Naval Research Laboratory

Stennis Space Center, MS 39529-5004

NRL/MR/7320--15-9580

Initial Evaluations of a U.S. Navy Rapidly Relocatable Gulf of Mexico/Caribbean Ocean Forecast System in the Context of the Deepwater Horizon Incident

EDWARD D. ZARON
Portland State University
Portland, Oregon

PATRICK J. FITZPATRICK
Mississippi State University
Stennis Space Center, Mississippi

SCOTT L. CROSS
National Coastal Data Development Center
Stennis Space Center, Mississippi

JOHN M. HARDING
Mississippi State University
Stennis Space Center, Mississippi

FRANK L. BUB
Naval Oceanographic Office
Stennis Space Center, Mississippi

JERRY D. WIGGERT
University of Southern Mississippi
Stennis Space Center, Mississippi

DONG S. KO
Ocean Dynamics and Prediction Branch
Oceanography Division

YEE LAU
Mississippi State University
Stennis Space Center, Mississippi

KATHARINE WOODARD
University of Southern Mississippi
Stennis Space Center, Mississippi

CHRISTOPHER N.K. MOOERS
Portland State University
Portland, Oregon

May 6, 2015

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
06-05-2015	Memorandum Report	

4. TITLE AND SUBTITLE

Initial Evaluations of a U.S. Navy Rapidly Relocatable Gulf of Mexico/Caribbean Ocean Forecast System in the Context of the Deepwater Horizon Incident

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0602435N

6. AUTHOR(S)

Edward D. Zaron,[1] Patrick J. Fitzpatrick,[2] Scott L. Cross,[3] John M. Harding,[4] Frank L. Bub,* Jerry D. Wiggert,[5] Dong S. Ko, Yee Lau,[4] Katharine Woodard,[5] and Christopher N.K. Mooers[1]

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
73-6669-05-5

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Oceanography Division
Stennis Space Center, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/MR/7320--15-9580

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Reseach
One Liberty Center
875 North Randolph Street, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)

ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

[1]Portland State University, Department of Civil and Environmental Engineering, Portland, OR
[2]Mississippi State University, Geosystems Research Institute, MSU Science & Technology Center, Stennis Space Center, MS
[3]National Coastal Data Development Center, Stennis Space Center, MS
[4]Mississippi State University, Northern Gulf Institute, MSU Science & Technology Center, Stennis Space Center, MS
[5]University of Southern Mississippi, Department of Marine Science, Stennis Space Center, MS
*Retired. Formerly at Naval Oceanographic Office, Stennis Space Center, MS

14. ABSTRACT

In response to the Deepwater Horizon (DwH) oil spill event in 2010, the Naval Oceanographic Office deployed a nowcast-forecast system covering the Gulf of Mexico and adjacent Caribbean Sea that was designated Americas Seas, or AMSEAS, which is documented in this manuscript. The DwH disaster provided a challenge to the application of available ocean-forecast capabilities, and also generated a historically large observational dataset. AMSEAS was evaluated by four complementary efforts, each with somewhat different aims and approaches: a university research consortium within an Integrated Ocean Observing System (IOOS) testbed; a petroleum industry consortium, the Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP); a British Petroleum (BP) funded project at the Northern Gulf Institute in response to the oil spill; and the Navy itself. Validation metrics are presented in these different projects for water temperature and salinity profiles, sea surface wind, sea surface temperature, sea surface height, and volume transport, for different forecast time scales. The validation found certain geographic and time biases/errors, and small but systematic improvements relative to earlier regional and global modeling efforts. On the basis of these positive AMSEAS validation studies, an oil spill transport simulation was conducted using archived AMSEAS nowcasts to examine transport into the estuaries east of the Mississippi River. This effort captured the influences of Hurricane Alex and a non-tropical cyclone off the Louisiana coast, both of which pushed oil into the western Mississippi Sound, illustrating the importance of the atmospheric influence on oil spills such as DwH.

15. SUBJECT TERMS

Gulf of Mexico
Deepwater Horizon
Ocean forecasting
Skill assessment

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dong Shan Ko
Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	61	**19b. TELEPHONE NUMBER** *(include area code)* (228) 688-5448

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

CONTENTS

I. INTRODUCTION

A. Rationale

The U.S. Navy has been a leader in developing and operating ocean forecast systems for more than a decade (e.g., Rhodes et al., 2002). Through its research and development (R&D) and operational oceanography arms (Naval Research Laboratory [NRL] and Naval Oceanographic Office [NAVOCEANO], respectively), the Navy began deploying global-scale nowcast-forecast systems in 2000. Since 2008, Navy operational capabilities have included a rapid-response modeling capability that allows the deployment of nested regional forecast models within a global nowcast/forecast system (Peggion et al., 2007). These nested models may be spun up rapidly from climatology or the global nowcast, and used to address specific Navy needs on a short lead time.

As a response to the Deepwater Horizon oil spill event in 2010 (DwH), NAVOCEANO deployed a high-resolution data-assimilating nowcast-forecast system covering the Gulf of Mexico and adjacent Caribbean Sea, nested within the operational global Navy Coastal Ocean Model (Global NCOM). This new regional model domain came to be designated Americas Seas, or AMSEAS. After a short spin-up and initial evaluation, AMSEAS model forecasts began to be released to the public and became a part of NOAA's official spill-trajectory forecast process for DwH, along with several other operational or quasi-operational ocean prediction systems (MacFadyen et al., 2011).

The AMSEAS evaluation process continued throughout the initial DwH response and beyond with four complementary efforts, each with somewhat different aims and approaches. These groups include a university research consortium on behalf of an Integrated Ocean Observing System (IOOS) coastal modeling testbed; a petroleum industry consortium; BP-funded university research through the Northern Gulf Institute; and the Navy itself. Here, these evaluations are summarized within the common context of oil-spill response, and some observations on the state-of-the-art with respect to operational ocean prediction at the time of the DwH incident are offered.

The past decade has seen particularly rapid advances in operational ocean-prediction capabilities. At the same time, demand has grown for nowcast and forecast information to support ocean operations and resource-management activities such as search-and-rescue,

Manuscript approved October 28, 2014.

safe management of offshore oil and gas platforms, oil spill mitigation, marine weather forecasting, fisheries and ecosystem management, and adaptive observing-system deployments. As users begin to apply ocean predictions, questions naturally arise about the relative accuracy and uncertainty of their various products, especially among those products considered operational. Skill assessments need to be broad to accommodate the wide range of potential user applications and requirements. Ideally they will treat qualitative as well as quantitative aspects of model performance, addressing properties of various prognostic fields (e.g., sea surface temperature and sea surface height), as well as synoptic, dynamical features (e.g., Loop Current position and strength) depending on the intended application. The R&D and the operational prediction communities have a joint interest in these skill assessments although their motivations, priorities, metrics and standards may differ. The R&D community is able to provide new dynamical and statistical insights and methodologies for the operational community and vice versa.

B. Overview of Navy Operational Ocean-Prediction Activities

The foundation of the Navy's ocean prediction capabilities is a collection of data-assimilating nowcast/forecast systems. In these systems, model output from a prior forecast cycle is combined with recent observations in a statistical analysis to form the initial condition for the present forecast cycle. A global-scale system provides boundary conditions for nested regional models that have higher spatial resolution. In 2010, the global-scale host model was the Global Navy Coastal Ocean Model (GNCOM, Rhodes et al., 2002); since replaced by the Navy Global Hybrid Coordinate Ocean Model (GHYCOM; Chassignet et al., 2009). The dynamical core of the GNCOM system was NCOM, a four-dimensional, primitive-equation, free-surface, hydrostatic ocean model that uses hybrid (sigma- and z-level) coordinates in the vertical (Martin, 2000; Barron et al., 2007). GNCOM had a nominal horizontal resolution of $1/8^\circ$ at mid-latitudes, with 40 depth levels (19 sigma levels in shallow water and 21 z-levels below 137 m depth). Atmospheric forcing was provided by the global-scale Navy Operational Global Atmospheric Prediction System (NOGAPS; Rosmond et al., 2002). Observations were assimilated using the Navy Coupled Ocean Data Assimilation (NCODA) system using multivariate optimum interpolation (MVOI) as described by Cummings (2005). GNCOM assimilated remotely-sensed sea surface temperature

and in situ temperature and salinity data as well as synthetic temperature and salinity profiles derived from satellite altimetry; see (Rhodes et al., 2002) for a fuller description of the GNCOM data assimilation process. Barotropic tidal elevations and currents from the quarter-degree resolution Oregon State University (OSU) tidal model were linearly added after the GNCOM run. At 1/8° resolution, GNCOM resolved ocean features of the order of 1/2° without aliasing or significant time-stepping errors.

In 2002, NRL developed as a research prototype an NCOM-based regional-scale prediction system called the Intra-Americas Sea Nowcast Forecast System (IASNFS), covering the Gulf of Mexico and Caribbean regions (Ko et al., 2003; Arnone et al., 2007; Haltrin et al., 2007; Green et al., 2008; Ko et al., 2008; D'Sa and Ko, 2008; Mendoza et al., 2009; D'Sa et al., 2011). IASNFS has run in real-time mode at NRL since 2003 and produces a nowcast and 72-h forecast daily. The nowcast/forecast has been used to support many operations such as the Navy's Haiti earthquake relief effort. In addition, IASNFS has served as a host to embedded higher resolution coastal models and applied to several R&D efforts (e.g., Chassignet et al., 2005; Arnone et al., 2007; Haltrin et al., 2007; Green et al., 2008; D'Sa and Ko, 2008; Mendoza et al., 2009; Arnone et al., 2010; D'Sa et al., 2011). The horizontal resolution of IASNFS is nominally 1/24° (4.6 km) and, as in GNCOM, there are 40 layers in the vertical. Data assimilation is accomplished by the Navy's Modular Ocean Data Assimilation System (MODAS), which combines sea-surface elevation from satellite altimeters (Jacobs et al., 2002) and sea-surface temperature from the Advanced Very High Resolution Radiometer (AVHRR) as well as available in situ surface and profile temperatures. As with GNCOM, the surface wind, air pressure and heat fluxes are supplied by NOGAPS. The lateral open boundary conditions of sea surface height, temperature, salinity and current are taken from 1/8° global GNCOM. As a mature ocean prediction system of proven utility, the IASNFS provides a useful benchmark against which to judge newer Navy (and other) systems, and some basic intercomparisons are performed in Sections II B and II C, below.

Building on the IASNFS prototype, NRL developed the relocateable NCOM (RNCOM) capability with horizontal resolutions of 1/36° or 3 km for a downscaling ratio of 1:5 relative to GNCOM. Even higher resolution coastal NCOM domains may be nested within the RNCOMs with 500-m (1/220°) resolution for a further downscaling of 1:6. The dynamical core of RNCOM is identical to GNCOM. The RNCOM assimilation scheme is a version of NCODA.

Unlike IASNFS, the RNCOM applications receive their atmospheric forcing from regional versus global atmospheric forecasts; RNCOM domains are forced by momentum and heat fluxes from the operational, high-resolution (6-15 km) Fleet Numerical Meteorological and Oceanography Center's Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS; Hodur, 1997; Hodur et al., 2002), where COAMPS is run in atmosphere prediction mode only. GNCOM boundary conditions of temperature, salinity, perpendicular and tangential currents, and surface elevation are applied at the domain boundaries. The barotropic tidal elevations are inserted at the boundaries as anomalies relative to a GNCOM mean field.

C. Deepwater Horizon Application within RNCOM AMSEAS Domain

The DwH event accelerated the initial implementation of a new operational RNCOM applied to the Gulf of Mexico. Named for the semi-enclosed seas connecting North, Central and South America, the AMSEAS model covers the Gulf of Mexico and Caribbean Sea (Fig. 1) and was initiated shortly after the April 2010 DwH oil spill. AMSEAS required fewer than two weeks of model time to spin up to a stable state from a GNCOM initial condition. The eastern boundary at 55°W was placed in deep water east of the Windward Islands and the northern boundary at 32°N is north of the Bahamas. The western and southern boundaries are confined by land. The AMSEAS grid has a horizontal resolution of 1/32° (3 km) and 55 vertical layers, with sigma levels down to 550-m and z-levels below that to 5000-m. Model output is interpolated to a regular grid in the horizontal and 40 standard levels in the vertical, with 3-h outputs saved as NetCDF files.

NAVOCEANO made preliminary results available for all three ocean models beginning early May 2010 via the NOAA OceanNOMADS web portal (`http://ecowatch.ncddc.noaa.gov/`; Harding et al., 2013). During the DwH event, GNCOM, IASNFS and AMSEAS were three of the forecast systems that provided daily input to the NOAA Office of Restoration and Response for their operational oil spill trajectory predictions in support of the Coast Guard and the Unified Command (MacFadyen et al., 2011). Table I provides a general summary of the attributes of the three ocean models described above.

The DwH disaster in 2010 provided a real-world, urgent challenge to the application of

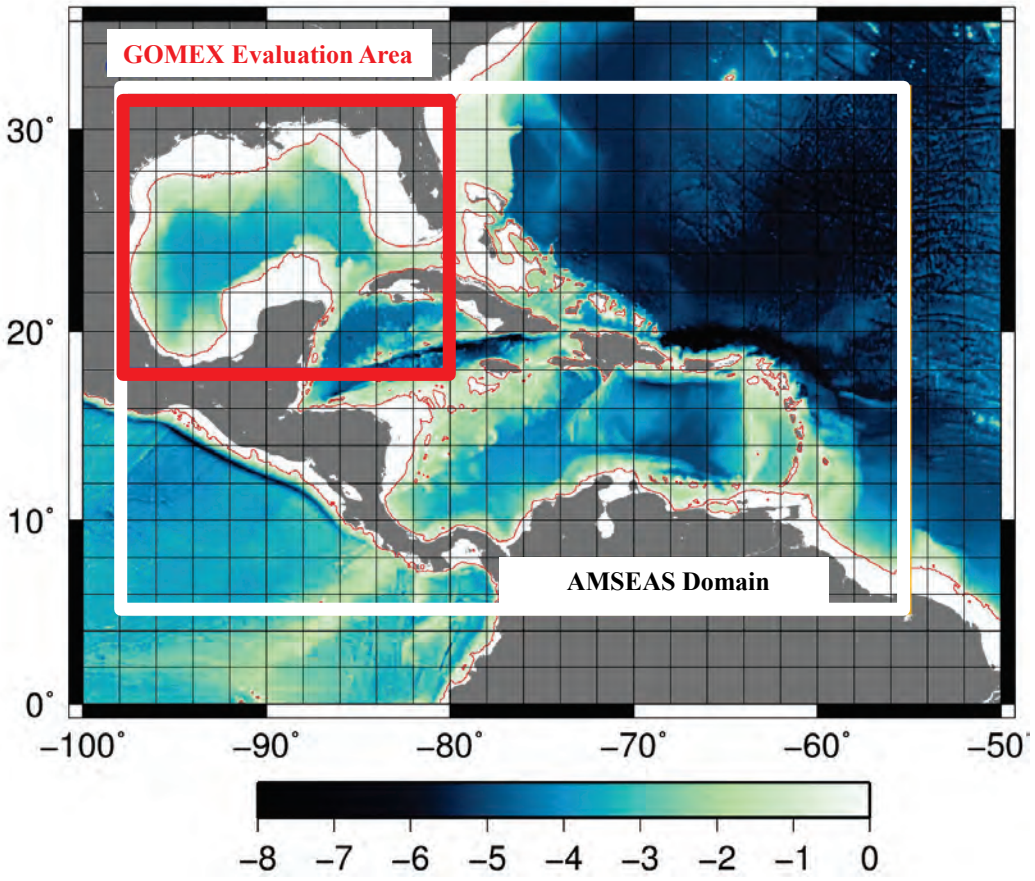

FIG. 1. AMSEAS domain with the insert outlining the Gulf of Mexico evaluation area. Depth scale in km.

available, accurate ocean-forecast capabilities. At the same time, DwH led to a historically large observational dataset. The works of Liu et al. (2011) and Lubchenco et al. (2012) supply an initial compilation of some of the DwH-related opportunities as well as important lessons learned by the research and operational communities.

D. Oceanographic Context for Evaluations

1. Gulf water masses

The Caribbean Sea serves as a conduit between the waters of the equatorial Atlantic and the downstream Gulf of Mexico. The warm North Atlantic waters of the westward flowing North Equatorial Current (NEC), supplemented by freshwater outflow from the South American Coast, flow into the eastern Caribbean Sea. South Atlantic Water is aperiodically

TABLE I. Summary of general forecast model attributes for the operational Global Navy Coastal Ocean Model (GNCOM; Rhodes et al. 2002), and two regional models nested within GNCOM covering the Gulf of Mexico and Caribbean regions. One model is a research and development NCOM-based regional-scale prediction system called the Intra-Americas Sea Nowcast Forecast System (IASNFS; Ko et al. 2003) with atmospheric forcing by the Navy Operational Global Atmospheric Prediction System (NOGAPS). The second was operationally implemented as a response to the 2010 DwH oil spill event, designated as Americas Seas (AMSEAS) with atmospheric forcing by the Coupled Ocean-Atmosphere Prediction System (COAMPS). See text and references for details on these models and their data assimilation schemes.

Ocean Model	Domain	Nominal Resolution	Atmospheric Forcing	Data Assimilation	Status
GNCOM	Global	15 km	NOGAPS (50 km)	NCODA (MVOI)	OPS
IASNFS	Gulf of Mexico & Caribbean	5 km	NOGAPS (50 km)	MODAS	R&D
AMSEAS	Gulf of Mexico & Caribbean	3 km	COAMPS (15 km)	NCODA (MVOI)	OPS

carried into the region in large anti-cyclonic eddies that break off from the retroflection of the North Brazil Current and move westward along the coast of South America. The Windward and Leeward Islands act as land and sill obstacles for Atlantic waters entering the Caribbean. This archipelago results in a highly variable westward extension of the NEC and limits the entry of intermediate and deep Atlantic water (Wilson and Johns, 1997). The core of the relatively weak westward current across the Caribbean Sea separates a loosely organized series of anticyclonic eddies to the north and cyclonic eddies to the south.

Nearly all of the Caribbean waters exit through the deep (2000 m) Yucatan Channel that connects to the Gulf of Mexico. The usual path is northward to form the Loop Current (LC), a semi-permanent, anticyclonic flow that can penetrate several hundred kilometers into the east-central Gulf. The LC then exits the Gulf, enters the Straits of Florida as the Florida Current between the Florida Keys and Cuba, and eventually evolves into the Gulf Stream to the north. With the Straits of Florida sill at about 800 m, much of the

Gulf of Mexico water between this depth and its 4000-m bottom depths remains trapped in the Gulf. In the northern Gulf, freshwater sources, including the Mobile River and especially the Mississippi and Atchafalaya Rivers, create a semi-permanent salinity front evident 70 to 150 km offshore (Morey et al., 2003). These rivers provide nutrients from the continent that result in the large annual hypoxic or dead zone in shelf waters off Texas and Louisiana (Rabalais et al., 2001), as well as sporadic hypoxic events in the Mississippi Bight (Brunner et al., 2006).

2. Gulf dynamics

The LC, its meanders, the large anticyclones it sheds, plus the smaller frontal cyclones constitute the major dynamical features of the Gulf. Though the northern Gulf freshwater inflow has profound influence on the circulation of the shelf waters, it is of secondary importance in the overall Gulf circulation. In contrast, the LC, when extended to its northernmost state, can play a critical role in entraining freshwater plumes into the central Gulf. The LC circulation and thermo-haline structure responds to the large-scale, seasonal atmospheric circulation and synoptic-scale weather systems, especially summertime easterly waves and tropical cyclones and wintertime cold fronts and extratropical cyclones. These events have much greater impact on the circulation and mixing of the shelf waters of the northern Gulf than elsewhere.

The warm-core LC can reach surface speeds in excess of 2 ms^{-1} and extend to depths greater than 500 m. The LC structure varies as it extends into the Gulf with small cyclonic eddies developing several times per year along the cyclonic side of the inflow where the LC impinges the western continental slope of the Yucatan Channel.

These cyclonic eddies translate around the cyclonic edge of the LC and may play a role in cutting it off, diverting the Yucatan inflow directly eastward into the Straits of Florida and releasing a large anti-cyclonic eddy into the Gulf (Schmitz Jr. et al., 2005). The eddy separation process is sporadic and broadband (once per six months to two years) and may involve a number of reconnections before the fully separated anticyclonic eddy finally migrates westward at 10 km/d. Several of these older eddies may co-exist in the Western Gulf where they eventually dissipate. During the 2010 DwH event in particular, the presence and evolution of "Eddy Franklin" played a major role in the transport (and trapping) of oil

in the northeastern Gulf region (Liu et al., 2011).

Understanding the evolution of the LC and mechanisms for the formation of LC eddies is hampered by the lack of a comprehensive ocean observational network in the Gulf, but the evolving realism of ocean models is leading to improved diagnostic studies of LC processes within them (Xu et al., 2013). The extent of the LC intrusion into the Gulf is thought to be determined by the northward mass and potential vorticity fluxes through the Yucatan Strait (Lugo-Fernández and Leben, 2010; Chang and Oey, 2011). Variability of the LC is influenced by the strength of the easterly trade winds in the Caribbean which set the stratification and potential vorticity of water passing through the Yucatan Strait (Chang and Oey, 2013). Unlike the Gulf Stream, which continuously sheds eddies through hydrodynamic instability, it appears that the LC eddy shedding process is triggered or modulated by external forcing. Different mechanisms may be responsible for triggering the formation of a LC eddy, but momentum balance dictates that the LC intrusion cannot remain steady (Pichevin and Nof, 1997; van Leeuwen and De Ruijter, 2009). The hypothesis has been advanced that Atlantic ocean variability propagates upstream through the Straits of Florida and initiates the separation (Sturges et al., 2010); however, recent analysis of models suggests tight coupling of LC dynamics to deep processes, the upstream conditions, and winds (Chang and Oey, 2011).

