4DVAR data assimilation in the Intra-Americas Sea with the Regional Ocean Modeling System (ROMS)

B.S. Powell [a,*], H.G. Arango [b], A.M. Moore [c], E. Di Lorenzo [d], R.F. Milliff [e], D. Foley [f,g]

[a] Institute of Marine Sciences, University of California, Santa Cruz, CA 95064, USA
[b] Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ, USA
[c] Department of Ocean Sciences, University of California at Santa Cruz, CA, USA
[d] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA
[e] NWRA, Colorado Research Associates Division, Boulder, CO, USA
[f] Joint Institute for Marine and Atmospheric Research, University of Hawai'i, Manoa, HI, USA
[g] NOAA Southwest Fisheries Science Center, Environmental Research Division, Santa Cruz, CA, USA

ARTICLE INFO

ABSTRACT

We present the background, development, and preparation of a state-of-the-art 4D variational (4DVAR) data assimilation system in the Regional Ocean Modeling System (ROMS) with an application in the Intra-Americas Sea (IAS). This initial application with a coarse model shows the efficacy of the 4DVAR methodology for use within complex ocean environments, and serves as preparation for deploying an operational, real-time assimilation system onboard the Royal Caribbean Cruise Lines ship *Explorer of the Seas*. Assimilating satellite sea surface height and temperature observations with in situ data from the ship in 14 day cycles over 2 years from January 2005 through March 2007, reduces the observation-model misfit by over 75%. Using measures of the Loop Current dynamics, we show that the assimilated solution is consistent with observed statistics.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The Intra-Americas Sea (IAS) encompasses the Gulf of Mexico and Caribbean Sea of the North Atlantic. The region (Fig. 1) is semi-enclosed and the circulation is strongly constrained by transport through the Antilles island passages and Florida Straits, and the IAS circulation is an important part of the western boundary current system of the North Atlantic sub-tropical gyre. The northward Antilles current flows to the east of the Antilles island chain and forms part of the western boundary current (Lee et al., 1996). The remainder is formed within the Caribbean consisting of water masses from numerous sources including the North Atlantic as well as waters from the South Atlantic and further afield (e.g. the Southern, Pacific, and Indian oceans). Recent and comprehensive reviews of the circulation in the region include Mooers and Maul (1998) and Schmitz et al. (2005).

The Caribbean Current is fed by ocean transport through the Antilles island passages, forming two westward jets centered, respectively at 13°N and 15°N, separated by a weak eastern flow (Morrison and Nowlin, 1982; Richardson, 2005). South of Hispani-

ola, these two currents meet and flow to the northwest as the Yucatan Current. Proceeding northwards, the Yucatan Current forms the Loop Current (LC), the energetic jet entering the Gulf of Mexico (GOM) through the Yucatan Channel and separating from the Campeche Bank. The LC is the dominant source of energy, variability, and momentum that drives much of the circulation within the Gulf of Mexico (Ohlmann et al., 2001; Schmitz et al., 2005). The Loop Current exits through the Florida Straits, forming the Florida Current, before proceeding north as the Gulf Stream.

Apart from the strong mean circulation of the GOM, there exists a richly varied dynamical structure of small-scale motions, very strong mesoscale turbulence, eddies, and jets (Leben, 2005; Schmitz, 2005). The bathymetry of the IAS is composed of five basins set apart by shallow sills less than 2000 m in depth. Due to the complex dynamics of the IAS region, it is often considered an excellent test for inverse modeling studies (e.g. Roemmich, 1981, 1983; Wunsch and Grant, 1982). Extensive reduced-gravity model studies were performed by Hurlburt and Thompson (1980), Smith and O'Brien (1983), and Arango and Reid (1991). With the advance of computers and general circulation models (GCM), comprehensive studies of the GOM include Sturges et al. (1993), Dietrich and Lin (1994), Oey (1995, 1996), Romanou et al. (2004) and IAS studies by Chérubin et al. (2005). For a complete review of numerical models in the GOM region, the reader is encouraged to peruse Oey et al. (2005b).
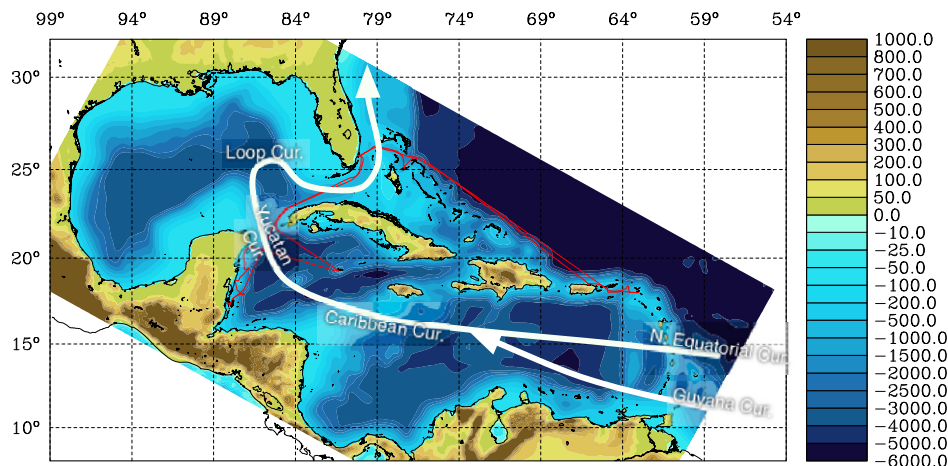
Fig. 1. The model domain encompasses the entirety of the IAS and portions of the North Atlantic. The bathymetry of the region is shown with illustrations of the main flows. The red tracks illustrate the western and eastern cruise tracks followed by the RCCL *Explorer of the Seas*.

Within the region, several of these models have used data assimilation techniques to adjust the numerical model to both *in situ* and remote observations (e.g. Wang et al., 2003; Oey et al., 2005a; Kantha et al., 2005). A review of previous work is available from Chassignet et al. (2005). We present a new modeling study using four-dimensional data assimilation based upon adjoint methods.

In preparation for the real-time data assimilation and prediction study performed aboard the Royal Caribbean Cruise Lines (RCCL) ship *Explorer of the Seas* from Winter through Fall of 2006–2007, we performed a series of experiments to explore data assimilation in the IAS. This paper has three purposes: present the method of four-dimensional data assimilation in Sections 2 and 4 as implemented in the Regional Ocean Modeling System along with the data used in Section 3; to test the minimization of the assimilation system with respect to the observations in Section 5; and, to present a state-of-the art data assimilation system that we are using as the basis of a real-time prediction system. It will be clear to the reader that our model does not fully resolve all dynamical scales of the region, and there is evidence that the sub-mesoscale has a significant influence on the dynamics of the LC (Chérubin et al., 2005); however, the observations used are of similar resolution and suffer from a similar limitation. Our purpose is to present the background and methodology of four-dimensional data assimilation with results of its application in the IAS.

2. Regional Ocean Modeling System (ROMS)

The Regional Ocean Modeling System (ROMS) is a free-surface, hydrostatic, primitive equation model discretized with a terrain following vertical coordinate system (Shchepetkin and McWilliams, 2005). The model has multiple sub-gridscale parameterizations of vertical mixing along with many options for open boundary conditions. Time-splitting of barotropic and baroclinic motions enables efficient time integration. ROMS has been used to model many regions of the world ocean (see http://www.myroms.org/papers), and it compares favorably with other terrain coordinate models such as the Princeton Ocean Model (Ezer et al., 2002).

In the present study, ROMS has been configured for the IAS region at both 1/3° and 1/6° horizontal resolution with 30 sigma-levels. This configuration is based upon the forward experiment of Haidvogel et al. (2000). The relatively coarse horizontal resolution is a necessary compromise because of the computational expense of the data assimilation procedure and the hardware constraints

of operating the assimilation system in real-time while onboard the ship. We cannot accurately resolve the first baroclinic mode of the Rossby radius of deformation at 1/3°; however, we can capture the large (200–300 km) Loop Current eddies in the Gulf of Mexico and many of the large energetic eddies in the Caribbean Sea. The current 1/3° coarse configuration forms part of a multiple grid approach to data assimilation which is under development using ROMS with progressively higher horizontal resolutions, namely 1/6° and 1/12° in the IAS. The current configuration recovers the large-scale IAS circulation that precondition the mesoscale. The model uses third-order horizontal advection with the generic length scale vertical mixing scheme (Warner et al., 2005) using k–kl mixing coefficients (corresponding to Mellor-Yamada Level 2.5).

The bathymetry of the region can present a challenge to terrain following coordinate models. Large changes in bathymetry between neighboring grid cells are capable of generating horizontal pressure-gradient force errors (PGE) that result in anomalous velocities. The so-called "r-factor" ($r = |\Delta H|/2H$) determines the maximum slope parameter or the ratio of the change in height (ΔH) versus total height (H) in adjacent cells, and it provides an indication of the potential for PGEs. Using an algorithm based upon the pressure Jacobian, Haidvogel et al. (2000) suggests that the "r-factor" not exceed 0.2. To minimize the PGE, a number of algorithms exist within ROMS, and we use the Density-Jacobian with Monotonized Cubic Polynomial Fits (Shchepetkin and McWilliams, 2003) which allows for higher "r-factors". Using the test proposed in Shchepetkin and McWilliams (2003), we integrate a homogenous, at-rest ocean for seven days without forcing or mixing to determine the PGE velocities in the coarse grid ($r = 0.35$). Any generated motions will be due to PGE. We found that the average spurious velocity along slopes greater than 150 m per cell was 0.5 ± 0.3 cm/s (max: 1.8 cm/s) in the GOM and 1.3 ± 1.1 cm/s (max: 8.2 cm/s) in the Caribbean. Comparing each point on the slope around the basins, we find that the growth of bottom PGE errors would account for 6.3% and 12.6% of the model velocity climatology in the GOM and Caribbean, respectively. Much higher r-factor values and PGEs have been used depending on the domain as discussed in Ezer et al. (2002).

There are two open boundaries in the IAS domain; the northern boundary, which spans part of the North Atlantic, and the eastern boundary, which includes the equatorial Atlantic and northeast Brazilian Current retroflection. At both open boundaries, the model is configured to conserve volume with a free-surface Chapman condition, a Flather condition for the 2D momentum, and clamped

for the 3D momentum and tracers (Marchesiello et al., 2001). The data for the clamped boundary conditions were derived from a long North Atlantic ROMS simulation run at 1/6° with NCEP forcing (Haidvogel et al., 2000). Using a 6 year time series of data, monthly climatological boundary conditions were computed for the IAS model. Climatological boundaries were used because the North Atlantic ROMS simulation does not coincide with the assimilation period and because during forecast periods boundary data are not available.

Ocean forcing is derived from the NCEP/NCAR reanalysis product (Kistler et al., 2001), at 2° spatial resolution every 6 h. From mid-1999 onwards, we use the QSCAT/NCEP blended ocean winds at 1/2° resolution every 6 h Milliff et al., 2004. The ocean surface heat flux is computed via the COARE algorithm (Fairall et al., 1996) using standard 2 and 10 m forcings from the NCEP/NCAR reanalysis. To allow for known uncertainties in the net surface freshwater flux and the lack of fresh water discharge data, the model E-P flux is adjusted such that the model sea surface salinity is consistent with the monthly Levitus climatology (Levitus et al., 1994).

The forward model was run for 16 years from 1 January, 1990 through 31 December, 2006. This model run serves as the reference for the mean and variance of the model state. A North Atlantic ROMS simulation at 3/8° performed by Haidvogel et al. (2000) describes the model physics and performance and it serves as the basis (as well as the boundary conditions) for the coarse model. A detailed comparison of the 1/6° ROMS IAS model found that EOF modes of SSH and temperature were consistent with observations (Mancilla Rojas, 2007). Because the LC is the dominant dynamical structure in the region, we will use it as a proxy for comparing the coarse model statistics. Over the 16 years, the LC eddy shedding interval of the coarse model ranges from 2.5 to 12.9 months with an average of 8.2 months. This irregular eddy shedding period is consistent with the observations and model results presented in Table 2 of Romanou et al. (2004). The average south-westward propagation speed of the model in the Eastern gulf is 5.5 ± 1.4 km day^{-1}. This speed is fast compared to observed translation speeds (Glenn and Ebbesmeyer, 1993; Hamilton et al., 1999) and modeling studies Oey, 1996; Romanou et al., 2004. One reason for the discrepancy may be due to the coarseness of the model which results in a semi-permanent anti-cyclonic gyre in the Western gulf rather than an eddy field. Eddies are shed from the LC and propagate to the southwest where they quickly merge into the model created gyre in the west.

The LC response to the eddy shed events is to retreat before beginning another intrusion/shedding cycle. Over the 16 years of simulation, the mean of the LC maximum intrusion latitude before eddy shedding is 25°N with a maximum intrusion to 28.2°N and a minimum retreat to 23°N. These results are consistent with the observed values found by Leben (2005).

