Review

Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting

S.-Q. Peng, L. Xie *

Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, P.O. Box 8208, Raleigh, NC 27695-8208, USA

Received 2 September 2005; received in revised form 22 March 2006; accepted 23 March 2006

Available online 2 May 2006

Abstract

A tangent linear model and an adjoint model of the three-dimensional, time-dependent, nonlinear Princeton Ocean Model (POM) are developed to construct a four-dimensional variational data assimilation (4D-Var) algorithm for coastal ocean prediction. To verify and evaluate the performance of this 4D-Var method, a suite of numerical experiments are conducted for a storm surge case using model-generated “pseudo-observations”. The pseudo-observations are generated by a nested-grid high-resolution numerical model which is coupled to an inundation/drying scheme that is not included in the original POM.

The 4D-Var algorithm based on POM is tested thoroughly for both code accuracy and the potential application in storm surge forecasting. The assimilation cycles lead to effective convergence between the forecasts and the “observations”. Assimilating water level alone or together with surface currents both lead to significant improvements in storm surge forecasts within and several hours beyond the data assimilation window. It is worth noting that, assimilating water level alone produces improvements in storm surge forecasts that are comparable to those by assimilating both water level and surface currents, suggesting that optimizations of water level and surface currents are linked through the 4D-Var assimilation cycles. However, it is also worth noting that, the benefit resulting from the reduction of initial error in water level and/or surface currents through data assimilation decreases rapidly in time outside the assimilation window. This suggests that determining initial conditions of water level and/or surface currents via data assimilation is only effective within and a few hours beyond the assimilation window for storm surge forecasting. Thus, alternative data assimilation approaches are needed to improve the accuracy and lead time in operational storm surge forecasting.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Data assimilation; Storm surge; Adjoint method; Ocean prediction; Numerical modeling

Contents

1. Introduction ................................................................. 2
2. The nonlinear POM and its tangent linear and adjoint model development .............................................. 3

* Corresponding author. Tel.: +1 919 5151435; fax: +1 919 5157802.
E-mail address: xie@ncsu.edu (L. Xie).

1463-5003/S - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2006.03.005
1. Introduction

In the past half century, tremendous progress has been made in numerical prediction of storm surge (Jelesnianski et al., 1992; Xie et al., 2004; Peng et al., 2005) and coastal ocean circulation (Blumberg and Mellor, 1987). Although improvements have been made, substantial prediction errors still exist (Houston et al., 1999). Numerical ocean predictions are never exact solutions of the real world ocean. Instead, they are only an approximation of the real ocean both in terms of dynamics and physics. The errors or uncertainties of model prediction come from two main sources: (1) dynamical simplifications and physical parameterizations; and (2) initial and boundary conditions. Therefore, oceanic prediction can be improved either by improving the dynamical approximations and physical parameterizations or by improving initial and boundary conditions. Significant improvements have already been made in modeling the ocean dynamics and physics. With the development of large ocean observing systems and remote sensing techniques, more and more oceanic data are becoming available. This provides a promising prospect for improving the model initial conditions through data assimilation. As a result, data assimilation has become widely used in meteorological and oceanographic predictions in recent years.

Among all data assimilation methods, four-dimensional variational data assimilation (4D-Var) is one of the most effective and powerful approaches developed over the past two decades. It is an advanced data assimilation method which involves the adjoint technique and has the advantage of directly assimilating various observations distributed in time and space into the numerical model while maintaining dynamical and physical consistency with the model. Thus, 4D-Var has been widely applied to meteorological and oceanographic data assimilation, sensitivity studies, and parameter estimation. Sasaki (1970) first described the adjoint approach with the governing equations as strong constraints. The application of the adjoint method in meteorological data analysis and assimilation was first implemented by LeDimet and Talagrand (1986) and was further explored extensively by many subsequent studies (e.g., Derber, 1987; Talagrand and Courtier, 1987; Courtier and Talagrand, 1990; Thepaut and Courtier, 1991; Navon et al., 1992; Rabier and Courtier, 1992; Zou et al., 1993; Andersson et al., 1994; Courtier et al., 1994; Zou and Xiao, 1999; Peng and Zou, 2002, 2004).