II. EVALUATIONS

A. Introduction to Evaluations and Classes of GODAE Metrics

The four evaluation efforts that are summarized here each derive from somewhat different motivations, but together provide a broad characterization of AMSEAS. Table II provides a summary of each of the evaluation efforts including the models evaluated; general locations of the evaluation; time period of each particular evaluation; variables that were compared; time and space scales of interest; and the specific purpose of each study. Fig. 1 shows the AMSEAS computational domain as well as the Gulf of Mexico subdomain that is the subject of these evaluations.

To provide logic to the presentations of model evaluations, they are organized according to the following groupings proposed by the GODAE project (Hernandez et al., 2009):

- Class 1 metrics are instantaneous views of the ocean state, to give a qualitative impression of the realism of the results.

- Class 2 metrics are direct comparisons of model outputs with in situ observations to quantitatively assess the goodness-of-fit and accuracy of the model and its forcings.

- Class 3 metrics are comparisons of derived quantities, e.g., volume transports, with observations.

- Class 4 metrics are skill assessments of model forecasts.

The evaluations begin with a series of Class 1 metrics computed within the Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP), which was initiated (coincidentally) in the same time frame as DwH. This project was sponsored by the Research Partnership to Secure Energy for America with the aim of supporting safe and efficient drilling operations as part of the drive for energy independence. The GOMEX-PPP work considered the seasonal time scales, and evolving mesoscale structures, that are applicable to the offshore energy industry, concentrating on the location and energetic currents of the LC and its Loop Current eddies (LCE). GOMEX-PPP contributes Class 1 intercomparisons of nowcasts from IASNFS and AMSEAS, as well as Class 2 and Class 3 metrics for AMSEAS using altimeter data and Florida Current transport.

A second group of evaluations was undertaken as part of the IOOS Coastal Ocean Modeling Testbed (COMT; Luettich et al., 2013), a multi-institutional effort to improve and accelerate the transfer of coastal ocean modeling R&D results to the operational prediction community. The COMT evaluations consist of Class 2 intercomparisons of buoy data with synoptic-scale surface wind products used to force the Navy forecasts. Additional Class 2 comparisons were performed to evaluate the vertical structure of temperature and salinity over the shelf.

A third evaluation, consisting of Class 3 comparisons, was performed to examine the efficacy of using AMSEAS surface currents for oil spill modeling. In this BP-funded effort, Lagrangian particle trajectories were used to examine atmospheric influences on the DwH oil spill, with a particular focus on pollution transport into the estuaries east of the Mississippi River.

The final set of evaluations presented were performed by NAVOCEANO as a part of

TABLE II. Summary of each evaluation effort including the models evaluated; general locations of the evaluation; time period of each particular evaluation; variables that were compared; time and space scales of interest; and the specific purpose of each study. The Naval Oceanographic Office (NAVOCEANO) evaluation was part of the Navys formal operational testing. The second was part of the Integrated Ocean Observing System (IOOS) Coastal Ocean Modeling Testbed (COMT). A third effort was associated with the Gulf of Mexico 3-D Operational Ocean Forecast System Pilot Prediction Project (GOMEX-PPP). Both the Navy and COMT efforts focused on Navy needs in the realm of nowcasts and few-day forecasts of temperature and salinity structure at representative depth levels over the shelves and deeper waters of the Gulf; an additional component of the COMT evaluated the synoptic-scale surface wind products used to force the Navy forecasts. The GOMEX-PPP work considered the seasonal-time-scale needs and mesoscale structures that are especially applicable to the offshore energy industry, concentrating on the location and energetic currents of the Loop Current and its Loop Current eddies. The fourth evaluation, funded by BP, examines the efficacy of using hindcast AMSEAS for oil spill modeling using a Lagrangian particle tracker.

Effort	NAVOCEANO	COMT Met.	COMT Ocean	GOMEX-PPP	BP
Model	AMSEAS	COAMPS	AMSEAS	AMSEAS & IASNFS	AMSEAS & COAMPS
Location	Gulf of Mexico	Northern Gulf Mexico	Gulf of Mexico	Loop Current & Fl. Straits	Northern Gulf Mexico
Time Period	Jun 2010 – Mar 2011	Jun/Jul 2010 & Dec/Jan 2011	Jun 2010 – Oct 2011	May 2010 – Dec 2010	20 Jun – 10 Jul 2010
Time Scale	00-72 h fcst	00-24 h fcst	00-96 h fcst	Daily averaged nowcast	n/a
Purpose	Operational QC/QA	SURA Evaluation	SURA Evaluation	Model Intercomparison	Oil spill simulations

the Navy's formal operational testing. Comparisons with in situ temperature and salinity measurement are utilized in the Navy's unique forecast skill assessment system (Class 4 metrics) tailored to its operational needs.

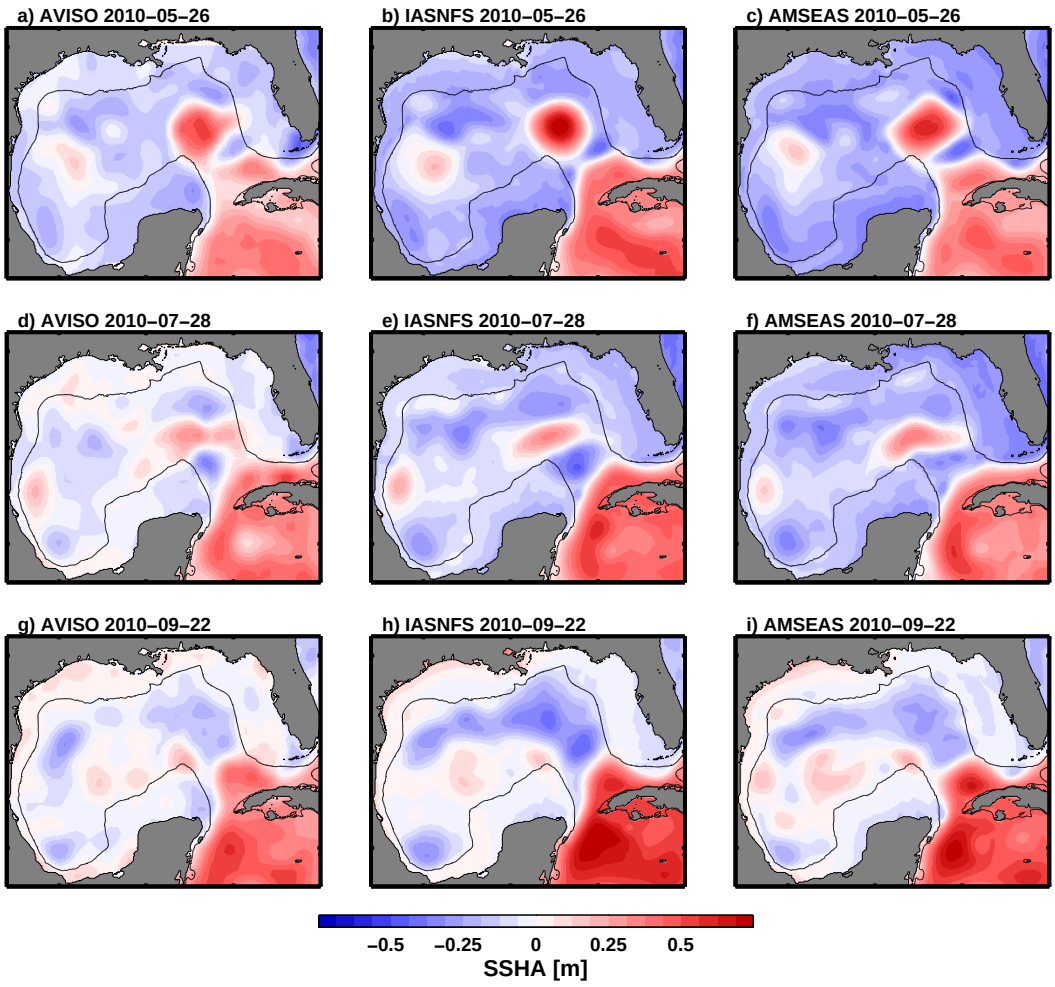

FIG. 2. Comparison of sea-surface height anomaly (SSHA) in AVISO hindcast (a, d, g – derived from satellite altimetry), IASNFS nowcast (b, e, h), and AMSEAS nowcast (c, f, i), for three dates in 2010. The solid contour represents the 500-m isobath.

B. Class 1: Overview of IASNFS and AMSEAS (GOMEX-PPP)

The evolution of the LC during the DwH time frame has been extensively described elsewhere (Liu et al., 2011). Here, snapshots of IASNFS and AMSEAS are used from this time period to illustrate the general similarity of the nowcasts and forecasts from these systems.

Fig. 2 shows snapshots of sea surface height (SSH) at two-month intervals from late May, July, and September of 2010, and compares the AVISO multi-satellite hindcast analysis product (`http://www.aviso.oceanobs.com/en/`) with the IASNFS and AMSEAS nowcasts for the same dates. The May snapshot (top row) shows the general similarity of the models

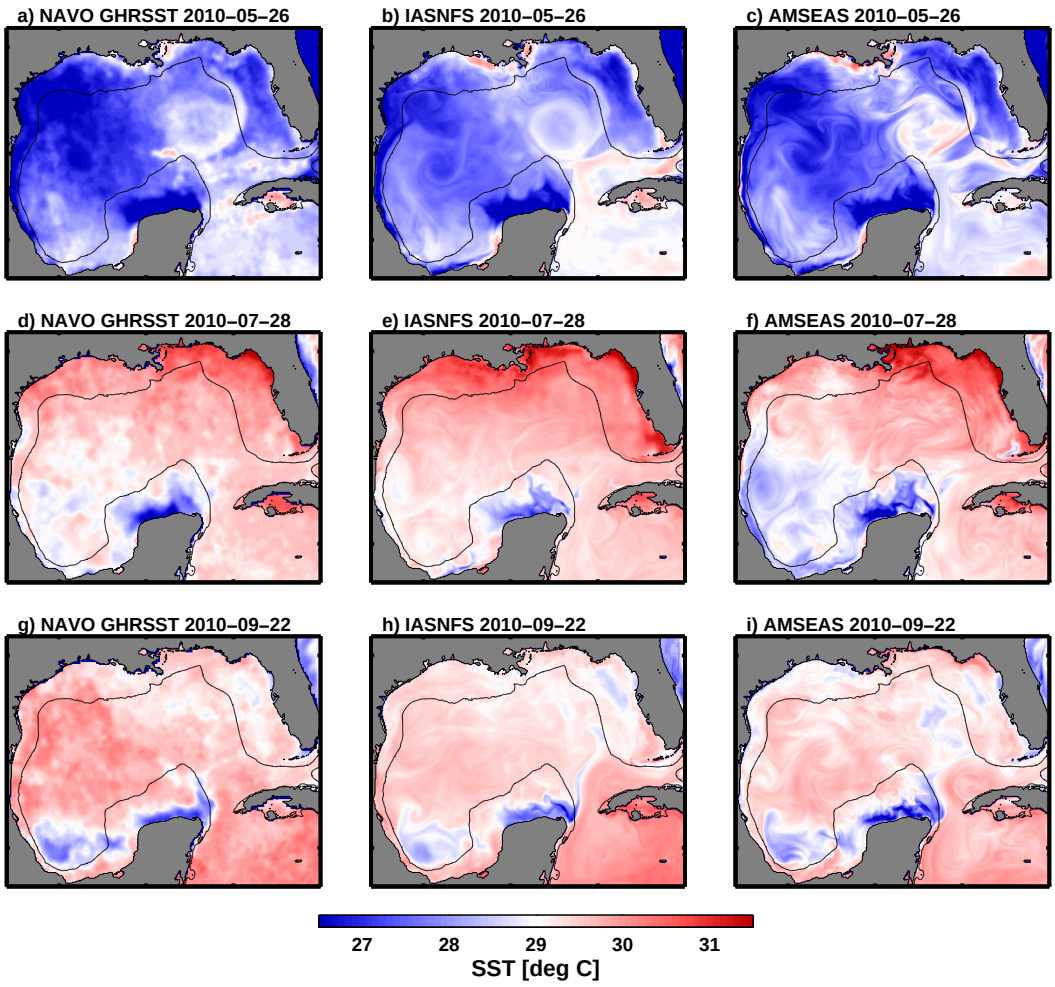

FIG. 3. Comparison of sea-surface temperature (SST) in NAVO GHRSST analysis (a, d, g – derived from multiple satellite measurements), IASNFS nowcast (b, e, h), and AMSEAS nowcast (c, f, i), for three dates in 2010. The solid contour represents the 500-m isobath.

and altimetric representations the LC and Eddy Franklin. Beyond the basic similarity of patterns; however, there are many differences in details. The shape of the large anticyclone (Eddy Franklin) is less symmetric in AMSEAS and AVISO, as compared to IASNFS, and it is bordered by cyclones to the northeast and southeast. In late July the SSH expression of the large anticyclone has been deformed and reduced in amplitude. By September the SSH signal is dominated by LC's path from the Yucatan to the Straits of Florida, with a cyclonic circulation pattern emerging in much of the central Gulf.

The sea-surface temperature fields shown in Fig. 3 reveal more detailed spatial strucure than do sea-surface height fields. In this case the models are compared with the Global High

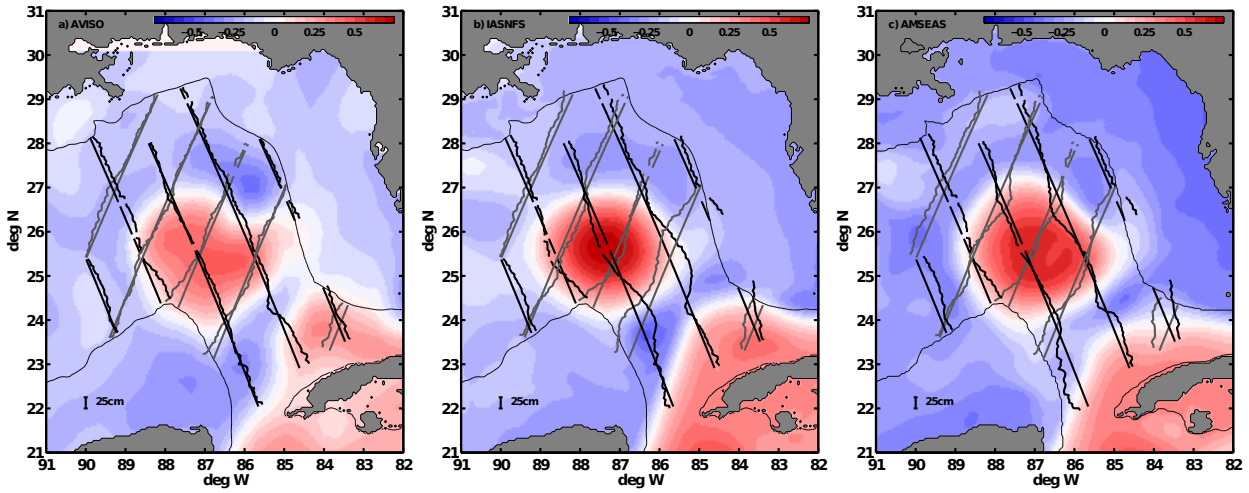

FIG. 4. Sea Surface Height (SSH) on June 10 2010. Three estimates of SSH (anomaly with respect to average sea level west of 90°W) are shown together with along-track data from the JASON-1 and 2 satellite altimeters. The AVISO hindcast (left) is produced in delayed mode from multiple satellite altimeters using objective analysis. The IASNFS (center) and AMSEAS (right) fields are nowcasts valid at 12 UTC on the given date. Observed minus model residuals along the ascending (gray) and descending (black) ground tracks during the 10-day window from 5 June to 15 June 2010 are projected perpendicular to, and northward of, the corresponding ground tracks.

Resolution Sea Surface Temperature analysis product from the Naval Oceanographic Office (NAVO Level 4 K10 GHRSST). The May 26 snapshot (top row) shows that the AMSEAS model represents the folding and stirring of small scale temperature gradients by the cyclonic eddies on the periphery of Eddy Franklin. Surface thermal gradients associated with the LC are reduced in July (middle row) and September (bottom row), but the models and observations show persistent upwelling of cooled water north of the Yucatan Peninsula.

Relative to IASNFS, the AMSEAS model output contains more small-scale structure, much of which appears to agree qualitatively with observations. The following sections explore the quantitative accuracy of AMSEAS in more detail.

C. Class 2: Satellite Altimetry (GOMEX-PPP; May 2010 to December 2010)

A comparison of AMSEAS with along-track altimetry data from the Jason-1 and Jason-2 satellite altimeters was performed to assess the fidelity of the dynamically-important pres-

sure field within the model (altimeter data extracted from the Radar Altimeter Database System; Naeije et al., 2002). Standard corrections were applied for the orbit altitude, dry troposphere, wet troposphere, ionosphere, inverse barometer, solid earth tide, ocean tide, load tide, and sea-state bias. Model output was linearly interpolated to the satellite ground-tracks at the times and locations of the satellite passes, omitting gaps in the satellite data. SSH values were referenced to mean sea level west of 90W, where dynamic height variations are much smaller than in the eastern Gulf, in order to correct for the unknown offset in absolute sea level in the models. Quantitative comparisons were restricted to the region (90°W,22°N) to (82°W,30°N), in water depths greater than 500 m, in order to exclude shelf variability and focus on processes connected with the LC and LCE (Fig. 4). (Note: the Jason data used in this assessment were assimilated by both AMSEAS and IASNFS; however, the data were not assimilated directly as SSH observations. Rather, they were converted to synthetic profiles of temperature and salinity versus depth before assimilation (Rhodes et al., 2002). Thus, the direct comparison of observed and nowcast SSH is a measure of the relationships between SSH and water properties used to infer the profiles, the assimilation methodology overall, and the latency of the near real-time data stream.)

The AVISO multi-satellite objective-analysis, IASNFS, and AMSEAS SSH fields are illustrated for a date in Fig. 4, together with along-track altimeter data. Eddy Franklin is a relatively asymmetric feature flanked by cyclonic eddies to the north and south at 86W. As noted in Fig. 2, IASNFS contains a stronger, more symmetric, rendition of Eddy Franklin than either AVISO or AMSEAS, and the arrangement of cyclonic eddies is different, with the southern eddy prominent. AMSEAS SSH in the eddy core agrees better with AVISO, but the configuration of cyclonic vorticity around the eddy is different from either AVISO or IASNFS. Compared to AVISO, the root-mean-square error of IASNFS and AMSEAS are nearly identical, between 0.11 m and 0.12 m, with AMSEAS slightly better than IASNFS in the period studied.

The ascending (gray) and descending (black) ground tracks of the JASON data within 10-days of the nowcasts illustrate the paucity of routine SSH data relative to the scale of the LC and its associated eddies (Fig. 4). Differences between the models and observations are sometimes greater than 0.25 m over length scales of 100 km, and overlap with the length scale of the cyclonic eddies. Geostrophic currents perpendicular to the ground tracks may be inferred from along-track sea-surface slope (Powell and Leben, 2004). Root-mean square

error of approximately 0.2 ms^{-1} is found for both IASNFS and AMSEAS. Anomaly correlation coefficients of 0.66 vs. 0.71, respectively, indicate a slight improvement in AMSEAS relative to IASNFS.

D. Class 2: COAMPS Winds (COMT; Jun/Jul 2010 & Dec/Jan 2011)

To support the NAVOCEANO AMSEAS transition effort, the COMT included a limited examination of wind forcing used by AMSEAS. COAMPS winds were compared to measurements from moored buoys and NOAA Coastal-Marine Automated Network (C-MAN) stations in the northern Gulf of Mexico using standard bias and absolute error metrics, as well as vector correlations. The data were obtained from the National Data Buoy Center (NDBC; `http://www.ndbc.noaa.gov`). For consistent comparison to COAMPS, the buoy/C-MAN measured wind speeds were converted to a standard 10-m height and 1-min averages. The 10-m adjustment was performed assuming a Charnock roughness length relationship (Charnock, 1955) and logarithmic wind profile. Anemometer heights for most moored buoys and C-MAN stations range from 5 to 14 m, requiring little vertical adjustment. But the C-MAN station near the mouth of the Mississippi River (BURL1) has an anemometer height of 31 m, and three offshore drilling platforms are used with anemometer heights of 122 m. However, while the C-MAN station observations represent near 1-min winds (they are 2-min winds), moored buoys and drilling platforms measure 8-min winds and were converted to 1-min winds with a 9% increase (Powell et al., 1996).

Several vector correlation schemes were tested (Breaker and Gemmill, 1994; Kundu, 1976) but the Hanson et al. (1992) scheme was chosen since it provides parametric coefficients, is invariant under rotation, provides an angle of rotation and a scaling factor, and is analogous to linear regression with a correlation coefficient between -1 and 1 as well as least-square fit coefficients. Experiments with hypothetical scenarios demonstrate that the scale factor can capture a uniform wind bias. For example, if an independent vector magnitude is consistently 50% less than the dependent vector magnitude, the scale factor will be 2.0, and for the same scenario the independent is 200% more, then the scale factor will be 0.5. Likewise, if an independent vector direction is consistently offset by 30°, Hanson's routine gives an angle of rotation of -30°. Additional experiments with different wind-speed magnitudes provided the same values. But for typical wind vector observations versus model winds which exhibit

both speed and direction biases over a long period for a range of wind magnitudes, the vector correlation provides the best measure of fit. As will be shown, this property provides a different validation perspective when analyzing summer (equivalent barotropic) and winter (baroclinic) differences.