As we will show in the next section, the assimilation relies upon the model physics to match the observations; therefore, the forward model must be capable of resolving the mean dynamics. We will primarily assimilate surface values, and we compare the mean surface values with the observations over two years, 2005–2007. Fig. 2 illustrates that the mean sea surface height compares well with the observations during the same period. The root mean square difference between the two mean fields is 2.4 cm. The time variability of the model does not correspond with the observations because other than wind forcing, there is no time-dependent connection with the actual circulation. Because the regional mean and the GOM statistics of the coarse model are consistent with other studies, the assimilation procedure is used to correct the time-dependent flow to better approximate the observations.

2.1. Data assimilation

Data assimilation is the procedure for combining a model with observational data to yield an estimate of the ocean circulation that should be more useful than either alone. There are several methods available including: Kalman filters, optimal interpolation (OI), and adjoint-based inverse techniques (see Ghil and Malanotte-Rizzoli, 1991; Wunsch, 1996; Bennett, 2002; Wikle and Berliner, 2007). In data assimilation schemes, we seek to dynamically minimize the difference between the model and available observations. Time-dependent variational techniques (4DVAR) utilize the full dynamical model in contrast to OI and 3DVAR that assume persistent dynamics. We utilize the incremental strong constraint four-dimensional variational data assimilation (IS4DVAR) method proposed by Courtier et al. (1994) widely used in operational numerical weather prediction (NWP). In this section, we briefly present the ROMS implementation of IS4DVAR, which follows that of Weaver et al. (2003).

Consider the ROMS non-linear model (NLM) with ocean state vector, Ψ, comprised of the model prognostic variables at all grid points,

$$\frac{\partial \Psi_b}{\partial t} = \mathcal{N}(\Psi_b) + \mathbf{f}(t), \tag{1}$$

where the solution, $\Psi_b(t)$, corresponds to a time-dependent state determined by the initial condition, $\Psi_b(0)$, and boundary conditions, $\Psi_{b_Q}(t)$, and forcing vector, $\mathbf{f}(t)$. ROMS uses 5 standard prognostic variables $\mathcal{N}(\zeta, u, v, T, S)$, where ζ represents the dynamic height (m), u and v the zonal and meridional velocities (m s^{-1}), T the potential temperature (°C), and S salinity. In IS4DVAR, we assume that the NLM is sufficiently "close" to the observations and that only small, linear increments to $\Psi_b(0)$, $\Psi_{b_Q}(t)$, and $\mathbf{f}(t)$ are required to deviate from the non-linear trajectory to follow a new trajectory, $\Psi(t) = \Psi_b(t) + \delta\psi(t)$ that is "closer" to the observations.

(a) Model Mean SSH anomalies   (b) Observed Mean SSH anomalies
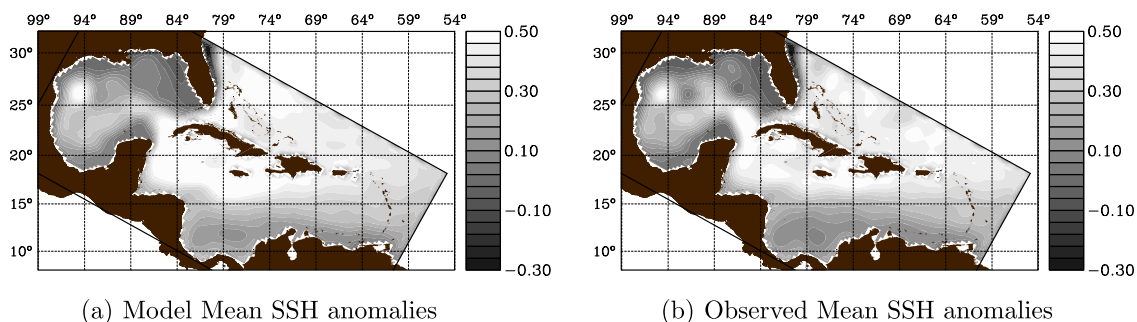
Fig. 2. Comparisons of the mean sea surface height (SSH) anomaly fields from the forward model and observations.

The increment trajectory, $\delta\psi(t)$, is found by integration of the tangent-linear (TLM) equation (Lacarra and Talagrand, 1988):

$$\frac{\partial\delta\psi}{\partial t} = \mathbf{A}(t)\delta\psi + \delta\mathbf{f}(t), \tag{2}$$

where $\mathbf{A}(t) = \partial\mathcal{N}/\partial\psi|_{\Psi_b}$ is obtained by linearizing (1) about $\Psi_b(t)$, also referred to as the background state. There exists a unique operator (referred to as the propagator), $\mathbf{R}(t_1, t_2)$, that connects solutions of (2) at time t_2 to the increment, $\delta\psi(t_1)$ (Le Dimet and Talagrand, 1986):

$$\delta\psi(t) = \mathbf{R}(0, t)\delta\psi(0) + \int_0^t \mathbf{R}(\tau, t)\delta\mathbf{f}(\tau)d\tau. \tag{3}$$

Increments to the boundary conditions, $\Psi_{b_o}(t)$, are part of the solution to (2) and are included in the propagator, $\mathbf{R}$, by definition.

Consider an innovation vector that represents the difference between the NLM and the observations: $\mathbf{d}_i = \mathbf{y}_i - \mathbf{H}_i\Psi(t_i)$, where $\mathbf{y}_i$ is the ith observation in space and time, $\mathbf{t}_i$ is the observation time, and $\mathbf{H}_i$ is an operator that samples the NLM trajectory at the observation location. Likewise, the difference between the observation and the perturbed circulation is given by $\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i$, where $\mathbf{H}'$ is the linear-form of $\mathbf{H}$. In the present study, we are using a strong-constraint and only the initial conditions $\Psi(0) = \Psi_b(0) + \delta\psi(0)$ are adjusted, while the surface forcing and boundary conditions are unperturbed over the assimilation interval, $\mathbf{t}$.

The goal is to—in a least-squares sense—perturb the circulation to minimize the difference between the observations and model circulation. We define a quadratic function:

$$J_o = \frac{1}{2}\sum_{i=0}^{N_{obs}}(\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i)^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i), \tag{4}$$

where $\mathbf{O}$ is the observation error and error of representativeness covariance matrix described in Section 3.4. There are never sufficient observations to uniquely constrain all dimensions of Ψ_0, so an additional constraint, $J_b = \frac{1}{2}\delta\psi(0)^{\mathrm{T}}\mathbf{B}^{-1}\delta\psi(0)$ is required. The increments are applied to the initial conditions, so $\mathbf{B}$ represents the background error covariances described in Section 4.1. We seek to minimize the full cost function, $J = J_b + J_o$, by equating the gradient to zero: $\nabla_{\delta\psi(t)}J = \mathbf{B}^{-1}\delta\psi(0) + \sum_{i=0}^{N_{obs}}\mathbf{H}_i'^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i) = 0$.

For any linear operator $\mathbf{A}$, there exists an adjoint, $\mathbf{A}^\dagger$ (Courant and Hilbert, 1962) such that, $\langle\psi^\dagger, \mathbf{A}\psi\rangle = \langle\mathbf{A}^\dagger\psi^\dagger, \psi\rangle$. It is easy to show that the adjoint yields $\nabla_{\delta\psi(0)}J$ when forced by $\mathbf{H}_i'^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i)$ (Talagrand and Courtier, 1987). The adjoint model (ADM) of ROMS is described in Moore et al. (2004).

Minimization of J is achieved through a Gauss–Newton iteration sequence of linear least-squares steps using an inner- and outer-loop structure (refer to Tshimanga et al. (2008) for an excellent overview). During the outer-loops (denoted by n), the NLM trajectory is updated using the increment $\delta\psi^{n-1}(0)$ from the previous inner-loop ($\delta\psi^{n-1}(0) = 0$ if $n = 1$) by $\Psi^n(0) = \Psi_b(0) + \sum_{k=1}^{n-1}\delta\psi^k(0)$. This initial condition is integrated by the NLM model to generate the trajectory $\Psi^n(t)$. The minimum J for $\Psi^n(t)$ is solved within the inner-loops.

In the first step of inner-loop m, the TLM initial condition is updated using the increment $\delta\psi_m^n(0)$ ($\delta\psi_m^n(0) = 0$ if $m = 1$) and integrated through the TLM linearized about $\Psi^n(t)$. Next, $\nabla_{\delta\psi(0)}J(0)$ is computed by the ADM forced by $\mathbf{H}_i'^{\mathrm{T}}\mathbf{O}^{-1}(\mathbf{H}_i'\delta\psi(t_i) - \mathbf{d}_i)$ and linearized about $\Psi^n(t)$. A conjugate gradient (CG) method is used to identify the minimum of J, and in ROMS is based upon the CON-GRAD algorithm developed at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Fisher, 1998). In order to achieve faster convergence, the cost-function is preconditioned by $\mathbf{B}^{-\frac{1}{2}}$ via a change of variable as detailed in Weaver et al. (2003). The inner-loops continue until a minimum of J for $\Psi^n(t)$ is found, or until the user-specified maximum for m is reached,

whichever comes first. Once the inner-loops are complete, $\delta\psi^n(0)$ is set to $\delta\psi_m^n(0)$ for the subsequent outer-loop.

The user specifies the number of outer, n, and inner, m, loops prior to execution. The reason for executing multiple outer-loops is that it is sometimes necessary to update the NLM trajectory in order to account for non-linear effects that arise in the flow during the assimilation period. It is important to note that the innovation vector, $\mathbf{d}_i$, is always computed relative to the initial background trajectory no matter the outer-loop number. In Section 4.4 we investigate the choices of n and m for our domain.

3. Observational data

A practical objective of this study was to utilize quality controlled data that would be available in the IAS in real-time for forecast purposes. For this reason, we are limited primarily to remotely sensed surface observations including sea surface temperature (SST) and sea surface height (SSH). In addition, acoustic doppler current profiler (ADCP) data collected by the RCCL *Explorer of the Seas* was also used. For the analysis presented herein, we limited ourselves to these three sources of data.

3.1. Satellite sea surface temperature

Attempts to routinely provide complete fields of SST adequate to resolve important coastal ocean features are often limited by the obscuring effects of clouds on infrared measurements (Stowe et al., 1999) and the coarse resolution and lack of near shore coverage from microwave measurements (Wentz et al., 2000). Numerous schemes for producing dense sea surface temperature fields have been in development for some time, but none of these data are readily available for real time data assimilation. Accordingly, a blended sea surface temperature product was developed for assimilation into the ROMS model. This product combines both infrared and microwave-based measurements from a variety of platforms, which eliminates clouds in many cases. In general, data from four platforms were used: the Imagers aboard the NOAA Geostationary Operational Environmental Satellites (GOES); the Advanced Very-High Resolution Radiometers aboard the NOAA Polar-orbiting Operational Environmental Satellites (POES); and the Moderate resolution Imaging Spectrometer (MODIS) aboard the NASA Aqua spacecraft; the Advanced Microwave Scanning Radiometer (AMSR-E), aboard the Aqua spacecraft.

The data were combined using 5-day mean values, inversely weighted by the nominal errors associated with each instrument to create a 0.1° map. For a given pixel centered at longitude Φ_{ij} and latitude Θ_{ij} at 0.1° spatial resolution, we calculate the mean temperature, T_{ij}, according to: $T_{ij} = [\sum_n\alpha_n(\sum_k T_{ij}^n(t_k))](\sum_n\alpha_n)^{-1}$, for all k such that the time, t_k is within a given 5-day period, and for all n independent measurement devices, where α_n gives the weight applied to each measurement based on the nominal error for sensor, n (Table 1). This results in a composite, 5-day map at 0.1° resolution.

The total error for the blended product is, of course, substantially lower than the individual errors for each instrument. As

Table 1
SST products used to create blended SST for assimilation

Data set	Spacecraft/sensor	Error (°C)	Source	Reference
AVHRR	POES/AVHRR	0.5	NOAA/NESDIS	Walton et al. (1998)
MODIS	Aqua/MODIS	0.4	NASA/GSFC	Brown and Minnett (1999)
AMSR	Aqua/AMSR-E	0.8	RSS, Inc.	Wentz et al. (2003)
GOES	GOES/Imager	1.0	NOAA/NESDIS	Maturi et al. (2004)

shown in (Reynolds et al., 2002), these errors can be dramatically reduced in combinations. Following Zanifé et al. (2003), the spatial power spectrum of the blended SST product is used to estimate a noise floor of less than 0.2 °C. We prescribe a minimum error of representativeness of 0.25 °C in the IS4DVAR implementation, but this error is spatially varying as detailed in Section 3.4.

Satellite SST measurements are based upon the radiance of the upper thin layer (skin temperature) of radiative exchange between the ocean and atmosphere, and are then regressed to 1 m of depth (bulk temperature) based on *in situ* measurements. In our model, the upper layer averages over 2 m in thickness with the shallowest being 37 mm. There may be slight temperature differences between the bulk SST measurement and the center of the model upper layer. In this paper, no attempt has been made to adjust the operator $\mathbf{H}_i$ to account for the difference between the model layer and bulk temperature; although, we are working on this problem currently.

