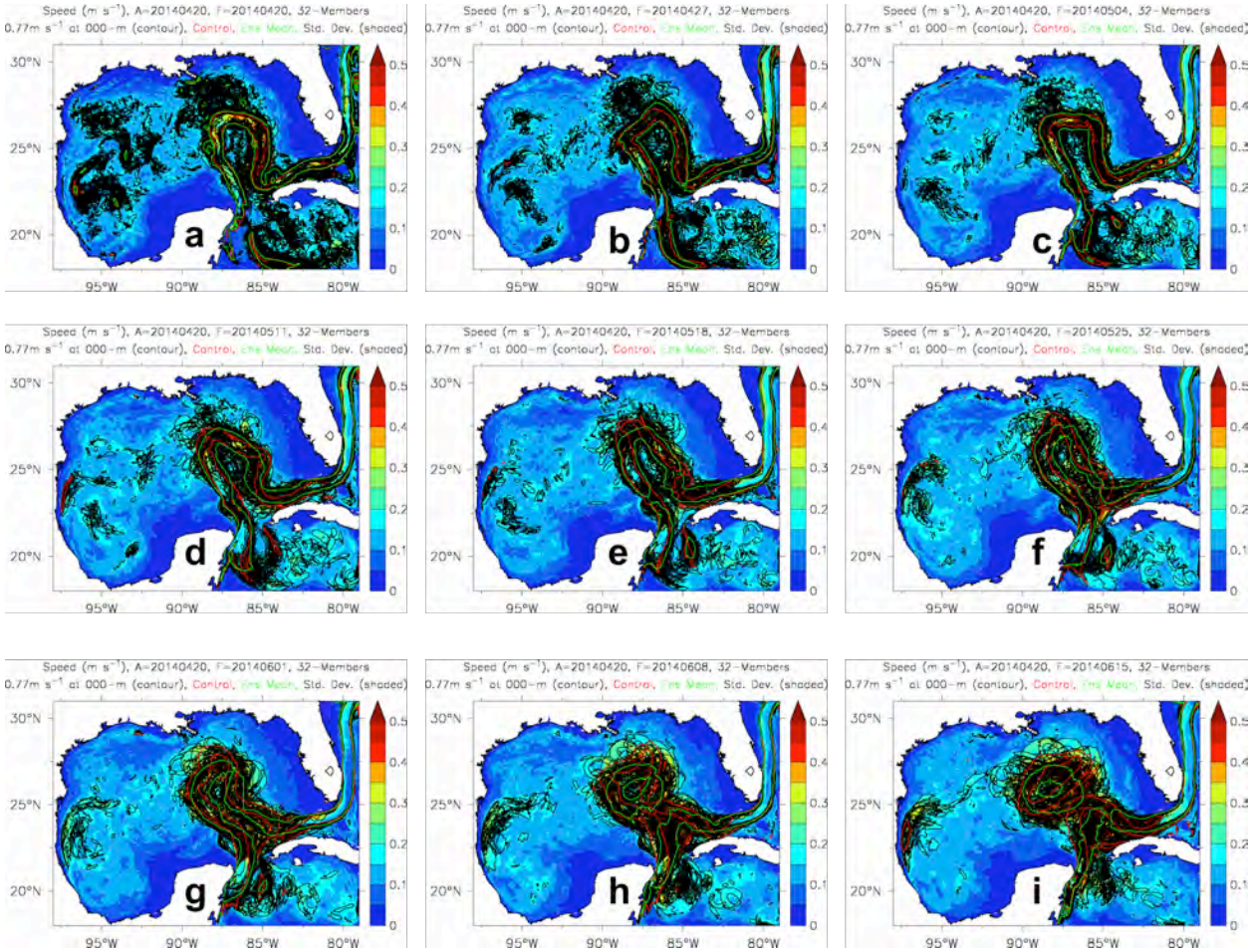


Figure 5. Ensemble spread of the surface speed with the 1.5 kt isotach from the individual members superimposed. Shaded region is the standard deviation of speed across the 32 ensemble members. Panels cover same dates as Fig. 2.

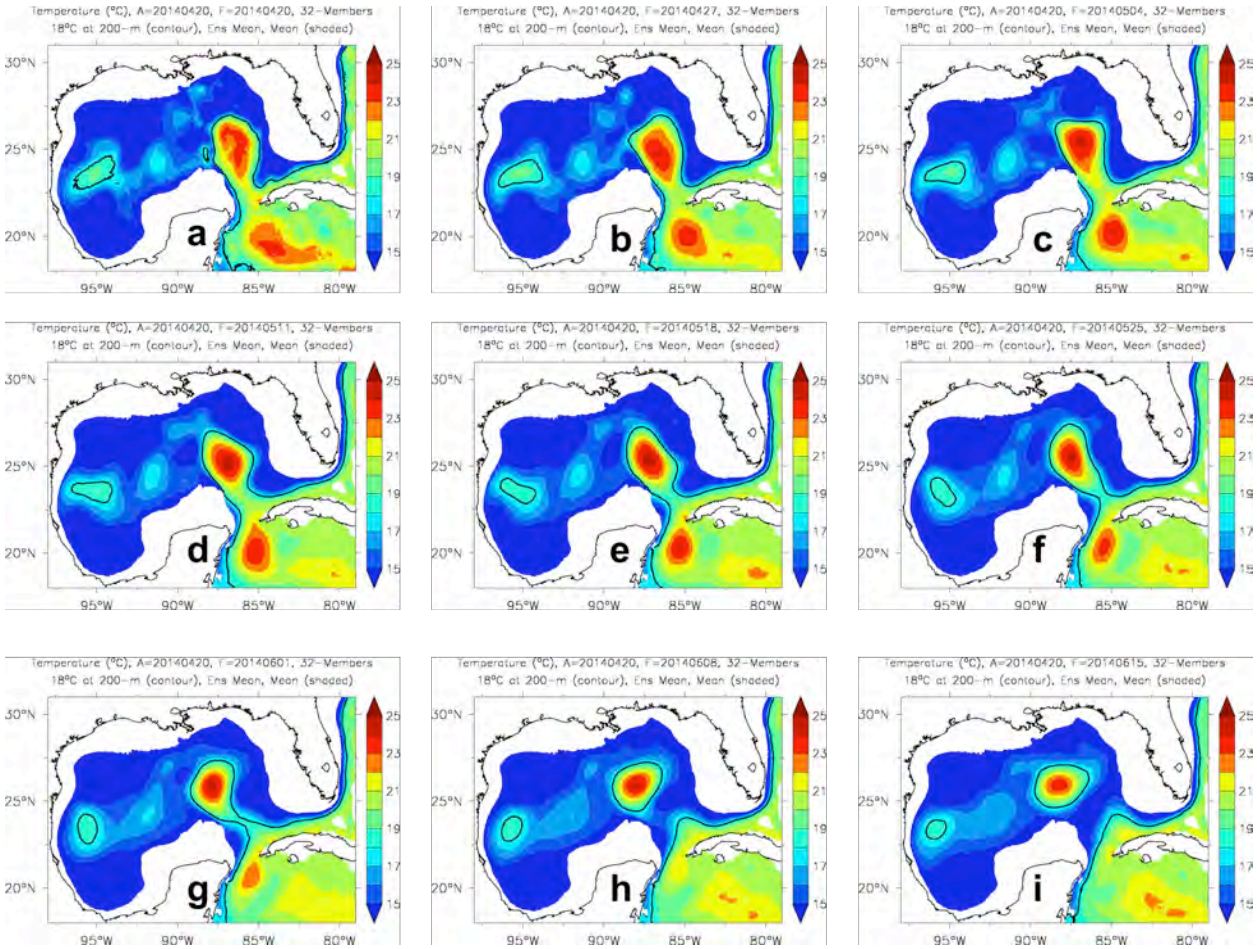


Figure 6. Ensemble mean temperature at 200m with the 18C isotherm superimposed. Panels cover same dates as Fig. 2.