Because this study also coincided with DwH, the focus region was the northern Gulf of Mexico and the observation dataset includes five moored NDBC buoys (42003, 42012, 42039, 42040, 42360), three offshore drilling platforms also designated as moored buoys (42362, 42363, 42364), and fifteen C-MAN stations (AMRL1, BURL1, DPIA1, GDXM6, LABL1, LKPL1, LUML1, MBLA1, MHPA1, NWCL1, PCLF1, SHBL1, TAML1, TRBL1, and WYCM6). Two periods were chosen for the analysis: 20 June 2010 to 10 July 2010 during DwH, and 1 December 2010 to 15 January 2011 for a winter comparison. The statistics were generated three ways: seasonal summaries, individual buoy metrics, and daily plots. Seasonal information is the most useful to ocean circulation modelers to isolate long-term errors and will be discussed next. Examples of the individual buoy plots and daily analyses, allowing future assessment for local and specific weather situations, will then be described.

Table III shows metrics for analyses for 12-h and 24-h forecasts of COAMPS winds. Forecasts of 6- and 18-h were also performed but contained no distinguishing results from the 00- to 12-h periods. The metrics include bias, absolute error, average squared vector correlation, average rotation angle, and average scale factor. Bias in these atmospheric evaluations is computed as COAMPS minus buoy. For the vector correlations, COAMPS winds are the independent variable. In the overall context including all 23 buoys, bias and absolute errors for speed are small, wind direction bias is small and absolute errors for wind direction are generally reasonable. However, seasonal and platform-type patterns are noticeable. Even though the overall bias is small (-0.7 ms^{-1} to -0.1 ms^{-1}), COAMPS consistently under predicts wind speed. When the moored (offshore) and C-MAN (coastal) observations are separated, this bias is associated with offshore winds in which 6-8 of the 8 moored buoys are consistently under- predicted for both seasons and for nowcast and forecast periods. The under-prediction of wind speed can be understood as the inability of the 15-km resolution COAMPS grid to adequately resolve pressure gradients from which the surface winds derive. In contrast, COAMPS coastal winds, with expected lower wind speeds than offshore, have little bias and smaller absolute error. However, wind direction absolute

TABLE III. Validation (with observation numbers) of COAMPS winds versus all buoys, moored buoys, and C-MAN stations for COAMPS winds analyses, 12-h forecast fields, and 24-h forecasts fields in the northern Gulf of Mexico during a summer and winter period. The summer dataset is in the period 0000 UTC 20 June 2010 to 0000 UTC 10 July 2010. The winter dataset is in the period 0000 01 December 2010 to 0000 15 January 2011. Bias is computed as COAMPS minus buoy observations.

	Analysis		12-h Forecast		24-h Forecast	
Statistical parameters	Winter	Summer	Winter	Summer	Winter	Summer
			All buoys ($n = 23$)			
Speed bias (ms^{-1})	-0.7	-0.1	-0.4	-0.7	-0.8	-0.2
Speed absolute error (ms^{-1})	1.8	1.4	1.7	1.6	1.8	1.4
Direction bias (°)	-2.6	2.4	2.6	-3.6	-1.8	4.4
Direction absolute error (°)	26.0	31.6	21.8	37.4	26.2	33.1
Vector correlation squared (%)	75.8	49.4	80.0	51.7	75.3	49.0
Scaling factor	1.00	0.82	0.96	0.89	1.01	0.83
Rotation angle (°)	-0.7	-12.1	6.7	-4.2	0.5	-5.0
			Moored buoys ($n = 8$)			
Speed bias (ms^{-1})	-1.4	-0.5	-1.2	-1.6	-1.3	-0.7
Speed absolute error (ms^{-1})	2.1	1.4	2.0	2.1	2.0	1.6
Direction bias (°)	1.2	2.8	1.4	-8.6	1.5	6.7
Direction absolute error (°)	21.9	29.9	19.5	28.1	22.8	32.2
Vector correlation squared (%)	78.4	47.1	85.4	65.1	80.2	48.2
Scaling factor	1.08	0.83	1.09	1.2	1.08	0.88
Rotation angle (°)	0.3	-12.6	3.9	-3.8	2.1	-1.1
			C-MAN stations ($n = 15$)			
Speed bias (ms^{-1})	-0.3	0.2	0.0	-0.1	-0.5	0.0
Speed absolute error (ms^{-1})	1.7	1.4	1.6	1.4	1.7	1.4
Direction bias (°)	-4.6	2.2	3.2	-0.9	-3.6	3.1
Direction absolute error (°)	28.2	32.5	23.0	42.4	28.0	33.5
Vector correlation squared (%)	74.5	50.6	77.0	44.4	72.6	49.4
Scaling factor	0.96	0.81	0.90	0.72	0.98	0.80
Rotation angle (°)	-1.1	-11.8	8.2	-4.4	-0.4	-7.2

errors are larger along the coast. This result may again be due to the COAMPS inability at 15-km resolution to adequately resolve coastal topography. Summer wind direction absolute direction errors are also larger than winter absolute errors (31-37°compared to 22-26°). An examination of daily wind nowcasts shows the larger absolute error is due to the weak pressure gradients that favor variable winds during the summer. An interesting feature is that moored buoys have less bias than C-MAN in the winter while the opposite is true

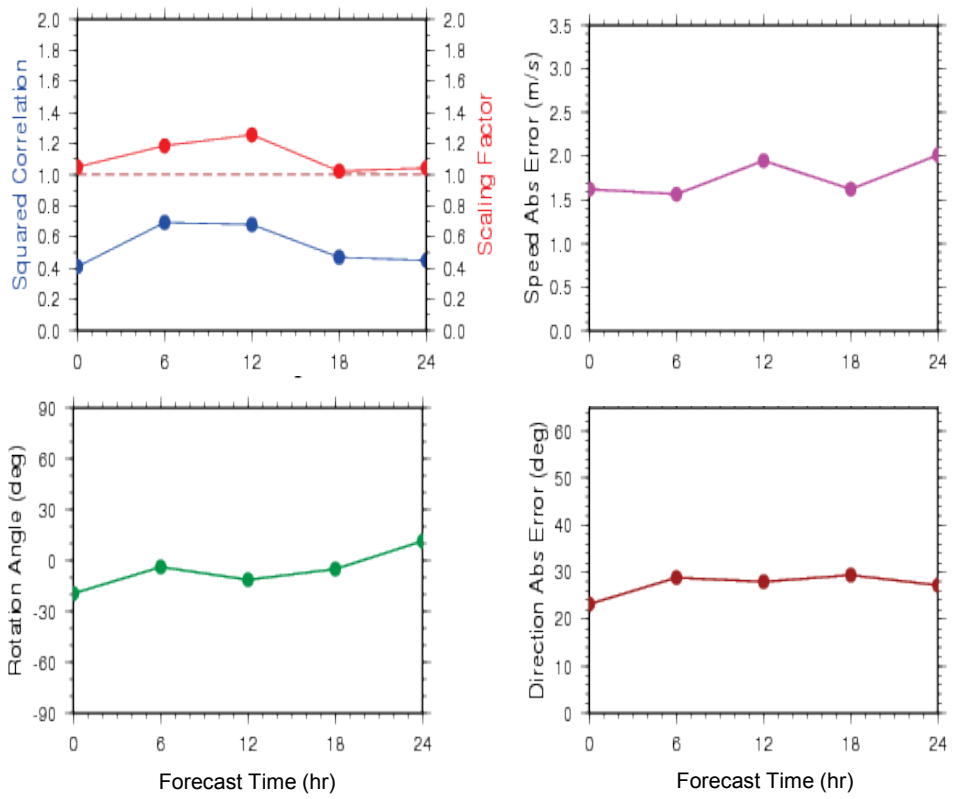

FIG. 5. Example of vector correlation squared, scaling factor (upper left) and rotation angle (lower left) based on methodology of Hanson et al. (1992) for COAMPS wind initialization and forecast interpolated to buoy 42003 during the period 0000 UTC 20 June 2010 to 0000 UTC 10 July 2010. The dashed line corresponds to the maximum possible squared correlation of 1. Time series of absolute errors for wind speed (upper right) and speed (lower right) for COAMPS at buoy 42003.

in the summer. Within the 24-h forecast period, no increasing prediction-error trends are apparent.

The vector correlation metrics provide an alternative viewpoint relative to standard statistics. Because speed and direction both contribute to errors with alternating contributions by platform and season, no consistency is evident with the scaling factor or rotation angle. However, note that the variance explained is 75-80% for winter versus 49-52% for summer, showing an obvious sensitivity to wind direction errors since speed errors are less in the equivalent barotropic conditions. A slight majority of COAMPS forecasts at moored buoys are associated with higher vector correlations than at C-MAN locations.

To facilitate coastal ocean modelers in examining the results further, platform statistics

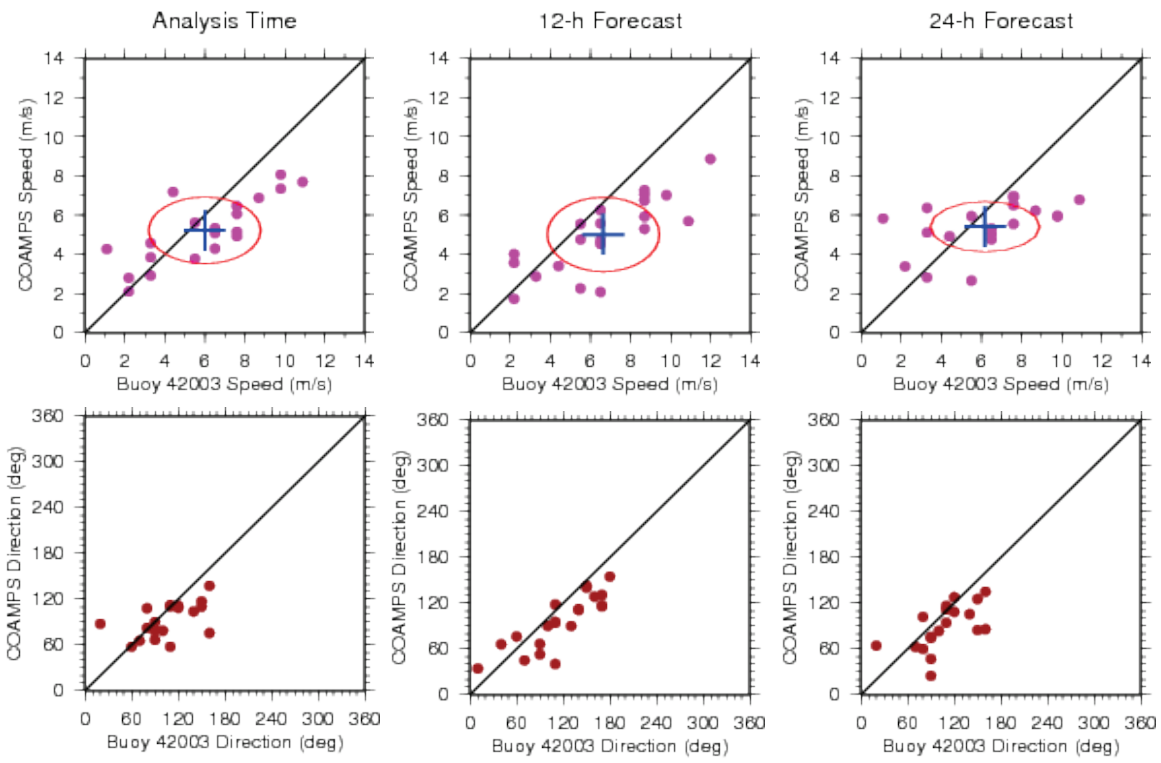

FIG. 6. Example of scatterplots for COAMPS wind initialization as well as 12- and 24-hr forecasts for speed (upper) and direction (lower) interpolated to buoy 42003 during the same period as Fig. 5. Wind speed plots also include ovals representing one standard deviation of each dataset; circular plots indicate both the model and buoys have the same data ranges and elliptic plots indicate one dataset has less range than the other.

and daily plots were generated for the COAMPS validation and archived at the SURA Web site at `http://testbed.sura.org/node/403`. Figs. 5 and 6 present examples of two types of station plots for Buoy 42003 in the central Gulf. Fig. 5 provides a forecast time series of vector correlation information and absolute errors. Fig. 6 provides scatterplots of COAMPS versus buoy 42003. In general, the wind direction has a reasonable positive linear correlation. However, while the wind speed scatter plots have a positive linear trend, a negative speed bias is evident. Furthermore, this bias increases with wind speed.

Ovals representing one standard deviation of COAMPS and buoy wind speeds are also displayed in Fig. 6, centered about their respective mean values. Circular plots indicate both the model and buoys have the same variability ranges, and elliptical plots indicate one dataset has less range than the other. The one-one lines reveal whether the model is generally

above or below (or well aligned with) the state of the natural system as represented by the buoy measurements. Cases where a consistent offset between the model and observations exists can be easily identified and quantified with these scatter plots and thus used to suggest model shortcomings, faulty instrumentation, or perhaps a standardization problem with the observations. Fig. 6 shows generally left-to-right elliptical patterns indicating that the model under-represents the observed wind speed ranges, and that the ellipse center is to the right of the straight line illustrating the COAMPS negative speed bias.

Daily plots were also performed to assess COAMPS initialization fields. A typical summertime example for surface wind direction and wind speed is displayed in Figs. 7 and 8 for 0000 UTC 22 June 2010. The top left displays the buoy observations and the top right displays the model field. The bottom two plots display the model errors for each buoy, presented in two different ways. On the bottom left error is presented as raw values (model minus observation). On the bottom right, error is presented as percentage within a defined error-tolerance limit, defined in keeping with Navy practice as 2 ms^{-1} speed and 40° direction. Shading the within-tolerance values shaded gray allows a researcher or an operational forecaster to focus on the more significant areas of discrepancy between model and data. Larger wind direction errors on the Mississippi and Alabama coasts are evident in this example, consistent with the COAMPS coastal topography resolution issue discussed earlier. Such plots are archived at the SURA testbed Website to assist with finding subtler model wind initialization and forecast errors in future studies.

In summary, the examination of winter and summer seasons presents an interesting contrast in validation metrics of vector quantities. Even though wind biases and absolute errors tend to be less in the summer, the vector correlation methodology shows that COAMPS provides more relative accuracy in the winter, and the individual plots clarify regional biases and whether COAMPS is capturing all the wind variability. It also illustrates why researchers need to include the examination of summary statistics, platform-based statistics, seasonal trends, and case studies for proper model validation.

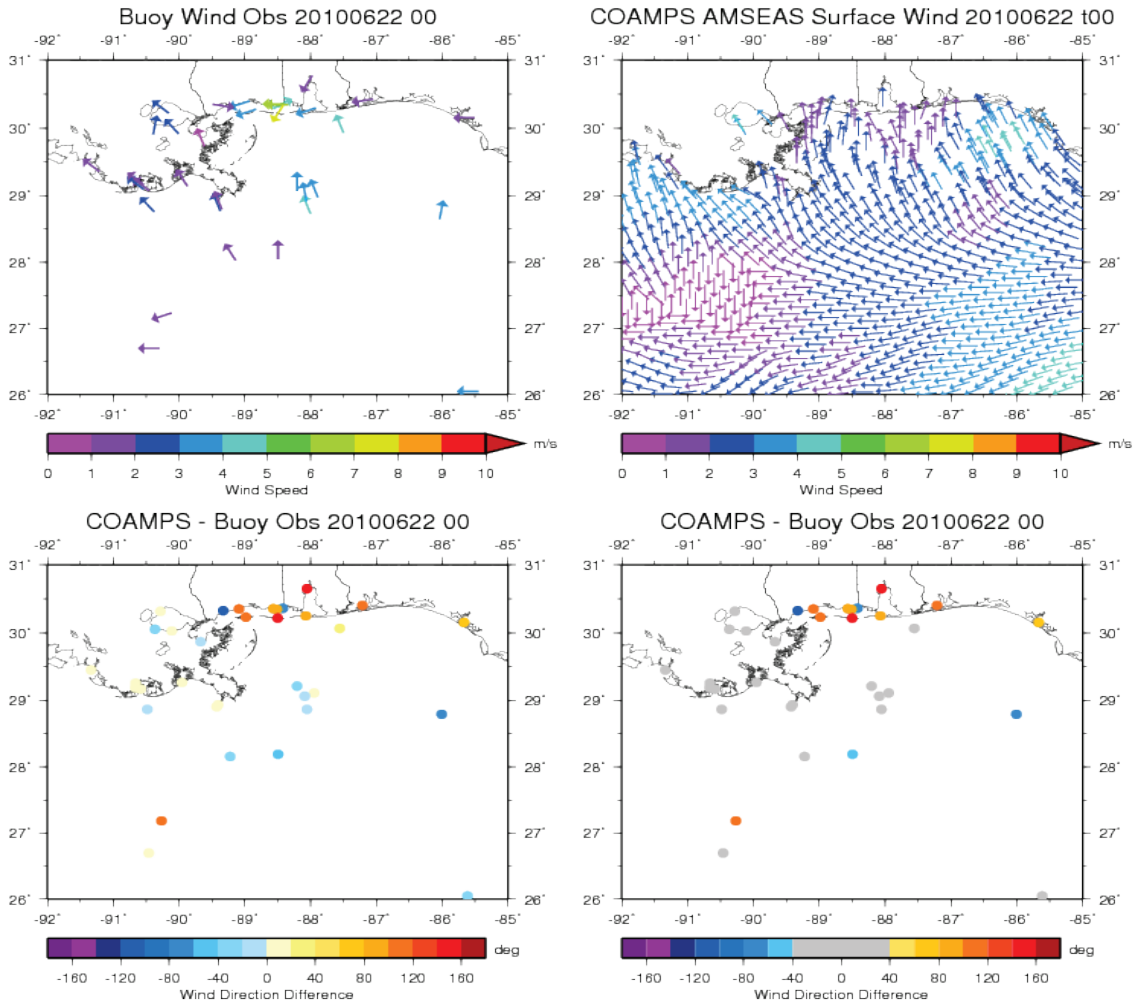

FIG. 7. Example of a plot for daily analyses for COAMPS initialization for 0000 UTC 22 June 2010. Observed buoy vectors, color shaded by wind speed (upper left). COAMPS wind vectors color shaded by wind speed (upper right). Color coded wind direction differences at the buoy locations (lower left). As on lower left but with buoy station direction difference within a 40° wind direction tolerance level shaded grey (lower right).

E. Class 2: Surface Temperature and Currents (COMT; June 2010 to October 2011)

For this component of the AMSEAS model assessment, time series of surface temperature and current velocity were obtained at a number of sites within the AMSEAS domain that featured moored instrumentation. Seventeen NDBC sites were identified within the AMSEAS domain (Table IV, Fig. 9) where the observed time series had data returns $> 90\%$ over

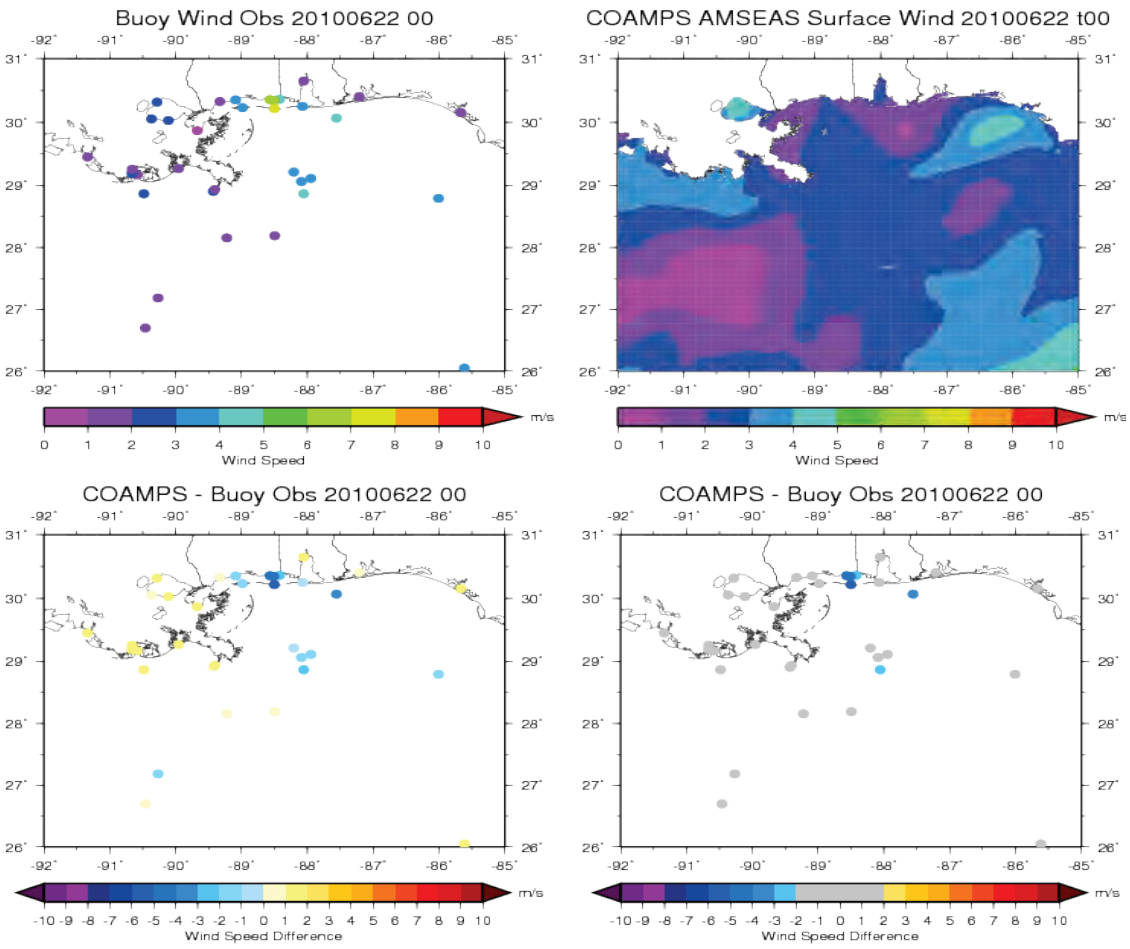

FIG. 8. Example of a plot for daily analyses for COAMPS initialization for 0000 UTC 22 June 2010. Observed buoy wind speeds, shaded by wind speed (upper left). Contours of COAMPS wind speed (upper right). Wind speed difference (lower left). Same as lower left with stations within the 2 ms^{-1} tolerance level shaded grey (lower right).

the assessment period (June 2010 – October 2011). Two of these sites (42001 and 42021) had a persistent temperature bias between the model and observations, and are considered to have suspect thermistor calibrations. These sites were retained in the analysis since they represent a scenario where the model can be leveraged to reveal data-quality anomalies.