3.2. Satellite sea surface height

For the dynamic SSH measurements, we used the altimeter merged SSH anomaly produced by Ssalto/Duacs and distributed by AVISO, with support from Centre national d'études spatiales (CNES). These anomalies are relative to the CNES dynamic SSH mean; although, the geoid and steric signals remain. The spatial resolution of the data is 0.3°, and alternates every 3 and 4 days until early July, 2006 when it is available every day. The merged product uses all available altimetry data (TOPEX/Poseidon, Jason-1, Envisat, and GFO); although, TOPEX/Poseidon entered safehold in September 2005 before being shutdown in January 2006. Pascual et al. (2007) found that with two satellites, the average error was just over 2.2 cm RMS; however, the actual value was found to be larger than the mapping error alone (Le Traon et al., 2001). The minimum observational error of representativeness assigned to the AVISO product for IS4DVAR was 4 cm.

In order the assimilate the SSH into ROMS, the anomalies must be converted into total dynamic height consistent with the model. First, the steric signal was removed from the AVISO anomalies using the Willis et al. (2004) database. Second, a mean dynamic height from the model was generated from the 16 year, forward model run from 1990 through 2006 subject only to NCEP surface forcing and initialized by the North Atlantic ROMS solution. The AVISO SSH anomalies with steric signal removed, were added to the mean model height, and these values are assimilated.

When using a higher resolution model, one would ideally use the along-track altimetry data to both increase the resolution and decrease the error associated with the measurement (Leben and Powell, 2003, see). In our case, with a 1/3° model, it is preferable to use the gridded multi-satellite product because it provides

a "smooth" field in time and space that more closely resembles the physics captured by the coarse model.

3.3. In situ data

The subsurface data used in this study were collected aboard the RCCL *Explorer of the Seas*, a ship of opportunity operating through a novel collaborative effort between RCCL, the Rosenstiel School of Marine and Atmospheric Science (RSMAS) at the University of Miami, and NOAA's Atlantic Oceanographic and Meteorological Laboratory. Between May 2001 and mid-2006, the *Explorer of the Seas* operated in the IAS domain along two primary cruise legs: a "western" cruise track and an "eastern" cruise track as illustrated in Fig. 1.

The ADCP units aboard the ship are Teledyne RD Instruments "Ocean Surveyors." There are two instruments operating at 38 kHz and 150 kHz that penetrate up to 1300 m and 300 m of water depth, respectively. All data used within this study are derived solely from the 38 kHz instrument. The heading information was provided by a Sperry NAVIGAT 2100 fiber-optic gyrocompass. The ship operates at an average speed of over 20 knots, and provides a noisy (bubble contaminated) and difficult (high-speed with shallow port calls) platform from which to operate an ADCP. The observation errors in the 38 kHz time series are upwards of 10 cm s^{-1} (Beal et al., 2008). Despite heading errors and bubble contamination, Beal et al. (2008) have shown this data to be consistent with previous Florida Current submarine cable measurements of transport. We elect to use the maximum error, 10 cm s^{-1} as the minimum error of representativeness for the IS4DVAR implementation.

3.4. Processing of observations

Over the range of latitudes in the domain, the ADCP and SST observational data sets are at a higher spatial resolutions than the model, so the observations are averaged into bins corresponding to the grid cells of the model and assimilated as "super" observations. This has two advantages: First, the variance of the observations relative to the super observation in each model grid cell provides an estimate of the error of representativeness for the observation. This variance is used in $\mathbf{O}$ (4) unless specified minimum error for a particular observation is higher. Second, averaging removes small-scale features captured in the observations but unresolved by the model. Tidal signals are small and inertial signals are not removed because the ADCP is sampled from a relatively fast moving platform.

A monthly average error of representativeness was computed for each instrument from January, 2005 through March, 2007 at each grid cell, and applied during each assimilation cycle. Fig. 3a shows the mean error of representativeness normalized by the an-
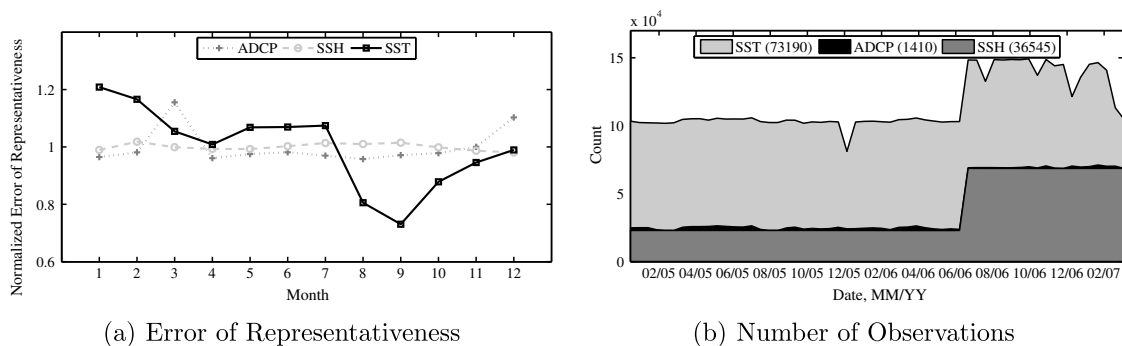

(a) Error of Representativeness

(b) Number of Observations

Fig. 3. (a) Monthly mean variance for each observed variable normalized by the annual mean value. (b) Total number of observations by type for each assimilation cycle.

nual mean associated with each observation type for each month. The error in SST exhibits a pronounced seasonal cycle due to mixed layer shoaling during the summer when the surface expression becomes homogenous revealing little variance in the observed SST. The number of available data used for each assimilation cycle is shown in Fig. 3b, showing the dramatic rise in SSH observations when AVISO began providing daily coverage.

4. Data assimilation setup

There are a number of steps required when implementing IS4-DAR in ROMS. Because the system was to be deployed on a stand-alone computer onboard a ship, a trade-off was necessary between model resolution and the CPU time required to complete an assimilation and ensemble forecasting cycle. In numerical weather prediction, the usual practice is to assimilate in the inner loops with reduced model resolution and/or physics (Trémolet, 2004). Outer-loop calculations and forecasts are then made with the full model, which serves to make the assimilation feasible by reducing CPU time and increasing the time-scale over which the tangent-linear assumption is valid. Following this approach, we assimilate at $1/3°$. In future applications, the forecasting step and outer-loop will be run at the $1/6°$ horizontal resolutions using increments from the lower-resolution model. At $1/3°$ resolution, the model does not capture the sub-mesoscale eddy field, and as the first-baroclinic mode of the Rossby radius is as low as 20 km in this region (Chelton et al., 1998), we cannot fully resolve the ocean mesoscale. However, we are capable of capturing the large LC and energetic Caribbean eddies and major ocean current systems that dominate the large-scale circulation. In the following, we focus upon the validity and results of the $1/3°$ assimilation model.

4.1. Background error covariance matrix

The background error covariance matrix, **B**, was introduced in Section 2.1 as a regularization term that acts to constrain the resulting circulation estimate by penalizing departures of $\Psi(0)$ from $\Psi_b(0)$. There are several issues with determining **B** including its structure and size, $O(10^7) \times O(10^7)$ in the current study. Estimating **B** is possible through factorization (Derber and Rosati, 1989; Weaver and Courtier, 2001), such that:

$$\mathbf{B} = \mathbf{K}_b \Sigma \Lambda L_v^{\frac{1}{2}} L_h L_v^{\frac{1}{2}} \Lambda \Sigma \mathbf{K}_b^{\mathrm{T}}. \tag{5}$$

The terms $\mathbf{K}_b$ are the covariance operators of the balanced dynamics, Σ and Λ are diagonal matrices of the background-error standard deviations and normalization factors, respectively, and L_v and L_h represent univariate (each prognostic state variable is unconstrained by the others) correlation matrices in the vertical and horizontal directions, respectively. Multivariate covariance relations are prescribed via $\mathbf{K}_b$; however, this feature is not currently implemented in ROMS and only the univariate covariance is considered. It should be noted that the TLM couples the dynamics of different increment variables; however, we are only considering the univariate statistical coupling in the statistics of **B**.

The background error standard deviations, Σ, were initially calculated from the 16 year forward run of the IAS forward model described in Section 3.2. When the first assimilation cycle begins, $\Psi_b(0)$ will be the result of the long model integration with surface forcing only. The uncertainty in $\Psi_b(0)$ will be similar to the difference between randomly chosen model states (i.e., the model climatological variance). As the assimilation cycles proceed, the efficacy of $\Psi_b(0)$ will improve so it is necessary to re-evaluate **B** as is addressed in Section 5.2; however, continually reducing **B** must be taken with care as it may be reduced to the point that the

assimilation procedure believes there are no errors in the model. The correlation matrices, L_v and L_h, with their associated normalization factors, Λ, were computed as solutions of diffusion equations following Derber and Rosati (1989) and Weaver and Courtier (2001). The length-scales chosen for L_v and L_h represent the decorrelation scale from a typical increment, and are currently assumed to be homogenous and isotropic. In the vertical, L_v is set to 100 m corresponding to a depth greater than the average mixed layer in the GOM (Müller-Karger et al., 1991). Utilizing a semi-variogram technique (Milliff et al., 2003), representative horizontal length-scales of dominant circulation features in the GOM and Caribbean are found to be approximately 100 km and 350 km, respectively.

Since J_b acts as a constraint on J, the effect of **B** is to spatially smooth the analysis increments $\delta\psi(0)$. This is a desirable feature when the observations are sparse, since it prevents the tendency for large, localized, and non-physically realizable increments in $\delta\psi(0)$. When the observations are closely spaced, as for satellite data, the smoothing influence of **B** results in a loss of observed information at scales less than the assumed decorrelation lengths. Because the vast majority of the observed data used is surface data, we wish use a minimal smoothing distance (thus, allowing the observations more weight) in the horizontal, but tightly constrain in the vertical. Currently the ROMS implementation of IS4DVAR does not allow spatially varying length scales within the domain, which requires that we use the shortest throughout the domain. Experiments for the IAS reveal that a horizontal decorrelation length scale of 80 km is an optimal compromise between smoothing and analysis error. This is slightly smaller than suggested by the semi-variogram analysis of the near surface circulation; however, it has the advantage of being slightly larger than two adjacent grid cells.

4.2. Validity of the tangent linear assumption

For 4DVAR methods, the assimilation time interval should not exceed the time for which the tangent linear approximation of (2) is valid. Periods longer than 14 days are deemed too long as compared against the e-folding time for the IAS region. We performed a 14 day assimilation for the period 7/20/2005–8/3/2005. This was a quiet period between hurricane events with a shallow mixed layer. During the assimilation of our test period, each ith pass through the inner-loop yields a different increment for the initial condition, $\delta\psi_i(0)$. Each sign, p (positive and negative), of these perturbations was integrated using the TLM (2) linearized about $\Psi_b(t)$ to create $\delta\psi_i^p(t)$. Likewise, a NLM trajectory is created for each as $\Psi_i^p(t)$ using perturbed initial conditions $\Psi_b(0) + \delta\psi_i^p(0)$. To provide a test of validity, the difference $\Delta\Psi_i^p(t) = \Psi_i^p(t) - \Psi_b(t)$ is compared to the corresponding TLM solutions, $\delta\psi_m(t)$; $m = [1, 50]$. This creates an ensemble of trajectories to compare the integration of the perturbations through both the NLM and TLM models.

Fig. 4 shows the ensemble mean of the perturbations at the near surface for temperature and SSH at the initial time. The largest perturbations in temperature and velocity (not shown) occur in the western GOM and the upwelling region along the coasts of Venezuela and Colombia. The Western GOM is a highly variable region into which LC eddies propagate, and the upwelling region varies in response to short-timescale variations in atmospheric forcing. The variance is also quite large in the region where the Gulf Stream exits the domain; however, this lies outside of the IAS region of interest and will not be considered. In SSH, the greatest perturbations and variance are in the LC and its northern extent where eddies are shed.

The perturbation growth over 14 days for SSH and temperature over the upper 100 m are compared in Fig. 5 in terms of normalized
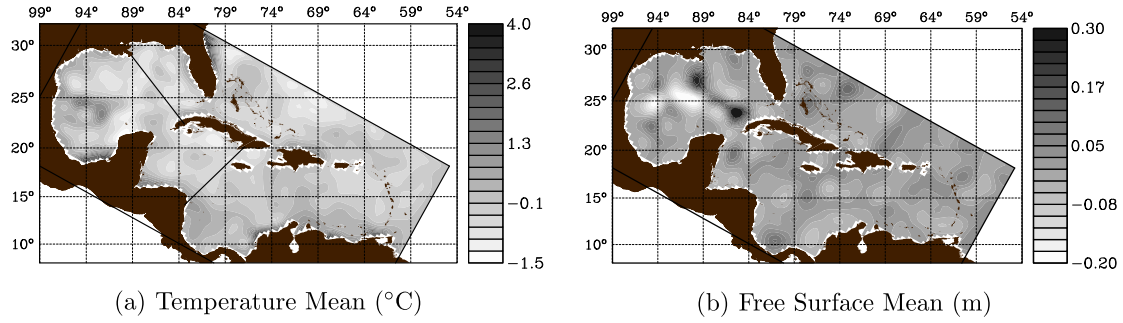
(a) Temperature Mean (°C)   (b) Free Surface Mean (m)

Fig. 4. Ensemble average of the perturbations added to the NLM and TLM initial conditions.