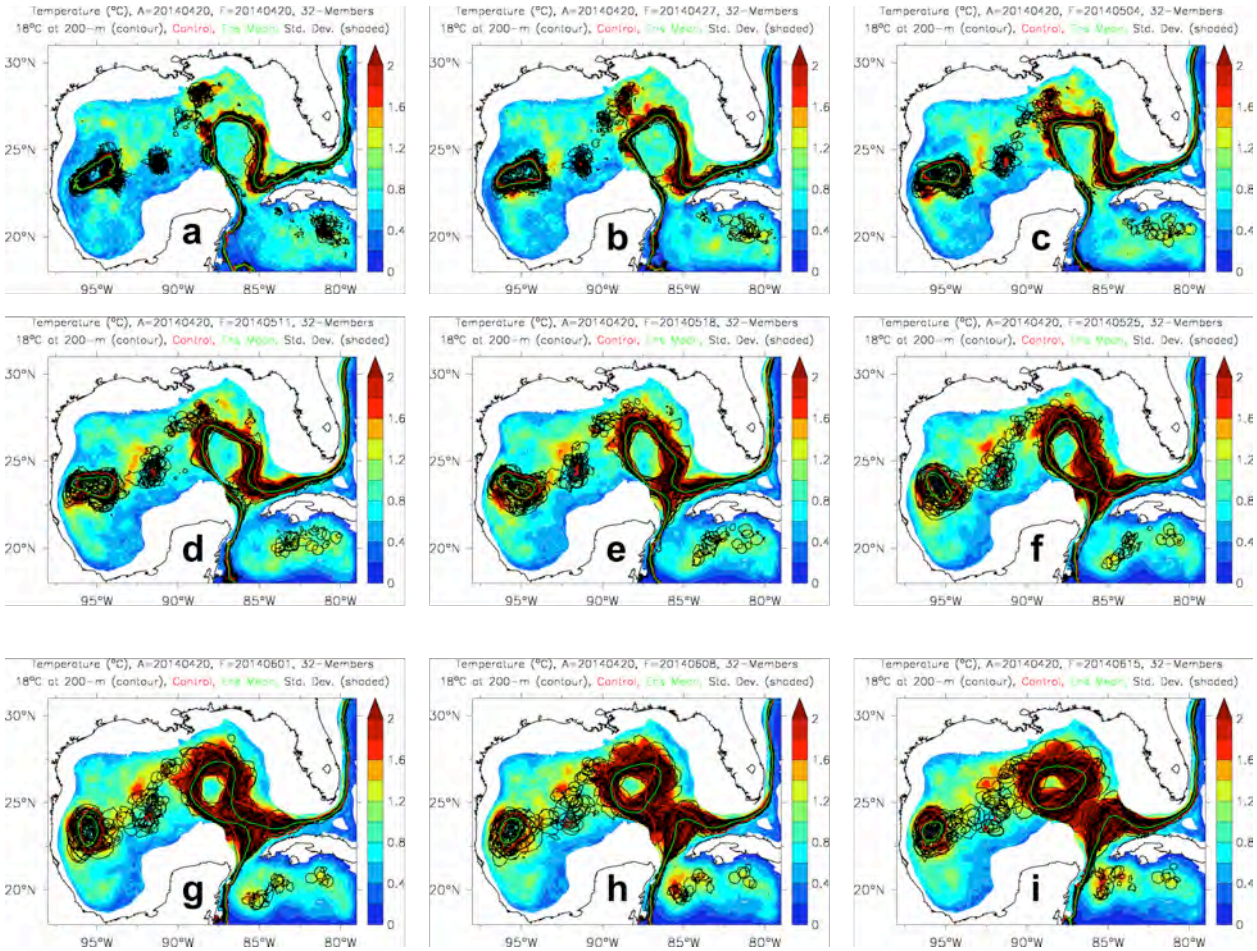


Figure 7. Ensemble spread of temperature at 200m with the 18C isotherm from individual members superimposed. Standard deviation of temperature at 200 m is shaded. Panels cover same dates as Fig. 2.

Ensemble Mean/Spread vs. Likely Mode Estimate

April 20, 2014 Analysis **June 15, 2014 Forecast**

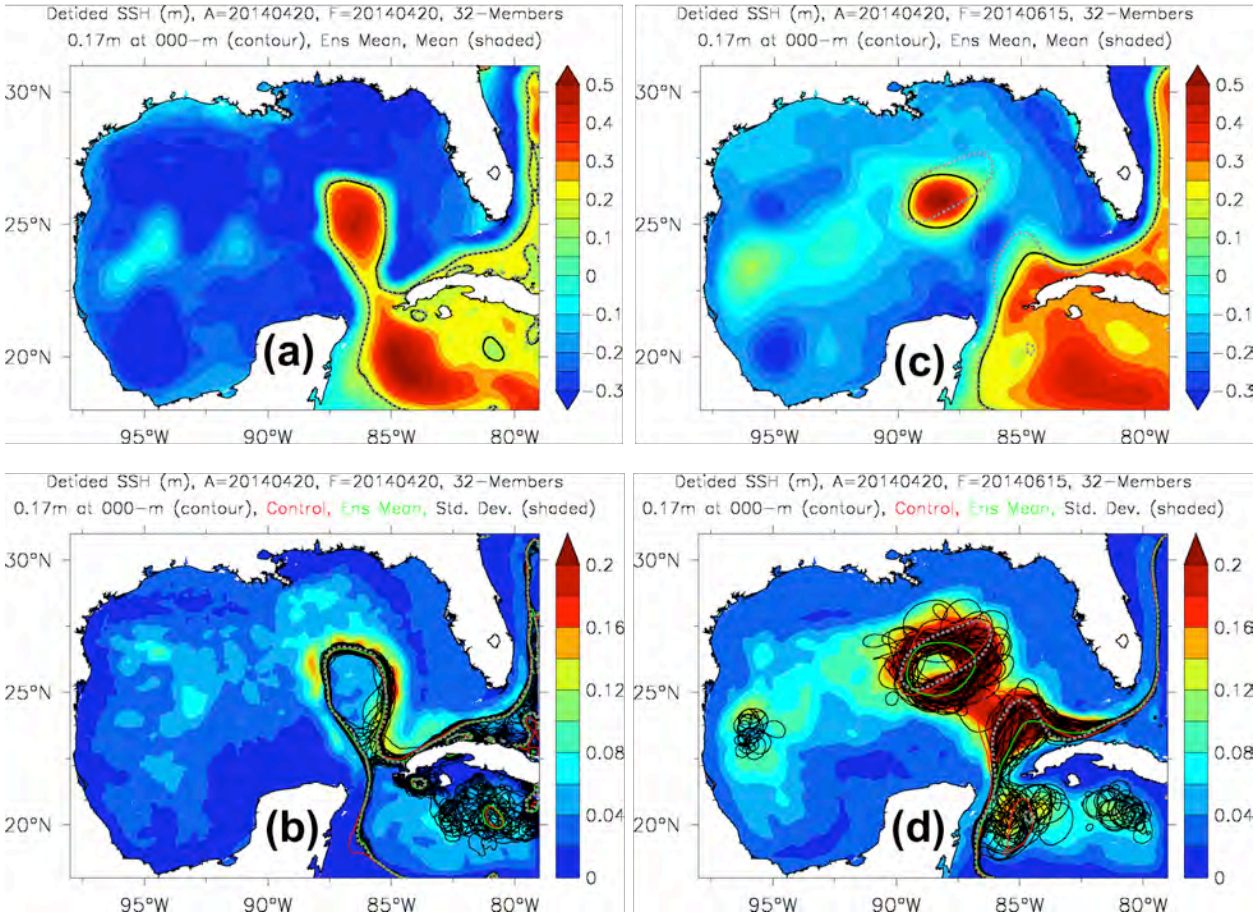


Figure 8: Example of likely mode estimation in the April 20/June 15, 2014 analysis/forecast cycle. The mean de-tided SSH for the analysis and forecast (color background) are shown in (a) and (c), and the corresponding SSH variance or ensemble spread in (b) and (d), respectively. Similar to Figures 2 and 3, the 17 cm contours are highlighted. Here, the likely mode estimate for the 17 cm SSH is depicted by the dashed grey contour. For clarity, it has more of an NE-SW orientation than the mean SSH, and is easiest to see in the mean SSH forecast (c).