While the adjoint approach has been applied to global ocean analyses and predictions using real observations and realistic Ocean General Circulation Models (OGCM) (e.g., Martel and Wunsch, 1993; Egbert et al., 1994; Stammer et al., 1997; Marotzke et al., 1999; Oldenborgh et al., 1999; Bonekamp et al., 2001; Weaver et al., 2003; Vialard et al., 2003; Köhl and Stammer, 2004), the application of the adjoint approach to the short-range coastal ocean forecasting remains more theoretical, with the use of simplified ocean models and model-generated psuedo-observations, than practical operational applications. It is worth noting some pioneering studies in this area. The adjoint approach was first employed in some simple (one-dimensional or two-dimensional) tidal models to estimate the initial conditions, boundary conditions, wind stress drag coefficients, oceanic eddy viscosity and bottom friction coefficients (Bennett and McIntosh, 1982; Yu and O’Brien, 1991, 1992; Das and Lardner, 1991). Lardner and Song (1995) implemented the adjoint method to determine the viscosity and friction coefficients in a quasi-three-dimensional tidal model. Seiler (1993) used the adjoint method to determine the open boundary conditions in a quasi-geostrophic ocean model. Lu and Hsieh (1997, 1998a,b) developed an adjoint model for a simple equatorial ocean model which was coupled to a simple atmospheric model and investigated the potential benefit of the variational data assimilation method in determining the model parameters and in estimating the initial conditions of coupled models. More recently, Heemink et al. (2002) assimilated tide gauge and altimeter data into a 3-D shallow water model to estimate the harmonic constants in the open boundary conditions, the friction and viscosity parameters and
water depth. Zhang et al. (2002, 2003) estimated the wind drag coefficient and the lateral tidal open boundary conditions by assimilating pseudo-observations of subtidal water level and predicted coastal tidal elevations into a two-dimensional Princeton Ocean Model using the adjoint data assimilation technique. The results of all of these studies have indicated that the adjoint variational approach is quite effective and robust in regional ocean data assimilation, sensitivity studies and parameter estimation and may benefit real-time prediction of oceanic circulation and sea state.

The studies mentioned above, however, mainly focused on investigating the feasibility of the adjoint method in ocean research and forecasting under simplified dynamics and numerics. The models used in these studies were simplified either physically (e.g., quasi-geostrophic approximation, shallow water approximation, or reduced gravity approximation) or numerically (e.g., one-dimensional or two-dimensional assumptions) due to computational considerations. The questions are

1. With the advent of more powerful supercomputing capabilities, is it practical to develop a 4D-Var algorithm for three-dimensional primitive equation coastal ocean models now or must we still resort to physical or numerical simplification to reduce computing cost?
2. What is the best strategy for data assimilation, as more and more coastal observing systems are deployed, various observations from moorings, buoys, ships, coastal radars and satellites become available in near real-time? Should we assimilate all available data simultaneously or identify a small subset of data sets which are most effective?
3. Storm surge forecasting often relies on a cold start of the numerical model (i.e., without initialization). Is this an acceptable practice? How important or effective is initialization in storm surge forecasting?

To answer these questions, an adjoint model of the three-dimensional, nonlinear primitive equation, Princeton Ocean Model (POM) (Mellor, 2003; Blumberg and Mellor, 1987) is developed. A 4D-Var algorithm based on the POM and its adjoint model is configured as the first step of our goal of performing real-time data assimilation and operational oceanic nowcasting/forecasting using the adjoint variational approach. Experiments with pseudo-observations generated by a different version of the POM with higher resolution are conducted to verify and evaluate the 4D-Var algorithm.

The paper is organized as follows: the next section describes the POM and the development of its tangent linear model (TLM) and adjoint model. Section 3 describes the experiment set-up for evaluating the 4D-Var algorithm. The results of the experiments are presented in Section 4. Discussions and a brief summary of the conclusions are presented in Section 5.