(a) Temperature   (b) Temperature
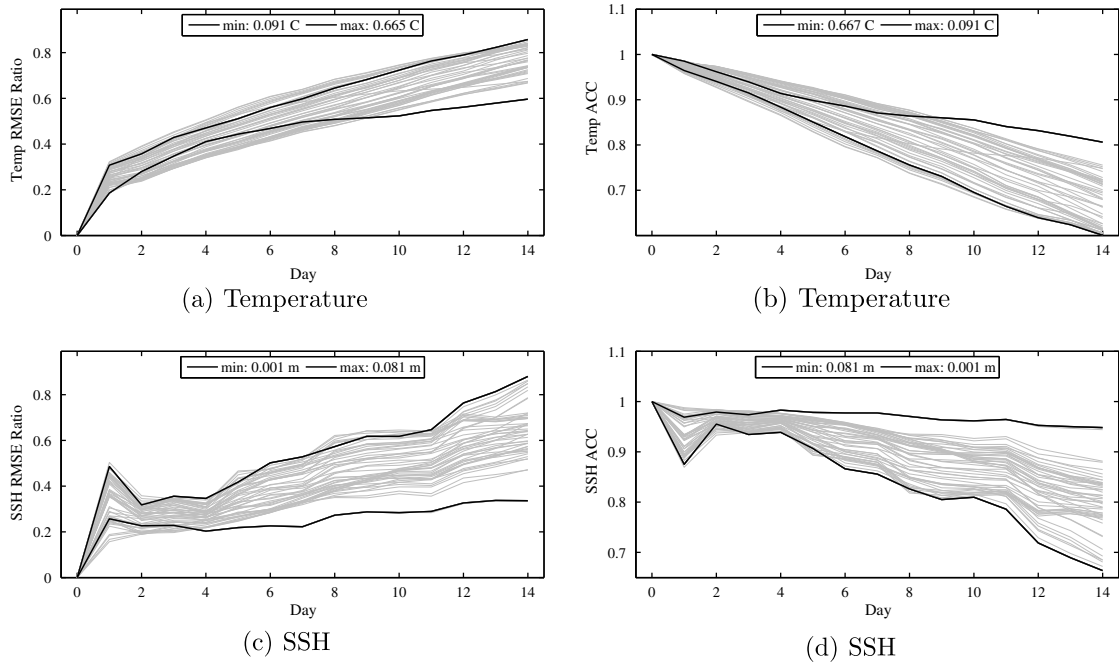
(c) SSH   (d) SSH

Fig. 5. Time-series of the normalized RMSE between the NLM and TLM solutions and their associated ACC for the upper 100 m of temperature and SSH. The values of the extreme initial perturbations are highlighted with their initial RMS values listed.

root mean square error (RMSE) and the anomaly correlation coefficient (ACC) for spatial correlation comparison. The normalized RMSE is given by $\mathrm{rms}(\Delta\Psi_m(t) - \delta\psi_m(t))/\mathrm{rms}(\Delta\Psi_m(t))$ for $m = [1, 50]$. The ACC for two vectors is defined as:

$$\mathrm{ACC}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})\left[\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2\right]^{-\frac{1}{2}}. \qquad (6)$$

Because the typical TLM solutions contain small-scale linear instabilities, the TLM fields, $\delta\psi$, were spatially smoothed using a second-order Shapiro filter over two grid cells to isolate longer length-scales of the circulation of interest (e.g., mesoscale and larger). This corresponds to the horizontal length-scale that was chosen in Section 4.1 when assimilating data. Each panel in Fig. 5 shows the minimum and maximum RMSE and ACC values for the set of perturbations used. Although the TLM assumption may begin to break down at short-scales for longer periods, we are interested in capturing signals of 100 km and greater (as identified by the semi-variogram analysis). This approach has been used successfully before by Weaver et al. (2003) with a 30-day interval in the tropical Pacific providing a long, continuous circulation for analysis. The

tangent-linear assumption was valid for the tropical instability waves they were interested in assimilating, but invalid for short periods and small spatial scales.

Fig. 5 shows that the envelopes bounding RMSE in SSH, and the ocean temperature in the upper 100 m, grow with time. The envelope bounding SSH RMSE ratio is 0.3–0.8 by day 14. The envelope bounding the upper 100 m of temperature is 0.5–0.9 and for the upper 100 m of velocity is 0.6–1.0 (not shown). Similarly, ACC envelopes spread with time as the ACC mean values decrease (Figs. 5b and d). The ACC envelope at day 14 in SSH is 0.6–0.95, in the upper 100 m of temperature is 0.6–0.85, and in the upper 100 m of velocity is 0.4–0.8 (not shown). Overall, Fig. 5 indicates that while RMSE of the TLM perturbations can exceed 50%, the pattern correlation is maintained over the typical 14-day assimilation interval; therefore, the amplitudes may diverge, but the features remain. Calculations of the unsmoothed TLM fields yield similar conditions with a slightly faster growth in RMSE.

Fig. 6 shows an example of the evolution of temperature perturbations over the full water column along the GOM and Caribbean transects indicated in Fig. 4a. Evolution of the large perturbations related to the LC show some substantial differences between the
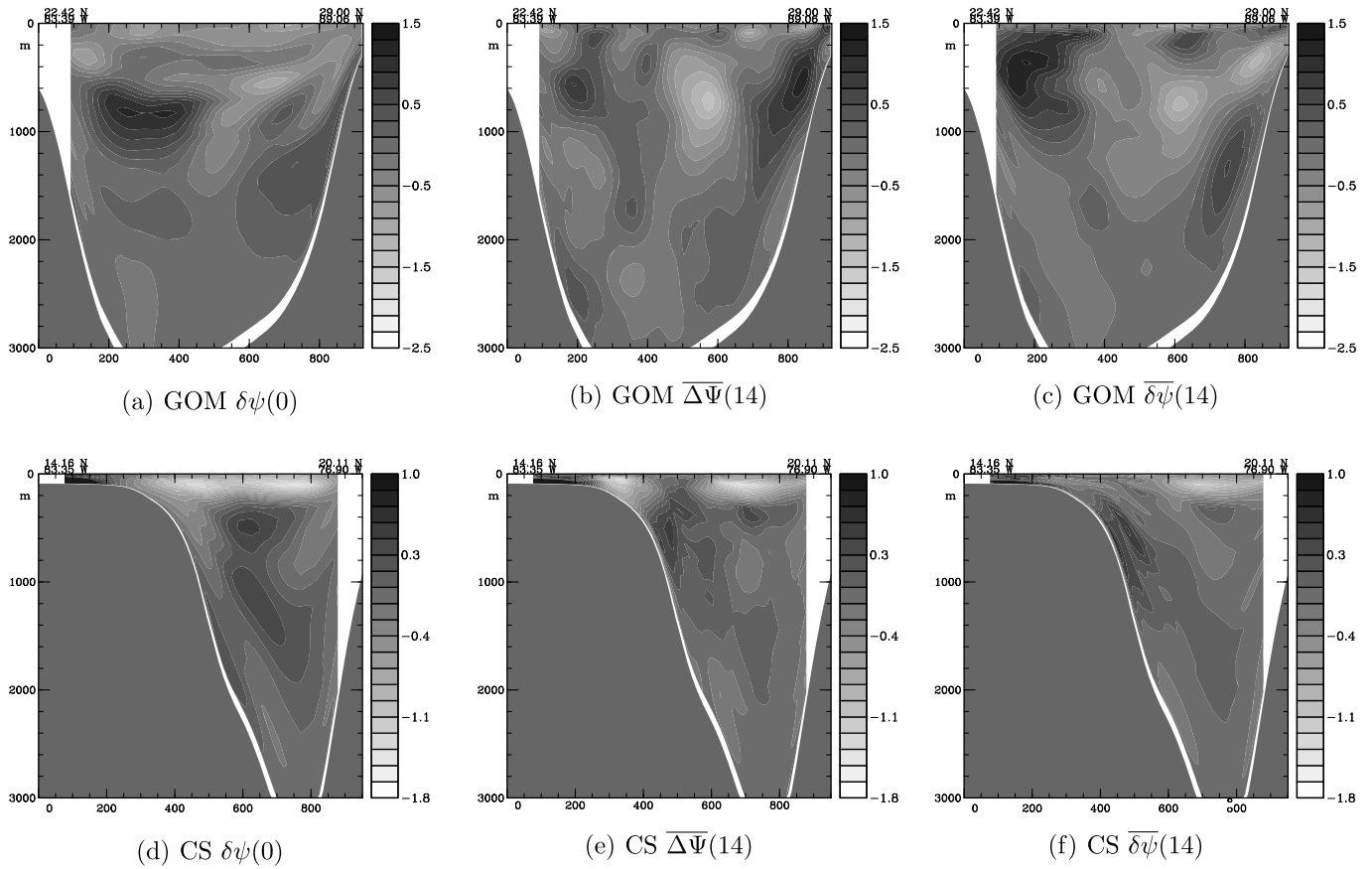
Fig. 6. Cross sections showing the 14-day evolution of temperature perturbations (°C) along the GOM and CS transects shown in Fig. 4a. The GOM scale varies from 1.5 °C (black) to −2.5 °C (white) and the CS scale from 1.0 °C to −1.8 °C.

TLM and NLM by day 14. We stress that the GOM is a worst-case scenario and the CS case (among others not shown) shows a stronger correlation with smaller differences. The statistical comparisons of Fig. 5 give us reason to be comfortable using a 14-day assimilation in the 1/3° model; although, this probably represents the upper bound of the assimilation interval for this configuration.

4.3. Assimilation interval

In this section, we explore the impact of the assimilation interval on the model solutions. One 14-day, two 7-day, and four 3-day plus a single 2-day assimilations are performed over the test period of the tangent-linear verification. Each case is run with a single outer loop and 25 inner loops, with 100 m and 80 km vertical and horizontal length scales, and assimilation of all available data during each period. Fig. 7 shows the RMSE between the model and observations on each day. The 14-day interval exhibits the largest RMSE; however, the 7-day interval averages only a 13% improvement in SST and 20% in SSH. The 3-day interval SST RMSE values are 56% lower on average than the 14-day interval.

We plan to use the IS4DVAR circulation estimates in sequence to explore the dynamics of the IAS circulation; therefore, the optimal assimilation period is one that is dynamically adjusted to the observations over the longest possible continuous interval that does not violate the tangent-linear assumption. This is the key difference between 3D- and 4DVAR techniques: assimilating as many of the observations in one minimization rather than creating discontinuities at every observation time. The IS4DVAR method, by virtue of altering the initial conditions for the assimilation period, creates discontinuities between cycles. A 3-day assimilation inter-
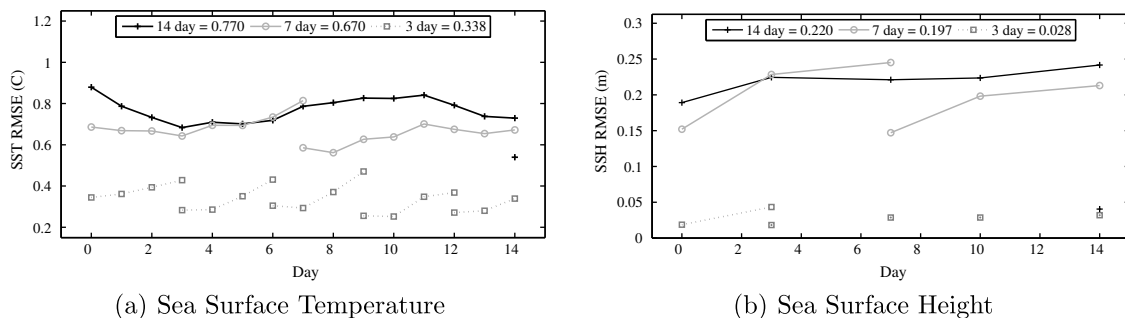

Fig. 7. Comparison of RMSE between model and observations for three separate assimilation intervals: 14-day, 7-day, and 3-day. The average RMSE for the entire time-period is shown in the legend.
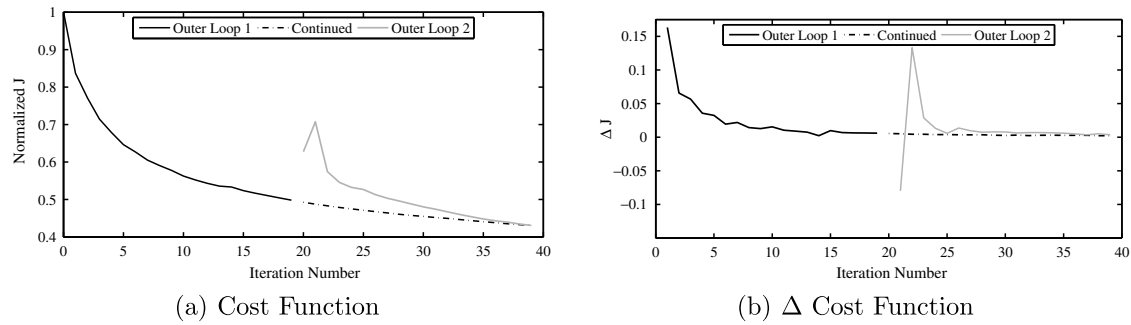
(a) Cost Function

(b) Δ Cost Function

Fig. 8. (a) The cost function, *J*, normalized by the its initial value (i.e. before assimilation) versus the number of inner-loops. The single outer-loop case continues its descent, while the updated outer-loop resets before achieving the same final value. (b) The change in cost function between consecutive loops.

val is clearly not desirable based upon this criteria. A 7-day assimilation interval would seem a reasonable choice; although, even a single interval yields large discontinuities between assimilation cycles, and performs only slightly better than a single 14-day cycle. The 14-day assimilation contains discontinuities (the first point of the subsequent 14-day assimilation is shown in Fig. 7) that are of similar magnitude to the 7-day. The 7-day interval has a lower average RMSE, but the benefits of a single 14-day assimilation for both the analysis and reduced computational requirements for the operational system make it the better choice.