8-Week Ensemble Forecast March – June 2013

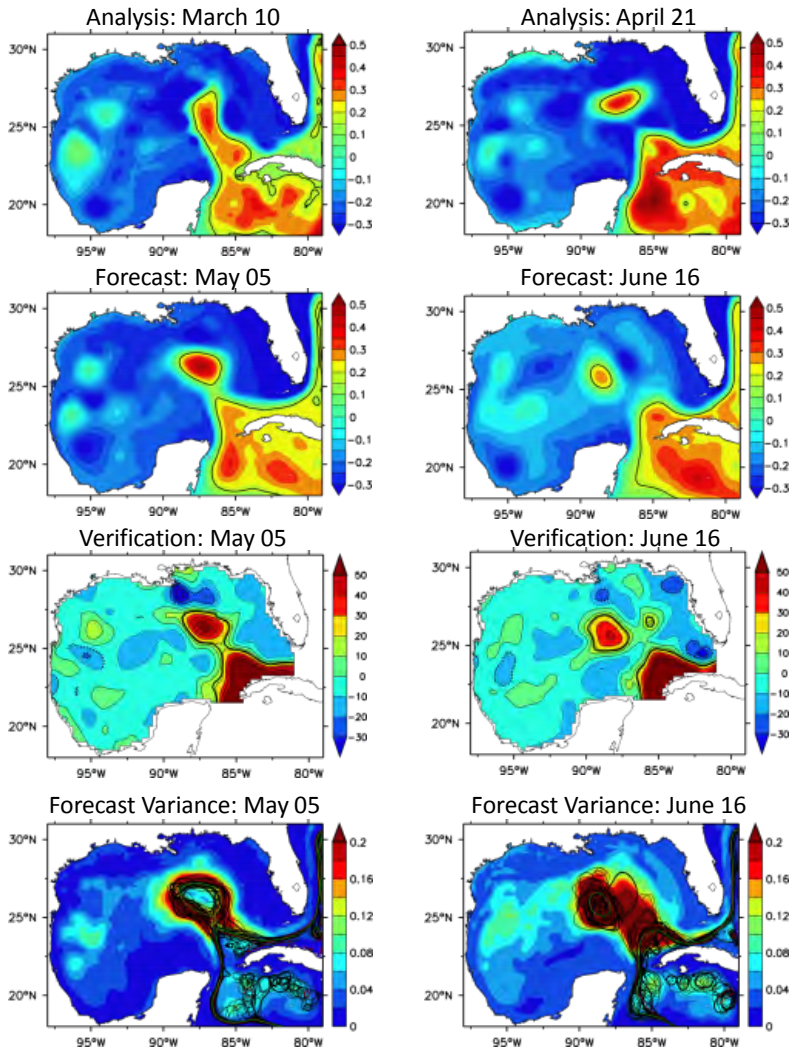


Figure 9. Demonstration of 60-day forecast skill of the Loop Current eddy separation. Left panel from top to bottom: Sea Surface Height (SSH) (17 cm contour) for the analysis on March 10, 2013, the corresponding 2-month forecast of the 32-member ensemble mean SSH (i.e. for May 05), SSH (anomaly) verification from real-time mesoscale altimetry from Colorado Center for Astrodynamics Research (CCAR), forecast variance of SSH calculated from 32 perturbed ensemble members (black contours are 17 cm SSH from each ensemble member). Right panels are for April 21, 2013 analysis, 2-month forecast (June 16), corresponding SSH anomaly verification from CCAR and forecast variance with 17 cm contours superimposed.

Ensemble SSH Analyses (April-May) and Forecasts (June-July) 2010

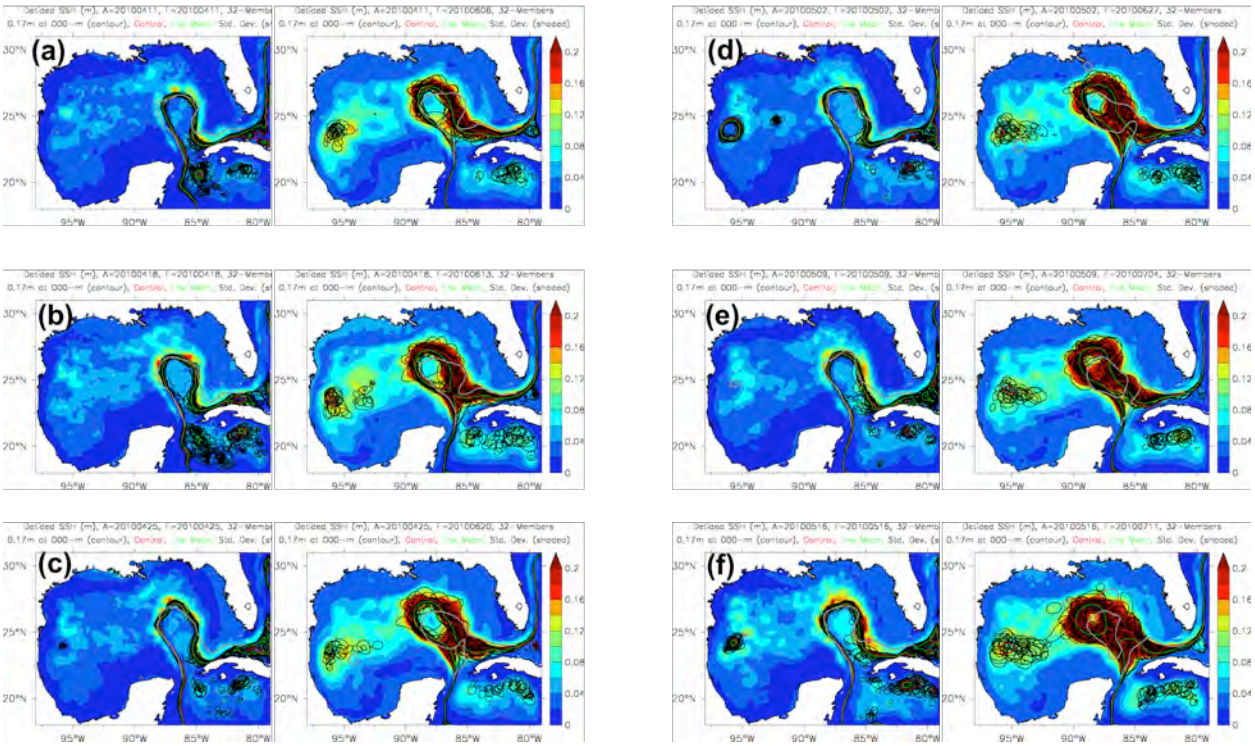


Figure 10. (The figure and caption are continued on the next page.)