2. The nonlinear POM and its tangent linear and adjoint model development

The Princeton Ocean Model (POM) 2002 version (referred to as pom2k) is used for the forward prediction model in this study. The POM is a three-dimensional, primitive equation ocean model which includes a simplified version of the level 2.5 Mellor and Yamada turbulent closure scheme (Mellor and Yamada, 1982, hereafter denoted as MY82, details can be found in Mellor, 1989). In this study, radiation open boundary conditions are applied to seaward boundaries. The bottom terrain-following, sigma coordinate equations of the POM as well as a full description of the POM can be found in Blumberg and Mellor (1987) and Mellor (2003). The governing equations of the POM can be written in a general form as

\[ \frac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}), \quad \mathbf{x}|_{t_0} = \mathbf{x}_0, \quad \mathbf{x}(t)|_F = \mathbf{y}(t), \]

where \( \mathbf{x} \) represents the vectors of model state variables, which includes current velocity \( U \) and \( V \), temperature \( T \), salinity \( S \), surface elevation \( \eta \), and turbulence quantities \( q^2 \) and \( q^2I \). \( \mathbf{x}_0 \) and \( \mathbf{y}(t) \) represent the initial condition (IC) at initial time \( t_0 \) and lateral boundary condition on \( F \), respectively.
The TLM of the POM can be obtained by linearizing the POM forecast model (1) about a nonlinear model trajectory $x(t)$ and $y$:

$$\frac{\partial x'}{\partial t} = \frac{\partial F(x)}{\partial x} x', \quad \text{(2a)}$$

$$x'|_{t_0} = x'_0, \quad \text{(2b)}$$

$$x'(t)|_{t_f} = y'(t), \quad \text{(2c)}$$

where prime represents perturbations of the corresponding variables. Integrating the POM TLM with perturbed IC $x'_0$ and/or LBC $y'_i$ ($i = 1, \ldots, m$), one obtains a perturbation solution $x'(t)$ which is accurate to the first-order approximation $O(\|x'_0\|^2)$ and $O(\|y'_i\|^2)$, i.e., the TLM solution $x'(t)$ should satisfy the following equation:

$$x(t)|_{(x_0 + x_0', y + y')} - x(t)|_{(x_0, y)} = x'(t) + O(\|x'_0\|^2) + O(\|y'_i\|^2). \quad \text{(3)}$$

A linear operator, $L^*$, is called the adjoint of $L$ if, for all $x$ and $z$ in a linear space $\mathcal{F}$,

$$\langle z, Lx \rangle = \langle L^*z, x \rangle, \quad \text{(4)}$$

where $\langle \cdot, \cdot \rangle$ represents an inner product. In Euclidean space, $L^* = L^T$. The adjoint model corresponding to (2) is

$$-\frac{\partial \tilde{x}}{\partial t} = \left(\frac{\partial F(x)}{\partial x}\right)^T \tilde{x}, \quad \text{(5a)}$$

$$\tilde{x}|_{t_0} = 0, \quad \text{(5b)}$$

$$\tilde{x}(t)|_{t_f} = 0, \quad \text{(5c)}$$

where $\tilde{x}$ represents the vector of adjoint variables and $t_f$ the ending time of forward integration. The negative sign in the left-hand side of (5a) indicates that the adjoint model integrates backward in time.

The adjoint model can be constructed by discretizing the continuous adjoint equation. However, as pointed out by Zou et al. (1997), using such a procedure to derive the adjoint model, although feasible for simple models such as shallow water models, is not straightforward for a complex three-dimensional primitive equation model for the following two reasons: (1) the derivation of the adjoint equations for a primitive equation model by the partial integration procedure can be very tedious, and this becomes even more difficult when various physics options with more than one expression are included; (2) the discretization of the adjoint equations may become inconsistent with the original model and the accuracy of the gradient will be limited to the accuracy of the finite-difference scheme used in discretization. Another way of constructing the adjoint model is to develop it directly from the discretized forward numerical model. This method can not only avoid the inconsistency due to non-commutativity of the adjoint model and discretization operations (Spitz, 1995), but can also simplify the procedure of constructing and debugging the adjoint model for a complex model. Therefore, we choose to develop the POM adjoint model directly from the original forward model code of POM. The TLM of POM is first developed by linearizing the nonlinear POM. The adjoint model of the TLM of POM is then developed. The procedure for developing the TLM and adjoint model is described in Appendix A.