4.4. Outer- and inner-loop configuration

The final choice to be made for the data assimilation system is the number of outer- and inner-loops. As mentioned in Section 2.1, minimization of *J* is achieved via a sequence of linear least-squares minimizations. Because the initial background state is typically close to the observations after several assimilation cycles, one or two outer-loops are typically sufficient. A choice must be made as to the maximum number of iterations that will be performed to obtain the estimate for the cost function minimum. The quadratic cost function will reach its theoretical minimum after N_{obs} iterations; however, this is not computationally feasible for real applications.

Using the test period with L_h = 80 km and 14-day assimilation period, two assimilation cycles were performed using a total of 40 loops: one with 1 outer-loop (40 inner-loops) and a second with 2 outer-loops (20 inner-loops). Fig. 8a shows *J* for each experiment and indicates that the final outcome is nearly identical for both cases. Fig. 8b shows the change in *J* between consecutive loops and indicates that after loop 20, the gradient of the descent algorithm is negligible in both cases and that updating the outer-loop causes a temporary increase in *J*. This "jump" of 21% in the cost function reveals a degree of non-linearity in the circulation, and although the final minimization value is similar between the two, we choose to update the non-linear circulation by using 2 outer-loops and 12 inner-loops, where the total 24 is chosen as a point at which the minimization has steadily slowed.

Inspecting the structure of the cost function, we find that SST dominates the cost function due to the higher number of SST observations. Similarly, the gradient norm of the cost function is dominated by temperature, which accounts for 93.5% of the conjugate gradient step. It is apparent that temperature is the driving factor in the assimilation, and it is the dominant variable we shall analyze in detail.

5. Results

There have been several studies assimilating altimetry and/or temperature data into primitive equation ocean models using var-

ious methods including: Kalman filtering (Fukumori et al., 1999), simple nudging with statistical vertical projection (Oey et al., 2005a), optimal interpolation schemes (Mellor and Ezer, 1991), 3DVAR (Ezer and Mellor, 1994), and 4DVAR (Weaver et al., 2003; Di Lorenzo et al., 2007; Chua and Bennett, 2001) including the Estimating the Circulation and Climate of the Ocean (ECCO) group (Stammer et al., 2002; Köhl et al., 2007). In the region of interest, there have been a variety of assimilation studies. The nudging method has been the most widely used: Oey et al. (2005a), Fan et al. (2004), Kantha et al. (2005) and Ko et al. (2007). Falkovich et al. (2005) describes a similar method adjusting modeled features to initialize a hurricane model. Lin et al. (2007) presents a combination of optimal interpolation and nudging techniques to assimilate a variety of data in the GOM. Ezer and Mellor (1994) used a 3DVAR technique to assimilate Geosat altimetry into the Princeton Ocean Model. The HYbrid Coordinate Ocean Model (HYCOM) uses optimal interpolation to assimilate a variety of data in the region (Chassignet et al., 2007). The representer method was tested in the GOM by Ngodock et al. (2007). For a more complete look at data assimilation in the Gulf of Mexico, we encourage the reader to review Chassignet et al. (2005). As variational schemes that rely upon the model dynamics, the IS4DVAR method (presented in this paper), Representer Method (Bennett, 2002; Di Lorenzo et al., 2007), and Ensemble Kalman Filtering (Evensen, 2003) are fundamentally different than either the nudging or statistical (optimal interpolation or 3DVAR) schemes.

A series of assimilation experiments summarized in Table 2 were performed to assess various aspects of the assimilation system. Based on the findings of Section 4, the following parameters were used unless noted otherwise: 14-day assimilation interval, 80 km horizontal and 100 m vertical length scales for the diffusion model of **B**, and 2 outer with 12 inner loops. For the remainder of this discussion, RMSE and ACC will be used as measures of agreement between the model and observations. Various sub-regions of the model domain are examined as shown in Fig. 9. We also examine the temporal correlation (TC) computed by (6) using a time-series of model gridpoint values and observations at the same location rather than a spatial comparison.

Table 2
Overview of experiments

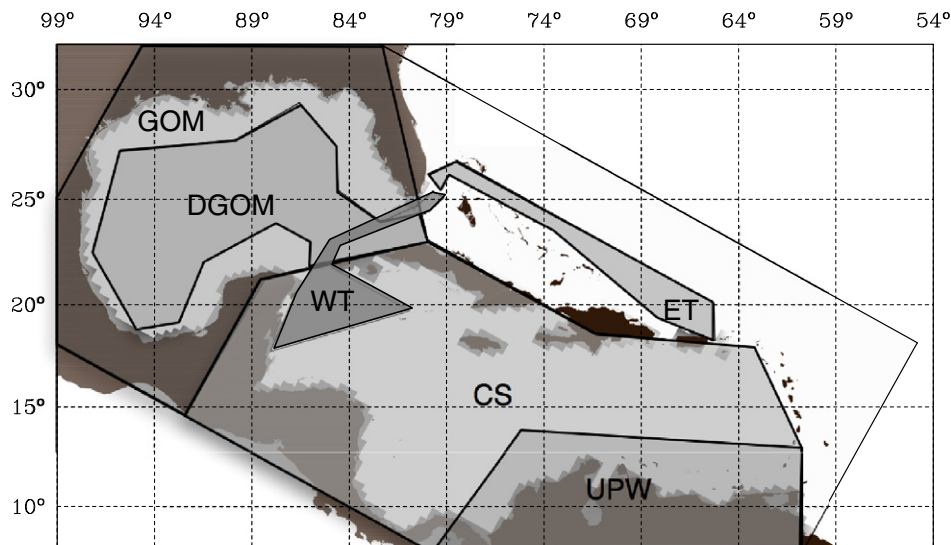| Name	Description	Time period
ExN	Forward Model with no assimilation	5 January 2005–28 March 2007
ExA	Assimilation of SSH, SST, and ADCP	5 January 2005–28 March 2007
ExANmc	Repeat of ExA using new background standard deviation estimate	11 May 2005–28 March 2007
ExAS	Repeat of ExA using only satellite data (no ADCP)	5 January 2005–4 January 2006

Fig. 9. Regional map showing sub-regions of interest: GOM—Gulf of Mexico (including DGOM); DGOM—Deep Gulf of Mexico; CS—Caribbean Sea (including UPW); UPW—upwelling region of South American north coast; IAS—composite region of GOM and CS; and, WT, ET—western and eastern cruise-tracks of RCCL *Explorer of the Seas*.

First, we will compare the effect of the minimization procedure on the cost function. We then look at how updating the uncertainty in the background estimate improves the minimization. This is followed by a comparison of the assimilated circulation with the original observations. Lastly, we look at changes made to the circulation by the assimilation procedure by comparing against other measures of the region.

5.1. Benchmark assimilation experiment

Our benchmark assimilation experiment (ExA) covers 1/5/2005–3/28/2007 providing two years of statistics before the shipboard program began on 1/6/2007. All experiments (except ExANmc) begin with initial conditions taken from the 16 year forward run of the IAS forward model described in Section 3.2 on 1/5/2005. Fig. 10 compares the time-series of 14-day observational cost function, J_o, between ExA and ExN. In the case of ExN, no increments are made and $J_b = 0$; therefore, we compare only J_o (on average, J_b is only 1.3% as large as J_o for ExA).

The minimization procedure reduces J_o by 72% in ExA. As shown in Fig. 3b, the number of SSH observations is roughly equal to the

SST observations after 7/5/2006 when daily SSH is available. Over this period, SST and SSH account for 60% and 39% of J_o with the remaining contribution from ADCP measurements. The GOM is responsible for 59% of J_o despite comprising only 43% of the IAS region. The DGOM (Fig. 9) is responsible for only 34% of the total J_o; hence, the shallow shelves in the GOM are large contributers to J_o. There are several factors that could explain this behavior: (i) the SST data has a higher error of representativeness along the coast; (ii) the sub-grid vertical mixing on the shelf is not properly capturing the highly dynamic nature present in the Louisiana-Texas and Florida shelves; or, (iii) the low horizontal resolution of the model makes it difficult to separate the coastal and deep regimes that are present in the GOM.

5.2. Updating **B**

As discussed in 4.1, **B** provides a constraint that penalizes deviations of $\delta\psi(0)$ from $\Psi_b(0)$. The statistics of **B**, specifically Σ in (5), are computed from the 16 year forward model run without assimilation. After several assimilation cycles, Ψ_b should be closer to the observations in which case, Σ needs to be updated to reflect smaller errors in $\Psi_b(0)$. In numerical weather prediction, the "NMC method" is often used (Parrish and Derber, 1992) to compute **B** by comparing model forecasts initialized at different times though valid on the same day. We use the ExA ocean state estimations from 5/25/2005 through 5/24/2006 to create a new estimate of Σ_{nmc}. The final day of each assimilation cycle is used as the initial state for a 7-day hindcast and the RMS difference between the hindcast and the first 7 days of the subsequent assimilation cycle gives us an estimate of Σ_{nmc}. The full assimilation is repeated for the period from 5/25/2005 through the 3/28/2007 with Σ_{nmc} to create ExANmc.

As **B** is reduced (due to the decrease in Σ), the minimization procedure will have less freedom and must focus on reducing the error in J_o. with the new background estimate, the total cost function, J, is reduced by 11.8%; however J_b only accounts for 3.3% of J for ExANmc. Σ_{nmc} has decreased the error in surface temperature, and SST (SSH) now accounts for 35% (64%) of J.

Following Bennett (2002) and Weaver et al. (2003), the theoretical minimum value of J is $N_{obs}/2$, but is achieved only if the prior estimates of the covariances **B** and **O** are correct. If the errors are
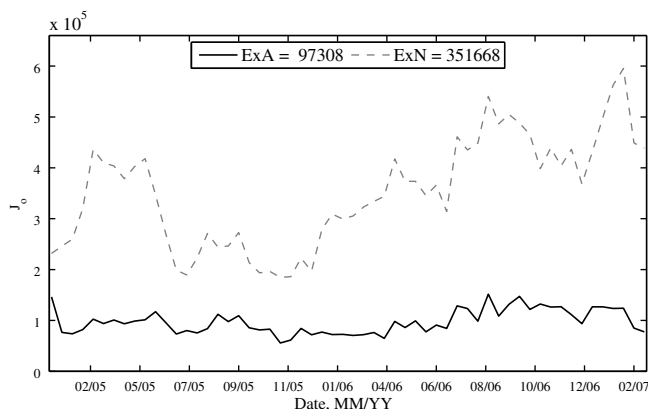

Fig. 10. Time series of the total observational cost function, J_o, for the IAS from ExA and ExN. The assimilation reduces the NLM error by an average of 72% for all observations.
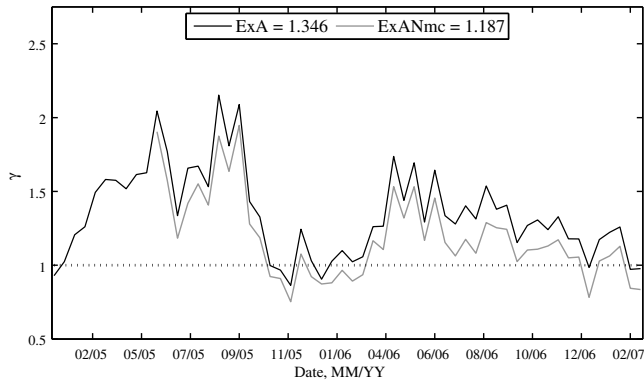
Fig. 11. Time-series of measure of fit, γ, for ExA and ExANmc. The dashed line represents $1 \pm \sqrt{2/N}$, the uncertainty in the ideal fit.