Sites that feature temperature time series are by far the most numerous, with thirteen time series from 1-m depth and two time series from 2-m depth. There are three Texas Automated Buoy System stations (TABS, `http://tabs.gerg.tamu.edu`) with current-meter time series. The 1-m temperature data consistently exhibit more complete returns and

TABLE IV. Moored sites accessed for assessment of AMSEAS model surface temperature and velocity. For each mooring, its NDBC ID and location are provided. The type of data obtained is indicated in the measurement column (T=Temperature, U, V = zonal, meridional velocity). Unless otherwise noted, the instrument depth is 1 m. The station caretaker and the sites water depth are provided.

ID	Measurement	Caretaker	Water Depth [m]	Longitude	Latitude
41009	T	NDBC	44	-80.17	28.52
41010	T	NDBC	873	-78.47	28.91
41012	T	NDBC	37	-80.53	30.04
42001	T	NDBC	3365	-89.66	25.89
42003	T	NDBC	3283	-85.61	26.04
42021	T	COMPS	-	-83.31	28.31
42035	T	NDBC	14	-94.41	29.23
42036	T	NDBC	307	-86.01	28.79
42039	T	NDBC	307	-86.01	28.79
42040	T	NDBC	165	-88.21	29.21
42044	T (2 m)	TABS (Station J)	21	-97.05	26.19
42045	U, V (2 m)	TABS (Station K)	62	-96.50	26.22
42049	U, V	TABS (Station W)	22	-96.01	28.35
42050	T, U, V (2 m)	TABS (Station F)	24	-94.24	28.84
42055	T	NDBC	3566	-94.00	22.20
42056	T	NDBC	4684	-84.86	19.80
42099	T	Scripps	94	-84.25	27.34

higher quality. Not surprisingly, the NDBC-maintained sites returned the most complete data.

AMSEAS forecasts were generated by NAVOCEANO beginning 25 May 2010 and are available from OceanNOMADS, as noted in Section I C, using Open-source Project Data Access Protocol (OPenDAP; `http://www.opendap.org/`). This access capability mitigates the need to use local storage to aggregate daily AMSEAS output (5.7 GB day^{-1} when compressed), a particular benefit when only narrowly prescribed spatial extraction from the model forecasts is required.

Monthly comparisons of AMSEAS/NDBC time series were generated as the NDBC quality-controlled data became available and then combined to form full comparisons over the 17-month assessment time frame. Examples of the full-period (June 2010 – October 2011) comparisons of temperature (site 42039) and current speed (site 42045) for forecast day-1 are shown (Fig. 10). (These 17-month comparisons as well as the monthly scatter plots

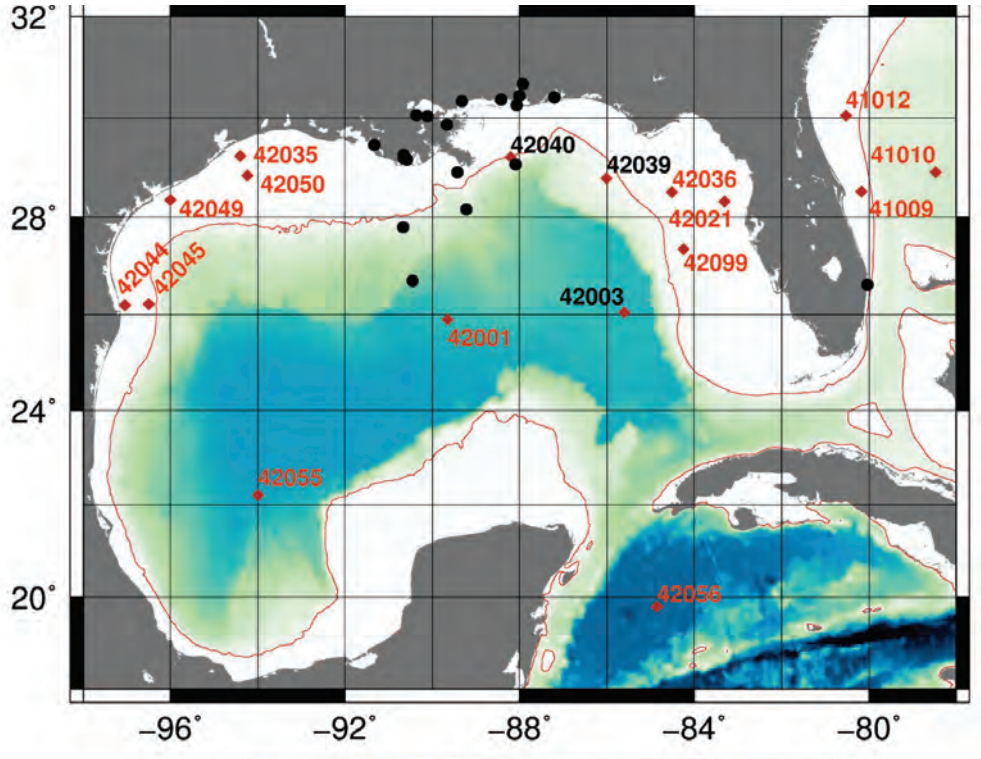

FIG. 9. Location of mooring sites from which time series data were acquired from the NDBC website. Parameter(s) obtained from a given location are indicated in Table IV. All parameters are not available from all locations. Bottom topography obtained from the NOAA's National Geophysical Data Center. Black dots represent station and buoy sites used in COAMPS evaluations. Black station numbers represent buoy overlap between COAMPS and oceanographic evaluations.

described in the next paragraph are available at `http://testbed.sura.org/node/580` for all stations.)

Scatter plots with a one-one line were generated to illustrate how well individual points match up for each month. The 3-h resolution of the stored AMSEAS output nominally results in 240 model – data match ups per month. For the locations shown in Fig. 10, the model – data scatter plot is shown from March 2011 (Fig. 11) with a red ellipse mean ± one standard deviation and a one-one line on each plot (see Section II D for their interpretation). Also included as part of these reported metrics were the percentages of points that lie above, within or below a prescribed tolerance. The tolerance band is indicated in Fig. 11 by the green lines. The tolerance limit applied for temperature (±0.5°C) coincides with that used in the NAVOCEANO analysis reported in Section II H. The tolerance limit applied for surface

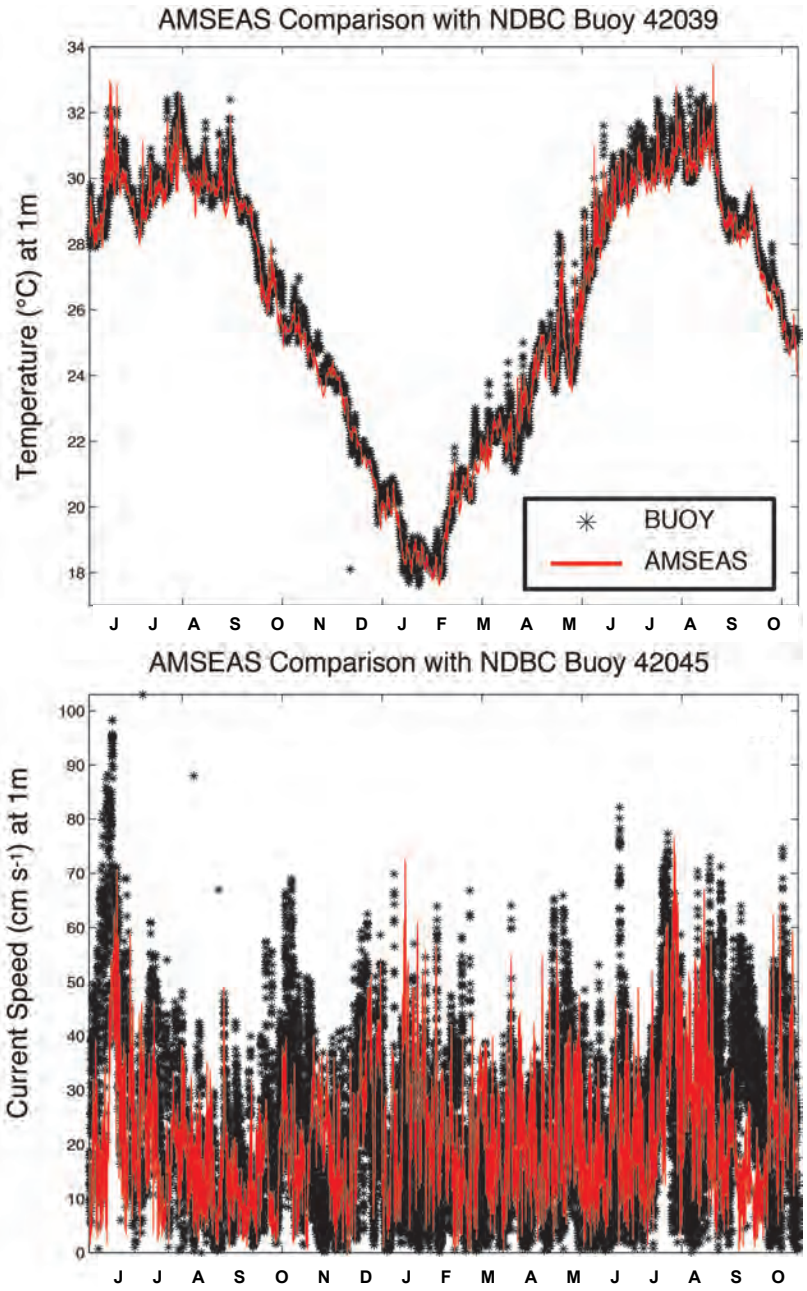

FIG. 10. Time series for the full AMSEAS-NDBC comparison period (June 2010 – October 2011) from two sites. The data shown are 1-m temperature at NDBC buoy site 42039 (upper) and 1-m current speed at TABS site 42045 (lower) for forecast day-1.

current speed (± 0.2 ms^{-1}) was chosen based on the values for (model – data) standard deviation (one of the scatter plot metrics), which was typically ≤ 0.18 ms^{-1} over all forecast days of the 17-month assessment.

Scatter plots for each month and forecast day have been generated for every good data

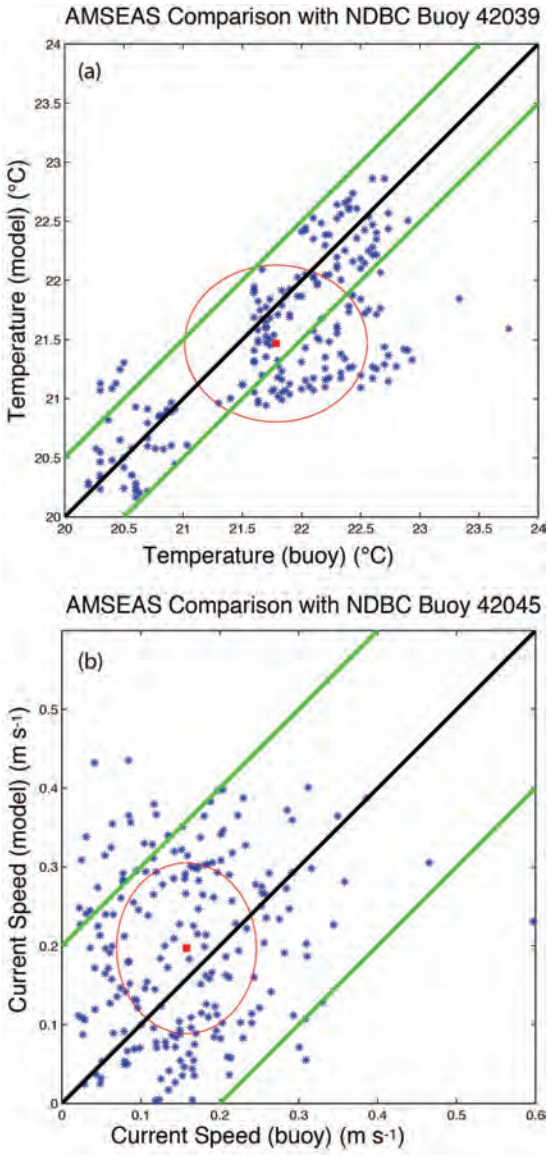

FIG. 11. Scatter plots of model vs. data (AMSEAS vs. Mooring) for the March 2011 portion of the data shown in Fig. 10 at NDBC buoy 42039 for 1-m temperature (upper) and at NDBC buoy 42045 for surface current speed (lower). The red ellipse represents one standard deviation around the mean of the model (y-axis) and observations (x-axis). The tolerance limits (green lines) determine the percentage of values above/below an acceptable error, are 0.5°C and 0.20 ms^{-1} for temperature and current speed, respectively. For these examples the mean and standard deviation of the model – data differences are -0.32 ± 0.53°C for temperature and 0.039 ± 0.13 ms^{-1} for current speed. The correlation coefficient, coefficient of determination and RMSD are 0.73, 0.54, 0.62 for temperature and 0.16, 0.026, 0.13 for current speed.

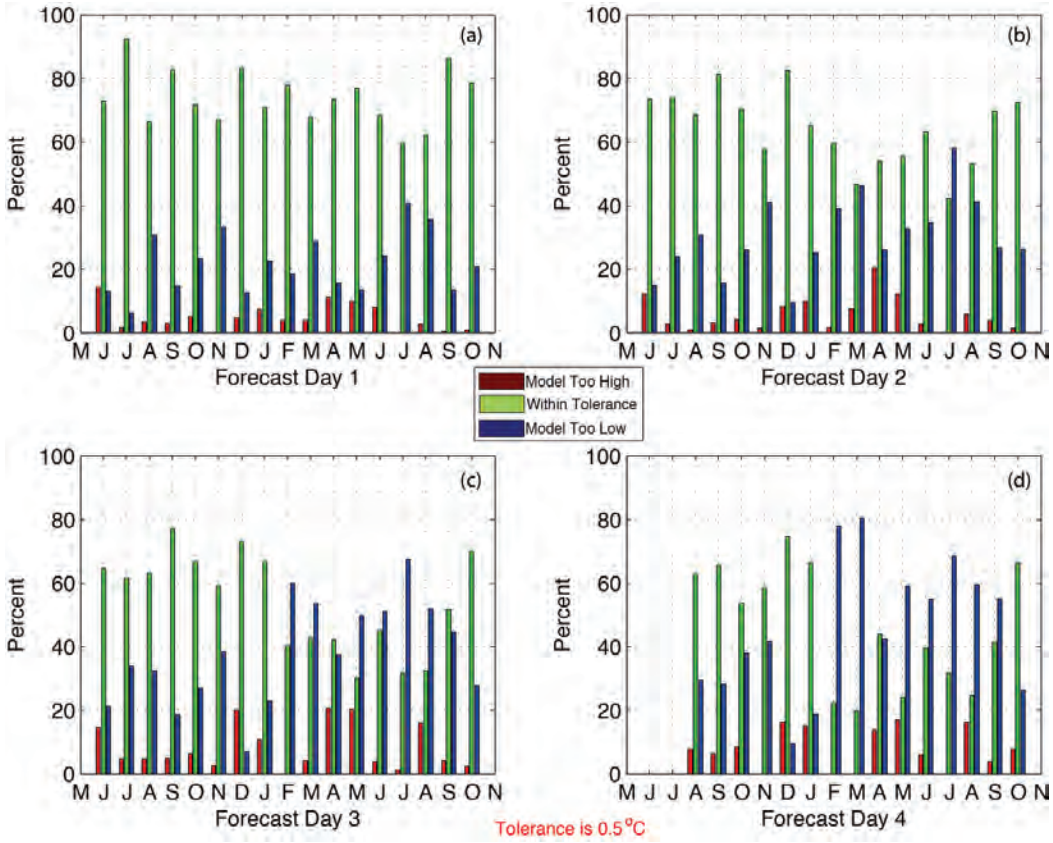

FIG. 12. Bar plot time series of 1-m temperature at NDBC buoy site 42039 showing the percentage of AMSEAS forecasts that are in the above/good/below tolerance bins for each month of all four forecast days over the full AMSEAS-NDBC buoy comparison period (June 2010 – October 2011). For forecast day-4, there is no June or July 2010 result since 4-day AMSEAS forecasts were not implemented until mid-July 2010. Green bars are within tolerance. Red bars are too high (model too warm); blue bars are too low (model too cold).

location listed in Table IV. To visualize the percentage of low, high and good points, as determined by applying the specified tolerances, time series bar graphs have been generated. For NDBC site 42039, which provided the temperature time series in Fig. 10a, the bar graph time series of low/good/high percentage are shown for all four forecast days (Fig. 12). This mooring site is located on the Florida Shelf southeast of Pensacola. These results show that for the first forecast day (Fig. 12a), the AMSEAS model is within tolerance at least 60% of the time throughout the assessment period and that the model tends to under-predict temperature when it lies outside of the tolerance range. For a given forecast day the bar

graph time series reveal seasonality in AMSEAS skill, which for this example reveals that August 2010 and July-August 2011 exhibit the three lowest within-tolerance percentages (Fig. 12a).

Trends were evident in the within-tolerance percentages over the course of the AMSEAS forecasts (nominally these are 4-day forecasts, except for June and July 2010). The complete set of bar graph time series for NDBC station 42039 is shown in Fig. 12. For this site on the Florida shelf, these metrics suggest that, as the model's skill degrades, it trends toward under-predicting surface temperature. By the third and fourth forecast day there is a transition from the model being primarily within tolerance to being primarily too cool in February. This under-prediction is not relieved until the next September in the day-3 forecast. In the day-4 forecast, the model under-prediction persists through September with the model only reestablishing itself within tolerance results ($> 50\%$ of the time) in October 2011.

For a broader perspective of seasonality in model skill and as a means of gaining insight into the spatial context, time series of the above/good/below tolerance percentages for all surface ocean sites in Table IV are aggregated in Fig. 13-15. Fig. 13 and 14, in particular, feature the above/good/below tolerance percentages for the surface temperature sites for forecast days 1 and 4, respectively. Fig. 13 provides an interesting comparison to the Section II H analysis (below) that indicates a warm bias for SST from October 2010 to March 2011. This trend only manifests itself for the fixed-buoy stations in the northern Gulf near the Texas/ Louisiana border (42035, 42050) and off western Florida (42021, 42036). This result suggests that the present observation network may not be adequate to represent the larger Gulf temperature trends. Alternatively, these few stations may be dominating the overall SST statistics of Section II H. This result recommends a future investigation into the spatial characteristics of the temperature bias for the observations discussed in Section II H.