It should be noted that, due to the high nonlinearity and discontinuity of vertical turbulence, it is still an open issue as to whether it is meaningful to linearize the turbulence closure scheme in an atmospheric or oceanic model. It is known that the linearization of a highly nonlinear and discontinuous physical process may produce nonphysical noise and lead to numerical instability, thus considerable caution should be taken (Errico et al., 1993; Xu, 1996; Zou, 1997; Errico and Reader, 1999). While a numerical implement scheme (e.g., semi-implicit scheme) is unconditionally stable for the nonlinear model, it may become conditionally stable for the TLM. Zhu and Kamachi (2000) showed that using a smaller time step in the model integration can eliminate the instability of TLM with the higher computational cost. Alternatively, a smooth approximation to the original nonlinear scheme is a common method to obtain a stable TLM of a parameterization scheme of a highly nonlinear physical process (Janiskova et al., 1999; Mahfouf, 1999; Laroche et al., 2002; Weaver et al., 2003). In this study, a great deal of effort has been made in developing the tangent linear and adjoint
codes of MY82. As a result, the derived tangent linear and adjoint codes of MY82 both have sufficient accuracy in their correctness check based on formulas (A.7) and (A.8) at every time step in a 3-h test. However, the accuracy in the correctness check for the whole TLM of the POM with the MY82 tangent linear code incorporated is only five-digit for some three-dimensional variables (such as current, temperature and salinity) and

Fig. 1. Water level fields from the differences between two nonlinear model run (left column (a)–(d)), the C-MY TLM run with C-MY (middle column (e)–(h)), and the TLM run with S-MY (right column (i)–(l)) at different integration time of 3 h, 6 h, 12 h, and 24 h from the top to the bottom (unit: mm).
three-digit for the turbulent kinetic energy variables in a 3-h model run. The accuracy in the correctness check for the whole adjoint model of the POM with the MY82 adjoint code incorporated is only eight-digit accuracy in a 1-h model run and six-digit in a 3-h model run. This loss of accuracy is mainly due to the direct linearization of the highly nonlinear and discontinuous turbulence scheme which generated some nonphysical noises during the TLM integration. A simple but efficient way of avoiding the noise problem is to neglect the variation of the vertical diffusion coefficients in the linearization of the vertical turbulence scheme, which has been employed in many studies (e.g., Mahfouf, 1999; Weaver et al., 2003). Specifically, this can be realized by prescribing the values of the vertical diffusion coefficients $K_m$ and $K_h$ in the TLM and adjoint model for every time step by a pre-run of the nonlinear POM model with MY82. In such a configuration, the correctness checks for the whole TLM and adjoint model of POM both have satisfactory accuracy (hereafter we call the configuration with pre-generated $K_m$ and $K_h$ as S-MY and the one without such a simplification as C-MY). To see how the TLM behaves in response to a realistic perturbation, we run the TLM starting with a perturbation $\Delta x = 0.1 x_0$ and compare the output of TLM with the difference between two nonlinear model runs which started with perturbed IC ($x_0 + \Delta x$) and non-perturbed IC ($x_0$), respectively. Here $x_0$ is the output of a 12-h spin-up of the nonlinear POM, which includes water level, current (including 3-D current and 2-D vertically integrated current), temperature, salinity, and bottom stress. The storm surge case, model domain and resolution are the same as those described in Section 3. Fig. 1 displays the evolution of water level. Both the TLM with C-MY and the TLM with S-MY behaved well with the patterns of amplitude the same as those of the difference between the two nonlinear run. The same results are found in the current fields (not shown). Thus, both the TLM with C-MY and the TLM with S-MY are valid for storm surge application, though the TLM with C-MY does not have satisfactory accuracy of correctness check based on formula (A.7). As we will show in Section 4, the 4D-Var scheme with C-MY for storm surge study works as well as that with S-MY. The results shown in Fig. 1 and the similar data assimilation results from C-MY and S-MY shown in Section 4 indicate that the impact of the linear approximation of the vertical turbulence closure scheme on the storm surge during 4D-Var is minor and thus negligible.

