Title:

Towards an integrated observation and modeling system in the New York Bight using variational methods, Part I: 4DVAR Data Assimilation

Authors: Weifeng G. Zhang[a,1], John L. Wilkin[a] and Hernan G. Arango[a]

[a] Institute of Marine and Coastal Sciences, Rutgers, The State University of New Jersey

 71 Dudley Road, New Brunswick, New Jersey, 08901, USA

Emails: wzhang@whoi.edu, wilkin@marine.rutgers.edu, arango@marine.rutgres.edu

Corresponding Author:

Weifeng G. Zhang

Tel.: +1 (508) 289 2521

Fax: +1 (508) 457 2194

Email: wzhang@whoi.edu

[1] *Present address*: Woods Hole Oceanographic Institution, Bigelow 410, MS#09, Woods Hole, Massachusetts, 02543, USA

Abstract

Four-dimensional Variational data assimilation (4DVAR) in the Regional Ocean Modeling System (ROMS) is used to produce a best-estimate analysis of ocean circulation in the New York Bight during spring 2006 by assimilating observations collected by a variety of instruments during an intensive field program. An incremental approach is applied in an overlapped cycling system with 3-day data assimilation window to adjust model initial conditions. The model-observation mismatch for all observed variables is reduced substantially. Comparisons between model forecast and independent observations show improved forecast skill for about 15 days for temperature and salinity, and 2 to 3 days for velocity. Tests assimilating only certain subsets of the data indicate that assimilating satellite sea surface temperature improves the forecast of surface and subsurface temperature but worsens the salinity forecast. Assimilating in situ temperature and salinity from gliders improves the salinity forecast but has little effect on temperature. Assimilating HF-radar surface current data improves the velocity forecast by 1-2 days yet worsens the forecast of subsurface temperature. During some time periods the convergence for velocity is poor as a result of the data assimilation system being unable to reduce errors in the applied winds because surface forcing is not among the control variables. This study demonstrates the capability of 4DVAR data assimilation system to reduce model-observation mismatch and improve forecasts in the coastal ocean, and highlights the value of accurate meteorological forcing.

Keywords: Data assimilation; 4DVAR; ROMS; Ocean prediction; New York Bight; River plume

1. Introduction

The New York Bight (NYB) lies in the center of the Mid-Atlantic Bight (MAB) adjacent to the coasts of Long Island and New Jersey. The mean depth-averaged flow in the MAB, driven primarily by a large scale along shore pressure gradient, is southward, alongshelf, and increases gradually with increasing water depth (Beardsley and Boicourt, 1981; Lentz, 2008). The circulation in the NYB is influenced by this remotely forced mean flow and also local forces, including river discharge, wind and variable bathymetry (Castelao et al., 2008; Chant et al., 2008; Choi and Wilkin, 2007; Johnson et al., 2003; Münchow and Chant, 2000; Tilburg and Garvine, 2003; Wong, 1999; Yankovsky et al., 2000). The alongshelf mean flow has its strongest influence on the mid- and outer-shelf (Zhang et al., 2009a), while it is mainly the interaction between buoyancy and wind that drives the circulation on the inner-shelf (Münchow and Chant, 2000; Tilburg and Garvine, 2003; Wong, 1999; Yankovsky, 2003; Yankovsky and Garvine, 1998; Yankovsky et al., 2000). Using a model, Zhang et al. (2009a) showed that unsteady wind is the main driver dispersing Hudson River source water onto the mid- and outer-shelf, and the presence of the Hudson Shelf Valley intensifies the freshwater outflow bulge recirculation in the apex of the NYB.

The Hudson River is a major source of nutrients, contaminants and other biogeochemically active tracers in the NYB (Adams et al., 1998; Howarth et al., 2006) and has been observed to impact local biogeochemistry (Moline et al., 2008; Schofield et al., 2008). Numerical ecosystem modeling has proven useful for examining aspects of biogeochemical cycling in this region (Cahill et al., 2008; Fennel et al., 2008; Fennel et

al., 2006), but to achieve skillful simulation of biogeochemical events within the NYB on short time and space scales requires quite accurate estimates of the ocean physical state – developing physical ocean state estimates for such a purpose is an objective of this study.

Being one of the best observed coastal areas in the world, the NYB has been the target of pioneering deployments of new observing instruments including autonomous underwater vehicles (gliders) (Schofield et al., 2007), surface current measuring High-Frequency (HF) radar (Kohut et al., 2006a), and a cabled observatory (Glenn et al., 2000). In spring 2005 and 2006, interdisciplinary process studies of the Hudson River plume (the Lagrangian Transport and Transformation Experiment, LaTTE) were conducted (Chant et al., 2008) using observations from satellites, HF radar, a fleet of gliders, moorings, surface drifters, and instruments aboard the R/V Cape Hatteras and R/V Oceanus to monitor the NYB. Simulation of the NYB using ROMS (Regional Ocean Model System, http://myroms.org) complemented observations in real-time. The large multi-instrument data sets available from LaTTE, and the on-going operation of much of the instrumentation on a quasi-continuous basis, make the NYB an attractive location to explore the integration of advanced observation and modeling capabilities for the purposes of implementing coastal ocean data assimilation and optimal observation systems.

Data assimilation (DA) is to use observations in conjunction with models to better describe the ocean (Bennett, 2002; Evensen, 2007; Wunsch, 2006) and it is currently an active research area in applied oceanography. The time-dependent variational method (4DVAR) is one DA approach and it takes the linearized dynamical model into consideration while adjusting model control variables to fit observations. The 4DVAR

method proceeds by iteratively minimizing a cost function defined as the weighted mismatch between the observations and the model state at the observation location and time, plus additional constraints such as the size of the permitted adjustment to the model control variables. In principle, the control variables to be adjusted can be anything imposed external to the model, such as initial conditions, boundary conditions, and forcing, or aspects internal to the model such as vertical mixing parameters and missing physics.

Assuming the model physics is "perfect" (i.e. applying the so-called "strong constraint" of (Talagrand and Courtier, 1987)) there are two approaches to the minimization problem: (i) Incremental Strong-constraint 4DVAR (IS4DVAR) (Courtier et al., 1994), and (ii) representer-based 4DVAR (Bennett, 2002) (also called 4D-PSAS (Courtier, 1997)). These approaches minimize the cost function in different spaces: IS4DVAR in model space, and representer-based 4DVAR in observation space. In IS4DVAR an iterative scheme is used to minimize the cost function based on information provided by the Tangent Linear and Adjoint models. In representer-based 4DVAR the minimum search is achieved through looking for the coefficients of the observational representers that give the model-observation mismatch. Courtier (1997) proved that the equations of representer coefficients are equivalent to the equations of control variable adjustment in IS4DVAR, and to solve the equations of representer coefficients is essentially to minimize a objective function whose Hessian has the same condition number as the cost function in IS4DVAR. More detailed discussion of the merits of these approaches can be found in the literature (Broquet et al., 2009; Di Lorenzo et al., 2007; Kurapov et al., 2007; Powell et al., 2008; Weaver et al., 2003).

Both 4DVAR algorithms have been applied to studies of ocean variability at different scales. On large scales 4DVAR has been applied for the purposes of methodology development (Vialard et al., 2003; Vossepoel et al., 2004; Weaver et al., 2003; Weaver et al., 2005) and improving estimate of ocean states and air-sea fluxes (Stammer et al., 2004; Stammer et al., 2002; 2003; Wunsch and Heimbach, 2007). On regional and coastal scales 4DVAR algorithms have been developed and applied mostly in either simple models or idealized configurations (Di Lorenzo et al., 2007; Kurapov et al., 2007; Ngodock et al., 2007; Scott et al., 2000) with a few exceptions (Broquet et al., 2009; Hoteit and Köhl, 2006; Powell et al., 2008; Smith and Ngodock, 2008).   Other coastal ocean DA studies have been conducted with other methods, such as nudging or melding (Lewis et al., 1998; Wilkin et al., 2005), optimal interpolation (Oke et al., 2002), 3DVAR (Li et al., 2008a; Li et al., 2008b), and the ensemble Kalman filter (Hoffman et al., 2008).

In this study we used the IS4DVAR system in ROMS, described comprehensively by Powell et al. (2008) and Broquet et al. (2009), to assimilate all available observations collected in the NYB in conjunction with the spring 2006 LaTTE field program. We describe a "pseudo-real-time" DA system, by which we mean a system that could have operated in real-time had we known then what we have learned here about practical issues of timeliness and quality control that must be addressed when assimilating observational data from the various platforms we used, and configuration of the IS4DVAR algorithm itself in a shallow inner-shelf region. This retrospective analysis with a comprehensive observational data set also allows us to evaluate, to a certain extent, the influences of differing observation sources on the performance of the forecast system.

Further evaluation and optimization of observational strategies using representer-based methods is explored in the Part II paper that accompanies this article.

This paper is organized as follows: Section 2 describes the data collected in spring 2006 and its quality control prior to assimilation; Section 3 describes the model configuration and Section 4 describes the DA system; Section 5 presents the results; and Section 6 summarizes the work.

2. Observational data

The 2006 LaTTE field campaign in the NYB targeted the spring freshet of the Hudson River, and was similar to that in spring 2005 described by Chant et al. (2008). In both LaTTE seasons, sustained observations were acquired from HF radar, gliders, and satellites by the Rutgers University Coastal Ocean Observation Laboratory (RUCOOL) (Glenn and Schofield, 2003).

Data were acquired from the Advanced Very High Resolution Radiometer (AVHRR) aboard the NOAA satellites by the RUCOOL L-Band ground station and processed to 1-km resolution Sea Surface Temperature (SST) data using the Multi-channel Sea Surface Temperature algorithm in SeaSpace Terascan software. We manually removed clouds from the individual satellite passes and spatially averaged the data to a 4 km resolution grid for assimilation. An example is shown in Figure 6a.

Gliders with a SeaBird CTD aboard observed profiles of temperature and salinity at a vertical resolution of 0.25 m. The conductivity cell on a glider is not pumped which introduces a thermal-lag in the response that can cause salinity data to have large errors

when a glider crosses a pronounced halocline/thermocline (Lueck, 1990) such as at the Hudson River plume boundary. We corrected the thermal-lag using the algorithm of Morison et al. (1994) and then vertically averaged the data to 1-m vertical resolution. The glider tracks are plotted in Figure 1, and an example of glider-measured temperature and salinity is given in Figure 7.

The HF radar data used here were derived from 5 antenna sites (Figure 1) to give 3-hour averaged sea surface current once per hour. The azimuth resolution of the raw data at all the sites is 5 degrees, and radial data from the sites was combined into total vectors using the algorithm described by Kohut et al. (2001) and Kohut et al. (2006b) and mapped to a 6-km resolution grid. An example of the HF-radar-measured surface current is given in Figure 6a. We choose 2.5 m to be the nominal depth of the HF radar measured currents according to the effective depth of the HF radar measurement (Stewart and Joy, 1974).