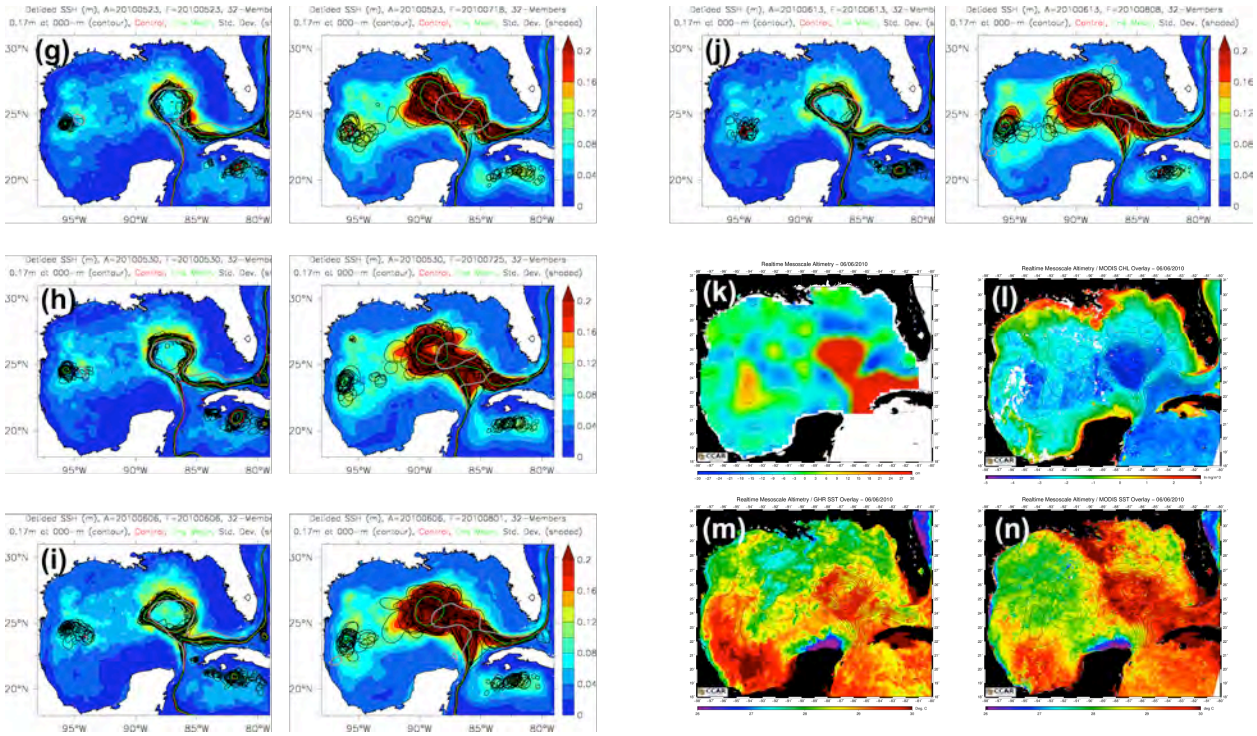


Figure 10. Weekly pairs of de-tided SSH analysis (left) and 8-week forecast (right) variance or ensemble spread (color) with the 17 cm contours from the individual members (black), the CCAR altimetry analysis (lavender), the control run (red), and the ensemble mean (green contour) overlaid. Figure 10a is the April 11/June 06 (analysis/forecast) pair, Figure 10b is the April 18/June 13 (analysis/forecast) pair, etc. The June 06 forecast in (a) corresponds with the analysis in (i), and the differences are due to forecast error. (k) is the CCAR altimetry analysis, (l) is the CCAR altimetry analysis (contours) with MODIS chlorophyll underlain in color, (m) is the same as (l) with GHR SST underlain in color, and (n) is the same as (l) with the MODIS SST underlain in color.

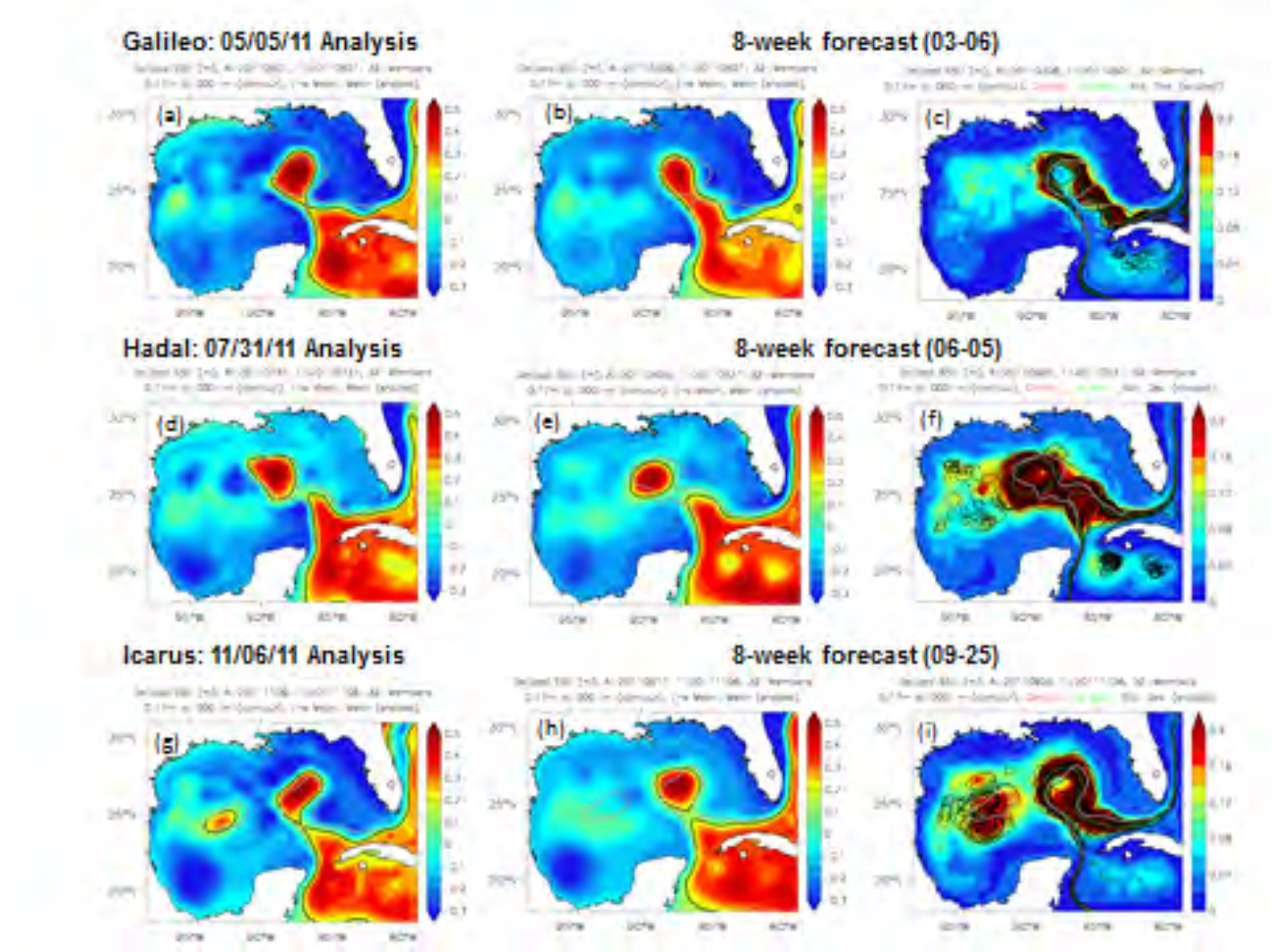


Figure 11: Ensemble mean SSH (de-tided) (color background) for (a) the May 05, 2011 analysis Eddy Galileo), (b) the May 05 forecast (relative to the March 03 analysis), and (c) the standard deviation of the ensemble SSH with the individual 17 cm SSh contours. (d) is the analysis for July 31 (Eddy Hadal), (e) is the forecast relative to the June 06 analysis, 2011 forecast, and (f) is the standard deviation of the SSH, which clearly shows the bi-modal distribution of Eddy hadal and the LC Extension. (g) shows the initial separation of Eddy Icarus, (h) is the 8-week forecast relative to the Sept. 11 analysis, and (i) is the corresponding SSH standard deviation.