Fig. 13 and 14 together reveal the degradation in model skill for surface temperature over the forecast run. For forecast day-1 (Fig. 13), there are five sites (41010, 41012, 42039, 42056 and 42099) for which the within-tolerance percentage (green line) remains the highest throughout the assessment period. These sites are widely distributed around the AMSEAS domain; two are on the eastern Florida shelf (41010, 42012), two are on the shelf in the eastern Gulf of Mexico (42039, 42099), and one is at a deep water site in the Caribbean (42056). For forecast day-4 (Fig. 14), only one of these sites (42056) retains this distinction

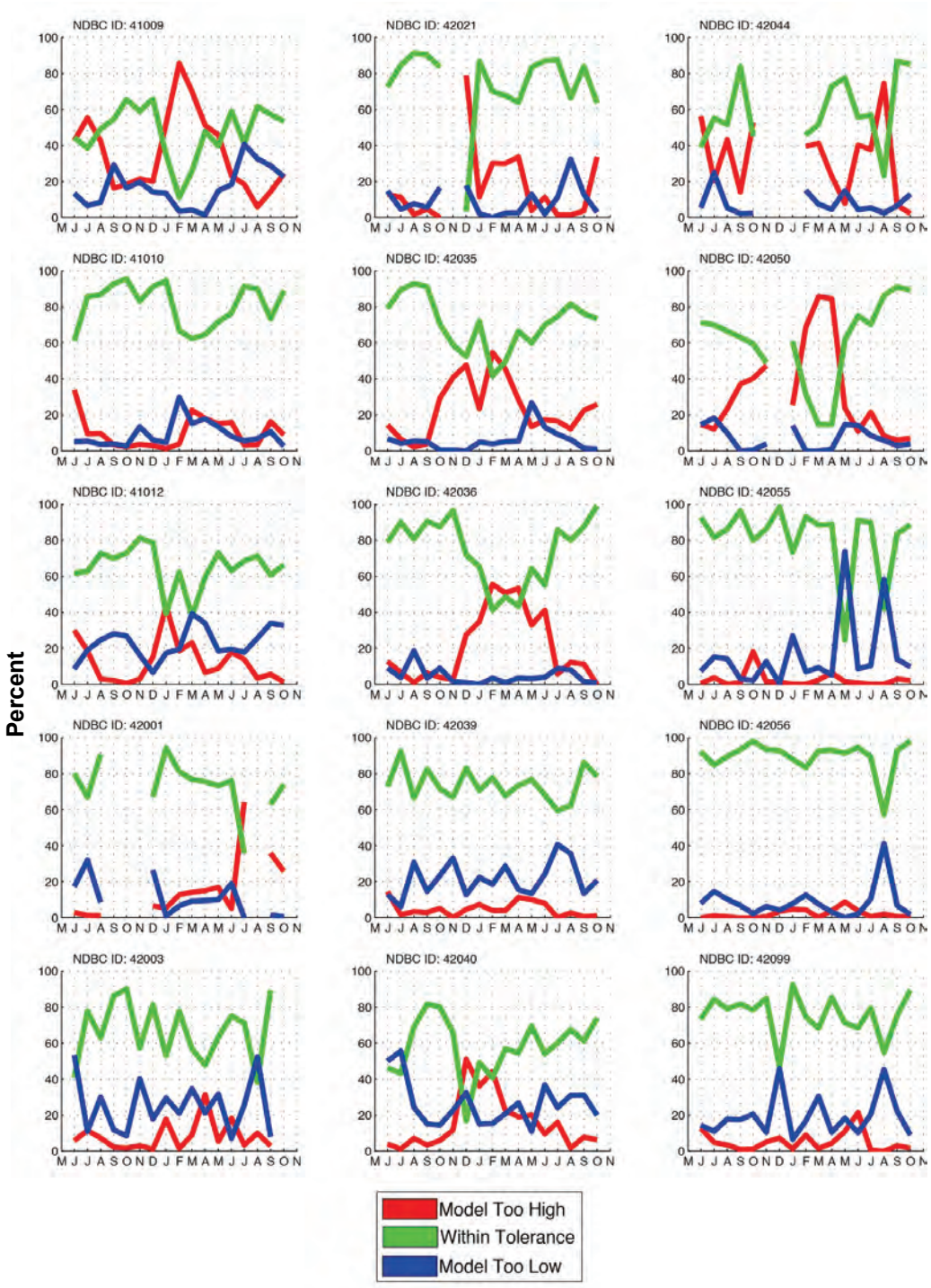

FIG. 13. Time series showing the percentage of AMSEAS 1-m forecasts that are in the above/good/below tolerance bins for each month of forecast day-1 over the full AMSEAS-NDBC buoy comparison period (June 2010 – October 2011). Results are shown for all buoy sites listed in Table IV. Temperature tolerance limits set at ±0.5°C.

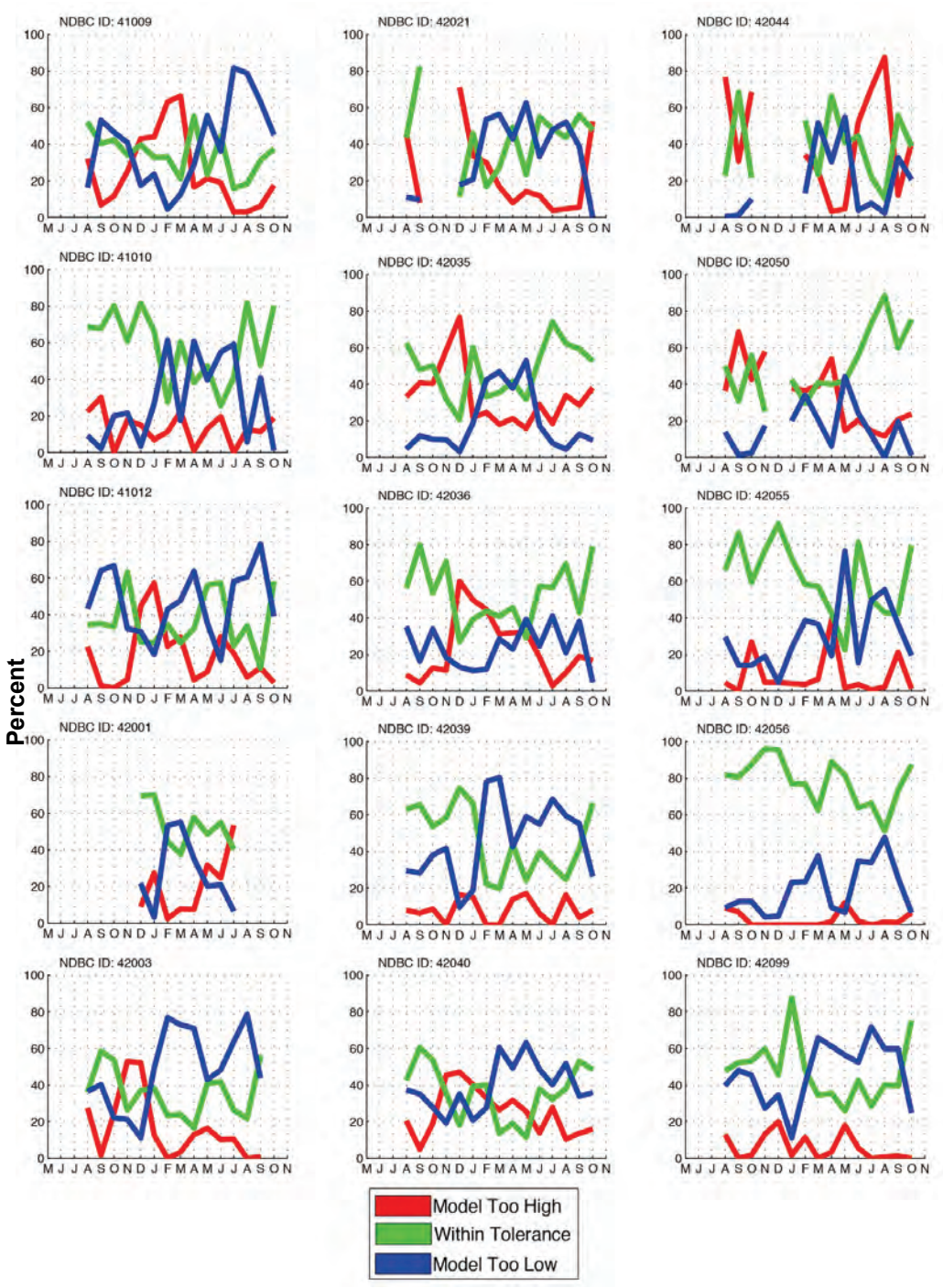

FIG. 14. Time series showing the percentage of AMSEAS 1-m temperature forecasts that are in the above/good/below tolerance bins for each month of forecast day-4 over the full AMSEAS/buoy comparison period (August 2010 – October 2011). No comparisons are possible in June and July 2010 since the AMSEAS forecasts were not extended to 4 days until mid-July 2010. Results are shown for all buoy sites listed in Table IV. Temperature tolerance limits set at $\pm 0.5^{\circ}$C.

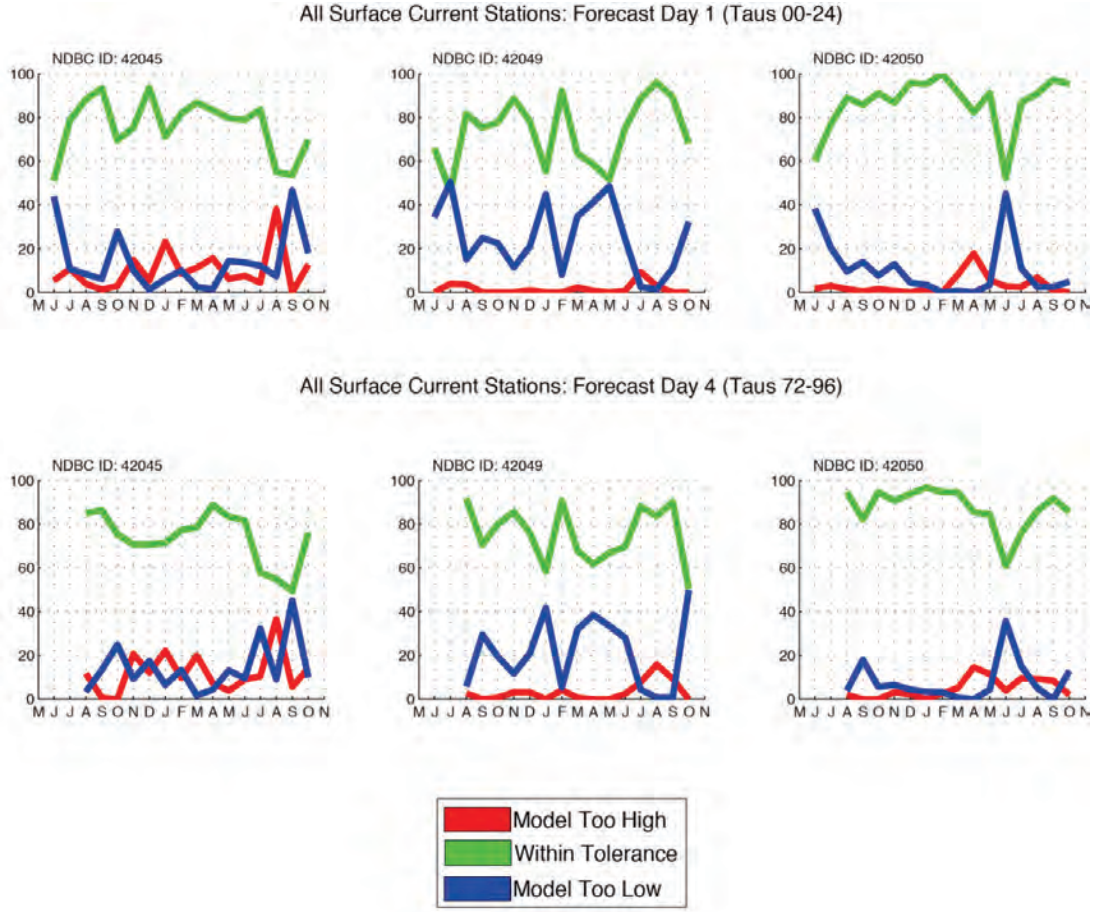

FIG. 15. Time series showing the percentage of AMSEAS surface current forecasts that are in the above/good/below tolerance bins for each month of forecast day-1 (above) and day-4 (below) over the full AMSEAS/ buoy comparison period (August 2010 – October 2011). No comparisons are possible in June and July 2010 since the AMSEAS forecasts were not extended to 4 days until mid-July 2010. Results are shown for all of the surface current measurement sites listed in Table IV. Current speed tolerance limits set at ± 0.2 ms^{-1}.

of maintaining the within tolerance percentages as highest.

The cool bias at station 42039 for forecast day-4, noted from Fig. 12, is clearly represented in Fig. 14. Two other stations (42055 and 42099) exhibit a cool bias only when the model temperature falls outside the 0.5°C tolerance (Fig. 14). Moreover, station 42056 can be seen to have this tendency as well, although as noted its within-tolerance percentage is always dominant. Thus, this tendency for a cool bias for the day-4 forecast is main-

tained at a diversity of locations in the AMSEAS domain, ranging over the eastern (42039, 42099) and western (42055) continental shelf of the Gulf and the Caribbean deep water site (42056). At the other stations, both cool and warm biases manifest for forecast day-4. None of the stations exhibit a consistent warm bias. Moreover when comparing between the forecast day-1 and forecast day-4 time series, a common attribute is that the percentages for temperature under-prediction commonly increase, whereas the percentages for temperature over-prediction exhibit no consistent adjustment.

An interesting feature manifests in August 2011 at a number of sites, most prominently in the forecast day-1 results and to lesser degree for forecast day-4. This feature consists of a pronounced decrease for within-tolerance percentage (green line) that is mirrored in the cool-bias percentage at five sites (42003, 42021, 42055, 42056 and 42099) and in the warm-bias percentage at one site (42044) (Fig. 13). The two most prominent stations from the cool-bias group are at the deep water sites (42055 and 42056). The passage of TS Harvey from the western Caribbean and over the Bay of Campeche (19-22 August 2011) was considered as a possible link; however, this possibility was ruled out since the storm passed more than a week prior to the model's cool-bias shift at these two sites (not shown). The range in the timing of the cool- and warm-bias occurrences at the other sites further suggests no simple spatial linkage, or remote forcing connection.

Fig. 15 shows the tolerance time series for forecast days-1 and -4 for current speed in the western Gulf obtained from three TABS moorings (42045, 420049 and 42050). At station 42045, the model current speed is commonly within tolerance at least 50% of the time. There is no clear tendency when it deviates from the tolerance bounds with shifts toward both low and high biases occurring. Based on unpublished assessments performed by NAVOCEANO, the guidance provided to the operational user community is that currents are commonly under-forecast by 10-20%. It is likely that the tendency for the model to exhibit overly energetic currents is related to the fact that the 42045 mooring site is situated in a relatively dynamic confluence region that is subject to notable seasonal current reversals, making it a challenging location with respect to model fidelity (Zavala-Hidalgo et al., 2003). At the other two sites (42049 and 42050), reductions in model skill are predominantly associated with low biases, with overly energetic forecasts (red line) almost always below 10% (Fig. 15).

To summarize, this analysis based on long-duration, gap-free, quality-controlled moored time series data identified seasonal patterns in model skill, and helps to characterize how

model skill evolves over the course of the 96-h forecast. There is a clear degradation over the forecast period in the model's skill at predicting temperature at all sites, which is typically associated with more pronounced cool bias by forecast day-4. In contrast, the current-speed forecast skill at the three TABS locations from the western shelf of the Gulf of Mexico does not reveal a clear degradation. In terms of contrasting the skill that is revealed by the surface temperature and surface current tolerance time series (Figs. 13-15), two key aspects should be considered. The first is that the AMSEAS model assimilates SST data but assimilation of surface current measurements has not been implemented; the second is that the percentages revealed in the tolerance time series are quite sensitive to the applied tolerance threshold. Trends over the course of the forecast runs suggest that the dynamical response to momentum forcing seems well captured by AMSEAS, whereas the COAMPS surface heat fluxes or how they are applied as surface boundary conditions warrant future investigation.

F. Class 3: Florida Current Transport (GOMEX-PPP; May 2010 to December 2010)

The LC is part of a larger current system along the western boundary of the North Atlantic which includes the North Brazil Current, the Antilles Current, the Yucatan Current, the Florida Current (FC), and the Gulf Stream. The transport of the FC has been monitored since 1982 in the Straits of Florida at 27°N (Larsen and Sanford, 1985; Shoosmith et al., 2005). The fidelity of the models for the FC is significant because of evidence which suggests a possible upstream trigger mechanism for the formation of LCE in the Gulf (Sturges et al., 2010), which may be related to the upstream (equatorward in the Straits of Florida) phase propagation of sea-level anomaly in coastal tide gauge stations and in the Navy's real-time global NCOM model noted by Mooers et al. (2005). The depth-integrated meridional transport from the IASNFS and AMSEAS models is here compared with the observed FC transport at 27°N (data from `http://www.aoml.noaa.gov/phod/floridacurrent/index.php`).

The meridional transport in the models are similar to each other (Fig. 16), with both carrying about 5 Sv less than the mean transport of 30.5 Sv estimated from observations. Differences in high-frequency variability between the models is apparent, perhaps due to

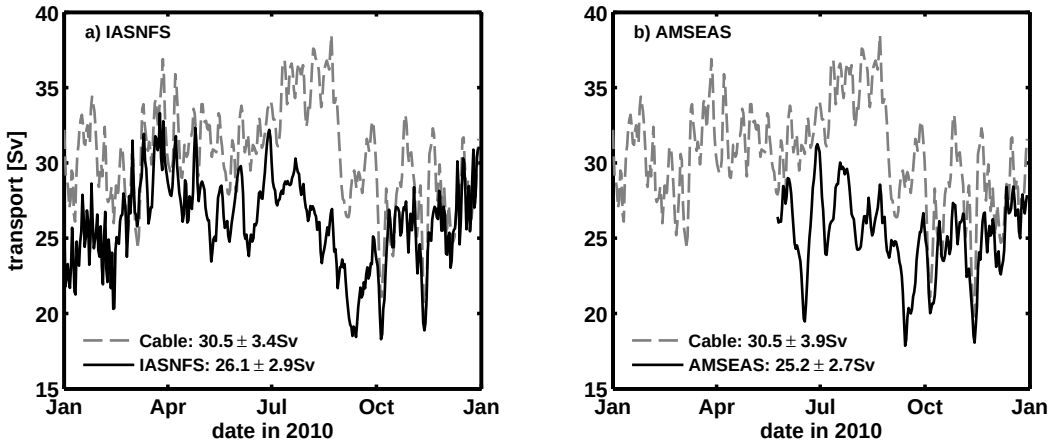

FIG. 16. 2010 Florida Current Transport: Observed (cable) and nowcast FC transport for IASNFS (left) and AMSEAS (right) indicates that the model transports are low by 4 to 5 Sv on average, but significantly larger error occurs during the June-September 2010 period. IASNFS and AMSEAS transports differ due to differences in atmospheric forcing, data assimilation, and other factors; although, the average transport and variability is similar in both models.

the differences in atmospheric forcing. The squared coherence between the observed and AMSEAS transport is approximately 0.6 for periods between 5 and 30 days, with essentially zero coherence at higher frequencies. Thus, the model captures a significant portion of LC dynamics as expressed in the FC transport. The statistics of the modeled and observed transport were not stationary in 2010 and large differences occurred during the June-September time period when Eddy Franklin was formed, suggesting a dynamical connection worthy of further study (Sturges et al., 2010).

G. Class 3: Lagrangian Trajectories for Oil Spill Modeling (BP; 20 June 2010 to 10 July 2011)

The DwH event provided an opportunity to study the effectiveness of using AMSEAS current-velocity data for oil spill dispersal modeling with a Lagrangian particle tracker. The role of synoptic weather feature interactions with ocean currents in transporting the oil spill could also be examined. Cyclones are known to significantly transport water pollutants with either beneficial or deleterious results. A mid-latitude atmospheric cyclone expanded the Exxon Valdez oil spill over a large region, while, in contrast, Hurricane Henri (1979), in

combination with a non-tropical low, cleansed the oil-polluted south Texas beaches (Gundlach et al., 1981). The previous sections provided validation metrics showing the ocean and atmospheric model fields provide reasonable skill to study the oil spill model trajectories.

The late June to early July 2010 timeline was identified as a period of interest since oil briefly impacted the Rigolets and western Mississippi coast, which represented the innermost penetration of oil pollution east of the Mississippi River Delta. Shoreline Cleanup and Assessment Technique (SCAT) data, assembled by NOAA and other governmental agency shoreline inspection teams are available at `http://gomex.erma.noaa.gov/erma.html`, provide a synthesis of this period. Fig. 17 compares shoreline oil pollution time periods for Lake Borgne/Rigolets versus its eastern marsh bordering Chandeleur Sound (known as the Biloxi Marsh). Also shown are the interior bays and the beaches west of the river, which experienced shoreline oiling during much of the DwH period. Areas west of the river experienced tarballs and light-to-heavy oiling throughout the period from Mid-May through September, while except for a brief 1-2 days in May, the eastern Biloxi Marsh was spared oil incursions until late June which then persisted through September. In contrast, Lake Borgne/Rigolets experienced oil in a brief late June to mid-July period. The data should not be taken literally, as the Louisiana SCAT surveys were not performed daily in all regions, but the trends can be noted. We also examined Mississippi Department of Environmental Quality reports (available at the oil spill links on `http://www.deq.state.ms.us`) and the Louisiana Bucket Brigade data (available at `http://labucketbrigade.org/`), which show a similar trend of oil not reaching the western Mississippi Sound until early July (not shown). An important component to understanding the oil transport during DwH is to distinguish the influences behind this apex moment. A Lagrangian particle tracker with random walk diffusion was implemented to simulate the oil spill from 20 June to 10 July 2010 (Hunter et al., 1993; Dimou and Adams, 1993). Input consisted of latitude and longitude parcel positions in the oil-contaminated area, wind, current, and an array of pseudo-random numbers.