The average power spectrum of the HF radar data in Figure 2 shows that tides dominate the observed surface current. Due to error in either the boundary conditions, propagation of tides within the model, or in the HF radar measurements themselves, the spatial patterns of modeled and observed tidal current harmonics differ (e.g. the comparison of M2 harmonic in Figure 3). The ROMS DA system implemented at the time of this study does not include the tidal harmonic boundary forcing as control variables, and therefore cannot adjust (or "tune") these directly to improve the model tidal harmonics solution. If we were to assimilate the 1-hourly interval HF radar data directly, systematic errors in the surface velocity due to this tidal harmonic mismatch could be interpreted by the DA system as requiring adjustments to the model initial

conditions, which is obviously not what we intend. To avert this potential problem, we filtered the tidal signal in the HF radar data with an harmonic analysis toolbox (Pawlowicz et al., 2002) and replaced it with the harmonic tidal signal derived from a control ROMS model simulation. This merger achieves consistency between the "modified observed" tidal currents and the model so that any mismatch in the tides will not dominate the cost function. By comparing to results when assimilating the original HF radar data we verified that assimilating the merged surface currents gives better velocity fit and forecast skill (not shown).

Further observations were gathered by moorings, drifters and ship surveys. Seven moorings (Figure 1), each with an Acoustic Doppler Current Profiler (ADCP) and two or three Conductivity/Temperature (CT) sensors at different depths, recorded data from the beginning of April to the end of June. Two surface drifters deployed between 4 and 8 May measured surface temperature (we did not attempt to use drifters for surface velocity data given the difficulty discussed above of adequately of accounting for the tides). Between 2 and 8 May surveys by the vessels R/V Cape Hatteras and R/V Oceanus measured temperature, salinity and velocity from towed undulating CTD and ship mounted ADCP instruments. All data from the towed undulating vehicle, ADCP, CTD, CT and drifters were averaged to resolutions of 2 m in the vertical, 5 km horizontally, and 12 minutes in time prior to assimilation. This reduced the scales resolved by the data to be comparable to those represented in the model; there is little point to keeping data at more fine resolution because the DA process would smooth them anyway.

Figure 4a indicates the availability of observations from different platforms on each day during the study period, and Figure 4b gives the total number of observations of each

ocean state variable on each day. We have in total 20,000 to 45,000 observations each day, with velocity being the most abundant and salinity the least. More than 60% of the velocity data is surface current measured by HF radar and the rest is measured by ADCP. More than 50% of the temperature data is satellite SST, about 13% is measured by gliders, and the rest is measured by moored and ship-borne CTD. About half of the salinity data is from gliders and half is from moorings and ships.

3. Model configuration

3.1 The Regional Ocean Modeling System

ROMS is a free-surface, hydrostatic, primitive equation ocean model using terrain-following vertical coordinates. Haidvogel et al. (2008) present an overview of the model design, embedded ecosystem, sediment transport and sea-ice modules, and examples of applications to estuarine, coastal and basin-scale studies. Shchepetkin and McWilliams (1998; 2003; 2005) describe in detail the ROMS computational kernel. The Adjoint and Tangent Linear models that are part of the ROMS code were developed by the ROMS Adjoint Group (Moore et al., 2004), and drivers have been developed that utilize the Adjoint and Tangent Linear models for Adjoint Sensitivity and Optimal Perturbation analysis, Optimal Observation design, and Incremental Strong Constraint 4-Dimensional Variational (IS4DVAR) and Weak constraint 4DVAR (W4DVAR) data assimilation (Broquet et al., 2009; Di Lorenzo et al., 2007; Moore et al., 2008; Moore et al., 2004; Powell et al., 2008; Zhang et al., 2009b).

3.2 New York Bight and New Jersey Inner Shelf Configuration

Our ROMS configuration for the NYB is similar to that of Zhang et al. (2009a). The

model domain (Figure 1) covers the NYB from south of Delaware Bay to eastern Long

Island and includes the two major rivers in this area, the Hudson and Delaware. Because

of the substantial computational demand of the IS4DVAR system (about 100-fold more

than the computation of a single forward model simulation), the horizontal resolution in

this study has been decreased from 1 km in Zhang et al. (2009a) to 2 km, but the vertical

resolution of (30 layers) is unchanged.

Model initial conditions prior to commencing DA on 5 April 2006 (when the LaTTE

observational program began in earnest) were obtained from the "full physics" simulation

in Zhang et al. (2009a). In all forward model simulations, Chapman (1985) and Flather

(1976) open boundary conditions are used for sea surface height and the barotropic

component of velocity on the model perimeter, respectively. These conditions impose

both a remotely-forced along-shelf mean flow computed from the water-depth/velocity

relationship deduced by Lentz (2008), and tidal harmonic variability from a regional tidal

simulation (Mukai et al., 2002). Gradient open boundary conditions are used for 3D

velocity and tracers. The Generic Length Scale method *k-kl* closure (Umlauf and

Burchard, 2003; Warner et al., 2005) is used for the vertical mixing; bottom drag is

quadratic ($C_D = 0.003$). In the forward simulations, air-sea fluxes of momentum and heat

are computed using bulk formulae (Fairall et al., 2003) with meteorological conditions

from the North American Mesoscale (NAM) model distributed by the National

Operational Model Archive and Distribution System (NOMADS) (Rutledge et al., 2006).

The river discharge data were obtained from USGS Water Data

(http://waterdata.usgs.gov/nwis) and modified to include ungauged portions of the

watershed. Figure 5 shows the river discharges and wind at the Hudson River mouth

over the DA experiment period.

In almost all respects, the model configuration here is identical to Zhang et al.'s

(2009a) model of the NYB; the differences are the reduced horizontal resolution – to

make the computational effort tractable – and the use of high resolution NAM operational

meteorological forecasts in place of lower resolution regional reanalysis products.

4. Data Assimilation System

4.1. IS4DVAR theory

We summarize briefly here the principles of IS4DVAR data assimilation for the

purpose of highlighting the choices to be made when implementing the method in a

practical coastal application. For a detailed description see (Courtier, 1997; Courtier et al.,

1994; Powell et al., 2008; Weaver et al., 2003).

The ROMS forward model can be represented as,

$$\begin{cases} \dfrac{\partial \mathbf{\Phi}(t)}{\partial t} = M(\mathbf{\Phi}(t)) + \mathbf{F}(t) \\ \mathbf{\Phi}(0) = \mathbf{\Phi}_i \\ \mathbf{\Phi}(t)\big|_{\Omega} = \mathbf{\Phi}_{\Omega}(t) \end{cases} \qquad (1)$$

where M is the model nonlinear operator; $\mathbf{\Phi}(t)$ is a state vector $[\mathbf{u}\ \mathbf{v}\ \mathbf{T}\ \mathbf{S}\ \zeta]^{\mathrm{T}}$ comprised of

the velocity, temperature, salinity and sea surface height at all model grid points at time t;

$\mathbf{F}(t)$ is the external forcing; $\mathbf{\Phi}_i$ the initial conditions; and $\mathbf{\Phi}_{\Omega}(t)$ are boundary conditions

along boundary Ω. We assume the model is "perfect," that is, no explicit account is made

for inadequacies in the forward model in the model-data misfit – this is the so-called *strong constraint* method, in which the model physics alone determine how the modeled ocean state evolves over time during the analysis time period for a given set of initial conditions, boundary conditions and forcing. In DA, the objective is to adjust the control variables (typically initial conditions, but also potentially boundary conditions and forcing) to minimize the mismatch between model and observations, denoted J_o, over some analysis interval:

$$J_o = \frac{1}{2} \sum_{n=1}^{N_{obs}} \left(H_n \mathbf{\Phi}(t_n) - \mathbf{y}_n \right)^{\mathrm{T}} \mathbf{O}^{-1} \left(H_n \mathbf{\Phi}(t_n) - \mathbf{y}_n \right), \tag{2}$$

where H is an operator that samples the nonlinear model state at the observation locations and times, $\mathbf{O}$ is the observational error covariance matrix, $\mathbf{y}$ is a vector of the observations, and N_{obs} is the number of observations in the analysis interval.

In the *incremental* formulation of 4DVAR (Courtier et al., 1994), we let $\mathbf{\Phi}_0$ denote a solution to the nonlinear problem (1) and assume $\mathbf{\Phi}_0$ is sufficiently close to the true ocean state that the adjustments to the control variables, $\boldsymbol{\varphi}_i = \delta \mathbf{\Phi}_i$ for initial conditions, $\boldsymbol{\varphi}_\Omega(t) = \delta \mathbf{\Phi}_\Omega(t)$ for boundary conditions, and $\mathbf{f}(t) = \delta \mathbf{F}(t)$) for forcing, will be small and can be described by a linearized model, the Tangent Linear model,

$$\begin{cases} \dfrac{\partial \boldsymbol{\varphi}(t)}{\partial t} = \left(\dfrac{\partial M}{\partial \mathbf{\Phi}} \right)\bigg|_{\mathbf{\Phi}_0} \boldsymbol{\varphi}(t) + \mathbf{f}(t) \\ \boldsymbol{\varphi}(0) = \boldsymbol{\varphi}_i \\ \boldsymbol{\varphi}(t)\big|_\Omega = \boldsymbol{\varphi}_\Omega(t) \end{cases}, \tag{3}$$

where $\varphi(t) = \Phi(t) - \Phi_0(t)$ is the perturbation state at time t. The mismatch between the model and observation, $\mathbf{d} = \mathbf{y} - H\Phi_0(t)$, is then small. The system can be linearized and the cost function is now defined as

$$J = J_o + J_b, \tag{4}$$

with

$$J_o = \frac{1}{2} \sum_{n=1}^{N_{obs}} \left(\mathbf{H}_n \varphi(t_n) - \mathbf{d}_n \right)^{\mathrm{T}} \mathbf{O}^{-1} \left(\mathbf{H}_n \varphi(t_n) - \mathbf{d}_n \right), \tag{5}$$

$$J_b = \frac{1}{2} \varphi_i \mathbf{B}_i^{-1} \varphi_i + \frac{1}{2} \varphi_\Omega \mathbf{B}_\Omega^{-1} \varphi_\Omega + \frac{1}{2} \mathbf{f} \mathbf{B}_f^{-1} \mathbf{f}, \tag{6}$$

where $\mathbf{H}$ is linearized H, and $\mathbf{B}_i$, $\mathbf{B}_\Omega$ and $\mathbf{B}_f$ are the assumed covariances of errors in initial conditions, boundary conditions and forcing, respectively. Equation (5) is the model-observation mismatch rewritten in the incremental formulation, and the background cost function J_b in Equation (4) penalizes adjustment of the control variables to ensure they are not too large.

In ROMS IS4DVAR the minimization of J is achieved iteratively in a so-called inner-loop using a Conjugate Gradient algorithm. The incremental formulation renders the system linear so that J is quadratic and the convergence of iterations is guaranteed. On each iteration the gradient of J with respect to the control variables, obtained from the Adjoint model forced by the model-observation mismatch, is used to compute the direction and step size of the minimum search. There is also an outer-loop to the iterative procedure in which the nonlinear forward model is rerun to update $\Phi_0(t)$ using the

adjusted control variables after the inner-loop has converged. The outer loop brings the nonlinearity into the system. In the end, corrected initial conditions, boundary conditions and forcing are obtained. Our choice for the number of inner and outer loops is discussed below.