A cost function which measures the misfit between the model forecast and the observations can be written as

$$ J(x_0) = \int_{t_0}^{t_1} H[x(t), t] dt, \tag{6} $$

where $H[x(t), t]$ is a scalar function measuring the distance between forecast $x(t)$ and the observations at time $t$. The procedure of finding an optimal IC, $x_0^*$, which minimizes the cost function $J$ using adjoint approach is

1. Starting from $x_0$, integrate the forward model [Eq. (1)] from $t_0$ to $t_1$. Store the values $x(t)$ ($t_0 \leq t \leq t_1$) which will make up the coefficients of the operator of the adjoint model. Calculate the cost function at the time $t_1$ when the observations are available.

2. Starting from $\bar{x}(t_1) = \nabla x H(t_1)$, integrate the adjoint model (Eq. (5)) backwards in time from $t_1$ to $t_0$, with the forcing term $\nabla x H(t)$ added to the currently computed solution at discrete observation time $t_i$. The final result at time $t_0$ is the gradient $\nabla x J$.

3. Use a minimization algorithm (such as limited-memory quasi-Newton method) to update $x_0$.

4. Repeat steps 1–3 until a convergence criteria is satisfied.

3. Experiment design

To verify and evaluate the performance of the 4D-Var algorithm based on the POM, we apply the algorithm to a storm surge case along the United States East Coast during hurricane Hugo, September 21–22, 1989. The “pseudo-observations” generated by a high resolution model which is described in detail in Xie et al. (2004) and Peng et al. (2004, 2005) were used. Using “pseudo-observations” in data assimilation studies has the advantage of providing a full suite of balanced data sets which can be assimilated into the forecast model (Zhang et al., 2002, 2003). In this study, the same wind field is used to generate the “pseudo-observations” and for the storm surge forecast, so the uncertainty associated with the wind forcing is minimized. This
allows us to focus on the effect of determining initial conditions on storm surge. Four numerical experiments are conducted:

**NoDA**: 9-h model run with original IC from a 12-h spin-up (control run);

**DA-1**: 9-h model run with optimal IC from 4D-Var of water level only;

**DA-2**: 9-h model run with optimal IC from 4D-Var of both water level and surface currents;

**DA-3**: same as DA-1, except that S-MY configuration was used during 4D-Var.

Fig. 2 shows a schematic illustration of the experimental set-up. The horizontal resolution for the four experiments is $2' \times 2'$ (about 3.1 $\times$ 3.7 km) with a total number of grid points of 136 $\times$ 106 and three vertical sigma levels. The time step is 3 min. Inflow boundary conditions are used for water level, radiation boundary condition for 3-D currents, and upstream advection boundary conditions for temperature, salinity and turbulent kinetic energy. In order to generate a large error in the control run so that improvements by 4D-Var can be seen clearly, a fixed boundary condition is applied to the vertically-averaged currents. The wind stress are calculated by using the Holland hurricane wind model (Holland, 1980) and updated every 10 min. The control run and the 4D-Var experiments started at 21Z September 21. A 12-h spin-up was run and the output from the spin-up was used as the initial condition of the NoDA experiment as well as the first guess field for the 4D-Var experiments (DA-1, DA-2 and DA-3). DA-1 and DA-2 used the C-MY approach for the turbulence closure model while DA-3 adopts the S-MY configuration.