New parcels were released at the location of the damaged Macondo rig at each hourly timestep. Twenty-five parcels were released at each position, and when combined with a 10 $m^2 s^{-1}$ diffusion coefficient, resulted in a trajectory spread with time. Initialization was based on NASA MODIS satellite imagery, SAR imagery from `http://www.cstars.miami.edu`, NOAAs Office of Response and Restoration oil trajectory maps at `http://response.restoration.noaa.gov`, and the NOAA/NESDIS Satellite Analysis Branch (SAB) exper-

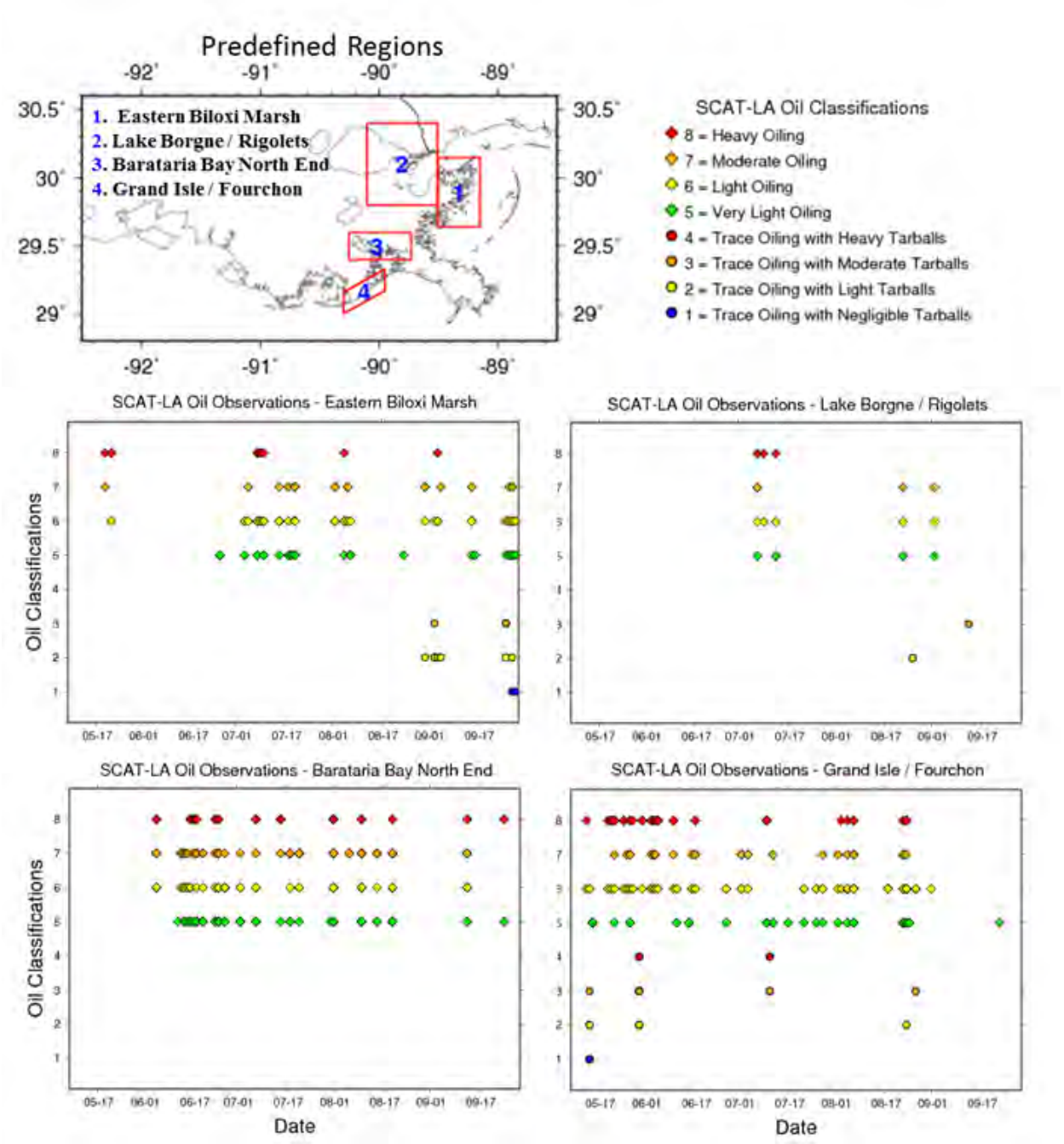

FIG. 17. Shoreline Cleanup and Assessment Technique (SCAT) data as tallied from 15 May to 20 September 2010. Shoreline oil pollution categories include designations for light, moderate, and heavy oiling; light, moderate, and heavy tarballs; and negligible tarballs. Plots are shown for Eastern Biloxi Marsh (region 1), Lake Borgne/Rigolets (region 2), North End of Barataria Bay (region 3), and the beach locations of Grand Isle/Fourchon (region 4).

imental surface oil analysis products at http://www.ssd.noaa.gov/PS/MPS/deepwater.html. The parcels were advected at 80% of the ocean current speed and at 3% of the wind speed. Bilinear interpolation was applied at each timestep to determine the currents and winds at each parcel position. AMSEAS supplied the near-surface currents, and COAMPS provided the 10-m winds. AMSEAS includes tidal components and a dynamic water surface which fluctuates from wind forcing, even capable of capturing storm surge events (D'Sa et al., 2011).

Fig. 18 shows four snapshots of the oil spill evolution simulated by the Lagrangian model for 20 June, 25 June, 30 June, and 5 July 2010, all at 0000 UTC. The first 8 days show two flow regimes: 1) east of the Mississippi River, oil moves northeast from the Macondo rig towards the Breton Sound islands, and the Alabama and west Florida coasts; and 2) west of the Mississippi River, a northwestward current impacting the west Delta Region, Sandy Point Beach, Barataria Bay, Terrebonne Bay, and the shorelines/estuaries further west ending in the vicinity of Atchafalaya Bay. Animations (not shown) include a pulsing action due to the diurnal tides common in this region. By the end of June, the simulation shows a sudden inward shift of the oil concentrations in western Mississippi Sound and Lake Borgne. A brief retreat occurs afterwards followed by a more prolonged inward penetration to these same sub-regions.

Synoptic data analysis clarifies the cause of these two events. We examined scatterometer data, satellite/radar imagery, high-frequency radar (HFR) currents, COAMPS wind fields, buoy data, and North American surface map analyses. The HFR data (not shown) indicated a switch from eastward to westward currents off Mississippi in late June, providing support that the AMSEAS ocean current changes were valid. An inspection of the weather maps shows a sequence of four distinct weather regimes (Fig. 19) that contributed to the two influxes of oil. A typical summertime pattern existed on 20 June, dominated by light winds and high pressure. Starting 25 June through 30 June, a tropical system affected the Gulf as a tropical wave entered the region and eventually became Hurricane Alex. The tropical wave became a depression by 1800 UTC 25 June about 148 km north-northeast of Puerto Lempira, Honduras, moved west-northwestward, became a tropical storm on 0600 UTC 26 June, and made its first landfall in the Yucatan Peninsula near Belize City around 0000 UTC 27 June. The weakened tropical storm then re-entered the southwest Gulf, strengthened to a category-2 hurricane, and made its final landfall near Soto la Marina, in northeastern Mexico

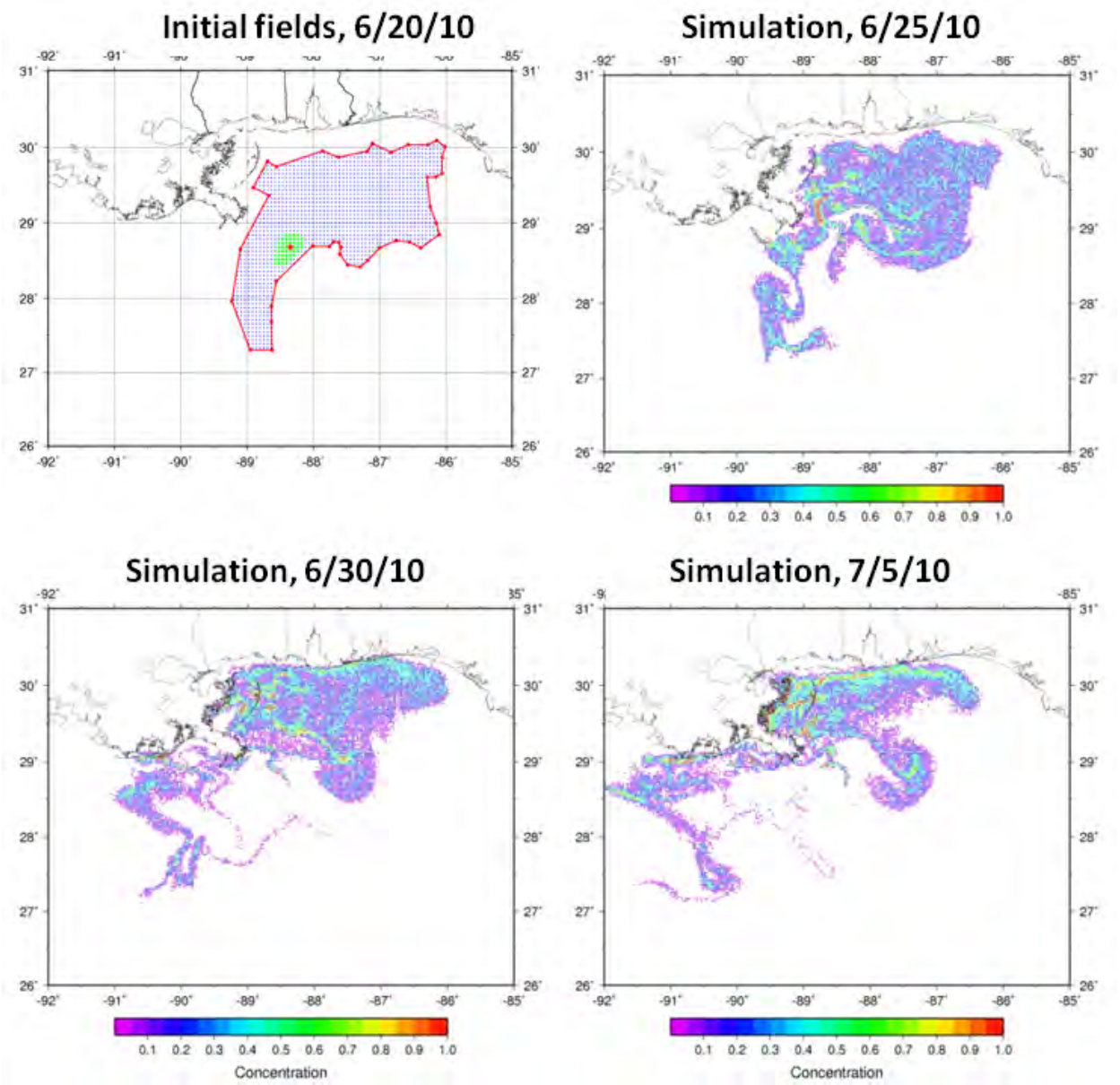

FIG. 18. Snapshot images of the DwH oil spill simulation from 0000 UTC 20 June 2010 to 0000 UTC 10 July 2010 in five-day increments. Note the inshore incursion into the Mississippi Sound and Lake Borgne regions starting in late June. Concentrations are computed as the ratio of parcels near a grid point divided by the number of parcels originally released at each point. In this simulation, each point has 25 releases at initialization, and then each trajectory is modified by a random number to mimic dispersion. Hence, concentrations in these runs are a fraction of 25. Concentration fields are shown from 0-100%, scaled from 0 to 1.

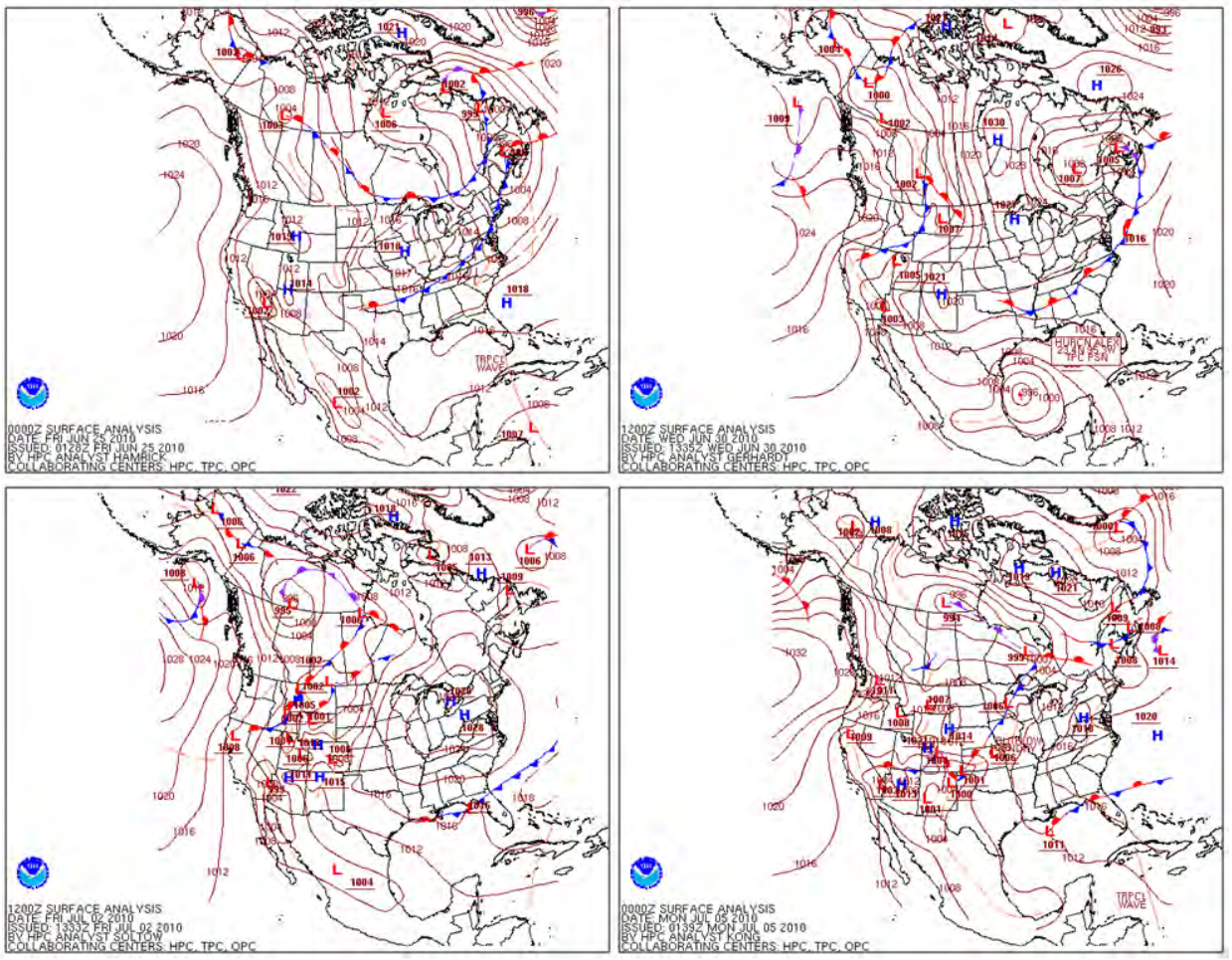

FIG. 19. Hydrometeorological Prediction Center (HPC) North American surface analysis for 0000 UTC 25 June 2010, 1200 UTC 30 June 2010, 1200 UTC 2 July 2010, and 0000 UTC 5 July 2010 (available at `http://www.hpc.ncep.noaa.gov/html/sfc_archive.shtml`). HPC is part of the NOAA/National Weather Service National Centers for Environmental Prediction.

around 0200 UTC 1 July. This period corresponds with the first inward oil incursion into the Lake Borgne region.

Afterwards, a cold front moved offshore into the eastern Gulf of Mexico (Fig. 19), creating a northerly wind flow in the northern Gulf Coast region. During this period, the oil retreated slightly. However, a non-tropical low pressure system formed on the western edge of this front, slowly moved westward, and stalled south of eastern Louisiana. This period was accompanied by a second oil incursion into the Mississippi Sound and Lake Borgne area.

The fringe effect of Alex, as well as the close proximity of the non-tropical low, not only switched alongshore westerward coastal currents (not shown) to an eastward direction, but

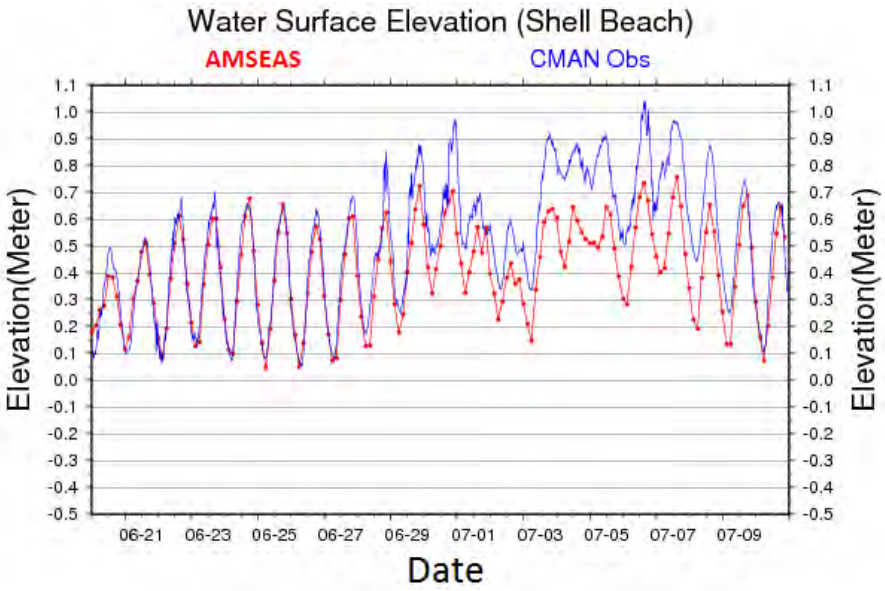

FIG. 20. Observed water level (blue) and predicted water AMSEAS water level (red) for Shell Beach CMAN station in Lake Borgne, LA, during 0000 UTC 15 June to 0000 UTC 15 July 2010 relative to NAVD88. Above average water elevation associated with Hurricane Alex and the non-tropical low pressure system are apparent on 29 June-1 July and 3-7 July, respectively.

also increased inland water levels by approximately 0.4 to 0.6 m above normal as mini-surge events. The Shell Beach C-MAN station (Fig. 20) located in Lake Borgne, LA, shows above normal water levels on 29 June to 1 July, followed by slightly above normal conditions as the front pushed through, then a more prolonged elevated water period for 4 to 7 July. C-MAN stations in Waveland, MS, and East Pascagoula, MS, display similar patterns (not shown). The closest AMSEAS gridpoint approximately 1.4 km away captured these two elevated water periods in Lake Borgne (Fig. 20), but the magnitudes are too low. This is probably because the model resolution cannot adequately capture the surge magnitudes this far inwards into the estuaries.

These results show cyclones can dramatically alter oil spill transport, even by fringe effects. The study also showed that this modeling formulation was capable of reproducing the oil spill transport. Much of the ocean current (not shown) south of the Mississippi River Delta was directed to the west, with oil impacting the Barataria Bay and Terrebonne Bay systems. To the east of the river system, the current moved towards Breton Sound, Alabama, and west Florida, and the oil spill was diplaced in a similar fashion. For the most

part, only these cyclonic events altered this pattern, which pushed the oil into the western Mississippi Sound and its marshes.

Because AMSEAS uses GNCOM for boundary conditions which has an approximate datum of zero mean sea level, but spin-up issues, gravity variations, and river input, datum difficulties can occur. For Fig. 20, the time series comparisons were phased together during benign weather conditions before Alex, then adjusted to NAVD88 using NOAA's Vertical Datum Transformation Tool (VDATUM; `http://vdatum.noaa.gov`), since C-MANs have no mean sea level datum option. Such subtle issues provide context for further study.

H. Class 4: Operational Tolerance Metrics (NAVOCEANO; June 2010 to March 2011)

Each new model improvement and implementation at NAVOCEANO goes through a formal and rigorous evaluation process before it can be declared "operational." This process culminates in an operational test (OPTEST) at NAVOCEANO to ensure the model meets specific Navy needs, which often focus on the three-dimensional temperature-salinity structure of the upper ocean. In the case of AMSEAS, DwH provided an unusual sense of urgency to the evaluation process, but at the same time provided an unusually rich observational dataset to enhance the evaluation. This section summarizes results from the AMSEAS OPTEST conducted at NAVOCEANO.

Liu et al. (2011) give an overview of the oceanographic observing effort in the northern Gulf of Mexico triggered by DwH. This included ship-based Expendable Bathythermograph (XBT) and Conductivity-Temperature-Depth (CTD) surveys; airborne XBT (AXBT) flights; Autonomous Profiler Explorer (APEX) profiling CTDs deployed in the international Argo field program; surface drifter deployments; and glider CTD surveys. Fig. 21 illustrates the total number of observations acquired in the GOMEX evaluation area between June 2010 and March 2011, at four depths (surface, 10 m, 100 m, and 500 m) that were used in the Navy's evaluation. Overall the number of measurements tapered off after the Macondo well was capped in July 2010. Subsurface measurements are fewer and decrease more rapidly with time than surface measurements owing to the network of moored – mostly coastal – buoys that continuously monitor surface properties.