4.2. Data Assimilation System setup

Our IS4DVAR analysis of NYB circulation is for the period 10 April to 5 June 2006 coinciding with the availability of in situ observations during LaTTE. The duration of each DA analysis window (the interval over which J_o is evaluated and the iteration on φ_i performed) was chosen to be 3 days based on tests in Zhang et al. (2009b) of the linearity assumption that underpins the incremental method. Following standard practice in real-time numerical weather prediction, we choose to overlap consecutive DA cycles: the beginning of the DA window advances one day from one cycle to the next, therefore creating a two-day overlap between consecutive cycles. The workflow is as follows: The first DA cycle starts at 0000 UTC 10 April 2006 with the first guess of the initial conditions taken from the control forward model simulation (a two-month continuous simulation prior to commencing DA). Assimilation of all the observational data within the 3-day period (0000 UTC 10 April – 0000 UTC 13 April) gives adjusted initial conditions for 0000 UTC 10 April from which an 18-day forward nonlinear model simulation is then launched. The model solution within the first 3 days is therefore an "analysis" result, being a fit to observations made at the same time, while the outcome for the subsequent 15 days is a forecast. The second DA cycle starts at 0000 UTC 11 April with the first guess of the initial conditions now taken from the analysis of the first DA cycle. Assimilation of observations in the window 0000 UTC 11 April to 0000 UTC 14

April then produces new adjusted initial conditions for 0000 UTC 11 April. Note that observations made on 11 and 12 April are assimilated in both the first and the second DA cycles. Another 18-day forward nonlinear model simulation is launched starting from the Cycle 2 adjusted initial conditions for 0000 UTC 11 April. We repeat this process, advancing 1 day each cycle, until the last DA cycle starts at 0000 UTC 3 June 2006. In total there are 55 overlapped cycles.

(Note: The meteorological forcing data set we use is a concatenation of the first 24 hours of each NCEP NAM forecast cycle, and therefore our ocean forecast is not strictly a true forecast because atmospheric observations in the "future" impact the prediction. Nevertheless, no *ocean* observations are utilized during the forecast, and our experiment is a faithful test of how IS4DVAR assimilation of ocean observations improves state estimation and prediction.)

At the time this study was conducted the IS4DVAR capability of ROMS allowed only for adjustments to the model initial conditions. The last two terms in equation (6) are therefore absent here, though it is certain that errors exist in the external forcing and boundary conditions. Given our relatively short 57-day analysis period and previous studies that show circulation in the NYB and New Jersey inner shelf is predominantly locally forced (Choi and Wilkin, 2007; Zhang et al., 2009a), we do not expect boundary conditions to play a significant role, but rather that initial conditions will dominate the evolution of oceanic tracers (temperature and salinity) and are therefore suitable control variables to adjust to reduce model-observation tracer mismatch. The adjoint sensitivity analysis of Zhang et al. (2009b) also emphasizes that SST in the immediate vicinity of the Hudson River plume has the greatest contribution to SST anomalies on the New Jersey

coast. However, this argument is not necessarily true for velocity as we will see later in this paper. The capability to adjust external forcing and boundary conditions has recently been added to the ROMS 4DVAR system and will be applied in future studies.

Within each DA cycle, 3 outer-loops and 11 inner-loops are used. Tests with different numbers of outer-loops and inner-loops prove this is a practical and effective combination in terms of system performance and affordability. Due to the strong nonlinearities embedded in the vertical turbulence closure and the air-sea flux bulk formulae, these features of the nonlinear forward model are not precisely linearized in the Tangent Linear and Adjoint models. Instead, space and time varying vertical viscosity and diffusivity coefficients, and surface heat and momentum fluxes computed in the first nonlinear model simulation of each cycle (corresponding to $\Phi_0(t)$) are stored and used by the Tangent Linear and Adjoint models in that cycle.

4.3. Error statistics

In Equation (5) the model-observation mismatch is weighted by observational error covariance. We assume the observations are independent of each other, and the observational error covariance matrix $\mathbf{O}$ is then a diagonal matrix. The error value assigned to each observation represents the combination of actual instrument accuracy, misrepresentation associated with processes unresolved by the model grid or absent from the model physics yet observed by the instruments (e.g. high frequency internal waves), and model error caused by inaccuracies in external forces that are not included in the assimilation control variables (surface forcing and open boundary conditions in this study). While the accuracy of each instrument (CTD, AVHRR, CODAR etc.) is

reasonably well known, the specification of observational error remains somewhat subjective and empirical because the misrepresentation associated with model resolution and physics is difficult to quantify *a priori*. We need to choose errors that realistically represent the extent to which the modeling system can fit the data; if not the DA cannot converge. The observational error standard deviations used in this study as listed in Table 1.

The background error covariance $\mathbf{B}_i$ takes into account the interconnection between the initial condition adjustment in a given state variable at neighboring locations (univariate), and between correlated adjustments in different variables (multi-variate) (Derber and Bouttier, 1999). It effectively spatially smoothes the initial condition adjustment inferred from the inner loop iterations of the Adjoint and Tangent Linear models. In 4DVAR it is impossible to explicitly form $\mathbf{B}_i$ given its size $(O(10^6) \times O(10^6)$ elements in this study), and instead is usually estimated based on an ensemble of model simulations (Li et al., 2008b; Parrish and Derber, 1992) or numerical simulation of diffusion equations (Weaver and Courtier, 2001). The latter approach is implemented in ROMS (Broquet et al., 2009; Powell et al., 2008). It separates $\mathbf{B}_i$ into a multi-variate balance operator (Weaver and Courtier, 2001), background error standard deviations, and a univariate correlation matrix. The correlation matrix is further separated into horizontal and vertical correlations, and each of them is inferred by solving a diffusion equation (2-dimensional for horizontal correlations and 1-dimensional for vertical).

The balance operator in ROMS is under development and was not used in this study, but this does not imply that correlations between state variables are neglected – we hasten to emphasize that dynamical connections between variables are embodied in the Tangent

Linear and Adjoint models. The background error standard deviations that scale the correlation matrix were calculated from a detided 3-month simulation reflecting an assumption that the corrections to the initial conditions should not exceed the magnitude of typical subtidal variability. The background error correlation scales we used in this NYB application were 20 km in the horizontal and 2 m in the vertical, chosen based on scales typical of observed spatial patterns in the region and with care not to over-estimate the scales lest we introduce spurious correlations and over-smoothing in the control variable increments. (Note that in ROMS IS4DVAR the length scale is defined as the distance at which the correlation is $2e^{-1}$.)

5. Results

Figures 6 and 7 show two examples of the DA results. In Figure 6, satellite-measured SST and HF-radar-measured surface current at 0700 UTC 20 April 2006 are compared to their equivalent in the control simulation, to the analysis given by the 10th cycle (3-day DA window commencing 0000 UTC 19 April) and to the forecast launched from the 6th cycle (DA window that ended 0000 UTC 18 April). SST in the control simulation is clearly too warm everywhere and most of the surface current vectors are pointing to the right of the observed. These model errors have largely been removed in the analysis by the DA (Figure 6c) as is to be expected: the SST bias is absent, the Hudson River plume sits at the right location, the warm patch along the New Jersey coast covers the correct area, and most the surface current vectors point to the same direction as observed. The 3-day forecast (Figure 6d) has SST and surface velocity closer to the observations than the

control simulation, but the forecast deviates more from the observations than the analysis, especially the surface current in the Hudson River plume area.

Between April 27 and 29 a glider deployed in the Hudson Shelf Valley (its track is shown in Figure 1) measured temperature and salinity cross-sections that are compared in Figure 7 to the control simulation, the analysis for the 3-day DA window that commenced at 0000 UTC 27 April, and the forecast launched from the DA that ended 0000 UTC 27 April. The control simulation shows about 1°C surface warm bias, 1°C subsurface cold bias, and 0.5 salinity bias at all depths. In the analysis the observed temperature and salinity patterns have largely been corrected (again, as expected, since 4DVAR is matching the solution to these data) except that the subsurface salty bias in the Hudson Shelf Valley becomes worse. In the 3-day forecast, large-scale biases are absent, and while details of spatial patterns depart from the observations (especially surface salinity) the forecast is still clearly superior to the control simulation.

These examples demonstrate that the IS4DVAR system implemented here is capable of bringing the model closer to the observations and giving somewhat improved forecasts compared to a control simulation without DA. Next, we examine statistical measures of the model performance; namely, the reduction of model-observation mismatch in analysis and forecast modes, respectively.

5.1. Model error reduction in analysis periods

Figure 8 shows the cost function (*J,* equation (4)) and cost function gradient norm on each iteration, for all 55 cycles. Each curve is normalized by the value at the beginning of the cycle. The cost function decreases on each iteration of the inner loop, but surges at

the beginning of an outer loop because the new nonlinear model trajectory changes the background state about which the Tangent Linear approximation is expanded and the previous inner loop solution is no longer optimal. As the minimization proceeds, the surge with each new outer loop becomes smaller, indicating the incremental method is converging. Most cycles show about 20% reduction in cost function after 33 iterations (3 outer loops times 11 inner loops), which may seem low but it must be recalled that, on average, 2/3 of the observations in each cycle have been assimilated by previous cycles because of the overlapping DA windows; notice that the normalized cost function curve in Figure 8a for the first cycle (the dashed curves) – when a full 3 days of new observations were assimilated for first time – reduces by more than 50% in 33 iterations. The cost function gradient norm in all cycles shows about 80%-90% reduction in 33 iterations which indicates that the conjugate gradient algorithm has found a minimum. For this application, the curves in Figure 8 and our overall experience suggest there is little added skill to be achieved by iterating to a set convergence tolerance instead of simply for a fixed number of outer and inner loops. Our choice of 33 iterations serves as a guideline for other applications, but may not be universally optimal.

To examine further the reduction of the model-observation mismatch, we compare all observations to the control simulation and the analysis for temperature, salinity and velocity (*u*-component only, the *v*-component results are similar) in Figure 9. The temperature comparison shows the warm bias in the control simulation has been removed in the analysis and the scatter around the diagonal has been reduced. RMS temperature error shows a reduction by 60%. The lowest salinities in the control run are much too fresh, and this is corrected in the analysis; RMS salinity error decreases by 30%. The

errors that remain in salinity occur mostly for ship-borne in situ salinity measurements made in the estuary. The model resolution is too coarse to resolve estuarine processes well so these residual errors are unsurprising. RMS error of the velocity u-component is reduced by 25% through DA, but clearly the scatter remains large. One reason for this is that the variability-to-span ratio of velocity is about 1, which is much larger than that of temperature and salinity. If we assume model error is somewhat proportional to the natural variability, then the ratio of model error to span (which is effectively what the scatter in Figure 9 depicts) would be larger for velocity than for temperature and salinity. A further reason for relatively larger errors in modeled velocity is error in the wind, which will be discussed in next section.