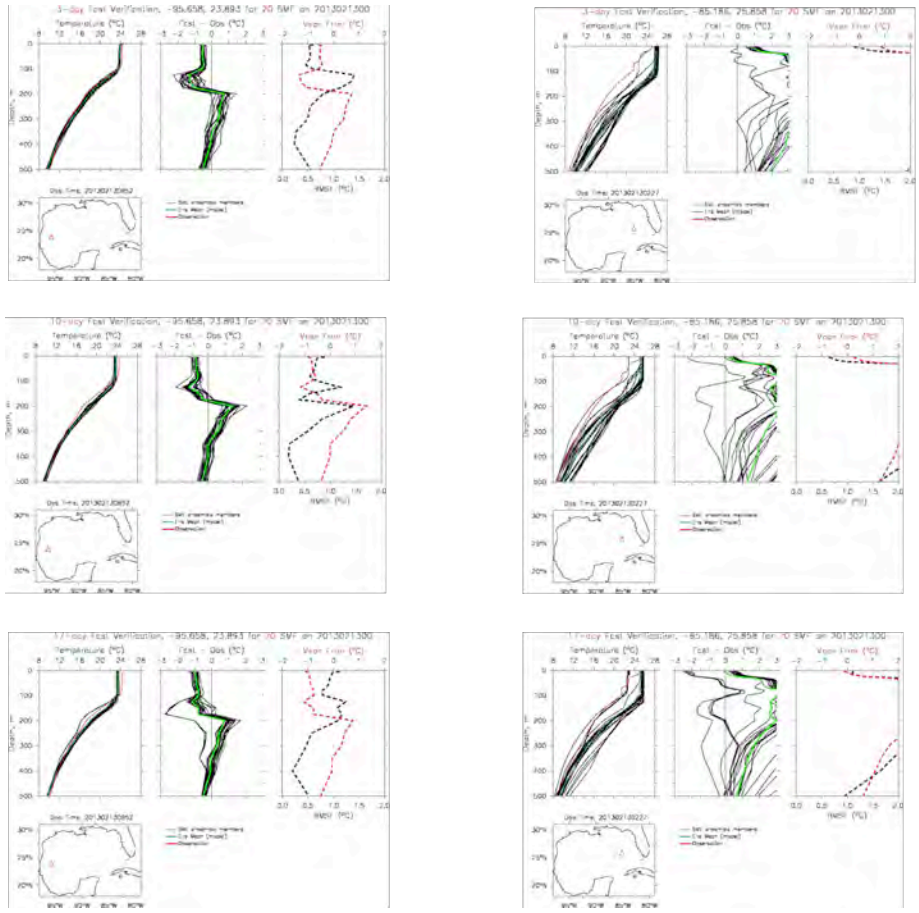


Figure 12. Validation of temperature profiles for two locations in the GOMEX on Feb. 13, 2013. The panels on the left (right) are for a location in western (eastern) Gulf. The left panels are temperature vs. depth (black are individual ensemble members, green is ensemble mean, red is observation). Middle panel is observation minus each ensemble member for 3 (top), 10 (middle), and 17 (bottom) day forecast. The right panels are the bias (mean error) (red) and RMS error (black).

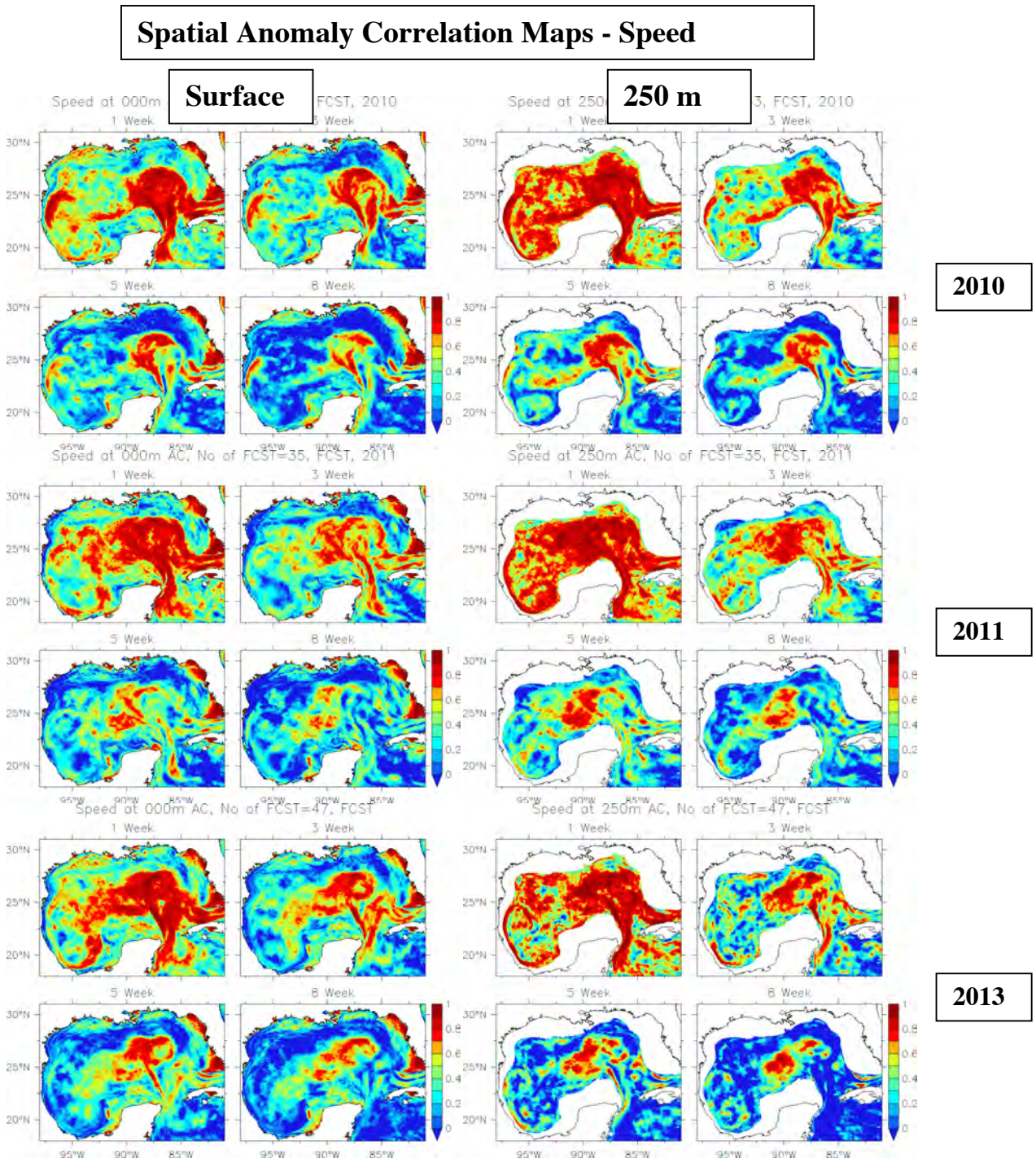


Figure 13: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) speed for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