The control variables in the 4D-Var experiments are water level, current (including 3-D current and 2-D vertically integrated current) field, temperature, salinity, and bottom stress. The pseudo-observations of water level and surface currents on each ocean grid point are generated by a different version of the POM running at double horizontal resolution (or half grid size) and with an inundation/drying scheme (Xie et al., 2004) which enables the coastal boundary to be time-dependent. A 3-h data assimilation window from 21Z September 21 to 00Z September 22 is set for the 4D-Var experiments with a 10-min sampling frequency. The cost function is defined as

$$ J(x_0) = (x_0 - x_b)^T B^{-1} (x_0 - x_b) + \sum_{r} \sum_{i} (P_{r,i} - P_{r,i}^{obs})^T R^{-1} (P_{r,i} - P_{r,i}^{obs}), $$

where $x_0$ represents the vector of control variables at the initial time $t_0$, and $x_b$ represents their background values. $P_{r,i}$ and $P_{r,i}^{obs}$ denote the model simulated values and the observations of water level or surface currents, respectively, at location $i$ and time level $r$. $N$ is the number of grid points over the ocean and $M$ the number of time levels of observations. $B$ and $R$ are the error covariance for the background vector and the observations, respectively.

The first term on the right-hand side of (7) is a simple background error term measuring the distance between the model’s initial condition $x_0$ (to be adjusted through an iterative minimization procedure) and the background field $x_b$, which is the output from the 12-h spin-up run. For simplicity, a diagonal matrix is used for $B$ in this study, which indicates that background errors for different variables at different locations or times are treated as uncorrelated. This background error covariance is obtained based on the differences between a 12-h model forecast and the initial state.

![Fig. 2. Schematic diagram of experiment set-up. $x_0$ denotes the initial condition for control run, and $x_0^*$ denotes the “optimal” IC from data assimilation. The dashed line represents the model integration, while the solid curve represents the observations.](image-url)
The second term on the right-hand side of (7) is the observation term for water level or surface currents, which is assimilated at 10-min intervals within the assimilation window. Since the pseudo-observations of water level or surface currents are model variables and available on each grid, no observation operators (such as spatial interpolation or transformation from model variables to observation quantities) are needed. When using observational data for real-time data assimilation, the observation error covariance $R$ may contain instrument error, representativeness error and the error caused by the observation operators such as spatial interpolation. In this study, these errors are neglected and $R$ is simply set to an arbitrary constant matrix for each data type, for instance, 1.0 ($m^2$) for water level in DA-1 and DA-3, and $1/100$ ($m^2$) and 1.0 ($m^2 s^{-2}$) for water level and surface currents in DA-2, respectively.

The limited memory Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton minimization algorithm (Liu and Nocedal, 1989) is employed in the minimization process to estimate the optimal control variables. To avoid the problem of ill-conditioning, we perform preconditioning on each control variable during the minimization process using the square-root of $B$, i.e., define a new vector of variables $w$ which satisfies

$$x^0 = Uw,$$

where $U$ is a diagonal matrix with positive diagonal elements $u$ that satisfies $b = u^2$. Recall that here $B$ is taken to be a diagonal matrix with diagonal elements $b$.

4. Results

All of the experiments are conducted on a single processor of the IBM Blade Center Linux Cluster at North Carolina State University. Fig. 3 shows the evolution of the norm (L-2 norm) of the gradient of the cost function with respect to the number of iterations for experiments DA-1 and DA-2. The norm of the gradient decreases by one order of magnitude in the norm during the first 10 iterations and approaches zero at the 25th iteration. Fig. 4 shows the evolution of the total cost function and each of its terms with respect to the number of iterations for Experiment DA-1. It can be seen that the evolution of the total cost function is similar to that of the observation term, which decreases rapidly during the first 10 iterations and then varies slowly. The values of the background term, which measures the difference between the background field and the initial condition at each iteration, is zero at the 0th iteration and increases with the number of iterations as the initial conditions are adjusted to fit the model trajectory to the “observations”. The evolution of the cost function with respect to the number of iterations for DA-2 is similar to that for DA-1.