Figure 10 presents the time series of the total cost function and the cost functions of temperature, salinity and velocity computed from the control simulation, the nonlinear model at the beginning of each cycle, and the analysis. Because the background cost function, J_b, is zero at the beginning of the minimization and about one order smaller than the observational cost function, J_o, at the end of each cycle, the time series of cost function in Figure 10 mainly reflects the change of J_o over the experiment period. Assuming the observational and background errors are Gaussian and their covariance **O** and $\mathbf{B}_i$ are described correctly, Chi-squared theory predicts that the minimum value of the cost function is $N_{obs}/2$ with variance $1/N_{obs}$ (Bennett, 2002; Powell et al., 2008; Weaver et al., 2003). It should be noted that because the definition of cost function in *Bennett* (2002) is different from the one in this study and other two referred studies by a ½-factor, the minimum value of cost function in *Bennett* (2002) is N_{obs}. The Chi-squared theory

predicted minimum cost functions are plotted with dashed lines in Figure 10a, but the small variances associated are neglected.

The cost function comparison in Figure 10a shows a big drop of the total cost function from the control simulation to the beginning of each cycle, which is the accumulated effect of the DA cycles. From the beginning to the end of the cycles, i.e. from Φ_0 to analysis, there is another drop of the cost function but to a smaller extent. As mentioned previously it is small because 2/3 of the observations in most of the cycles have been assimilated by previous cycles. In the end, the cost functions given by the analysis simulations are 2 to 3 times larger than the optimal minimum predicted by the Chi-squared theory. Given the fact that the IS4DVAR system used here corrects only initial conditions, and errors in surface forcing, boundary conditions and model physics are neglected, this result is unsurprising. As will be presented in next subsection, there are substantial errors in wind forcing. It certainly contributes to the discrepancies between minimized cost functions and their optimal minimum.

For all variables, DA decreases the mismatch, but as previously reported velocity mismatch decreases the least. This suggests that much of the velocity mismatch falls into the null-space of the DA system and cannot be corrected by adjusting model initial conditions. As we will see in next subsection, it is, at least partially, due to the spatially coherent errors in the wind forcing. But because surface forcing is not part of the control variables the DA system treats that part of the model-observation discrepancy as error in the observations.

5.2. *Effects of wind error*

23

A notable feature in Figure 10 is the spike in total cost function around 21 May that persists after DA. While DA eliminates the model-data mismatch in temperature it does not do so for velocity. This suggests that a significant part of the velocity model-observation mismatch falls into the null-space of the DA system, that is, there are either insufficient data to constrain the control variables or the control variable set is incomplete and cannot adjust the model trajectory in a manner required to match the data. The availability of a sizeable HF radar data set suggests it is unlikely to be the former problem. Suspecting potential errors in wind forcing, which would immediately impact surface currents but less directly affect surface temperature or salinity, we computed the difference between 20 m wind observed at the Ambrose Tower maintained by National Data Buoy Center (Figure 1) and the 10 m NCEP NAM wind used to force the model at the same location (Figure 11). The inconsistency of the height of the wind products is neglected here because the measurements required to compute the 10 m wind from 20 m wind are missing. The comparison between observed and modeled wind quantifies the wind errors only in the vicinity of Ambrose Tower, but should be indicative of the likely magnitude of errors elsewhere, and the potential skill to be derived from adding surface wind forcing to the DA control variables. The magnitude of the wind error averaged over each DA window is plotted in Figure 11 together with the velocity part of the J_o of the control simulation. To eliminate the effect of different numbers of observations, the J_o in Figure 11 is normalized by the number of velocity observations assimilated in each cycle and is therefore equivalent to the mean squared model-observation error.

The number of velocity observations each day varies relatively little (Figure 4b) and the normalized velocity cost functions in Figure 11a preserve the basic temporal variation

seen in Figure 10, including the spike around 21 May. The daily averaged NAM wind

error (Figure 11) shows a corresponding jump around 21 May. The overall correlation

between averaged NAM wind error and normalized velocity J_o is about 0.62 (significant

at 95% confidence level) suggesting that errors in the wind forcing are likely an

important contributor to the model-observation mismatch, especially for velocity.

The power spectrum of surface current measured by HF Radar (Figure 2) shows a

distinct peak within the inertial band, significant at the 95% level, which suggests that in

this area the wind could be an efficient driver of inertial oscillations. Recall that tidal

currents in HF radar data have been replaced with ROMS tidal currents, so we are

essentially assimilating detided, but not low-passed, surface current data. Model-

observation velocity mismatch in the inertial band could be a major part of the total

velocity mismatch. Given the wind forcing errors highlighted above, it is likely the

model reaction to erroneous winds will create velocity mismatch between model and

observation.

We present further evidence for this explanation with a forward model simulation

forced with winds corrected, somewhat, through a simple procedure. We assume error in

NAM model wind has a horizontal scale larger than our model domain (this is certainly

true of the wind fields themselves) and add the difference between NAM-modeled and

Ambrose-measured winds to the NAM winds everywhere in our model domain to obtain

an "improved" forcing wind field. We conducted a forward model simulation that was

identical to the control case except for these modified winds, and compare the normalized

observational cost functions for velocity for these two cases in Figure 11. The simple

wind correction substantially decreases the large model-observation misfit in the cycles around 15 April and 21 May.

This wind correction approach is simplistic and impractical for real-time forecasting. Nevertheless, the exercise clearly demonstrates the potential value of acquiring better forecast winds, or developing better methods of correcting the wind. A natural approach to this in IS4DVAR is to include surface forcing in the control variables of the DA system, and this is the subject of work to be reported in a future publication.

We make one further comment on the effect of errors in the NAM wind in the present DA system: Though the spikes in the velocity mismatch on 21 May remain (Figure 10d), the magnitude has been substantially reduced. This means that the IS4DVAR system has been able, by adjustment of the model initial conditions (the only control variables here), to reduce some of the mismatch that is presumably due to the wind error. Since the DA system cannot differentiate between the sources of the velocity mismatch. Potentially, this degrades the performance of the system, especially with respect to velocity forecast skill, and we will return to this issue in the next section.

5.3. Forecast skill

A primary objective of our study is development of a system suited to practical real-time ocean forecasting, so we present in Figure 12 a statistical measure of the skill of the DA system for each variable. The skill is defined as

$$S = 1 - \frac{\text{RMS}_{\text{afterDA}}}{\text{RMS}_{\text{beforeDA}}}, \tag{7}$$

where RMS is the root-mean-square of model-observation mismatch weighted by observational error, which is equivalent to the square root of J_o. Note that observational data within *forecast* windows are not assimilated by DA, and they are therefore independent data. With this definition, any skill value greater than zero represents an improvement of the model performance, and the maximum possible skill is one. Skill was computed for each day of each analysis and forecast window for all 55 cycles, and the ensembles of 1-day, 2-day etc. forecasts were averaged. The ensemble average and 95% confidence interval for each analysis and forecast day are plotted in Figure 12. The skill of the DA system that assimilates all available observations is denoted by the black curves. In order to diagnose the effect of different data sets, we formed three other DA systems in which we individually withdrew from the assimilation the HF-radar-measured velocity data, glider-measured temperature and salinity data, and satellite-measured SST data. To clearly distinguish these four DA systems we denote them All-data, No-HFradar, No-glider, and No-SST, respectively. The skill values in Figure 12 were computed from the comparison of each modeling system to all observational data irrespective of which data were withdrawn from the DA.

Figure 12a shows that model-observation mismatch in temperature during the analysis period is dramatically reduced (about 70% for the All-data, No-HFradar and No-glider systems, and 40% for the No-SST system) and temperature forecast is substantially improved in all DA systems. The All-data, No-HFradar and No-glider systems have comparable skill for temperature – starting from 0.6 at 1 day and gradually decreasing to 0 at about 14 days into the forecast window. The No-SST system has substantially less skill – starting at 0.4 and dropping to 0 at 5 days into the forecast window.

To further diagnose the impact of different data, we separately consider temperature

skill evaluated in terms of glider-only observations (throughout the full water column)

and satellite SST (surface only) as shown in Figure 13. Skill of subsurface temperature

for the All-data system and No-HFradar systems is 0.5 at 1 day and drops to 0 at 7 days.

Evidently, skill is better beyond 8 days if surface velocity from HF radar is not

assimilated. When glider observations are withdrawn, the subsurface temperature skill in

the analysis window is poor, but the model subsequently gains some skill in the first

week forecast. This gain is presumably caused by assimilating SST data, the information

in which propagates downward through the water column due to physics. That SST data

add subsurface skill is demonstrated by noting that when SST data are withdrawn (green

line in Figure 13a) the subsurface temperature skill drops for the entire forecast period

compared to the All-data case.

In Figure 13b, the All-data, No-HFradar, and No-glider systems show the same

performance at forecasting SST, with skill exceeding 0.4 for the entire forecast period.

When satellite-measured SST data are not assimilated, the DA system has no skill at

forecasting SST at all. SST observations are therefore vital to subsequently forecasting

SST, while other data sources are not. This suggests that ocean surface temperature is

mainly controlled by SST initial conditions and, presumably, surface heat exchange, with

initial subsurface temperature and other variables exerting relatively little influence.

Figure 13b also gives the persistence skill of the satellite SST observations, which is

the only persistence skill we can practically calculate given the available data set. The

persistence skill, a commonly used standard of reference for measuring the accuracy of

the forecasts (Di Lorenzo et al., 2007; Murphy, 1992), is computed by adopting a SST

map as the forecast of subsequent fields of SST. The persistence skill starts around 0.6 for 1 day and quickly approaches zero after 3 days into the forecast window. The comparison between the persistence skill and the skill of the All-data system in Figure 13b shows the IS4DVAR system is able to use the observations to correct the model dynamics and give a smooth model trajectory through the model integration, rather than just fitting the observations.

The relative performance of the DA systems with respect to salinity is largely similar to temperature, but with some differences. Figure 12b shows that the All-data DA system achieves a salinity model-observation skill of 0.4 during analysis, which is substantially less than the skill for temperature; salinity skill is about 0.3 at 1 day and about 0.1 at 3 days into forecast. Thereafter, skill stays around 0.1. Comparing skills of the different DA systems we see that assimilating glider-measured subsurface data improves salinity skill for the entire period, assimilating SST data actually degrades the salinity skill, and assimilating surface velocity data has minimal impact on salinity. The adverse impact that assimilating SST data has on the salinity forecast reflects the loss of freedom when the DA system adjusts initial conditions to fit more observations. When SST data are withdrawn, the DA system need not reconcile glider and satellite temperatures and has rather more freedom to adjust initial salinity to improve the salinity analysis – recall that all the variables are dynamically linked through the adjoint and tangent linear models. However, the associated loss of temperature skill is significant, and emphasizes that, on balance, diversity in the data sources is to be preferred.

We also note that the adjoint sensitivity analysis of Zhang et al. (2009b) showed temperature and salinity interact in subtle ways in this coastal circulation regime because

they affect density stratification and therefore baroclinic pressure gradients and vertical mixing. Through the adjoint, IS4DVAR modifies not only the tracer conditions upstream but also the transport dynamics. Having a variety of data types for assimilation helps constrain both influences, decreasing the null-space of the DA system which might otherwise impose increments to control variables that subsequently have a negative impact on the forecast.