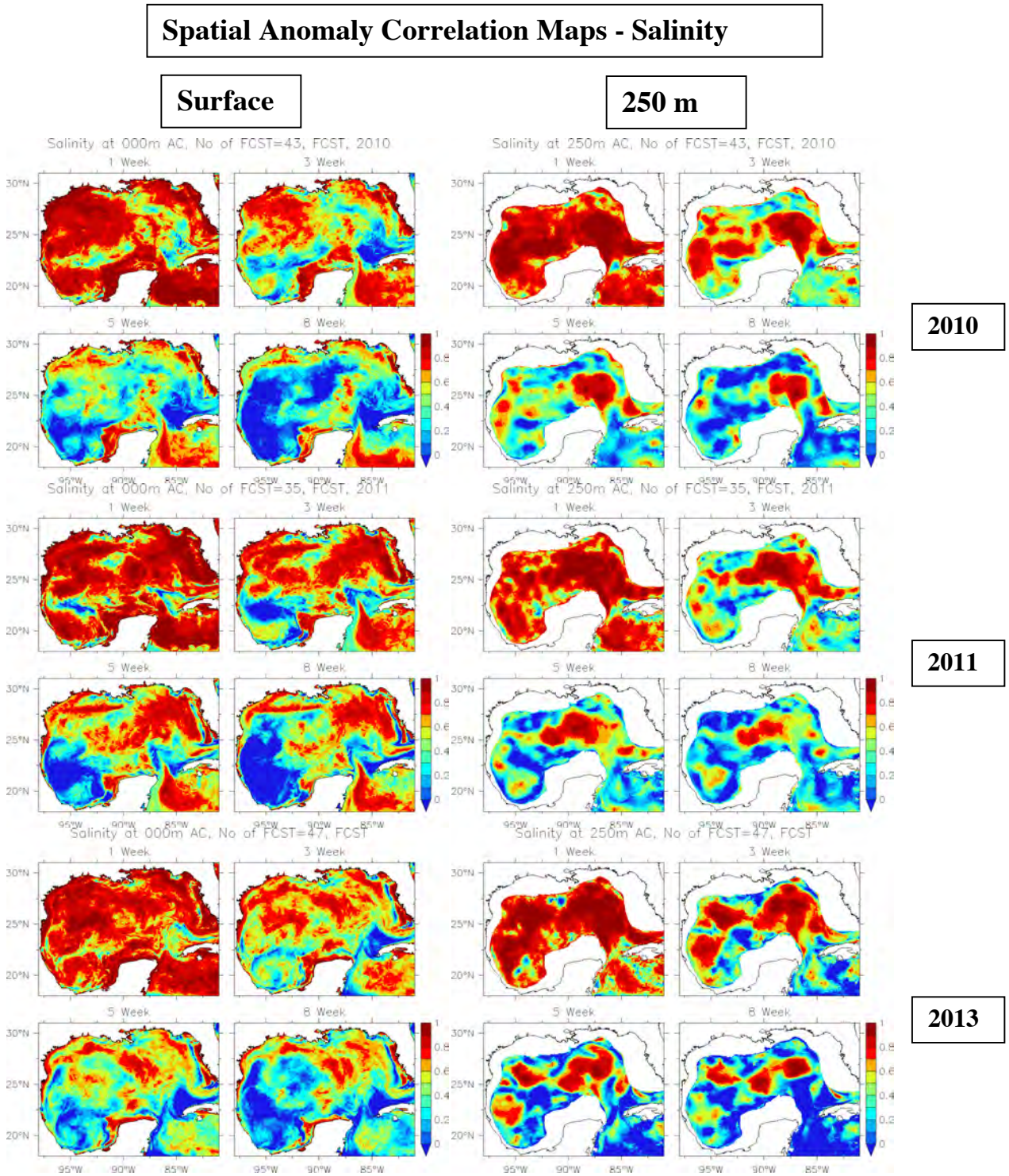


Figure 14: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) salinity for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

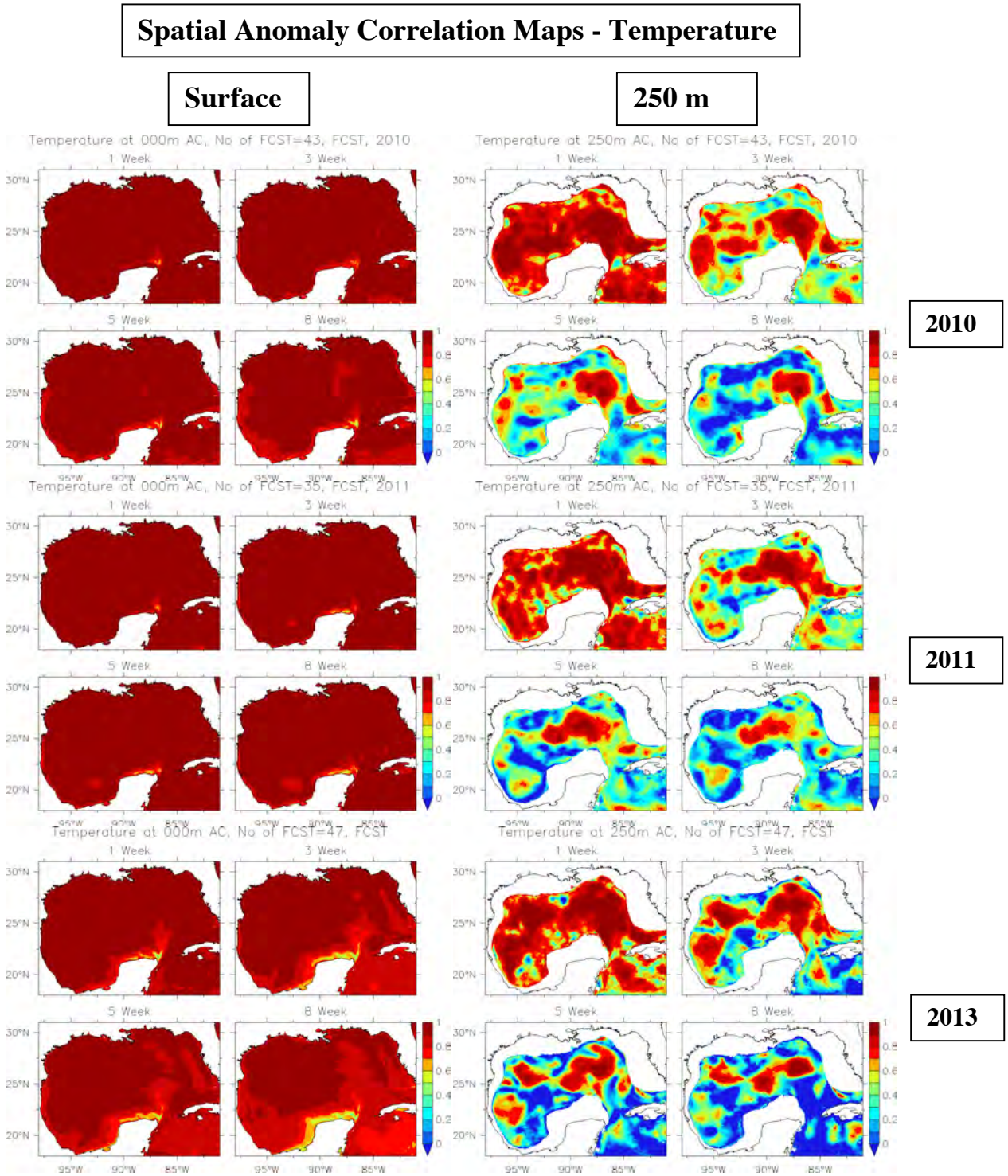


Figure 15: Spatial anomaly correlation (AC) maps of the surface (left) and 250 m (right) temperature for 2010, 2011, and 2013 at 1, 3, 5 and 8 weeks into the forecast period. The correlations reflect the predictability between the forecast and the corresponding best analysis with regions AC greater than 0.6 show good forecast skill and AC of 1 being the perfect forecast. Note that the number of 60-day forecast used for each year is different; that may lead to difference in the skill among the years. Red colors indicate regions of high skill, blue low skill.