The NAVOCEANO analysis was confined to ocean temperature and salinity. The assess-

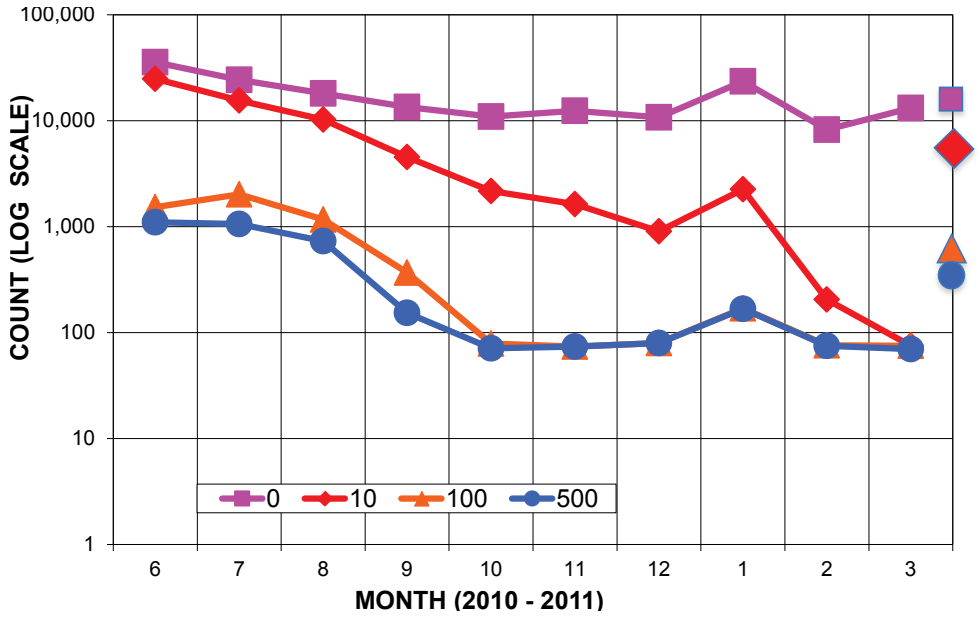

FIG. 21. Monthly count of observed temperature data at 0-, 10-, 100-, and 500-m depths. Note vertical scale is logarithmic. The 10-month means plotted at the right are reflected in Column 2 of Table V.

ment process is based on the NRL-developed program called AutoMetrics (Dykes, 2011). For each observation received, software finds and logs matching forecast data from concurrent ocean-model fields. Thus, before its assimilation into a subsequent model run, one observation can be independently compared with multiple AMSEAS forecasts produced one, two, three, or more days earlier. The observed properties are interpolated in time and space to the nearest model time step, grid-point, and depth (thus introducing interpolation errors) and stored in daily arrays. This procedure leads to the consistent comparison of observations and model products in model space.

Monthly metrics were computed to ensure that there are sufficient data to produce statistically significant results. Standard approaches, as outlined in Zhang et al. (2006, 2010), were used to calculate observed means, model means, model bias (modeled minus observed mean differences), correlation coefficients (Pearson method), and root mean square differences (RMSD; the term "difference" is used here instead of "error" to reflect the fact that observation errors are not defined for these data). In addition to these global mean measures, a "tolerance" metric, equivalent to central frequency in Zhang et al. (2006, 2010), was employed. This metric can be described as the percentage of model-minus-observed

TABLE V. Means of 10-month evaluation statistics for the Gulf of Mexico area of the AMSEAS NCOM for forecast day-1, for water temperature (°C) and salinity (psu). *OBS = observed, STD = Standard Deviation, CORR COEF = correlation coefficient, RMSD = Root Mean Square Difference. TOL = tolerance (see text).

	Depth (m)	# of Points	OBS MEAN	OBS STD	MODEL MEAN	MODEL STD	MODEL-OBS BIAS	BIAS STD	CORR COEF	RMSD	% IN TOL.
T	0	17051	23.99	2.19	24.23	2.11	+0.25	0.51	0.86	0.56	67.0%
	10	6241	24.33	1.30	24.34	1.26	+0.02	0.29	0.87	0.29	88.1%
	100	564	20.08	1.89	19.90	1.96	-0.18	0.71	0.86	0.73	57.5%
	500	358	8.63	1.12	8.47	1.12	-0.16	0.37	0.85	0.40	78.0%
S	0	17051	34.93	1.44	34.93	1.54	-0.01	0.76	0.78	0.74	41.9%
	10	6241	35.91	0.34	35.85	0.34	-0.06	0.18	0.51	0.17	78.4%
	100	564	36.45	0.07	36.41	0.08	-0.04	0.08	0.38	0.09	96.5%
	500	358	35.04	0.13	35.02	0.13	-0.02	0.05	0.83	0.05	100.0%

differences that lie within a specified objective for model accuracy. The tolerance metric is essentially a simplified approach to demonstrate probability distribution or data spread. The temperature and salinity tolerances set by Navy for AMSEAS were ±0.5°C and ±0.20 psu, respectively. These values relate to 2 ms^{-1} accuracy in sound speed, an important metric relevant to anti-submarine warfare. Table V summarizes these statistics for forecast day-1 for the 10-month period evaluated.

Examining these statistics by month reveals some interesting patterns. Considering only output from 24-h forecasts, at all levels, the monthly mean bias metrics for temperature and salinity are close to, or within, the AMSEAS tolerance goals (dashed lines, Fig. 22a & 22b). The very low surface-temperature biases within ±0.1°C from June to September reflect the assimilation of the extensive observations acquired during this period. The increasing warm bias beginning in October and peaking to an average of +0.6°C by February requires further investigation. Examination of atmospheric heat fluxes from COAMPS may provide useful clues to understanding this warming. The 10-month mean surface bias is +0.2°C. The 10-m temperature bias remains excellent at nearly 0.0°C until January when it starts to increase to +0.2°C peak by February. The 100-m temperature bias is near 0.0°C June to September, dropping to −0.7°C in December, and returning to 0.0°C by February, for a 10-month, 100-m mean bias of −0.2°C. At 500 m, the monthly bias remains slightly negative with a minimum of −0.4°C by December and an overall mean of −0.2°C. The 500-m model temperatures are

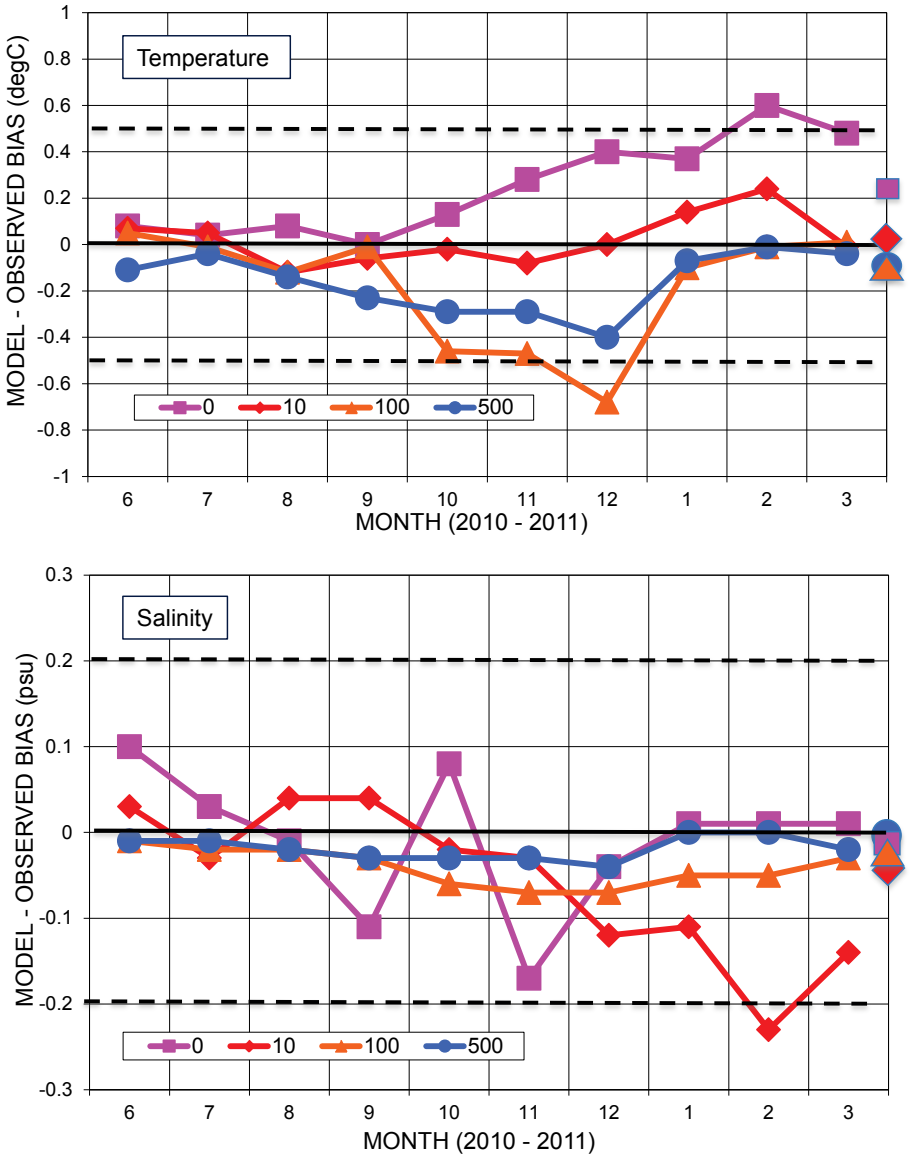

FIG. 22. Month-by-month mean model minus observed bias for temperature (upper) and salinity (lower) at the surface, 10-, 100-, and 500-m depths, for 24-h forecasts. Dashed lines represent tolerance goals of ±0.5°C for temperature and ±0.20 psu for salinity. The 10-month (June 2010 – March 2011) means at the right are reflected in Column 7 of Table V.

strongly influenced by climatology suggesting that actual ocean temperatures at this depth might be warmer than those of the climatology used. Salinity biases at all depths remain within the ±0.20 psu envelope.

The surface temperature RMSD, again for 24-h forecasts, rises throughout the period, reaching a maximum of 0.9°C by March (Fig. 23a). At 10 m, the 0.4°C RMSD in June

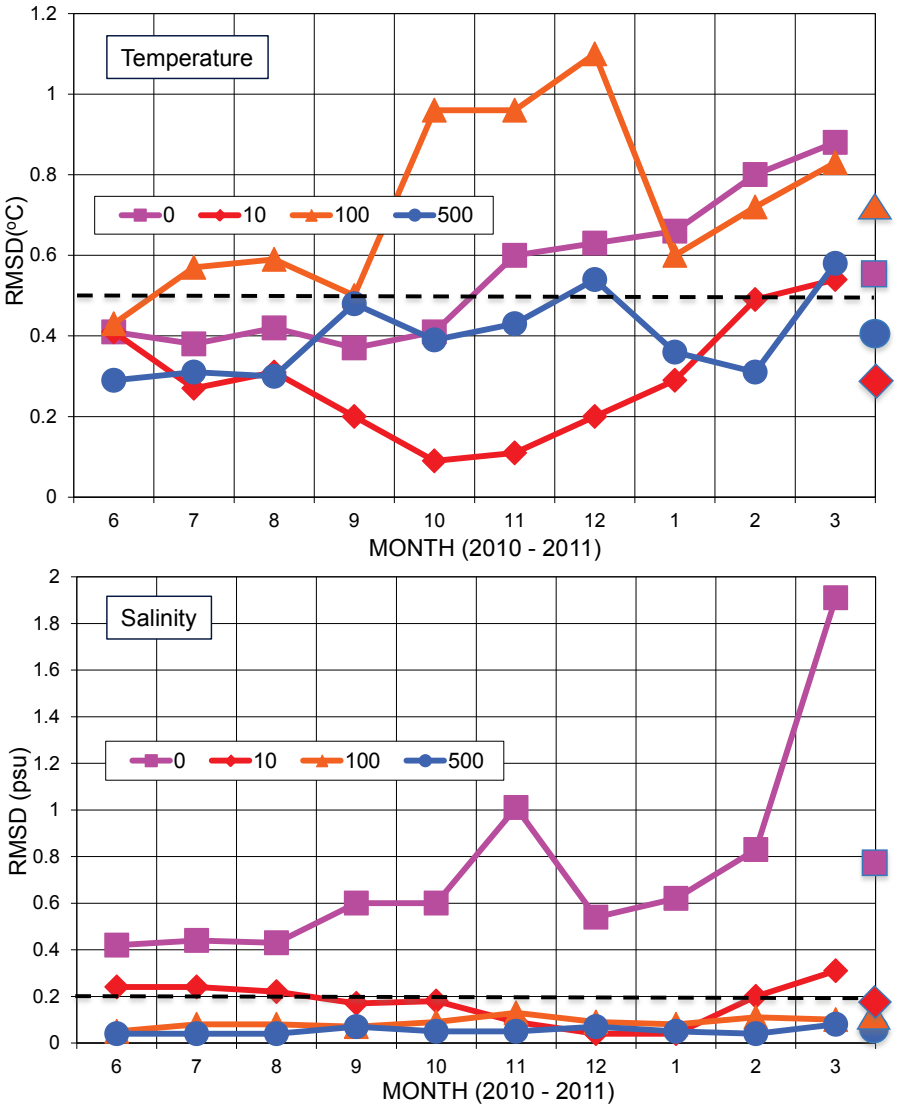

FIG. 23. Month-by-month mean temperature (upper) and salinity (lower) root mean square differences (RMSD) at the surface, 10-, 100-, and 500-m depths. Dashed lines represent tolerance goals of ±0.5°C for temperature and ±0.20 psu for salinity. The 10-month (June 2010 – March 2011) means at the right are reflected in Column 10 of Table V.

continually improves to 0.1°C by October, rising to 0.5°C by March, roughly opposite the 100-m temperature which has a period of larger RMSD from September to January. Given the unresolved internal wave fluctuations and strong vertical gradients at this depth, this result is not unexpected. The 500 m temperature RMSD ranges between 0.3°C and 0.5°C with a mean of 0.4°C.

The surface salinity RMSD ranges from 0.40 psu to 1.90 psu, peaking in March (Fig. 23b).

The 10-month surface mean is 0.75 psu. That these values are above the stated goal of 0.20 psu may result from of a number of factors including the climatological as opposed to real-time river runoff in this model, or from inaccurate atmospheric evaporation-minus-precipitation parameterization suggesting an additional area for future investigation. At 10 m, salinity RMSD is much closer to the designated tolerance limits, falling from 0.25 psu during the summer to near 0.05 psu December and January, and rising again to 0.30 psu by March. The 10-m mean RMSD is 0.20 psu. At 100 m and 500 m, the salinity RMSD values are 0.01 to 0.15 psu, for mean values of 0.10 psu at 100 m and 0.05 psu at 500 m, both of these within the tolerance limits.

Monthly mean bias and RMSD metrics might lead to an overly optimistic conclusion about AMSEAS performance. However, the tolerance plots show that while the mean measures appear quite good, there is substantial variability in individual results. Large percentages of the model-minus-observed differences fall outside the stated tolerance objectives (Fig. 24) as is especially true for the surface and 100-m temperatures and surface salinities.

The percentage of comparisons within tolerance stands near 80% for surface temperature from June to October, declining to 42% by February, with a mean over the period of 67%. Temperature percent within tolerance at 10- and 500-m ranges between 70% and 100% with overall means at 78% and 88%, respectively. At 100 m, temperature percent within tolerance is generally lower and more variable, ranging from 30% to 80%, with performance generally declining over time. Mean temperature percent within tolerance at 100 m over the period is only 58%.

Temporal patterns of salinity percent within tolerance are broadly similar to temperature with two important exceptions: first, 100-m salinity is much more accurate throughout the period, with desired percent within tolerance near 100%; second, while surface salinity percent within tolerance shows a similar pattern of declining performance with time, the overall percentages are much lower than for temperature, averaging only 42%. Table VI and Fig. 25 illustrate the decay in model skill between the first, second and third forecast days. For example the 100-m temperature tolerance percentage drops from 57.5% to 52.0% between day-1 and day-3, a relative 10% loss in skill over 72-h. The 100-m temperature RMSD increases 10% from 0.74 to 0.82°C between day-1 and day-3. Comparisons at other levels (except for salinity at 10-m) are broadly consistent suggesting that a 10% loss at 100-m is a reasonable estimate for the drop in model skill between the 24- and 72-h forecasts.

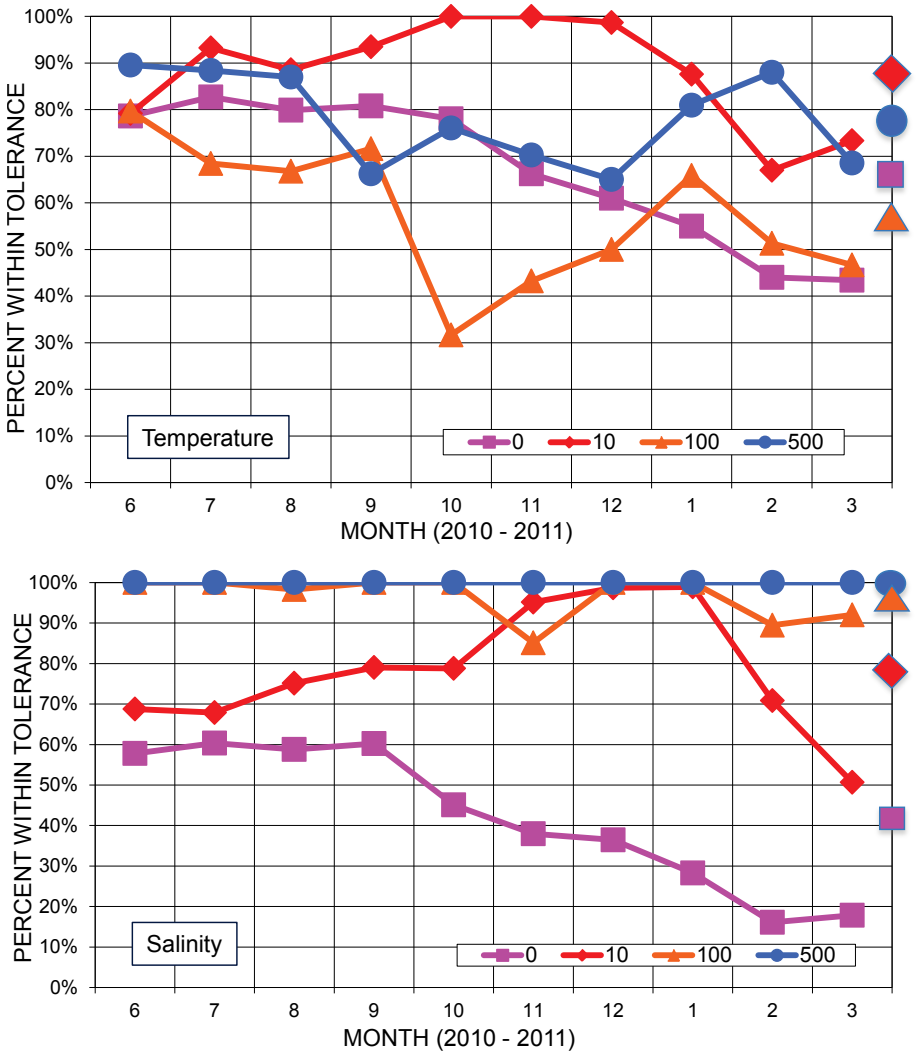

FIG. 24. Month-by-month mean percent within tolerance for temperature (upper) and salinity (lower) at the surface, 10-, 100-, and 500-m depths. The 10-month (June 2010 – March 2011) means at the right are reflected in Column 11 of Table V. Tolerance goals are ±0.5°C for temperature and ±0.20 psu for salinity.

In summary, temperature and salinity mean monthly biases are excellent as they generally remain within the stated modeling objectives of ±0.5°C and ±0.20 psu. Temperature RMSD metrics remained at or below the 0.5°C goal for all but the 100-m metrics, where an increased model error would be expected due to the strong vertical gradients that characterize the thermocline. Salinity RMSD is well below the 0.20 psu standard for all levels except the surface. While the means of bias and RMSD suggest excellent model skill, the tolerance

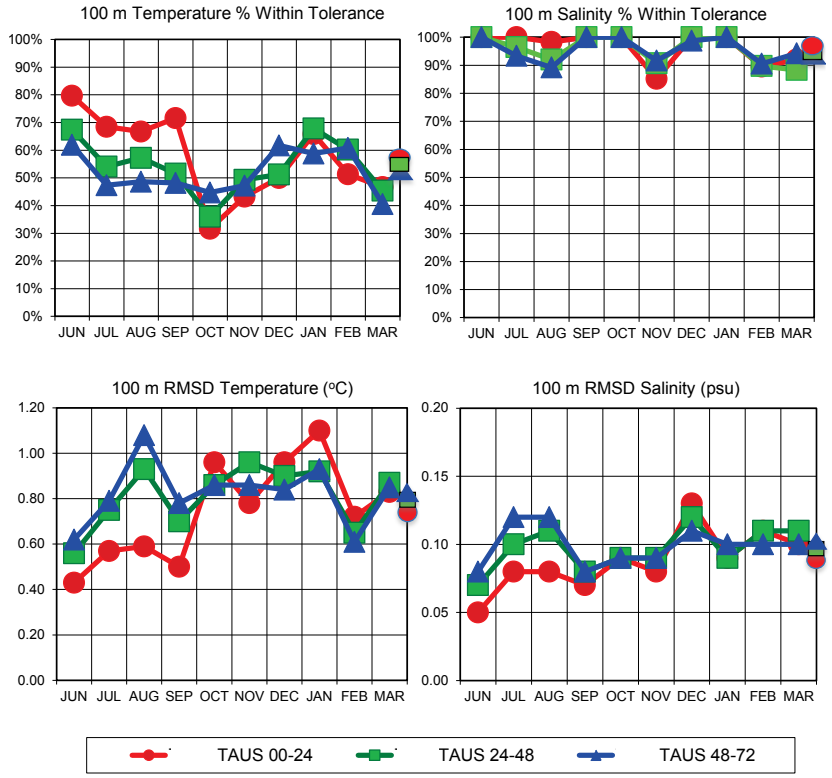

FIG. 25. Monthly means of percent within tolerance limits for 100-m temperature (upper left) and 100-m salinity (upper right) and RMSD for 100-m temperature (lower left) and 100-m salinity (lower right) for forecast days one (red), two (green) and three (blue). Tolerance limits are ±0.5°C for temperature and ±0.20 psu for salinity. The 10-month (June 2010 – March 2011) means are plotted at right on each plot.