Figures 12c,d show that the All-data system achieves a velocity skill of about 0.45 in the analysis window but has velocity skill above 0 only for 2-3 days into the forecast. A more rapid decline in skill for velocity compared to temperature and salinity is expected, given that autocorrelation timescales for velocity are always less than for tracers indicating they are inherently less predictable. Decreased velocity skill will also result from the errors in wind forcing noted in the previous subsection. The similarity in velocity skill in the All-data, No-glider, and No-SST systems suggests that assimilating temperature and salinity data contributes little to the improvement of the model's velocity prediction. The difference between All-data and No-HFradar systems, however, shows that assimilating HF-radar-measured surface currents does improve velocity predictability by 1-2 days.

To examine changes in skill over time, we plot the ratios after DA (the All-data system) to before DA of RMS error and cross-correlation error $(1-CC)$ for different variables (Figure 14). Both RMS and CC are obtained from the comparison of all available observational data on a given day to the relevant model realization. The results are plotted as a function of start date for each forecast cycle (abscissa), and days into the forecast window (ordinate). Each 45-degree tilted line therefore depicts a single DA cycle,

all values with ordinate less than 0 are within analysis periods, and all values at the same abscissa value represent different forecasts of the same date. For both quantities plotted, a ratio less than 1 means DA improves the model. In these plots, values consistently greater than 1 on the same date mean that date was never forecast well regardless of when the forecast was launched, whereas values greater than 1 following a 45-degree line mean that forecast cycle always gave poor results.

The ratios of RMS error and CC error for temperature are much smaller than 1 in the analysis window for almost all cycles. In the forecasts, the RMS error ratio remains less than 1 for most of the cycles except several days around 9 May and 3 June. The CC error shows similar performance but is a more critical skill metric and shows ratio greater than 1 more frequently. DA decreases RMS error for salinity for most of the time, though the period around 25 April is notably poor. No forecast launched prior to 25 April was able to produce a salinity prediction for 24-26 April that was better than the no-assimilation case, the salinity analysis itself for 24-26 April is poor, and the forecast launched from that analysis is not skillful. April 25 is a time of peak in Hudson River discharge (Figure 5) but this does not itself explain the lackluster model performance, because from previous studies (Zhang et al., 2009a) we expect the model to have some skill at simulating the river plume trajectory.

The occurrence of high ratios for RMS and CC error in salinity during some periods is a concern because it indicates the DA system might degrade the forecast compared to a conventional no-assimilation forward model, but interpretation may be affected by the sampling distribution for salinity which is not extensive, and is quite heterogeneous. Consider that *in situ* observations include ship-towed undulating CTD data during 2 to 8

May (Figure 4) – the time period when salinity appears to be consistently poorly forecast as judged by the CC error (Figure 14d). The vessel cruise track (Figure 1) samples regions where salinity is not observed by any other instruments during the experiment, and it is plausible that the introduction of these data to the forecast verification data set beginning 2 May reveals forecast errors that were previously unknown because of a lack of data to identify them. Inspection of Figure 14d suggests that after the towed CTD data have been incorporated by the DA analysis during 2–8 May, the forecasts launched thereafter do rather better. Thus the irregular space-time sampling pattern for salinity may be producing a misleading forecast skill assessment here if much of the data falls into the null-space of the DA system, i.e. where the unavailability of observations means the analysis step has had no opportunity to improve the model state and subsequent forecasts. We cannot rigorously test this conjecture until a more extensive in situ observation network is available.

For the velocity components, the RMS and CC error ratios are less than 1 in all analysis cycles, but rather quickly rise in the forecast period, consistent with the results above for overall forecast skill. The error ratios do not reach the extremes noted during some cycles for salinity, but this may simply indicate that when the model loses velocity forecast skill the error variances with and without assimilation are comparable and the ratio remains of order unity. The sampling distribution for HF-radar is more extensive and consistent that for SST and gliders.

6. Summary and conclusions

As part of a long-term project building an integrated observation and modeling system for the New York Bight for the purposes of coastal ocean prediction and observing system design, this study has evaluated 4-Dimensional Variational data assimilation using ROMS in a realistic and pseudo real-time setup. In an accompanying article (Part II) we describe further results of the project on observing system evaluation and design.

In this study we assimilated all available observations of temperature, salinity and velocity collected by variety of platforms in spring 2006 during a campaign of field observation targeting the Hudson River plume as it flows into the New York Bight and is dispersed across the New Jersey inner shelf. All observations were preprocessed for quality control and, where appropriate, binned or averaged to be comparable to model spatial scales. Errors in the observations were assumed to be independent, and an error standard deviation was assigned to each observation according to instrument accuracy, model representation of observed physical processes, and the convergence of the DA system. ROMS IS4DVAR was applied with a 3-day DA analysis window in an overlapped cycling system to adjust initial conditions for a new forecast every day. This mode of implementation is standard practice in Numerical Weather Prediction and represents a practical approach to formulating a real-time ocean forecast system. The background error covariance that is an important component of the IS4DVAR was assumed univariate with 20 km horizontal and 2 m vertical decorrelation scales.

System performance was evaluated by examining model-observation mismatch in the analysis and forecast periods. Time series of cost function and its gradient norms indicate that ROMS IS4DVAR successfully reduces the model-observation mismatch in analysis

periods and converges towards the cost function minimum. The overall mismatch reduction is about 60% for temperature, 30% for salinity and 25% for velocity. Time series of the cost functions show that DA was able to reduce the total cost function and cost functions of all variables for all 55 cycles, but the obtained minimum of total cost functions are 2-3 times larger than the optimal minima given by the Chi-squared theory. Because the IS4DVAR system used in this study corrects only initial conditions, with errors in other external input and model physics all neglected, this result is unsurprising.

A correlation was found between errors in wind forcing and model-observation mismatch in velocity, suggesting that a significant proportion of the velocity error might results from error in the winds used to force the model. Because surface forcing is not part of the control variables and is then assumed to be true, the DA system treats the associated surface velocity mismatch as "observational error". This "observational error" would then have a large spatial correlation scale if inertial oscillations are the physical process causing the velocity error from the wind error as suggested by the power spectrum of the surface currents. This may also contribute to the large cost function minimum relative to the optimal minimum given by the Chi-square theory because it breaks the assumption of independent error in the surface velocity observations and then causes the underestimation of the observational error covariance. An improved forward model simulation forced with winds partially corrected by a simple procedure supports the conjecture, and highlights the importance of having accurate wind forcing in coastal ocean DA, especially if predicting surface velocity is a key objective. This also exhibits the necessity of correcting surface forcing by adding it into the control variables of the DA system in order to have a good analysis product, especially of surface velocity.

Skill of the DA system in improving model forecasts was assessed by comparing 15-day forecasts of each cycle with unassimilated independent observations. Overall, the DA system has positive skill for about 15 days forecast for temperature and salinity and 2 to 3 days forecast for velocity. Three DA sub-systems that withdrew the observations from different instruments were used to reveal the effects different data sets have on the skill. Assimilating satellite-measured SST was shown to improve not only the surface temperature forecast but also the forecast of subsurface temperature. However, satellite SST assimilation evidently somewhat impairs the improvement of salinity forecast. Assimilating glider measurements significantly improves the salinity forecast but has little effect on the temperature forecast. Assimilating HF radar surface current data extends by 1 to 2 days the time period for which the velocity forecast is improved, even with the errors in the wind forcing. Assimilating HF radar currents somewhat impairs the forecast of subsurface temperature.

The meteorological forcing we use in this study is a concatenation of the first 24 hours of each NAM forecast cycle and is presumably superior to the true 72-hour forecast in a real-time system. Therefore, the results presented here provide an upper bound for the performance of a real-time ocean prediction system if the same machinery and set-up were used operationally.

This study demonstrates that ROMS IS4DVAR Data Assimilation has the capability to use a large and diverse set of observations of the type increasingly available from practical coastal ocean observing systems, reduce model-observation mismatch in the analysis period, and subsequently provide improved forecasts for 2 to 15 days depending on the forecast variable. It also reveals some of the practicalities of numerical ocean

prediction in real-time with data assimilation: (i) preprocessing of the observational data must be conducted in a timely manner and in a way consistent with model resolution and assumptions made about observational error, (ii) the meteorological conditions used to force ocean model ought to be as accurate as possible, especially for better prediction of current, and (iii) the choice of some of the parameters in the data assimilation system, such as observational error standard deviations, background error decorrelation scales and standard deviations should be based on a thorough understanding of the local physics and the model used in the data assimilation, while the choice of model resolution and number of inner and outer loops, will be dictated by the targeted oceanic processes and available computational resources.

The analysis here provides some general guidelines on the design of oceanic observing systems, which we consider further and from a different prospective in the Part II paper. The spatial scale of the errors in model and observations should be taken into account when considering the spacing of the observations. Although observations on scales smaller than those of modeled and observational errors provide detailed information about ocean physics, they are of little merit for data assimilation. Given limited resources, diversity in observed variables and large and stable coverage of observation in space ought to be emphasized because these both diminish the null-space and bolster the skill of the data assimilation system.

References

Adams, D. A., J. S. O'Connor, and S. B. Weisberg, 1998: Sediment quality of the NY/NJ harbor systemEPA/902-R-98-001.

Beardsley, R. C. and W. C. Boicourt, 1981: On estuarine and continental shelf circulation in the Middle Atlantic Bight. Evolution of Physical Oceanography, B. A. Warren and C. Wunsch, Eds., The MIT Press, 198-223.

Bennett, A. F., 2002: Inverse Modeling of the Ocean and Atmosphere. Cambridge University Press, 234 pp.

Broquet, G., C. A. Edwards, A. M. Moore, B. S. Powell, M. Veneziani, and J. D. Doyle, 2009: Application of 4D-Variational data assimilation to the California Current System. Dyn. Atmos. Oceans, in press.

Cahill, B., O. Schofield, R. Chant, J. Wilkin, E. Hunter, S. Glenn, and P. Bissett, 2008: Dynamics of turbid buoyant plumes and the feedbacks on near-shore biogeochemistry and physics. Geophys. Res. Lett., **35,** L10605, doi:10.1029/2008GL033595.

Castelao, R. M., O. Schofield, S. Glenn, R. J. Chant, and J. Kohut, 2008: Cross-shelf transport of fresh water on the New Jersey Shelf. J. Geophys. Res., **113,** C07017.

Chant, R. J., S. M. Glenn, E. Hunter, J. Kohut, R. F. Chen, R. W. Houghton, J. Bosch, and O. Schofield, 2008: Bulge formation of a buoyant river flow. J. Geophys. Res., **113,** C01017, doi:10:1029/2007JC004100.

Chapman, D. C., 1985: Numerical treatment of cross-shelf open boundaries in a barotropic ocean model. J. Phys. Oceanogr., **15,** 1060-1075.

Choi, B.-J. and J. L. Wilkin, 2007: The effect of wind on the dispersal of the Hudson River plume. J. Phys. Oceanogr., **37,** 1878-1897.

Courtier, P., 1997: Dual formulation of four-dimensional variational assimilation. Q. J. R. Meteorol. Soc., **123,** 2449-2461.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4DVAR using an incremental approach. Q. J. R. Meteorol. Soc., **120,** 1367-1388.