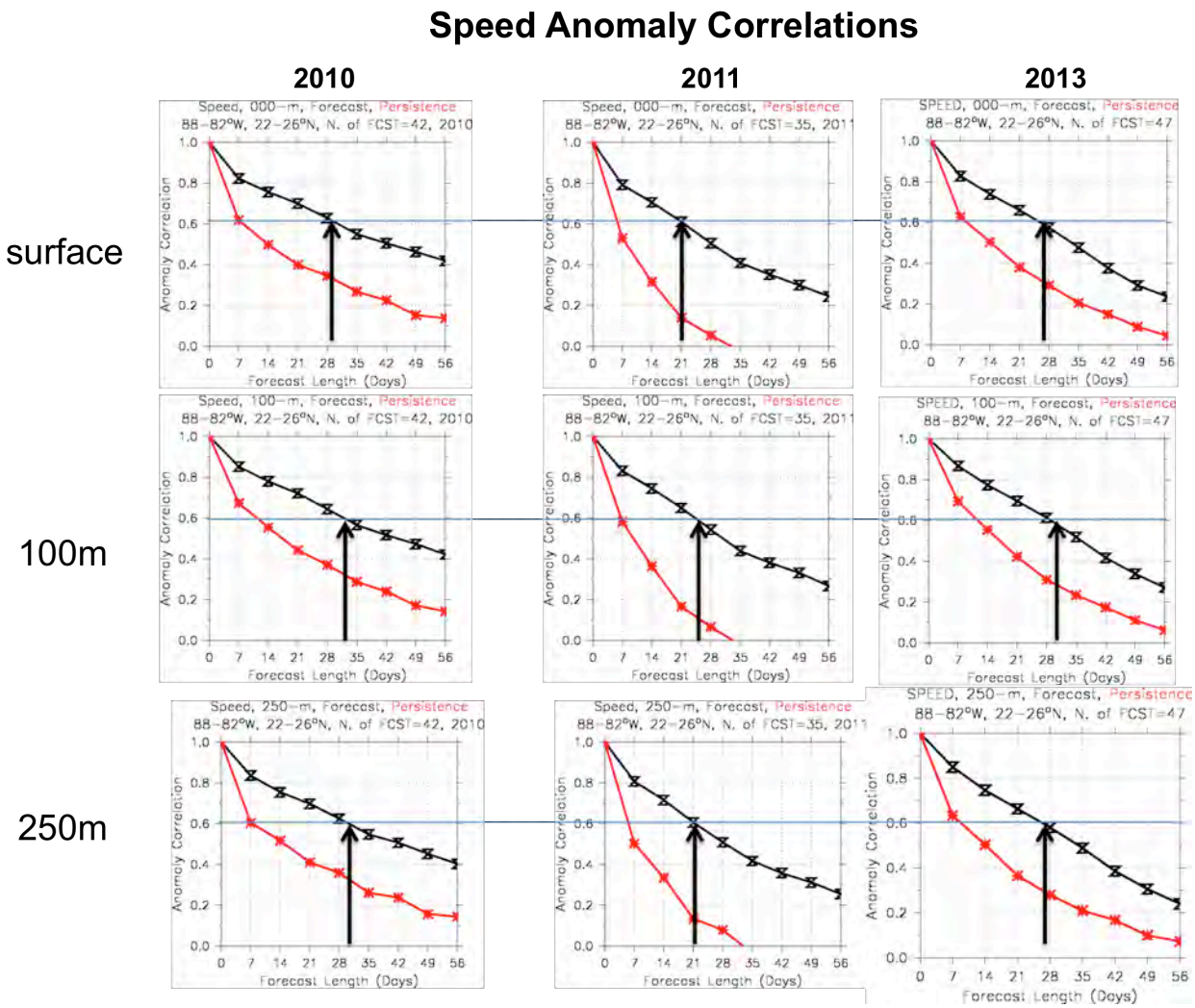


Figure 16: Anomaly correlation of surface, 100 m, and 250 m forecast of current speed with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Salinity Anomaly Correlations

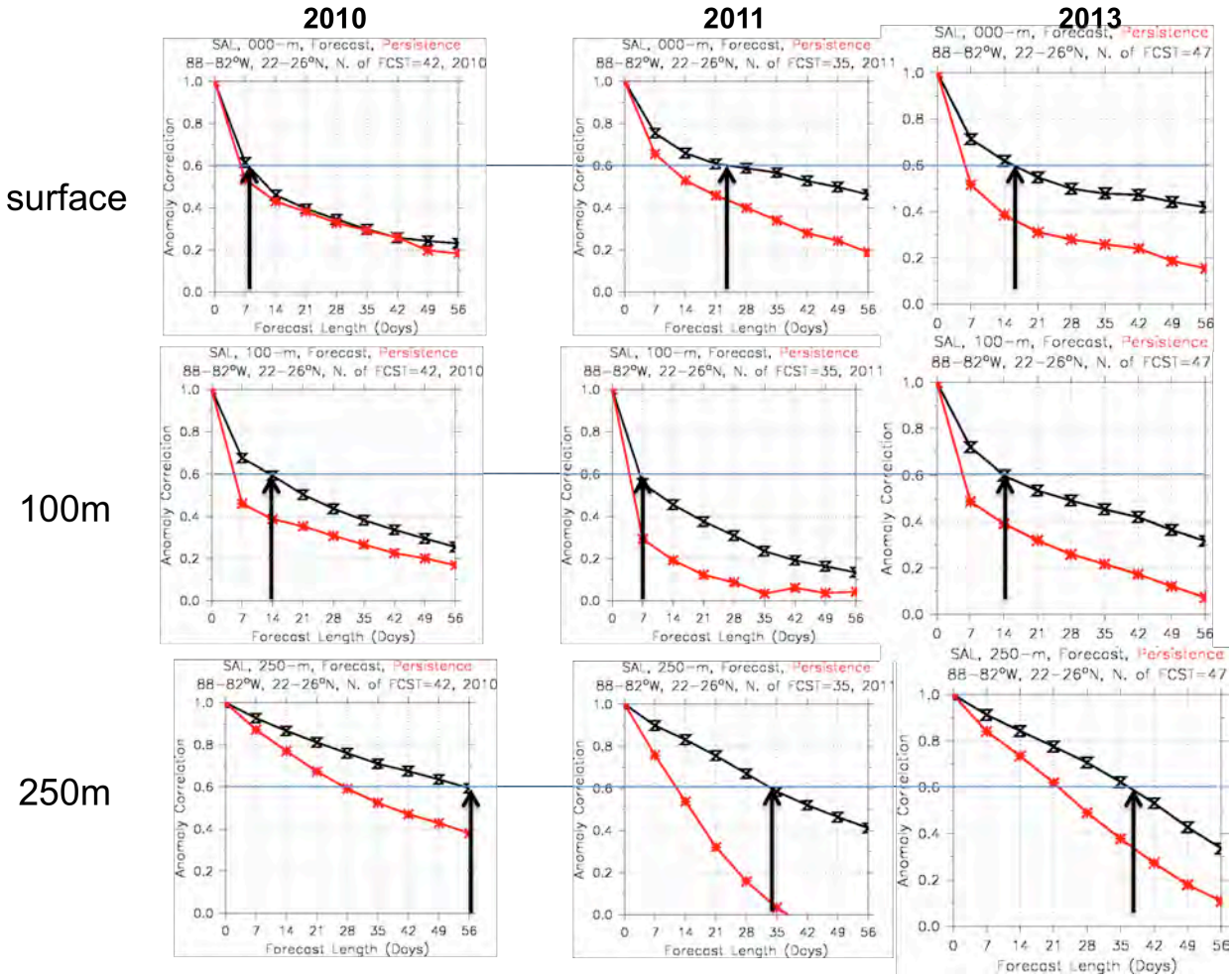


Figure 17: Anomaly correlation of surface, 100 m, and 250 m salinity with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Temperature Anomaly Correlations