Gaussian, then J is a χ^2 variable with variance equal to $1/N$. For a linear system and correct choices of **B** and **O**, we should expect that $\gamma = 2J/N_{obs}$ to reach a value of $1 \pm \sqrt{2/N}$, where γ is the "goodness of fit," and $\gamma > 1$ ($\gamma < 1$) represents an under(over)-estimate of the

error covariances. One must be careful in this estimation because both **B** and **O** could be chosen in a closed-form solution to minimize γ; however, the goal is not to minimize γ, but to use it as a measure of changes to **B**.

Fig. 11 shows γ for each assimilation cycle during ExA and ExANmc. Starting in May 2005 with Σ_{nmc}, γ decreases by an average of 12%. The optimal best fit to our data is $\gamma = 1.0 \pm 0.005$. The average value for ExA is $\gamma = 1.346 \pm 0.30$ and ExANmc $\gamma = 1.187 \pm 0.29$, which represents an underestimate of the errors in both cases. The parameter γ provides a useful measure of how well the error estimates allow us to approach the optimal fit. By this measure, we have improved the procedure by updating **B**. It should be noted that continual refinements to **B** could be made; however, one must exercise caution as **B** becomes too low, the minimization procedure will be penalized for deviating from the background as the model will be seen as "perfect".

5.3. Fit with observations

Having presented measures (J_o and γ) of IS4DVAR minimization performance, we turn our attention to more intuitive measures of skill. Fig. 12 shows the RMSE, model bias, and ACC between the
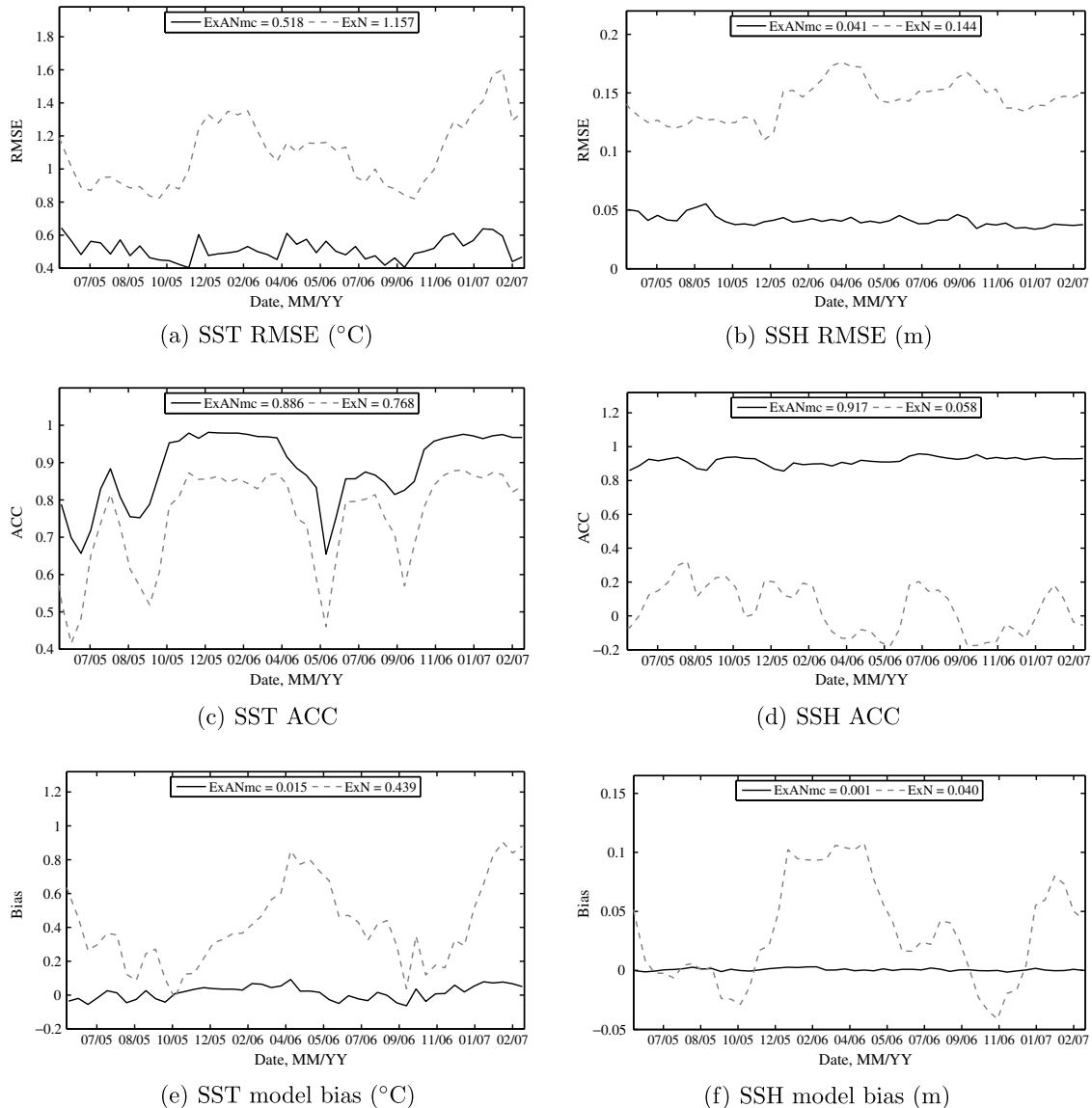

(a) SST RMSE (°C)

(b) SSH RMSE (m)

(c) SST ACC

(d) SSH ACC

(e) SST model bias (°C)

(f) SSH model bias (m)

Fig. 12. Comparisons of SST and SSH 14-day average RMSE, ACC, and model bias for ExANmc, and ExN in the IAS region.

model and observations for ExANmc and ExN computed over each 14-day assimilation cycle. The RMSE of ExANmc over the IAS in SST (SSH) is reduced 55% (72%) as compared to ExN, and averages 0.52 °C (4.1 cm) for the entire period. The SST RMSE trends downward during Spring and Summer then rises again in the Fall, and the ACC is opposite in phase. The improvements in the GOM were similar to the regional with an average RMSE of 0.65 °C (5.4 cm) in SST (SSH) an improvement of 57.5% (68%) over ExN. Likewise, the Caribbean RMSE averages 0.41 °C (3 cm).

There is a seasonal cycle in SST that negatively impacts the ACC during the summer months. As noted in Fig. 3a, the temperature variance of the observations is lower during the summer months when the mixed layer shoals from a winter minimum of 110 m to a summer maximum of 12 m. This compares with an observed winter minimum of 125 m and summer maximum of less than 20 m presented by Müller-Karger et al. (1991). Interestingly, the mixed layer depth is 25% deeper for ExN compared to ExANmc. In the future, we could utilize a variety of SST products as it has been shown that multiple products can be used to elucidate SST structures during the summer period Walker et al., 2003; however, in comparison with other SST products, the RMS differences between SST products is larger than the RMSE presented. ExN contains over 0.4 °C and 4 cm bias (Figs. 12e and f) that have been reduced to nearly zero for both cases by ExANmc. The SST bias values are slightly higher in the GOM at 0.05 °C as compared to 0.015 °C over the entire region. ExN exhibits a summer cycle in bias of both SST and SSH that has been removed by the assimilation. These results suggest that the model vertical mixing does not adequately capture the seasonal changes in the region. Like in RMSE, the results for the GOM and Caribbean are consistent with the regional values. The ACC in the GOM averages 0.827 (0.892) for SST (SSH) and averages 0.769 (0.937) in the Caribbean.

To better understand the source of the SST error in the GOM and the regions of seasonal variability, we examine maps of the average RMSE and time correlation (TC) over the entire assimilation period for ExN and ExANmc in Fig. 13. Without assimilation, ExN (Fig. 13a) shows a large RMSE throughout the domain with a poor TC in the LC intrusion and UPW regions. As indicated earlier, the worst region of RMSE for ExANmc (Fig. 13c) occurs on the shelves

around the GOM as well as in the region of LC intrusion. The RMSE in the western GOM is low considering the highly variable nature of the LC eddy "graveyard". This suggests that once the model has assimilated a separated LC eddy, the eddy propagation is consistent with the observations. The UPW region also exhibits a very localized area of relatively large RMSE.

For TC, the dominant problem region for all experiments is the Nicaraguan shelf. This is most likely the result of either overly smoothed bathymetry to fit the model grid or incorrect model physics (topography or mixing parameters) for that region. The UPW region is similar as it is an area of strong upwelling near a shallow coastal shelf. During the summer season, the Caribbean and southern GOM exhibit very low correlations, while TC is higher in the northern GOM; whereas, winter TC is high most everywhere except in the LC region and western GOM. During this period, the RMS of the SST observations is largest around the northern edges of the LC.

Maps of RMSE and TC for SSH are shown in Fig. 14 and reveal that ExANmc fits the observations well except for the LC where the observational SSH variance is largest.

ADCP measurements from the ship are known to be quite noisy (Beal et al., 2008); however, of greater concern is the low horizontal resolution of the model compared with the high spatial resolution of the data. As discussed in Section 3.4, all data within a single grid box were averaged together, which should help to reduce the amplitudes of the signals associated with inertial oscillations and tides. Aliasing errors will remain due to the unresolved dynamical processes in the model. As shown in Fig. 1, most of the data collected during the western cruise track lie within the strong Florida, Yucatan, and Caribbean currents, and the eastern track captures island inflows.

The assimilation averages a 27% improvement (with an average RMSE of 14.7 cm/s) over ExN in RMSE for the ADCP measurements along both cruise tracks, and there is no improvement in the ExAS case. There is a dramatic improvement in the ACC, particularly in the western cruise track of 31% (with an average ACC of 0.81). Although the RMSE is not improved in ExAS, the ACC is by 12%. This suggests that the adjustments made to the SSH are affecting the subsurface velocity profile positively. This relation is mutual as

(a) ExN SST RMSE (°C)

(b) ExN SST TC

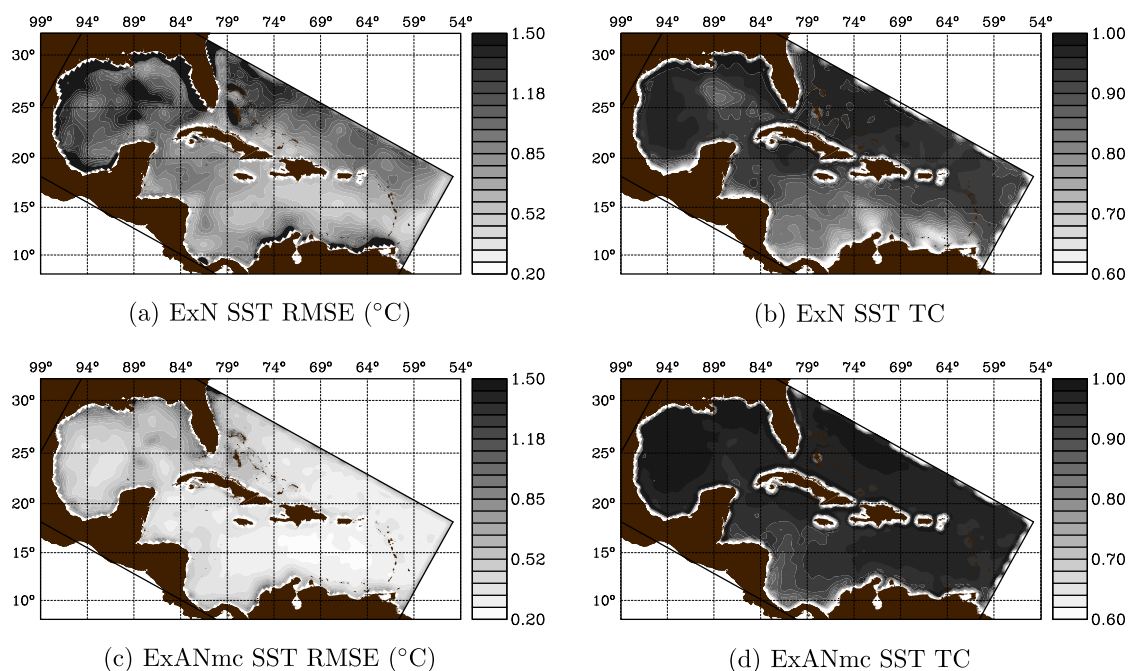(c) ExANmc SST RMSE (°C)

(d) ExANmc SST TC

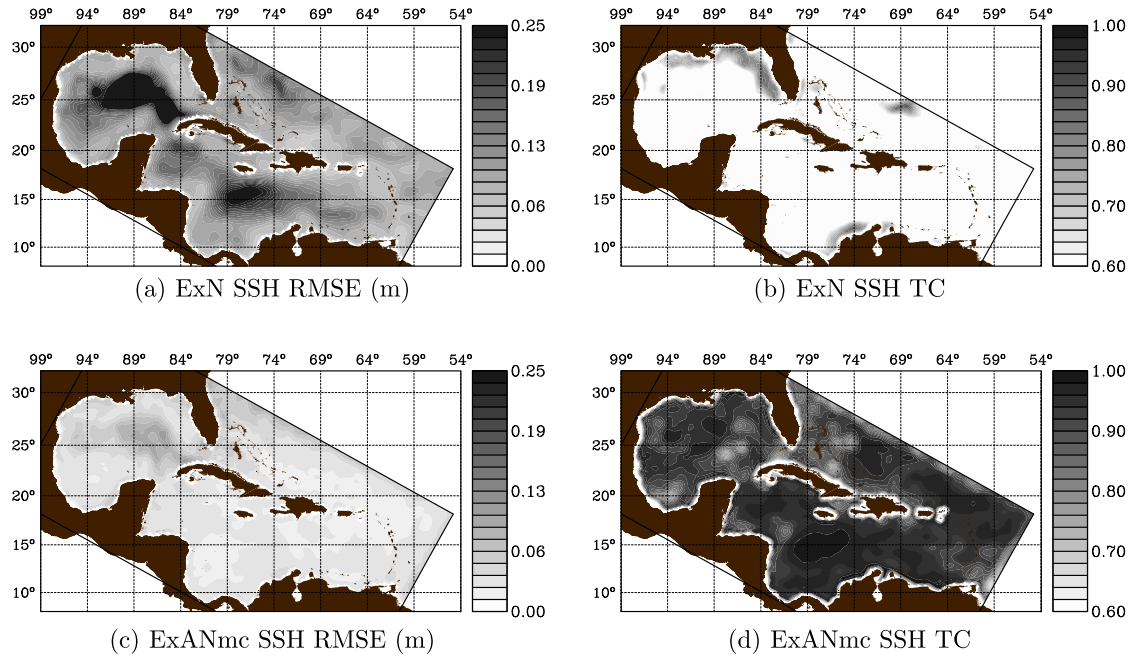Fig. 13. Spatial maps of SST RMSE and TC for ExN and ExANmc.