Derber, J. and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. Tellus A, **51,** 195-221.

Di Lorenzo, E., A. M. Moore, H. G. Arango, B. D. Cornuelle, A. J. Miller, B. Powell, B. S. Chua, and A. Bennett, 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 Model., **16,** 160-187.

Evensen, G., 2007: Data Assimilation: The Ensemble Kalman Filter. Springer, 208 pp.

Fairall, C. W., E. F. Bradley, J. E. Hare, A. A. Grachev, and J. Edson, 2003: Bulk Parameterization of Air–Sea Fluxes: Updates and Verification for the COARE Algorithm. J. Clim., **16,** 571-591.

Fennel, K., J. Wilkin, M. Previdi, and R. Najjar, 2008: Denitrification effects on air-sea CO_2 flux in the coastal ocean: Simulations for the Northwest North Atlantic. Geophys. Res. Lett., **35,** L24608.

Fennel, K., J. Wilkin, J. Levin, J. Moisan, J. O'Reilly, and D. Haidvogel, 2006: Nitrogen cycling in the Middle Atlantic Bight: Results from a three-dimensional model and implications for the North Atlantic nitrogen budget. Global Biogeochem. Cy., **20,** doi:10.1029/2005GB002456.

Flather, R. A., 1976: A tidal model of the northwest European continental shelf. Mem. Soc. Roy. Sci. Liege, Ser. 6, **10,** 141-164.

Glenn, S. M. and O. Schofield, 2003: Observing the oceans from the COOLroom: Our history, experience. Oceanography, **16,** 37-52.

Glenn, S. M., J. F. Grassle, and C. J. von Alt, 2000: A well sampled ocean: the LEO approach. Oceanus, **42,** 28-30.

Haidvogel, D. B., H. Arango, W. B. Budgell, B. D. Cornuelle, E. Curchitser, E. Di Lorenzo, K. Fennel, W. R. Geyer, A. J. Hermann, L. Lanerolle, J. Levin, J. C. McWilliams, A. J. Miller, A. M. Moore, T. M. Powell, A. F. Shchepetkin, C. R. Sherwood, R. P. Signell, J. C. Warner, and J. Wilkin, 2008: Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the regional ocean modeling system. J. Comput. Phys., **227,** 3595-624.

Hoffman, R. N., R. M. Ponte, E. J. Kostelich, A. Blumberg, I. Szunyogh, S. V. Vinogradov, and J. M. Henderson, 2008: A simulation study using a local ensemble transform Kalman filter for data assimilation in New York Harbor. J. Atmos. Oceanic Technol., **25,** 1638-1656.

Hoteit, I. and A. Köhl, 2006: Efficiency of reduced-order, time-dependent adjoint data assimilation approaches. J. Oceanogr., **62,** 539-550.

Howarth, R. W., R. Marino, D. P. Swaney, and E. W. Boyer, 2006: Wastewater and watershed influences on primary productivity and oxygen dynamics in the lower Hudson River estuary. The Hudson River Estuary, J. S. Levington and J. R. Waldman, Eds., Cambridge University Press, 121-139.

Johnson, D. R., J. Miller, and O. Schofield, 2003: Dynamics and optics of the Hudson River outflow plume. J. Geophys. Res., **108,** 3323, doi:10.1029/2002JC001485.

Kohut, J. T., S. M. Glenn, and D. E. Barrick, 2001: multiple HF-radar system development for a regional longterm ecosystem observatory in the New York Bight. American Meteorological Society: Fifth Symposium on Integrated Observing Systems, Albuquerque, New Mexco, 4-7.

Kohut, J. T., H. J. Roarty, and S. M. Glenn, 2006a: Characterizing Observed Environmental Variability with HF Doppler Radar Surface Current Mappers and Acoustic Doppler Current Profilers: Environmental Variability in the Coastal Ocean. IEEE J. Oceanic Eng., **31,** 876-884.

Kohut, J. T., S. M. Glenn, and J. Paduan, 2006b: The inner-shelf response to tropical storm Floyd. J. Geophys. Res., **111,** C09S91.

Kurapov, A. L., G. D. Egbert, J. S. Allen, and R. N. Miller, 2007: Representer-based variational data assimilation in a nonlinear model of nearshore circulation. J. Geophys. Res., **112,** C11019.

Lentz, S. J., 2008: Observations and a model of the mean circulation over the Middle Atlantic Bight continental shelf. J. Phys. Oceanogr.**,** 1203-1221.

Lewis, J. K., I. Shulman, and A. F. Blumberg, 1998: Assimilation of Doppler radar current data into numerical ocean models. Cont. Shelf Res., **18,** 541-559.

Li, Z., Y. Chao, J. C. McWilliams, and K. Ide, 2008a: A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System. J. Atmos. Oceanic Technol., **25,** 2074-2090.

——, 2008b: A three-dimensional variational data assimilation scheme for the Regional Ocean Modeling System: implementation and basic experiments. J. Geophys. Res., **113,** C05002.

Lueck, R. G., 1990: Thermal inertia of conductivity cells: theory. J. Atmos. Oceanic Technol., **7,** 741-755.

Moline, M. A., T. K. Frazer, R. J. Chant, S. Glenn, C. Jacoby, J. R. Reinfelder, J. Yost, M. Zhou, and O. M. Schofield, 2008: Biological responses in a dynamic, buoyant river plume. Oceanography, **21,** 70-89.

Moore, A. M., H. G. Arango, E. Di Lorenzo, A. J. Miller, and B. D. Cornuelle, 2008: An adjoint sensitivity analysis of the Southern California Current circulation and ecosystem. Part I: The physical circulation. J. Phys. Oceanogr.**,** in press.

Moore, A. M., H. G. Arango, E. Di Lorenzo, B. D. Cornuelle, A. J. Miller, and D. J. Neilson, 2004: A Comprehensive Ocean Prediction and Analysis System Based on the Tangent Linear and Adjoint of a Regional Ocean Model. Ocean Model., **7,** 227-258.

Morison, J., R. Andersen, N. Larson, E. D'Asaro, and T. Boyd, 1994: The correction for thermal-lag effects in Sea-Bird CTD data. J. Atmos. Oceanic Technol., **11,** 1151-1164.

Münchow, A. and R. J. Chant, 2000: Kinematics of inner shelf motion during the summer stratified season off New Jersey. J. Phys. Oceanogr., **30,** 247-268.

Mukai, A. Y., J. J. Westerink, R. A. Luettich, and D. Mark, 2002: Eastcoast 2001, A tidal constituent database for the western North Atlantic, Gulf of Mexico and Caribbean Sea, Tech. Rep. ERDC/CHL TR-02-24, 196 pp.

Murphy, A. H., 1992: Climatology, persistence, and their linear combination as standards of reference in skill scores. Weather Forecast., **7,** 692-698.

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

Oke, P. R., J. S. Allen, R. N. Miller, G. D. Egbert, and P. M. Kosro, 2002: Assimilation of surface velocity data into a primitive equation coastal ocean model. J. Geophys. Res., **107,** 3122.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. Mon. Weather Rev., **120,** 1747-1763.

Pawlowicz, R., B. Beardsley, and S. Lentz, 2002: Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Comput. Geosci., **28,** 929-937.

Powell, B. S., H. G. Arango, A. M. Moore, E. Di Lorenzo, R. F. Milliff, and D. Foley, 2008: 4DVAR data assimilation in the Intra-Americas Sea with the Regional Ocean Modeling System (ROMS). Ocean Model., **25,** 173-188.

Rutledge, G. K., J. Alpert, and W. Ebisuzaki, 2006: NOMADS: A climate and weather model archive at the National Oceanic and Atmospheric Administration. Bull. Am. Meteorol. Soc., **87,** 327-341.

Schofield, O., J. Kohut, D. Aragon, L. Creed, J. Graver, C. Haldeman, J. Kerfoot, H. Roarty, C. Jones, D. Webb, and S. M. Glenn, 2007: Slocum gliders: robust and ready. J. Field Robot., **24,** 1-14.

Schofield, O., R. Chant, B. Cahill, R. Castelao, D. Gong, A. Kahl, J. Kohut, M. Montes-Hugo, R. Ramadurai, P. Ramey, Y. Xu, and S. M. Glenn, 2008: The decadal view of the

Mid-Atlantic Bight from the COOLroom: Is our coastal system changing? Oceanography, **21,** 108-117.

Scott, R. K., J. S. Allen, G. D. Egbert, and R. N. Miller, 2000: Assimilation of surface current measurements in a coastal ocean model. J. Phys. Oceanogr., **30,** 2359-2378.

Shchepetkin, A. F. and J. C. McWilliams, 1998: Quasi-monotone advection schemes based on explicit locally adaptive diffusion. Mon. Weather Rev., **126,** 1541-1580.

——, 2003: A Method for Computing Horizontal Pressure-Gradient Force in an Oceanic Model with a Non-Aligned Vertical Coordinate. J. Geophys. Res., **108,** 3090, doi:10.1029/2001JC001047.

——, 2005: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. Ocean Model., **9,** 347-404.

Smith, S. R. and H. E. Ngodock, 2008: Cycling the Representer Method for 4D-variational data assimilation with the Navy Coastal Ocean Model. Ocean Model., **24,** 92-107.

Stammer, D., K. Ueyoshi, A. Köhl, W. G. Large, S. A. Josey, and C. Wunsch, 2004: Estimating air-sea fluxes of heat, freshwater, and momentum through global ocean data assimilation. J. Geophys. Res., **109,** C05023.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. N. Hill, and J. Marshall, 2002: Global ocean circulation during 1992-1997, estimated from ocean observations and a general circulation model. J. Geophys. Res., **107,** 3118.

——, 2003: Volume, heat, and freshwater transport of the global ocean circulation 1993-2000, estimated from a general circulation model constrained by World Ocean Circulation Experiment (WOCE) data. J. Geophys. Res., **108,** 3007.

Stewart, R. H. and J. W. Joy, 1974: HF radio measurements of surface currents. Deep-Sea Res., **21,** 1039-1049.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with adjoint vorticity equation. I: Theory. Q. J. R. Meteorol. Soc., **113,** 1321-1328.

Tilburg, C. E. and R. W. Garvine, 2003: Three-dimensional flow in a coastal upwelling zone: Convergence and divergence on the New Jersey Shelf. J. Phys. Oceanogr., **33,** 2113-2115.

Umlauf, L. and H. Burchard, 2003: A generic length-scale equation for geophysical turbulence models. J. Mar. Res., **61,** 235-265.

Vialard, J., A. T. Weaver, D. L. T. Anderson, and P. Delecluse, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part II: Physical validation. Mon. Weather Rev., **131,** 1379-1395.

Vossepoel, F. C., A. T. Weaver, J. Vialard, and P. Delecluse, 2004: Adjustment of near-equatorial wind stress with four-dimensional variational data assimilation in a model of the Pacific Ocean. Mon. Weather Rev., **132,** 2070-2083.

Warner, J. C., C. R. Sherwood, H. G. Arango, R. P. Signell, and B. Butman, 2005: Performance of four turbulence closure models implemented using a generic length scale method. Ocean Model., **8,** 81-113.