TABLE VI. Mean 1-, 2-, and 3-day model skills based on comparative tolerance and root-mean-squared differences averaged over 10 months for temperature (°C) and salinity (psu).

	Depth	TOL.	TOL.	TOL.	RMSD	RMSD	RMSD
	(m)	day 1	day 2	day 3	day 1	day 2	day 3
T	0	67.0%	64.1%	61.5%	0.54	0.57	0.61
	10	88.1%	83.8%	81.7%	0.29	0.34	0.39
	100	57.5%	54.0%	52.0%	0.74	0.81	0.82
	500	78.0%	76.4%	73.8%	0.43	0.40	0.42
S	0	41.9%	37.6%	36.4%	0.75	0.76	0.78
	10	78.4%	69.5%	63.8%	0.20	0.24	0.28
	100	96.5%	95.7%	95.8%	0.09	0.10	0.10
	500	100.0%	100.0%	100.0%	0.06	0.05	0.05

measures show substantial spread, indicating areas for future investigation. This conclusion is particularly true for surface and 100-m temperature and surface salinity tolerance metrics. The decay in model skill over the 3-day forecast is only about 10%, suggesting that the AMSEAS products can be used with confidence throughout the forecast period. Seasonal changes indicated by these skill metrics must be tempered by the knowledge that the rate of subsurface data acquisition tapered off substantially after the immediate response to DwH, with an associated reduction in statistical confidence.

III. CONCLUSIONS

Based on the above initial evaluations, especially the operationally-relevant tolerance metrics, AMSEAS provides a useful operational baseline nowcast/forecast capability for use by both the research and operational communities. By providing this baseline capability, AMSEAS represents a standard against which existing and future research capabilities in the Gulf of Mexico can be measured for operational implementation. In keeping with the COMT goals, the multiple evaluation methods and graphics outlined above also provided useful examples as challenges to the developing testbed. All validation metrics indicate AMSEAS produces skillful forecasts and small, but systematic, improvements compared to IASNFS.

Limitations identified in the above evaluations suggest multiple areas for future research and analysis to better understand and improve the present capability. As noted in Section 3.1, low day-1 SST biases from June to September 2010 associated with the high data availability in the post-DwH months developed into a subsequent warming bias from October 2010 to March 2011 as the amount of data available for assimilation decreased. These results recommend both an increased need for an expanded long-term observational network in the Gulf available for data assimilation, as well as an investigation into both surface mixing in AMSEAS and into the COAMPS heat fluxes used to force the model. Alternatives to the use of climatological river discharge as well as investigation of the evaporation and precipitation used as forcing are both suggested by the evolution of surface salinity away from designated tolerance levels.

Beyond the day-1 forecast, the 10-month mean within-tolerance levels for day-2 and day-3 appear to degrade by an acceptable 10% per forecast day. However, reviewing the monthly

evolution of the SST from August 2010 to October 2011 generally revealed a cooling bias from day-1 to day-4 forecasts for the moored buoys analyzed. This result again recommends future work investigating AMSEAS surface mixing and the COAMPS forecast heat fluxes.

The limited COAMPS study investigating AMSEAS wind forcing demonstrated that, in the northern Gulf area, COAMPS consistently under predicted wind speed for the summer and winter periods. Future work examining threshold metrics of winter versus summer wind regimes need to be performed since weaker equivalent barotropic wind regimes contain more directional variability while winter baroclinic wind regimes statistically contain more variance to explain but with larger absolute wind errors. Additional study of the use of vector correlation would be useful to bridge this gap. Coastal versus offshore wind comparisons revealed greater errors along the coast suggesting the 15-km COAMPS grid's inability to resolve coastal topography and/or land-sea temperature differences, both having the ability to significantly influence the coastal winds. Future work investigating the significance of the COAMPS grid resolution would help resolve this issue.

The subset of the GOMEX-PPP effort reported above focused on nowcast skill of IASNFS and AMSEAS related to the LC structure and FC transport in the Eddy Franklin timeframe of 25 May to 31 December 2010. Relative to the AVISO analysis, SSH RMSD is nearly identical for both models at 0.11 m and 0.12 m, respectively. Advances in observational networks, data assimilation, and increased model grid resolution are needed to better represent the cyclonic eddies influencing the LC and LCE. The temporal coincidence of the shedding of Eddy Franklin and the change in FC transport at this time, and the underestimation of transport by both IASNFS and AMSEAS, suggest two additional areas of future research. In summary, the evaluation of AMSEAS and IASNFS within the GOMEX-PPP project found differences between the nowcasts of the two forecast systems. In general, the differences point to systematic, if small, improvements from the older IASNFS to the newer AMSEAS. The comparisons also highlight the relative paucity of the present observational system for evaluating progress in the evolution of Gulf of Mexico modeling systems.

The AMSEAS validation studies provided confidence that an oil spill modeling effort could be performed, and a simulation was conducted for the period 20 June to 10 July 2010 using a Lagrangian particle tracker with random walk diffusion of archived AMSEAS data, with a particular focus on pollution pulses that penetrate into the estuaries east of the Mississippi River. The initial parcel locations were subjectively determined based on a combination of

NASA MODIS satellite thermal imagery, SAR imagery, NOAA oil trajectory maps, and the NOAA/NESDIS Satellite Analysis Branch (SAB) experimental surface oil analysis. This modeling formulation was capable of reproducing the oil spill transport, with ocean current south of the Mississippi River Delta directed to the west and impacting the Barataria Bay and Terrebonne Bay systems, while to the east of the river system, the current flowed towards Breton Sound, Alabama, and west Florida, with the oil spill displaced in a similar fashion. This modeling effort also captured the estuarine water inundation influences of Hurricane Alex and a non-tropical cyclone off the LA coast, both of which pushed oil into the western Mississippi Sound, Lake Borgne, the Rigolets, and vicinity inner marshes. The utilization and interpretation Lagrangian transport models for oil spill dispersal modeling is an area of active development (Dietrich et al., 2012; Le Hénaff et al., 2012).

The lack of a comprehensive, well-designed, operational observing system in the Gulf is a significant weakness in monitoring various environmental and ecological disasters and in supporting and validating data assimilative prediction systems. Other than the aggregation of coastal tide gauge, coastal meteorological station, and meteorological buoy networks, there is no in situ, operational (real-time, standardized, and sustained) observing system for the Gulf of Mexico. Satellite SSH and SST data streams are invaluable but can only partially substitute for time series of vertical profiles of field variables. This need is exemplified by the work of Shay et al. (2011) that yielded a set of nine synoptic maps of the upper ocean thermal structure at intervals of 7 to 10 days over the LC and Eddy Franklin between May and July 2010. Evaluating regional HYCOM hindcast simulations with these data, reductions of 30% in RMSD and 50% in bias were found when compared to simulations assimilating remotely-sensed data alone.

From another perspective, more than ten Gulf of Mexico circulation models exist internationally, yet there is no systematic and sustained activity to evaluate them vis-a-vis observations and document their capabilities, short falls, and improvements measured against community standards. Ironically, this lack exists at a time when ocean model forecast products need error estimates as well as field estimates to satisfy user needs, and when ensemble modeling (single model or multiple model), through, statistical analyses, estimate the evolving mean and variance of field variables through data assimilation. Thus, institutional and programmatic leadership are needed to advance operational ocean predictions accompanied with testbed functionality.

ACKNOWLEDGMENTS

The authors were supported in this work by NOAA Grant 2010-012 through the U.S. IOOS Program administered by the Southeastern Universities Research Association (Fitzpatrick, Harding, Ko, Wiggert); by BP through the Gulf of Mexico Research Initiative Award #10121360 to the Northern Gulf Institute (Fitzpatrick, Harding); by the U.S. Department of Homeland Security Award #2008-ST-061-ND 0001 to Mississippi State University (Fitzpatrick); and by the Department of Energy Research Partnership to Secure Energy for America Project #08121-2801 to Portland State University (Mooers, Zaron). The authors thank Jamie Davis for contributing to AMSEAS analyses and graphics development. Brian Hanson provided the FORTRAN code for the Hanson et al. (1992) algorithm (available at `http://hanson.geog.udel.edu/~hanson/hanson/Research.html`). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NOAA, BP, the U.S. Department of Homeland Security, the U.S. Department of Energy, or the U.S. Navy.

REFERENCES

Arnone, R. A., Casey, B., Ko, D. S., Flynn, P., Carrolo, L., Landner, S., 2007. Forecasting coastal optical properties using ocean color and coastal circulation models. Proceedings of SPIE - The International Society for Optical Engineering, 6680.

Arnone, R. A., Casey, B., Ladner, S., Ko, D. S., Gould, R. W., 2010. Forecasting the coastal optical properties using satellite ocean color. In: Barale, V., Gower, J. F., Alberotanza, L. (Eds.), Oceanography from Space. pp. 335–348.

Barron, C. N., Birol Kara, A., Rhodes, R. C., Rowley, C., Smedstad, L. F., 2007. Validation test report of the 1/8° Global Navy Coastal Ocean Model nowcast/forecast system. Tech. Rep. NRL/MR/7320–07-9019, Naval Research Laboratory, Stennis Space Center, MS, 144 pp.

Breaker, L. C., Gemmill, W. H., 1994. The application of a technique for vector correlation to problems in meteorology and oceanography. Journal of Applied Meteorology 33, 1354–1365.

Brunner, C. A., Beall, J. M., Bentley, S. J., Furukawa, Y., 2006. Hypoxia hotspots in the

Mississippi Bight. Journal of Foramineral Research 36, 95–107.

Chang, Y.-L., Oey, L.-Y., 2011. Loop Current cycle: Coupled response of the Loop Current with deep flows. Journal of Physical Oceanography 41, 458–471.

Chang, Y.-L., Oey, L.-Y., 2013. Coupled response of the trade wind, SST gradient, and SST in the Caribbean Sea, and the potential impact on Loop Current's interannual variability. Journal of Physical Oceanography 43, 1325–1344.

Charnock, H., 1955. Wind stress over a water surface. Quarterly Journal of the Royal Meteorological Society 81, 639–640.

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

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Barron, C. N., Ko, D. S., Rhodes, R. C., Shriver, J. F., Wallcraft, A. J., Arnone, R. A., 2005. Assessment of data assimilative ocean models in the Gulf of Mexico using ocean color. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models. Vol. 161 of Geophysical Monograph Series. American Geophysical Union, pp. 87–100.

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

Dietrich, J. C., Trahan, C. J., Howard, M. T., Fleming, J. G., Weaver, R. J., Tanaka, S., Yu, L., Luettich Jr., R. A., Dawson, C. N., Westerink, J. J., Wells, G., Lu, A., Vega, K., Kubach, A., Dresback, K. M., Kolar, R. L., Kaiser, C., Twilley, R. R., 2012. Surface trajectories of oil transport along the northern coastline of the Gulf of Mexico. Continental Shelf Research 41, 17–47.

Dimou, N. K., Adams, E. E., 1993. A random-walk, particle tracking model for well-mixed estuaries and coastal waters. Estuarine, Coastal, and Shelf Science 37, 99–110.

D'Sa, E. J., Ko, D. S., 2008. Short-term influences on suspended particulate matter distribution in the northern Gulf of Mexico: Satellite and model observations. Sensors 8, 4249–4264.

D'Sa, E. J., Korobkin, M., Ko, D. S., 2011. Effects of Hurricane Ike on the Louisiana-Texas coast from satellite and model data. Remote Sensing Letters 2, 11–19.

Dykes, J. D., July 2011. Implementation of the automated numerical model performance metrics system. Tech. Rep. NRL/MR/7320-11-9353, Naval Research Laboratory, Stennis Space Center, MS, 31 pp.

Green, R. E., Gould, R. W., Ko, D. S., 2008. Statistical models for sediment/detritus and dissolved absorption coefficients in coastal waters of the northern Gulf of Mexico. Continental Shelf Research 28, 1273–1285.

Gundlach, E. R., Finkelstein, K. J., Sadd, J. L., March 1981. Impact and persistence of Ixtoc I oil on the South Texas Coast. In: Proceedings: 1981 Oil Spill Conference (Prevention, Behavior, Control, Cleanup). Atlanta, GA, pp. 477–485.

Haltrin, V. I., Arnone, R. A., Flynn, P., Casey, B., Weidemann, A. D., Ko, D. S., 2007. Restoring number of suspended particles in ocean using satellite optical images and forecasting particle fields. Proceedings of SPIE - The International Society for Optical Engineering, 6615.

Hanson, B., Klink, K., Matsuura, K., Robeson, S. M., Willmott, C. J., 1992. Vector correlation: Review, exposition, and geographic application. Annals of the Association of American Geographers 82, 103–116.

Harding, J., Cross, S., Bub, F., M., J., Carleton, C., Tolman, H., Parsons, A. R., 2013. OceanNOMADS- an update: Real-time and retrospective access to operational U.S. ocean prediction products, 11th Symposium on the Coastal Environment at the 93rd Annual Meeting of the American Meteorology Society, Austin, TX.

Hernandez, F., Bertino, L., Brassington, G., Chassignet, E., Cummings, J., Davidson, F., Drévillon, M., Garric, G., Kamachi, M., Lellouche, J.-M., Mahdon, R., Martin, M. J., Ratsimandresy, A., Regnier, C., 2009. Validation and intercomparison studies within GODAE. Oceanography Magazine 22, 128–143.

Hodur, R. M., 1997. The Naval Research Laboratorys Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). Monthly Weather Review 125, 1414–1430.

Hodur, R. M., Hong, X., Doyle, J. D., Pullen, J., Cummings, J., Martin, P., Rennick, M. A., 2002. The Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). Oceanography 15, 88–98.

Hunter, J. R., Craig, P. D., Phillips, H. E., 1993. On the use of random walk models with spatially variable diffusivity. Journal of Computational Physics 106, 366–376.

Jacobs, G. A., Barron, C. N., Fox, D. N., Whitmer, K. R., Klingenberger, S., May, D.,

Blaha, J. P., 2002. Operational altimeter sea level products. Oceanography 15, 13–21.

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

Ko, D. S., Preller, R. H., Martin, P. J., 2003. An experimental real-time Intra Americas Sea ocean nowcast/forecast system for coastal prediction. In: Proceedings, AMS 5th Conference on Coastal Atmospheric & Oceanic Prediction & Processes. Seatlle, WA, p. 5.2.

Kundu, P. K., 1976. Ekman veering observed near the ocean bottom. Journal of Physical Oceanography 6, 238–242.

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

Le Hénaff, M., Kourafalou, V. H., Paris, C. B., Helgers, J., Aman, Z. M., Hogan, P. J., Srinivasan, A., 2012. Surface evolution of the Deepwater Horizon oil spill patch: Combined effects of circulation and wind-induced drift. Environmental Science & Technology 46 (13), 7267–7273.

Liu, Y., MacFadyen, A., Ji, Z., Weisberg, R., 2011. Introduction to monitoring and modeling the Deepwater Horizon oil spill. In: Liu, Y., MacFadyen, A., Ji, Z., Weisberg, R. (Eds.), Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise. Vol. 195 of Geophysical Monograph Series. American Geophysical Union, Washington, D.C., pp. 1–7.

Lubchenco, J., McNutt, M. K., Dreyfus, G., Murawski, S. A., Kennedy, D. M., Anastas, P. T., Chu, S., Hunter, T., 2012. Science in support of the Deepwater Horizon response. Proceedings of the National Academy of Sciences of the United States of America 109 (50), 20222–20228.

Luettich, R., Wright, D. L., Signell, R., Friedrichs, C., Friedrichs, M., Harding, J., Fennel, K., Howlett, E., Graves, S., Smith, E., Crane, G., Baltes, R., 2013. Introduction to special issue on The U.S. IOOS Coastal and Ocean Modeling Testbed: Overview of its motivation, goals, implementation and scope. Journal of Geophysical Research 118, 6319–6328.

Lugo-Fernández, A., Leben, R. R., 2010. On the linear relationship between Loop Current retreat latitude and eddy separation period. Journal of Physical Oceanography 40 (12), 2778–2784.

MacFadyen, A., Watabayashi, G., Barker, C., Beegle-Krause, C. J., 2011. Tactical modeling of surface oil transport during the Deepwater Horizon spill response. In: Liu, Y.,

MacFadyen, A., Ji, Z., Weisberg, R. (Eds.), Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise. Vol. 195 of Geophysical Monograph Series. American Geophysical Union, Washington, D.C., pp. 167–178.

Martin, P. J., 2000. Description of the Navy Coastal Ocean Model Version 1.0. Tech. Rep. NRL/FR/7322–00-9962, Naval Research Laboratory, Stennis Space Center, MS, 42 pp.

Mendoza, W. G., Zika, R. G., Corredor, J. E., Ko, D. S., Mooers, C. N., 2009. Developmental strategy for effective sampling to detect possible nutrient fluxes in oligotrophic coastal reef waters in the Caribbean. Journal of Operational Oceanography 2, 35–47.

Mooers, C. N., Meinen, C. S., Baringer, M. O., Bang, I., Rhodes, R., Barron, C. N., Bub, F., 2005. Cross validating ocean prediction and monitoring systems. EOS, Transactions of the American Geophysical Union 86, 269–273.

Morey, S. L., Martin, P. J., O'Brien, J. J., Wallcraft, A. A., Zavala-Hidalgo, J., 2003. Export pathways for river discharged fresh water in the northern Gulf of Mexico. Journal of Geophysical Research 108 (C10).

Naeije, M., Doornbos, E., Mathers, L., Scharroo, R., Schrama, E., Visser, P., 2002. Radar Altimeter Database System: Exploitation and extension (RADSxx). Tech. Rep. NUSP-2 report 02-06, NUSP-2 project 6.3/IS-66, Delft Institute for Earth-Oriented Space Research (DEOS), Delft, Netherlands, iSBN 90-5623-077-8.

Peggion, G., Barron, C., Rowley, C., 2007. A rapidly relocatable ocean prediction system. Tech. Rep. NRL/PP/7320-06-6212, Naval Research Laboratory, Stennis Space Center, MS, 5 pp.

Pichevin, T., Nof, D., 1997. The momentum imbalance paradox. Tellus 49, 298–319.

Powell, B. S., Leben, R. R., 2004. An optimal filter for geostrophic mesoscale currents from along-track satellite altimetry. Journal of Atmospheric and Oceanic Technology 21, 1633–1642.

Powell, M. D., Houston, S. H., Reinhold, T. A., 1996. Hurricane Andrews landfall in South Florida. Part I: Standardizing measurements for documentation of surface wind fields. Weather and Forecasting 11, 304–328.

Rabalais, N. N., Turner, R. E., Wiseman, Jr., W. J., 2001. Hypoxia in the Gulf of Mexico. Journal Environmental Quality 30, 320–329.

Rhodes, R. C., Hurlburt, H. E., Wallcraft, A. J., Barron, C. N., Martin, P. J., Smedstad, O. M., Cross, S., Metzger, E. J., Shriver, J., Kara, A., Ko, D. S., 2002. Navy real-time

global modeling system. Oceanography 15.

Rosmond, T. E., Teixeira, J., Peng, M., Hogan, T. F., Pauley, R., 2002. Navy Operational Global Atmospheric Prediction System (NOGAPS): Forcing for ocean models. Oceanography 15 (1), 99–108.

Schmitz Jr., W. J., Biggs, D. C., Lugo-Fernandez, A., Oey, L.-Y., Sturges, W., 2005. A synopsis of the circulation in the Gulf of Mexico and on its continental margins. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models. Vol. 161 of Geophysical Monograph Series. American Geophysical Union, Washington, D.C., pp. 11–30.

Shay, L. K., Jaimes, B., Brewster, J. K., Meyers, P., McCaskill, E. C., Uhlhorn, E., Marks, F., Halliwell Jr., G. R., Smedstad, O. M., Hogan, P., 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. Vol. 195 of Geophysical Monogrograph Series. American Geophysical Union, Washington, D.C., pp. 131–151.

Shoosmith, D. R., Baringer, M. O., Johns, W. E., 2005. A continuous record of Florida Current temperature transport at 27°N. Geophysical Research Letters 32, L23603.

Sturges, W., Hoffmann, N., Leben, R., 2010. A trigger mechanism for loop current ring separations. Journal of Physical Oceanography 40, 900–913.

van Leeuwen, P. J., De Ruijter, W. P., 2009. On the steadiness of separating meandering currents. Journal of Physical Oceanography 39, 437–448.

Wilson, W. D., Johns, W. E., 1997. Velocity structure and transport in the Windward Islands Passages. Deep Sea Research 44 (3), 487–520.

Xu, F.-H., Chang, Y.-L., Oey, L.-Y., Hamilton, P., 2013. Loop Current growth and eddy shedding using models and observations: Analyses of the July 2011 eddy-shedding event. Journal of Physical Oceanography 43, 1015–1027.

Zavala-Hidalgo, J., Morey, S. L., OBrien, J. J., 2003. Seasonal circulation on the western shelf of the Gulf of Mexico using a high-resolution numerical model. Journal of Geophysical Research 108 (C12), 3389.

Zhang, A., Hess, K. W., Wei, E., Myers, E., 2006. Implementation of model skill assessment software for water level and current in tidal regions. Tech. Rep. NOS CS 24, NOAA, 61 pp.

Zhang, A., W., H. K., Aikman III, F., 2010. User-based skill assessment techniques for operational hydrodynamic forecast systems. Journal of Operational Oceanography 3, 12–24.