Fig. 3. The evolution of the norm of the gradient of the cost function with respect to the number of iterations for DA-1.
Fig. 4. The evolution of each term of the cost function with respect to the number of iterations for DA-1.

Fig. 5. The water level fields at 00Z September 22 from (a) pseudo-observations; (b) NoDA; (c) DA-1 and (d) DA-2 (unit: m).
Fig. 6. As in Fig. 5, except at 01z September 22.

Fig. 7. Time series of the root mean square error (RMSE) of water level averaged over all ocean grid points for each experiment with respect to the pseudo-observations of water level starting from 21Z September 21 to 05Z September 22 (unit: m).
Fig. 5 shows the water level field from the “observations” and each experiment at 00z September 22 which is at the end of the assimilation window. Compared to the “observations” (Fig. 5a), the control run without data assimilation (NoDA, Fig. 5b) under-predicts the water level along much of the coastline north of the Georgia–South Carolina border (32°N). After assimilating the water level (Fig. 5c), the height of the water level over this area increases and is closer to the “observations”. Assimilating both water level and surface

Fig. 8. Maximum height of water level along line AB shown in Fig. 6d for the pseudo-observations and each experiment (the number 1 in X-axis corresponding to location A and 17 corresponding to B. unit: m).

Fig. 9. The evolution of water level over four selected points (6–9) located on the middle section of line AB (Fig. 6d) for pseudo-observations, NoDA and DA-2, starting at 00z September 22 (unit: m).
currents (Fig. 5d) has similar results as assimilating only water level at 00z. Similar improvements are evident at 01z September 22 which is outside the data assimilation window (Fig. 6a–d), with DA-2 slightly better than DA-1. Fig. 7 shows the time series (starting at 21Z September 21) of the root mean square error (RMSE) of water level averaged over all ocean grid points for each experiment with respect to the “observations” of water level. The model forecasting errors are reduced significantly by data assimilation within and a few hours beyond the assimilation window, with DA-2 slightly outperforming DA-1. However, the effect of data assimilation outside the assimilation window decreases as forecast time increases.

The height of the peak storm surge along the coast is often the quantity of interest during the threat of a tropical cyclone. Fig. 8 shows the peak surge at 17 locations evenly distributed along a line parallel to the coast (line AB in Fig. 6d) for the “observations” and each experiment. It indicates that data assimilation produces significant improvements in the estimation of peak surge along the southern section of line AB, but no improvement on the northern section of the line. It is worth noting that although the storm surge predicted by the stand-alone POM without data assimilation produced large errors north of 32°N as shown in Fig. 6, it is able to capture the peak surge that occurred near location 12. As a result, the improvement in peak surge is small near location 12. The large error in peak surge that occurred near the northern boundary (locations 15–17) is not effectively reduced by data assimilation. The error in this region is apparently less sensitive to initial conditions. This could be the result of model deficiencies, such as the lower resolution used, the lack of an inundation/drying scheme, and the fixed lateral open boundary conditions for the vertically-averaged 2-D current. The time series of water level of four points (7–10) located in the middle section of the line...
Fig. 11. As in Fig. 10, except for 01Z September 22.

Fig. 12. The evolution of each term of the cost function with respect to the number of iterations for DA-3 (solid lines) and DA-1 (dash lines).
AB for “observations”, NoDA and DA-2 are shown in Fig. 9a–d. The figures show that data assimilation leads to significant improvements in storm surge prediction during the simulation period.