Weaver, A. T. and P. Courtier, 2001: Correlation modelling on the sphere using a generalized diffusion equation. Q. J. R. Meteorol. Soc., **127,** 1815-1846.

Weaver, A. T., J. Vialard, and D. L. T. Anderson, 2003: Three- and four-dimensional variational assimilation with an ocean general circulation model of the tropical Pacific Ocean. Part 1: Formulation, internal diagnostics and consistency checks. Mon. Weather Rev., **131,** 1360-1378.

Weaver, A. T., C. Deltel, E. Machu, S. Ricci, and N. Daget, 2005: A multivariate balance operator for variational ocean data assimilation. Q. J. R. Meteorol. Soc., **131,** 3605-3625.

Wilkin, J. L., H. G. Arango, D. B. Haidvogel, C. S. Lichtenwalner, S. M. Glenn, and K. S. Hedström, 2005: A Regional Ocean Modeling System for the Long-term Ecosystem Observatory. J. Geophys. Res., **110,** doi:10.1029/2003JC002218.

Wong, K. C., 1999: The wind driven currents on the Middle Atlantic Bight inner shelf. Cont. Shelf Res., **19,** 757-773.

Wunsch, C., 2006: Discrete inverse and state estimation problems with geophysical fluid applications. Cambridge University Press, 371 pp.

Wunsch, C. and P. Heimbach, 2007: Practical global oceanic state estimation. Physica D, **230,** 197-208.

Yankovsky, A. E., 2003: The cold-water pathway during an upwelling event on the New Jersey shelf. J. Phys. Oceanogr., **33,** 1954-1966.

Yankovsky, A. E. and R. W. Garvine, 1998: Subinertial Dynamics on the Inner New Jersey Shelf during the Upwelling Season. J. Phys. Oceanogr., **28,** 2444-2458.

Yankovsky, A. E., R. W. Garvine, and A. Munchow, 2000: Mesoscale Currents on the Inner New Jersey Shelf Driven by the Interaction of Buoyancy and Wind Forcing. J. Phys. Oceanogr., **30,** 2214-2230.

Zhang, W. G., J. L. Wilkin, and R. J. Chant, 2009a: Modeling the pathways and mean dynamics of river plume dispersal in New York Bight. J. Phys. Oceanogr., **39,** 1167-1183.

Zhang, W. G., J. L. Wilkin, J. C. Levin, and H. G. Arango, 2009b: An Adjoint Sensitivity Study of Buoyancy- and Wind-driven Circulation on the New Jersey Inner Shelf. J. Phys. Oceanogr., **39,** 1652-1668.
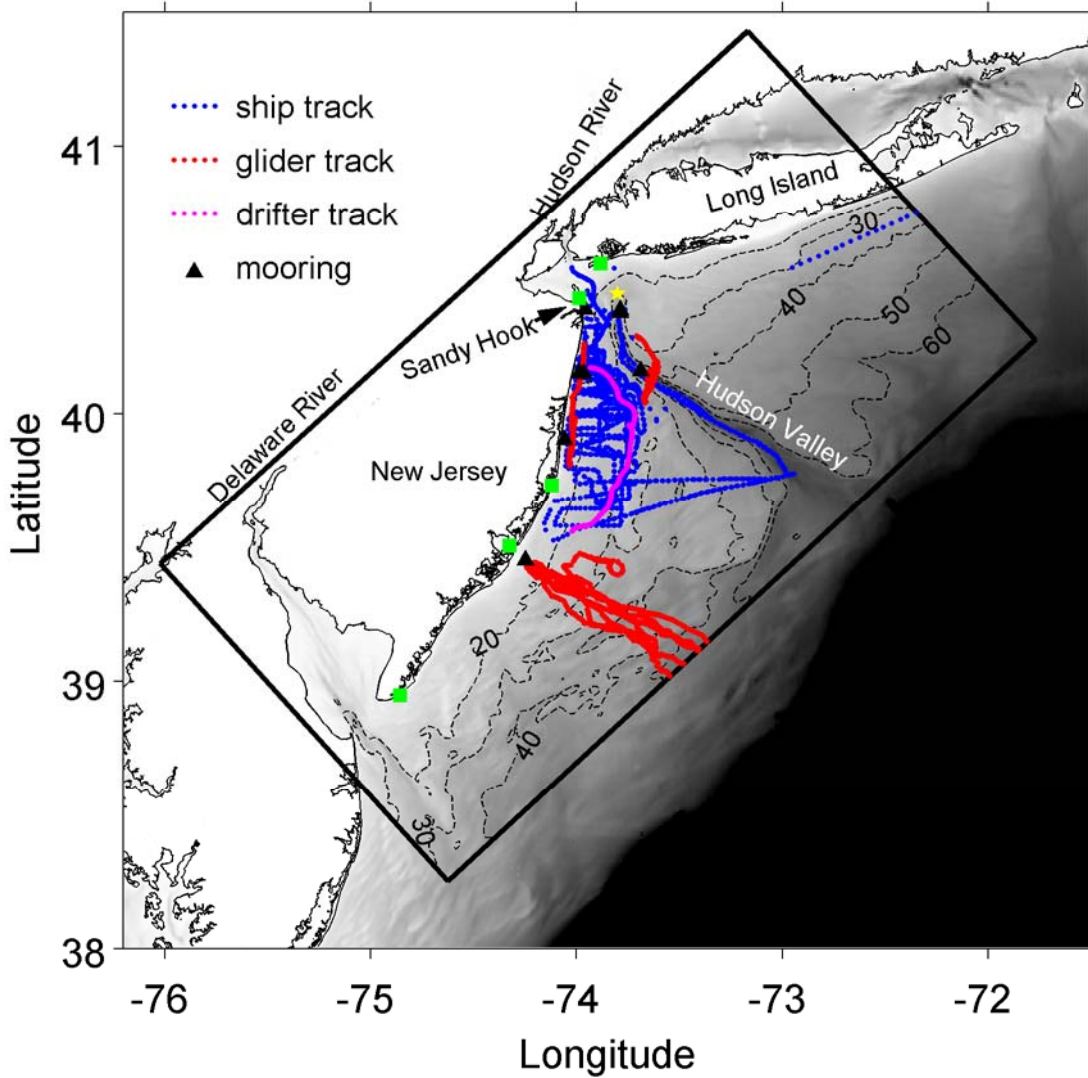
Figure Captions

Figure 1. The study domain and observation locations. The black frame indicates the model domain; Bathymetry of the New York Bight is in grayscale; Black dash lines are contours of model isobaths in meters; the yellow pentagram indicates the location of Ambrose Tower; the green squares indicate the locations of five HF Radar stations.

Figure 2. Averaged spectrum of HF-Radar-measured surface current. Dash lines indicate local inertial frequency band and the confidence limit applies to data within the inertial frequency band.

Figure 3. Comparison between HF Radar observed and modeled M_2 tide.

Figure 4 Types (a) and numbers (b) of observations over the data assimilation period (10 April - 5 June, 2006).

Figure 5. River discharges (a) and zonal (b) and meridional (c) components of the wind at the Hudson River mouth over the experiment period.

Figure 6. Comparison of observed and modeled sea surface temperature and current at 0700 UTC 20 April 2006.

Figure 7. Comparison of glider-measured and modeled temperature and salinity along a glider track between 27 and 29 April, 2006 (The red line across the Hudson Shelf Valley in Figure 1).

Figure 8. Normalized cost functions (a) and cost function gradient norm (b) at each iteration of all the 55 data assimilation cycles. The normalization is achieved through dividing the cost functions and cost function gradient norms by their value at the

beginning of each DA cycle. The dashed curves indicate the change of the cost function in the first DA cycle and the vertical doted lines separate the outer loops.

Figure 9. 2-D histogram of the comparison between observed and modeled temperature, salinity, and u component of the velocity for model before (control simulation) and after (analysis simulation) data assimilation. The color indicates the $\log_{10}$ of the number of observations.

Figure 10. Cost function of the control run, at the beginning of each cycle and of the analysis and the Chi-squared-theory-predicted optimal minimum of cost function of each cycle.

Figure 11. Magnitude of NAM wind error and normalized velocity model-observation misfit of the control run before and after the wind correction. All misfits are normalized by the number of corresponding observations assimilated in each cycle.

Figure 12. Ensemble average of the skill of different DA systems over analysis and forecast periods for individual forecast variables. Vertical bars on symbols indicate 95% confidence intervals. Vertical dashed lines denote the boundary between analysis window and forecast window.

Figure 13. Ensemble average skill of different DA systems over analysis and forecast periods for (a) glider-measured temperature and (b) satellite-measured SST. Vertical bars indicate 95% confidence. Vertical dashed lines denote the boundary between analysis window and forecast window.

Figure 14. Ratios after data assimilation to before data assimilation of RMS error and

cross-correlation error (1−CC) at each day of all cycles for the DA system assimilating all

observational data. Thick white lines are contours of value 1.

Table 1. Observational Error Representation

Observational platform	Satellite	HF Radar	Glider	Mooring	Drifters	Shipborne
Velocity (m s^{-1})	----	0.05	----	0.02	----	0.06
Temperature (^{o}C)	0.4	----	0.4	0.4	0.3	0.6
Salinity	----	----	0.4	0.4	----	0.6

Figure 1. The study domain and observation locations. The black frame indicates the model domain; Bathymetry of the New York Bight is in grayscale; Black dash lines are contours of model isobaths in meters; the yellow pentagram indicates the location of Ambrose Tower; the green squares indicate the locations of five HF Radar stations.

Figure 2. Averaged spectrum of HF-Radar-measured surface current. Dash lines indicate local inertial frequency band and the confidence limit applies to data within the inertial frequency band.

Figure 3. Comparison between HF Radar observed and modeled M_2 tide.

Figure 4 Types (a) and numbers (b) of observations over the data assimilation period (10 April - 5 June, 2006).

Figure 5. River discharges (a) and zonal (b) and meridional (c) components of the wind at the Hudson River mouth over the experiment period.

Figure 6. Comparison of observed and modeled sea surface temperature and current at 0700 UTC 20 April 2006.

Figure 7. Comparison of glider-measured and modeled temperature and salinity along a glider track between 27 and 29 April, 2006 (The red line across the Hudson Shelf Valley in Figure 1).

Figure 8. Normalized cost functions (a) and cost function gradient norm (b) at each iteration of all the 55 data assimilation cycles. The normalization is achieved through dividing the cost functions and cost function gradient norms by their value at the beginning of each DA cycle. The dashed curves indicate the change of the cost function in the first DA cycle and the vertical doted lines separate the outer loops.

Figure 9. 2-D histogram of the comparison between observed and modeled temperature, salinity, and u component of the velocity for model before (control simulation) and after (analysis simulation) data assimilation. The color indicates the $\log_{10}$ of the number of observations.