Fig. 14. Spatial maps of SSH RMSE and TC for ExN and ExANmc.

assimilating the ADCP (ExA) reduces the SSH RMSE by 7% as compared to the ExAS case. In similar comparisons, the tracer adjustments for SST assimilation has little apparent impact on the RMSE of the velocity measures and vice versa.

5.4. Loop Current measures

We seek to compare how well the assimilation model performs against measures of the dominant circulation features in the IAS. Leben (2005) has developed a set of indices for the LC that track a 17 cm isoline of SSH derived from adding observed SSH anomalies to a mean derived from a model. Because our model has a different mean dynamic height, the isoline that most closely matches the maximum gradient on the edge of LC eddies is 32 cm. Following the procedure of Leben (2005), we calculate the maximum western longitude and northern latitude, length, total area, and circulation of the identifiable LC. One would expect that after assimilation of the SSH data, the LC statistics would be comparable; however, both SST and ADCP are also assimilated and in adjoint-based methods, all observations are assimilated using the full physics of the model. This requires that a dynamic consistency be found between all measures. To match the LC, the full 4D circulation is adjusted and using the altimetry as a proxy may yield a measure of the circulation.

The results for each experiment are summarized in Table 3. We use the AVISO anomalies interpolated to the model grid as

a benchmark for comparison. Not surprisingly, the agreement between the model and AVISO is least for ExN. The length (and area) is increased by the assimilation cases in comparison to ExN. The SSH RMSE in the DGOM of ExANmc is 6 cm, and as shown in Fig. 14d, the region of greatest uncertainty is in the LC, which can affect the contour algorithm for determining the isoline of the LC. Even with such uncertainty, the assimilation was able to capture the larger circulation including the objective dynamics of the LC.

An important issue in the formation of LC eddies is the role of upstream influences (Oey et al. (2005a) presents an excellent review of the circulation of the GOM). The transport through the Yucatan Channel has a direct impact upon the LC (Murphy et al., 1999; Oey et al., 2003). Computing the mean transport over 2005–2007 from ExN and ExANmc we find values of 26.0 ± 4.4 Sv and 24.5 ± 7.4 Sv, respectively. From September, 1999 through May, 2001, the Mexican CANEK program (Sheinbaum et al., 2002; Ochoa et al., 2003) measured a mean transport of 22.8 ± 3.3 Sv. Although observed over an earlier period and potentially in a different phase of the transport, the model results are consistent with the observed values, but slightly lower than the models results of Romanou et al. (2004).

6. Conclusions

We have presented a study of the efficacy of 4DVAR data assimilation within the IAS using predominantly surface observations at a relatively coarse resolution. The limited resolution and surface observations used here were constraints imposed on the system so that it could be deployed aboard the RCCL *Explorer of the Seas* for sea trials beginning in January 2007.

Other data assimilation studies have been carried out in the GOM and/or IAS region with a detailed review in Chassignet et al. (2005). Some employ "nudging" methods (e.g., Oey et al., 2005a; Fan et al., 2004; Kantha et al., 2005) that relax a particular field to the observation over a specified relaxation interval. These methods do not adjust the model in a dynamically consistent way, but rather add an additional term to the right-hand side of the model equations to relax the value toward to the

Table 3
Summary statistics of the Loop Current measures

Exp.	Max. west long. (°W)	Max. north lat. (°N)	Length (km)	Area (km²)	Circulation (×10⁶ m² s⁻¹)
AVISO	88.4	26.3	1292	106,579	2.44
σ	±1.6	±1.2	±568	±49,997	±1.38
ExN	87.6	25.1	1186	82,724	2.14
σ	±1.8	±0.9	±528	±44,357	±1.02
ExA	88.7	26.4	1651	124,924	2.28
σ	±1.5	±1.2	±544	±50,678	±1.19
ExANmc	88.7	26.4	1641	126,526	2.40
σ	±1.6	±1.2	±541	±51,003	±1.17

observation. Ko et al. (2007) operates a real-time IAS assimilation model using the techniques described in his paper for the Portuguese coast. In the case of Fan et al. (2004), an optimal interpolation approach Mellor and Ezer, 1991 was used to assimilate satellite data and nudging for Lagrangian drifter data in the northeastern GOM. Oey et al. (2005a) shows the validity of using the satellite data from Fan et al. (2004) as an initialization for a predictive model.

Similar 4DVAR methods are being used by the ECCO group (see Stammer et al.,2002, 2003, 2004; Köhl et al., 2007) with success on a global scale (currently down to 1° horizontal resolution); although, the ECCO implementation is not based upon the incremental approach. Their approach adjusts the surface forcing such that the circulation fits the observations. The current ROMS incremental implementation adjusts the initial circulation without adjusting the surface forcing to match the observations. The IS4D-VAR method has also been implemented in the French Océan Parallelisé (OPA) ocean model and used successfully in the tropical Pacific by Weaver et al. (2003). ROMS also provides methods that assume an imperfect model and adjust the circulation during the assimilation period (as opposed to the initial conditions as presented here) as described in Di Lorenzo et al. (2007) based upon Bennett (2002).

We achieved a misfit RMSE between ExANmc and SST (SSH) that averages 0.52 °C (4.1 cm) compared to the observation standard deviation given by 0.27 °C (4 cm). The model bias was reduced to 0.015 °C (1 mm) and averaged a cross-correlation coefficient of 0.89 (0.92) with the observations. This compares with the forward model (ExN) RMSE of 1.16 °C (14.4 cm), a bias of 0.44 °C (4 cm), and an ACC of 0.77 (0.52). It is important to adjust **B** during the assimilation to better constrain the initial increments as the circulation improves. Updating **B** improved the RMSE in SST and SSH by 8% and 16%, respectively. Future work to improve the assimilation should include a multivariate **B** and increments applied to the surface forcing and boundary conditions.

Assimilation results from the 1/6° model are very similar to the coarse 1/3° model, without significant improvement in the assimilation despite twice the spatial resolution. This suggests that it may be beneficial to assimilate at a resolution that is similar to the observations as is performed operationally in NWP. Although the model presented cannot resolve sub-mesoscale energy, it has been shown to be representative of the larger-scale circulation present within the IAS. This result gives us the ability to deploy a real-time assimilation system aboard the ship in an operational capacity.

Additional experiments on the impact of assimilating in one subregion upon another subregion show that assimilation of data only within the CS contributes very little to the circulation estimate of the GOM as compared to the unassimilated observations in the GOM. Likewise, assimilation in the GOM did not impact the CS circulation. The impact was measured by the decrease in RMSE between a case with full assimilation and the case with assimilation only in the other region.

The primary objective of this work is to present a state-of-the-art data assimilation system that forms the basis of our real-time assimilation and prediction system. The circulation generated from this work serves as the initial starting point for a real-time assimilation and ensemble prediction system that is currently operating on-board the RCCL *Explorer of the Seas*. This ensemble prediction system will be discussed in a future article.

Acknowledgements

References

Arango, H.G., Reid, R.O., 1991. A generalized reduced-gravity ocean model. Atmos. Ocean 29, 256–287.

Beal, L.M., Hummon, J.M., Williams, E., Brown, O.B., Baringer, W., Kearns, E.J., 2008. Five years of Florida Current structure and transport from the Royal Caribbean Cruise Ship Explorer of the Seas. J. Geophys. Res. 113, C6–C06001.

Bennett, A.F., 2002. Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press.

Brown, O.B., Minnett, P.J., 1999. MODIS Infrared Sea Surface Temperature Algorithm Theoretical Basis Document, Ver 2.0. Tech. Rep., NASA Goddard Space Flight Center. URL: <http://modis.gsfc.nasa.gov/data/atbd/atbd_mod25.pdf>.

Chassignet, E.P., Hurlburt, H.E., Martin Smedstad, O., 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. American Geophysical Union, pp. 165–180.

Chassignet, E.P., Hurlburt, H.E., Martin Smedstad, O., Halliwell, G.R., Hogan, P.J., Wallcraft, A.J., Baraille, R., Bleck, R., 2007. The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. J. Mar. Syst. 65, 60–83.

Chelton, D.B., deSzoeke, R.A., Schlax, M.G., Naggar, K.E., Siwertz, N., 1998. Geographical variability of the first-baroclinic Rossby radius of deformation. J. Phys. Oceanogr. 28, 433–460.

Chérubin, L.M., Sturges, W., Chassignet, E.P., 2005. Deep flow variability in the vicinity of the Yucatan Straits from a high-resolution numerical simulation. J. Geophys. Res. 110. doi:10.1029/2004JC002280.

Chua, B.S., Bennett, A.F., 2001. An inverse Ocean Modeling System. Ocean Modell. 3, 137–165.

Courant, R., Hilbert, D., 1962. Methods of Mathematical Physics. Wiley-Interscience.

Courtier, P., Thépaut, J.N., Hollingsworth, A., 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. Quart. J. Roy. Meteorol. Soc. 120, 1367–1387.

Derber, J., Rosati, A., 1989. A global oceanic data assimilation system. J. Phys. Oceanogr. 19, 1333–1347.

Di Lorenzo, E., Moore, A.M., Arango, H.G., Cornuelle, B.D., Miller, A.J., Powell, B.S., Chua, B.S., Bennett, A.F., 2007. Weak and strong constraint data assimilation in the inverse Regional Ocean Modeling System (ROMS): development and application for a baroclinic coastal upwelling system. Ocean Modell. 16, 160–187.

Dietrich, D.E., Lin, C.A., 1994. Numerical studies of eddy shedding in the Gulf of Mexico. J. Geophys. Res. 99, 7599–7615.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 53, 343–367.

Ezer, T., Arango, H.G., Shchepetkin, A.F., 2002. Developments in terrain-following ocean models: intercomparisons of numerical aspects. Ocean Modell. 4, 249–267.

Ezer, T., Mellor, G.L., 1994. Continuous assimilation of Geosat altimeter data into a three-dimensional primitive equation Gulf Stream model. J. Phys. Oceanogr. 24, 832–847.

Fairall, C.W., Bradley, E.F., Rogers, D.P., Edson, J.B., Young, G.S., 1996. Bulk parameterization of air–sea fluxes for tropical ocean–global atmosphere coupled ocean-atmosphere response experiment. J. Geophys. Res. 101, 3747–3764.

Falkovich, A., Ginis, I., Lord, S., 2005. Ocean data assimilation and initialization procedure for the coupled GFDL/URI Hurricane prediction system. J. Atmos. Ocean. Technol. 22, 1,918–1,932.

Fan, S., Oey, L.-Y., Hamilton, P., 2004. Assimilation of drifter and satellite data in a model of the Northeastern Gulf of Mexico. Cont. Shelf. Res. 24, 1,001–1,013.

Fisher, M., 1998. Minimization algorithms for variational data assimilation. In: Seminar on Recent Developments in Numerical Methods for Atmospheric Modelling. ECMWF.

Fukumori, I., Raghunath, R., Fu, L.-L., Chao, Y., 1999. Assimilation of TOPEX/Poseidon altimeter data into a global ocean circulation model: How good are the results? J. Geophys. Res. 104, 25,647–25,665.

Ghil, M., Malanotte-Rizzoli, P., 1991. Data assimilation in meteorology and oceanography. Adv. Geophys. 33, 141–266.

Glenn, S.M., Ebbesmeyer, C.C., 1993. Drifting buoy observations of a loop current anticyclonic eddy. J. Geophys. Res. 98, 20,105–20,120.