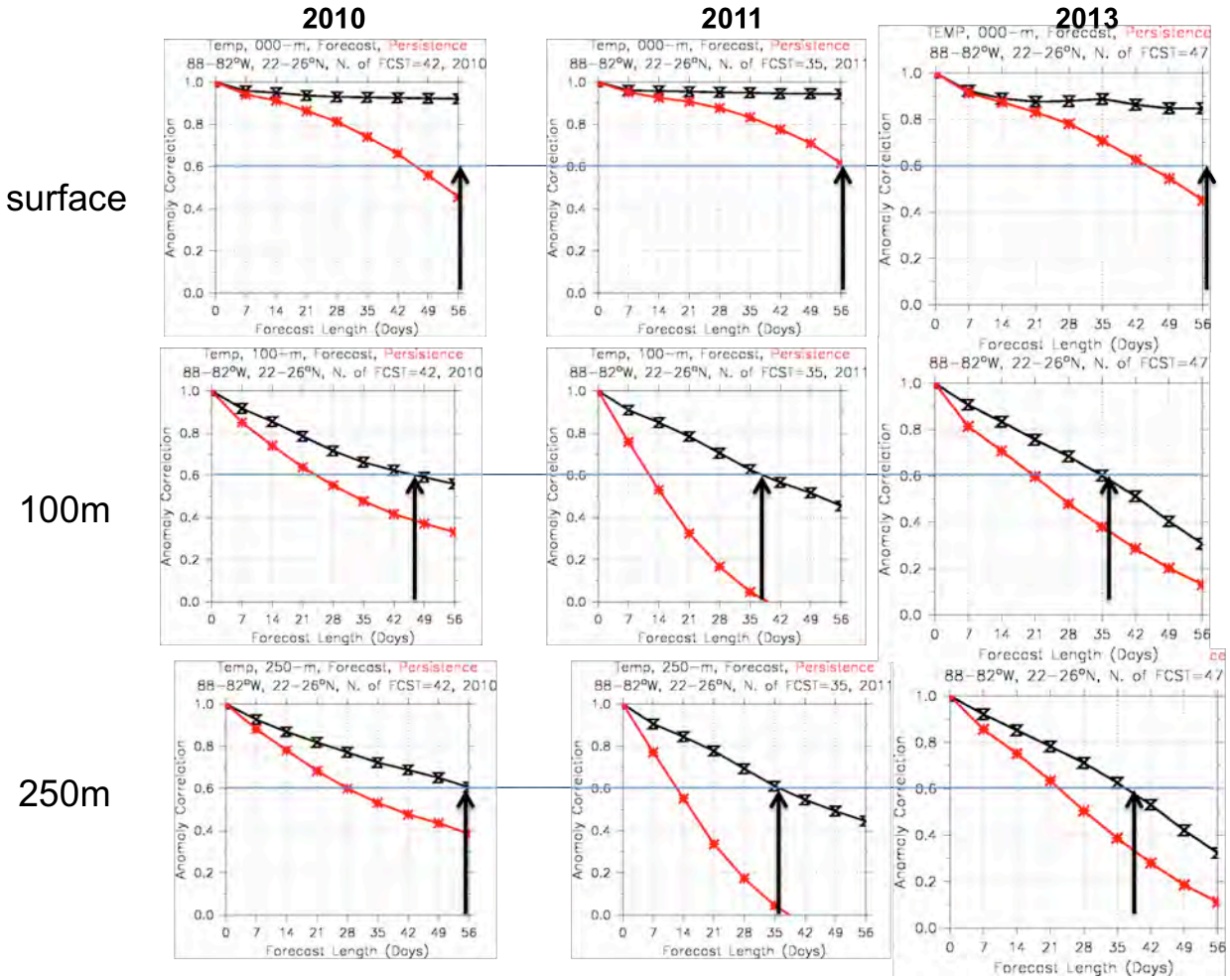
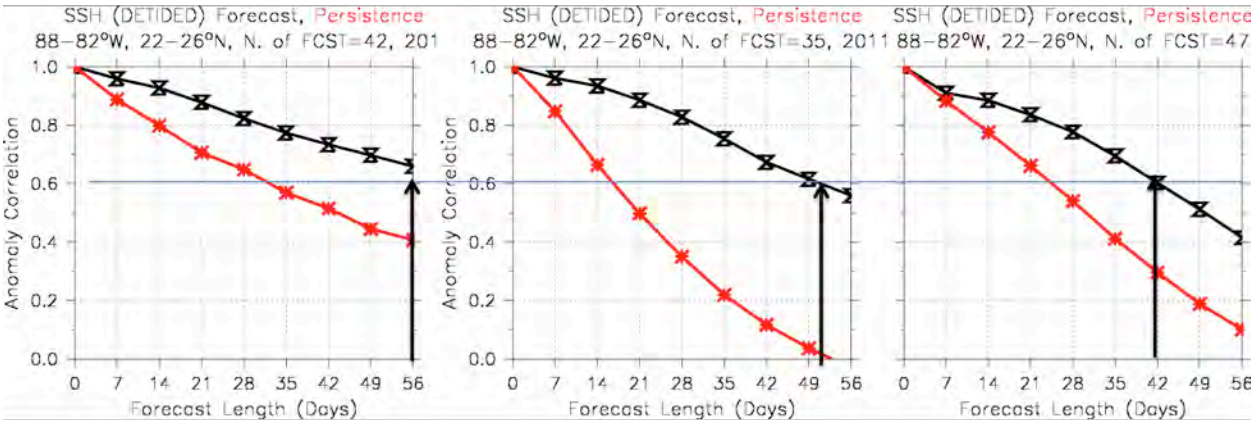


Figure 18: Anomaly correlation of surface, 100 m, and 250 m temperature with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

SSH Anomaly Correlation



2010 **2011** **2013**

Figure 19: Anomaly correlation of de-tided SSH with the corresponding best analysis (truth) for 2010, 2011 and 2013. The black curves correspond to the anomaly correlation computed around the loop current region (88-82W, 22-26N), using 42, 35, 47, 60-day forecasts for 2010, 2011 and 2013 respectively. The red curves show the anomaly correlation assuming a forecast of no change (persistence). The anomaly correlation is calculated based on the long-term analysis mean. An anomaly correlation of 0.6 or greater (indicated by the location of the black arrow) is considered to indicate forecast skill.

Risk Management RAC Codes		Probability			
		A >0.75	0.75>B>0.5	0.5>C>0.25	0.25>D>0.1
Severity	Level I – Very High	1	1	2	3
	Level II - High	1	2	3	4
	Level III - Medium	2	3	4	5
	Level IV - Minimal	3	4	5	5

1 – Critical 2 – Serious 3 – Moderate 4 – Minor 5 – Negligible

Figure 20: Risk Assessment Code. Red indicate a high probability of a severe event, but the severity of the event is application dependent and chosen by the user.

Ensemble Risk Assessment Products

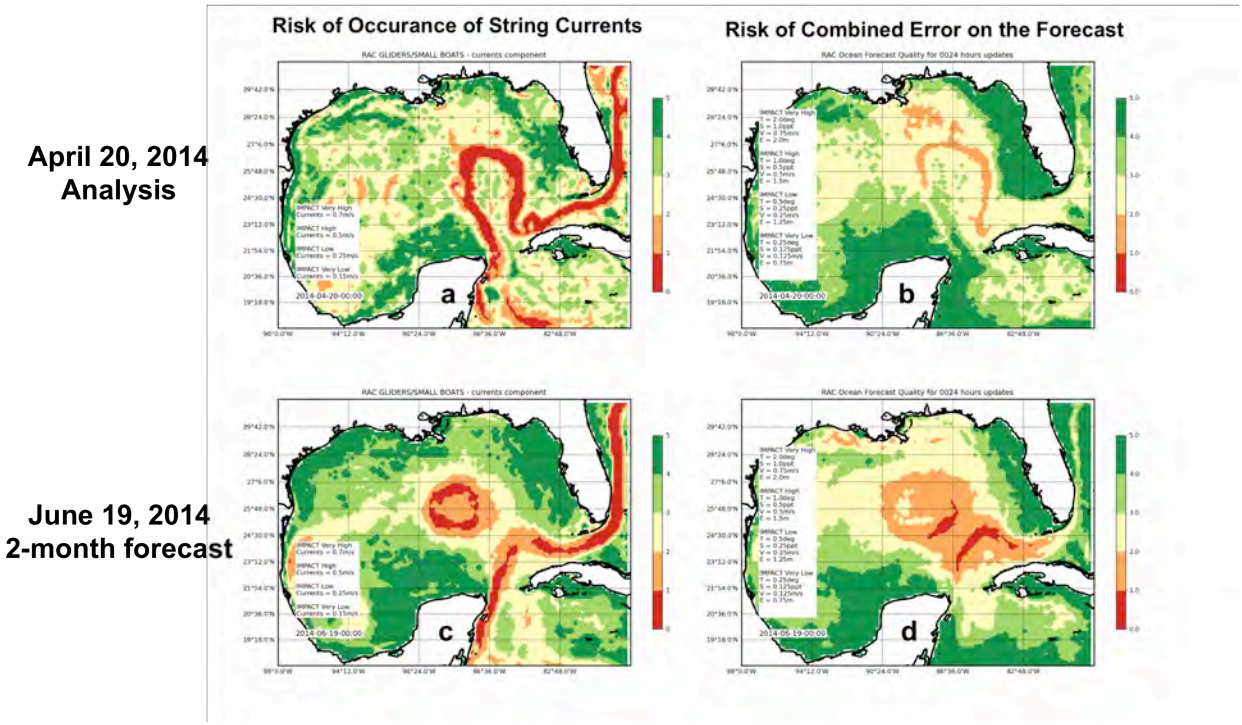


Figure 21: Examples of Risk Assessment Codes (RACs) in the Gulf of Mexico. (a) and (c) are the risk of occurrence of surface currents for specified thresholds (see text) for the analysis on April 20, 2014 and the forecast on June 19, 2014, respectively. (b) and (d) are the risk of combined error in the forecast for the same analysis/forecast. The risk of combined error also takes into account sea state (and thus surface winds) in addition to surface currents. Red areas are high probability of severe impact for specific thresholds.