Figure 10. Cost function of the control run, at the beginning of each cycle and of the analysis and the Chi-squared-theory-predicted optimal minimum of cost function of each cycle.
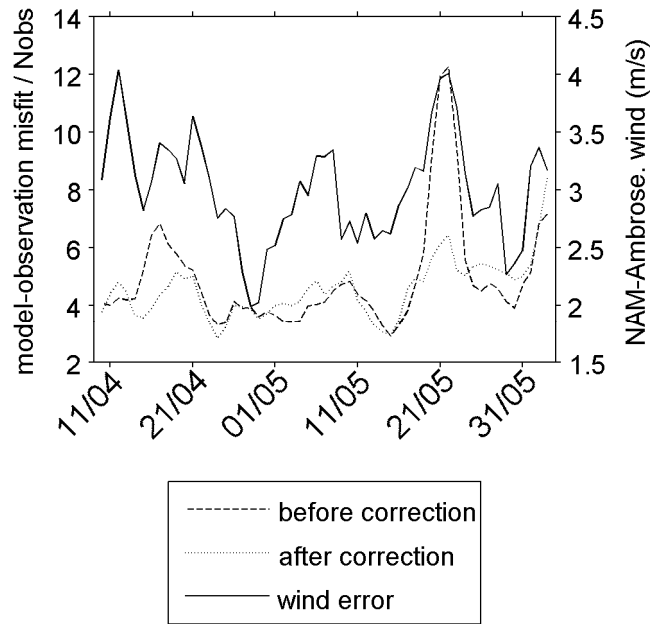
Figure 11. Magnitude of NAM wind error and normalized velocity model-observation misfit of the control run before and after the wind correction. All misfits are normalized by the number of corresponding observations assimilated in each cycle.
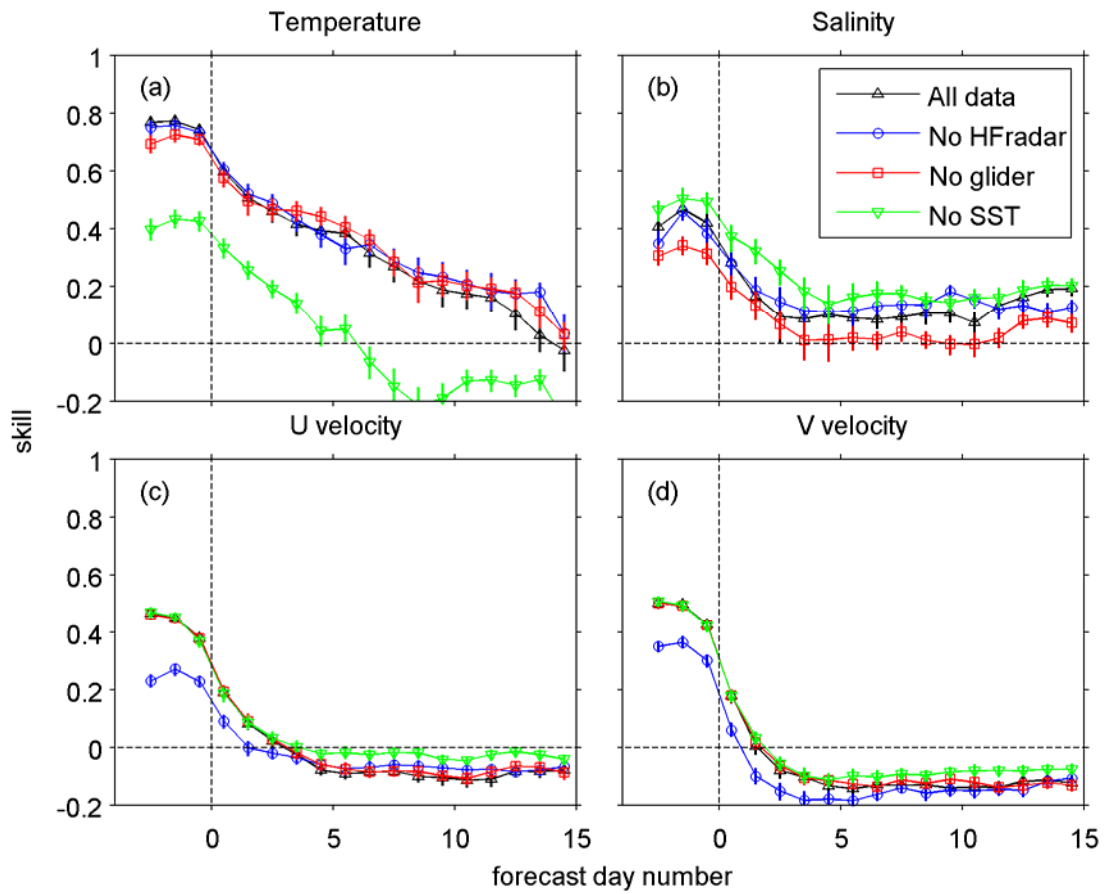
Figure 12. Ensemble average of the skill of different DA systems over analysis and forecast periods for individual forecast variables. Vertical bars on symbols indicate 95% confidence intervals. Vertical dashed lines denote the boundary between analysis window and forecast window.
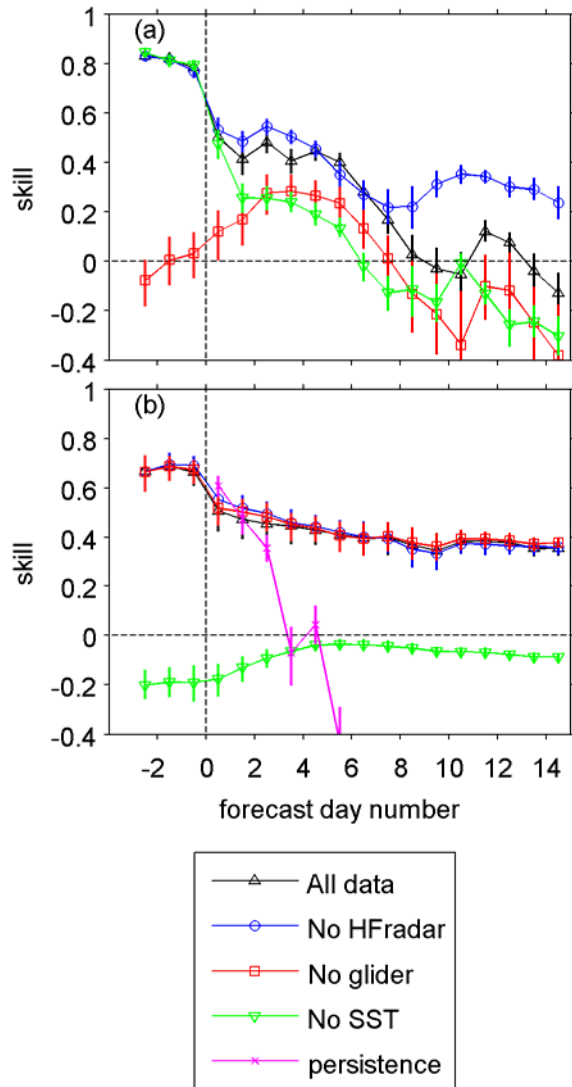
Figure 13. Ensemble average skill of different DA systems over analysis and forecast periods for (a) glider-measured temperature and (b) satellite-measured SST. Vertical bars indicate 95% confidence. Vertical dashed lines denote the boundary between analysis window and forecast window.
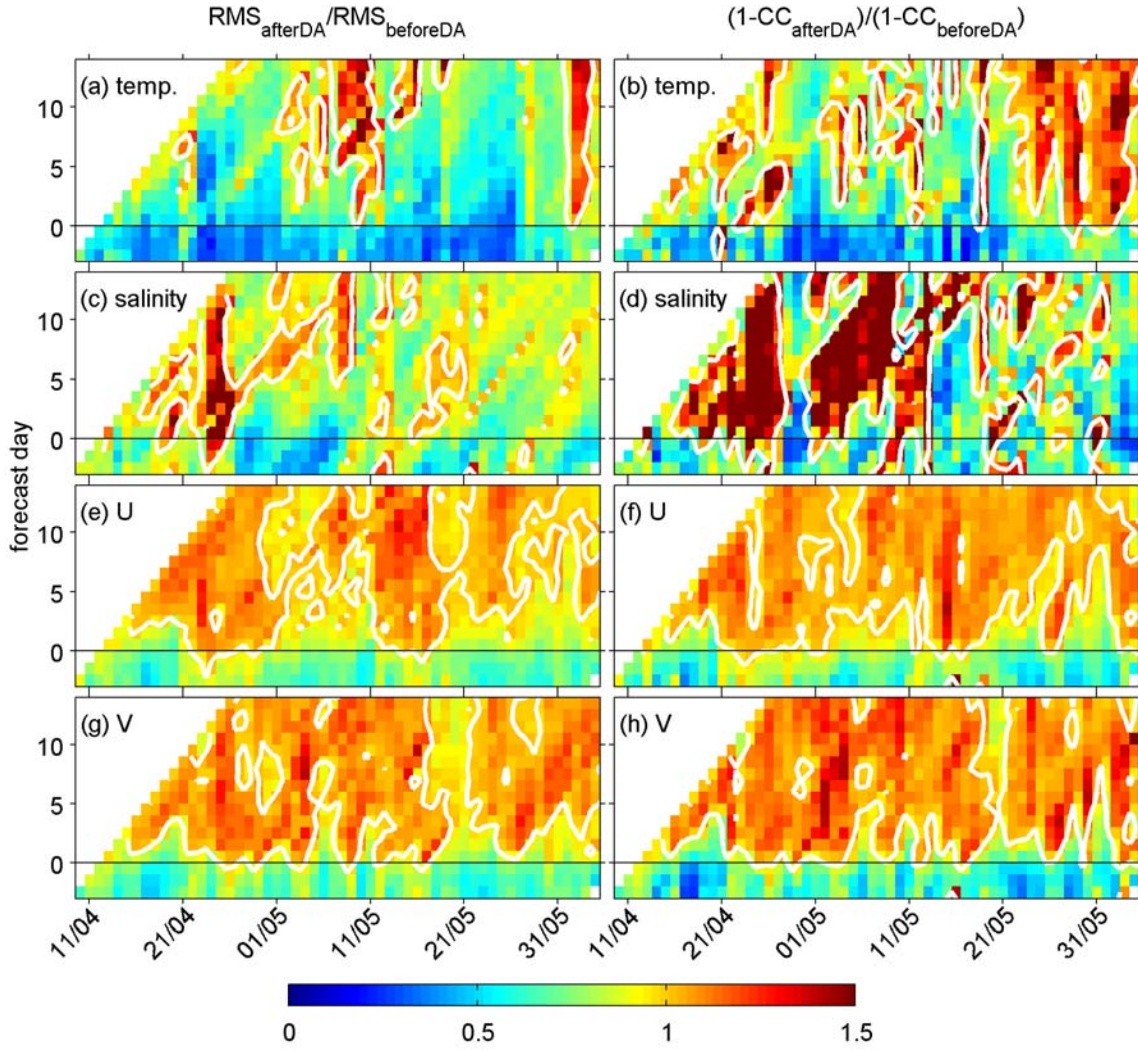
Figure 14. Ratios after data assimilation to before data assimilation of RMS error and cross-correlation error (1−CC) at each day of all cycles for the DA system assimilating all observational data. Thick white lines are contours of value 1.