Haidvogel, D.B., Arango, H.G., Hedstrom, K., Beckmann, A., Malanotte-Rizzoli, P., Shchepetkin, A.F., 2000. Model evaluation experiments in the North Atlantic Basin: simulations in nonlinear terrain-following coordinates. Dyn. Atmos. Oceans 32, 239–281.

Hamilton, P., Fargion, G.S., Biggs, D.C., 1999. Loop Current eddy paths in the western Gulf of Mexico. J. Phys. Oceanogr. 29, 1180–1207.

Hurlburt, H.E., Thompson, J.D., 1980. A numerical study of Loop Current intrusions and eddy shedding. J. Phys. Oceanogr. 10, 1611–1651.

Kantha, L.K., Choi, J.K., Schaudt, K.J., Cooper, C.K., 2005. A regional data-assimilative model for operational use in the Gulf of Mexico. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models, vol. 161. American Geophysical Union, pp. 165–180.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woolen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., Fiorino, M., 2001. The NCEP/NCAR 50-year reanalysis. Bull. Am. Meteorol. Soc. 82, 247–268.

Ko, D.S., Martin, P.J., Rowley, C.D., Preller, R.H., 2007. A real-time coastal ocean prediction experiment for mrea04. J. Mar. Syst.. doi:10.1016/j.jmarsys.2007.02.022.

Köhl, A., Stammer, D., Cornuelle, B., 2007. Interannual to decadal changes in the ECOO global synthesis. J. Phys. Oceanogr. 37, 313–337.

Lacarra, J.F., Talagrand, O., 1988. Short-range evolution of small perturbations in a barotropic model. Tellus 40A, 81–95.

Le Dimet, F., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. Tellus 38A, 97–110.

Le Traon, P.Y., Ducet, N., Dibarboure, G., 2001. Use of a high resolution model to analyze the mapping capabilities of multiple-altimeter missions. J. Atmos. Ocean. Technol. 18, 1277–1288.

Leben, R.R., 2005. Altimeter-derived Loop Current metrics. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models, vol. 161. American Geophysical Union, pp. 181–201.

Leben, R.R., Powell, B.S., 2003. Accuracy assessment of Jason-1 and TOPEX/Poseidon along-track sea surface slope. Mar. Geod. 26, 355–366. doi:10.1080/01490410390256754.

Lee, T.N., Johns, W.E., Zantopp, R.J., Fillenbaum, E.R., 1996. Moored observations of western boundary current variability and thermohaline circulation at 26.5°n in the subtropical North Atlantic. J. Phys. Oceanogr. 26, 962–983.

Levitus, S., Burgett, R., Boyer, T., 1994. World ocean atlas 1994 volume 4: Temperature. In: NOAA Atlas NESDIS 4. US Department of Commerce, 132 pp.

Lin, X.H., Oey, L.-Y., Wang, D.-P., 2007. Altimetry and drifter data assimilations of loop current and eddies. J. Geophys. Res. 112. doi:10.1029/2006JC003779.

Mancilla Rojas, M., 2007. Validation of a Simulation of ROMS Model in the Gulf of Mexico and Caribbean Sea. Ph.D. Thesis, CICESE.

Marchesiello, P., McWilliams, J.C., Shchepetkin, A., 2001. Open boundary conditions for long-term integration of regional oceanic models. Ocean Modell. 3, 1–20.

Maturi, E., Merchant, C., Harris, A., Li, X., Potash, B., 2004. Geostationary sea surface temperature product validation and methodology. In: Poster Presentation at AMS 13th Conference on Satellite Meteorology and Oceanography.

Mellor, G.L., Ezer, T., 1991. A Gulf Stream model and an altimetry assimilation scheme. J. Geophys. Res. 96, 8,779–8,795.

Milliff, R.F., Morzel, J., Chelton, D.B., Freilich, M.H., 2004. Wind stress curl and wind stress divergence biases from rain effects on QSCAT surface wind retrievals. J. Atmos. Ocean. Technol. 21, 1216–1231.

Milliff, R.F., Niiler, P.P., Morzel, J., Sybrandy, A.E., Nychka, D., Large, W.G., 2003. Mesoscale correlation length scales from NSCAT and Minimet surface wind retrievals in the Labrador Sea. J. Atmos. Ocean. Technol. 20, 513–533.

Mooers, C.N.K., Maul, G.A., 1998. Intra-Americas Sea circulation. In: Robinson, A.R., Brink, K.H. (Eds.), The Sea. John Wiley & Sons, pp. 183–208 (Chapter 11).

Moore, A.M., Arango, H.G., Lorenzo, E.D., Cornuelle, B.D., Miller, A.J., Neilson, D.J., 2004. A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint of a regional ocean model. Ocean Modell. 7, 227–258.

Morrison, J.M., Nowlin, W.D., 1982. General distributions of water masses within the eastern Caribbean Sea during the winter of 1972 and the fall of 1973. J. Geophys. Res. 87, 4207–4229.

Müller-Karger, F.E., Walsh, J.J., Evans, R.H., Meyers, M.B., 1991. On the seasonal phytoplankton concentration and sea surface temperature cycles of the Gulf of Mexico as determined by satellites. J. Geophys. Res. 96, 12645–12665.

Murphy, S.J., Hurlburt, H.E., O'Brien, J.J., 1999. The connectivity of eddy variability in the Caribbean Sea, the Gulf of Mexico, and the Atlantic Ocean. J. Geophys. Res. 104, 1,431–1,453.

Ngodock, H.E., Smith, S.R., Jacobs, G.A., 2007. Cycling the representer algorithm for variational data assimilation with a nonlinear reduced gravity ocean model. Ocean Modell. 19, 101–111.

Ochoa, J., Badan, A., Sheinbaum, J., Candela, J., 2003. CANEK: Measuring Transport in the Yucatan Channel. In: Nonlinear Processes in Geophysical Fluid Dynamics. Kluwer Academic Publishers, p. 376.

Oey, L.-Y., 1995. Eddy- and wind-forced shelf circulation. J. Geophys. Res. 100, 8621–8637.

Oey, L.-Y., 1996. Simulation of mesoscale variability in the Gulf of Mexico: sensitivity studies, comparison with observations, and trapped wave propagation. J. Phys. Oceanogr. 26, 145–175.

Oey, L.-Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., Fan, S., 2005a. An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico. Geophys. Res. Lett. 32. doi:10.1029/2005GL023253.

Oey, L.-Y., Ezer, T., Lee, H.-C., 2005b. Loop Current, rings and related circulation in the Gulf of Mexico: a review of numerical models and future challenges. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models, vol. 161. American Geophysical Union, pp. 31–56.

Oey, L.-Y., Lee, H.-C., Schmitz, W.J., 2003. Effects of winds and Caribbean eddies on the frequency of Loop Current eddy shedding: a numerical modeling study. J. Geophys. Res. 108. doi:10.1029/2002JC001698.

Ohlmann, J.C., Niiler, P.P., Fox, C.A., Leben, R.R., 2001. Eddy energy and shelf interactions in the Gulf of Mexico. J. Geophys. Res. 106, 2605–2620.

Parrish, D.F., Derber, J.C., 1992. The National Meteorological Center's spectral statistical-interpolation analysis system. Mon. Weather Rev. 120, 1747–1763.

Pascual, A., Pujol, M.-I., Larnicol, G., Traon, P.-Y.L., Rio, M.-H., 2007. Mesoscale mapping capabilities of multisatellite altimeter missions: first results with real data in the Mediterranean Sea. J. Mar. Syst. 65, 190–211.

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., Wang, W., 2002. An improved in situ and satellite SST analysis for climate. J. Climate 15, 1609–1625, doi:10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

Richardson, P.L., 2005. Caribbean Current and eddies as observed by surface drifters. Deep-Sea Res., Part II 52, 429–463.

Roemmich, D., 1981. Circulation of the Caribbean Sea: a well-resolved inverse problem. J. Geophys. Res. 86, 7993–8005.

Roemmich, D., 1983. The balance of geostrophic and Ekman transports in the Tropical Atlantic Ocean. J. Geophys. Res. 13, 1534–1539.

Romanou, A., Chassignet, E.P., Sturges, W., 2004. The Gulf of Mexico circulation within a high resolution numerical simulation of the North Atlantic Ocean. J. Geophys. Res. 109. doi:10.1029/2003CJ001770.

Schmitz, W.J., 2005. Cyclones and westward propagation in the shedding of anticyclonic rings from the Loop Current. In: Sturges, W., Lugo-Fernandez, A. (Eds.), Circulation in the Gulf of Mexico: Observations and Models, vol. 161. American Geophysical Union, pp. 241–262.

Schmitz, 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. American Geophysical Union, pp. 11–30.

Shchepetkin, A.F., McWilliams, J.C., 2003. A method for computing horizontal pressure-gradient force in an oceanic model with nonaligned vertical coordinate. J. Geophys. Res. 108. doi:10.1029/2001JC001047.

Shchepetkin, A.F., McWilliams, J.C., 2005. The Regional Oceanic Modeling System: a split-explicit, free-surface, topography-following-coordinate ocean model. Ocean Modell. 9, 347–404.

Sheinbaum, J., Candela, J., Badan, A., Ochoa, J., 2002. Flow structure and transport in the Yucatan Channel. Geophys. Res. Lett. 29 (3). doi:10.1029/2001GL013990.

Smith, D.C., O'Brien, J.J., 1983. The interaction of a two-layer isolated mesoscale eddy with topography. J. Phys. Oceanogr. 13, 1681–1967.

Stammer, D., Ueyoshi, K., Köhl, A., Large, W.B., Josey, S., Wunsch, C., 2004. Estimating air–sea flux estimates through global ocean data assimilation. J. Geophys. Res. 109. doi:10.1029/2003JC002082.

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., Adcroft, A., Hill, C.N., Marshall, J., 2002. The global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. J. Geophys. Res. 107. doi:10.1029/2001JC000888.

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., Adcroft, A., Hill, C.N., Marshall, J., 2003. Volume, heat and freshwater transports of the global ocean circulation 1993–2000, estimated from a general circulation model constrained by WOCE data. J. Geophys. Res. 108. doi:10.1029/2001JC001115.

Stowe, L.L., Davis, P.A., McClain, E.P., 1999. Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer. J. Atmos. Ocean. Technol. 16, 656–681.

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

Talagrand, O., Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. Quart. J. Roy. Meteorol. Soc. 113, 1311–1328.

Trémolet, Y., 2004. Diagnosis of linear and incremental approximations in 4D-Var. Quart. J. Roy. Meteorol. Soc. 130, 2,233–2,251.

Tshimanga, J., Gratton, S., Weaver, A.T., Sartenaer, A., 2008. Limited-memory preconditioners, with application to incremental four-dimensional variational data assimilation. Quart. J. Roy. Meteorol. Soc. 134, 751–769.

Walker, N., Myint, S., Babin, A., Haag, A., 2003. Advances in satellite radiometry for the surveillance of surface temperatures, ocean eddies and upwelling processes in the Gulf of Mexico using GOES-8 measurements during summer. Geophys. Res. Lett. 30, 1854. doi:10.1029/2003GL017555.

Walton, C.C., Pichel, W.G., Sapper, J.F., May, D.A., 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. J. Geophys. Res. 103 (C12), 27,999–28,012.

Wang, D.-P., Oey, L.-Y., Ezer, T., Hamilton, P., 2003. Nearsurface currents in Desoto Canyon. J. Phys. Oceanogr. 33, 313–326.

Warner, J.C., Sherwood, C.R., Arango, H.G., Signell, R.P., 2005. Performance of four turbulence closure models implemented using a generic length scale method. Ocean Modell. 8, 81–113.

Weaver, A., Courtier, P., 2001. Correlation modelling on the sphere using generalized diffusion equation. Quart. J. Roy. Meteorol. Soc. 127, 1815–1846.

Weaver, A.T., Vialard, J., Anderson, D.L.T., 2003. Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. Mon. Weather Rev. 131, 1360–1378.

Wentz, F.J., Gentemann, C., Ashcroft, P., 2003. On-orbit calibration of AMSR-E and the retrieval of ocean products. In: Proceedings of the 83rd AMS Annual Meeting. American Meteorological Society.

Wentz, F.J., Gentemann, C., Smith, D., Chelton, D., 2000. Satellite measurements of sea-surface temperature through clouds. Science 288, 847–850.

Wikle, C.K., Berliner, L.M., 2007. A Bayesian tutorial for data assimilation. Physica D 230, 1–16.

Willis, J.K., Roemmich, D., Cornuelle, B., 2004. Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales. J. Geophys. Res. 109. doi:10.1029/2003JC002260c.

Wunsch, C., 1996. The Ocean Circulation Inverse Problem. Cambridge University Press.

Wunsch, C., Grant, B., 1982. Towards the general circulation of the North Atlantic Ocean. Progr. Oceanogr. 11, 1–59.

Zanifé, O.Z., Vincent, P., Amarouche, L., Dumont, J.P., Thibaut, P., Labroue, S., 2003. Comparison of the Ku-band range noise level and the relative sea-state bias of the Jason-1, TOPEX, and Poseidon-1 radar altimeters. Mar. Geod., 201–238.