Fig. 10a and b show the surface currents from the “observations” and the control run (NoDA) and Fig. 10c and d show the differences of surface currents between the data assimilation experiments and the control run at 00z September 22. A strong vortex over the southeast corner of the domain and a northeastward current along the coast are seen in the “observations” (Fig. 10a). In the control run (Fig. 10b), however, the surface currents along the coast are opposite to that of “observations” and the vortex over the southeast corner is weak. Assimilating only water level (Fig. 10c) or assimilating both water level and surface currents (Fig. 10d) intensifies the vortex and produced onshore currents which led to an increase in the water level along the coast. At 01z September 22 (Fig. 11a–d), the vortex in NoDA became stronger and covered a larger area of the coast water. The current vectors from DA-1 and DA-2 (Fig. 11c and d) were toward the coast over the southwest section of the coastline and away from the coast over the northeast section of the coastline. This suppressed the over-predicted vortex over the coastal area. It is interesting to note that, although assimilating both water level and surface currents have larger impacts on the model simulation of surface currents (please note the different vector scales in different panels of Figs. 10 and 11), assimilating water level alone can produce a comparable improvement in storm surge forecasting. Because the 4D-Var approach adjusts all control model variables simultaneously through the model dynamics and physics when assimilating one or more types of observations, assimilating water level leads to significant impacts on the surface currents.

When smooth approximation is made in the linearization of MY82 turbulence scheme (i.e., S-MY in DA-3), the results are similar to those with C-MY (DA-1). Fig. 12 shows the evolution of the cost function with respect to the number of iterations for both DA-1 and DA-2. The difference between DA-1 and DA-2 is very small. The improvements of water level forecasts from DA-3 are almost identical to those from DA-1 (figures omitted). Therefore, the impact of the noises due to linear approximation of the turbulence closure scheme on the storm surge in 4D-Var is negligible.

It should be noted that the above results are obtained based on the choice of closed lateral boundary conditions for the vertically-averaged currents. When open lateral boundary conditions are used for the vertically-averaged currents, similar conclusions are reached, but the improvements by 4D-Var is smaller in absolute terms since the forecasting errors in the control run with open lateral boundary conditions is smaller (results are not shown here).

5. Summary and discussion

In this study, the tangent linear model and the adjoint model of the 3-D primitive equation coastal ocean circulation model (POM) are developed. Experiments for a storm surge case are conducted on a single Linux-based 2.8–3.2 GHz Xeon processor to evaluate the potential application of a 4D-Var algorithm based on the POM adjoint model in storm surge forecasting by assimilating pseudo-observations of water level and surface currents into the model. For a 12-h forecast with a 3-h data assimilation window (with 10 iterations of minimization) with the set-up used in this study, it requires approximately half an hour of CPU time. Thus, real-time ocean forecasting using a three-dimensional primitive equation model with 4D-Var data assimilation is not beyond reach. When a multi-processor parallel computer cluster is used, higher resolutions can be achieved to meet the requirements of today’s operational ocean forecasting needs.

The experimental results demonstrate that the 4D-Var data assimilation based on the developed POM adjoint model (either C-MY or S-MY) is able to find an “optimal” initial condition for the storm surge forecasting, with the values of the cost function which measures the difference between the model and “observations” reducing rapidly during the first 10 minimization iterations. Improvements on water level prediction are obtained both within and several hours beyond the assimilation window by assimilating water level “observations” alone or assimilating both water level and surface current “observations”. The added benefit of assimilating both water level and surface currents is relatively small since water level and current fields are adjusted in dynamical and physical consistency with the constraint of the model control equations and the cost function.

For the storm surge case studied in this paper, notable improvements are obtained by finding an “optimal” initial condition through data assimilation. However, the improvements decrease rapidly in time beyond the
assimilation window. The effect of data assimilation only lasts for several hours beyond the data assimilation window. Therefore, only changing the initial condition through data assimilation does not ensure an accurate forecast of storm surge with a long lead time.

Note that the same Holland wind model is used to generate the hurricane winds for the prediction model as well as the model which generates the “pseudo-observations”. Thus, the forecast errors are the result of the differences between the two models. The data-generation model is run at higher resolution and also includes an inundation scheme, whereas the forecast model is configured at a lower resolution and does not couple to an inundation model. Thus, the results presented above show that forecast errors due to deficiencies in model physics or numerics can not always be effectively corrected through improving initial conditions alone.

The fact that similar results were obtained from DA-1 and DA-3 indicates that, the errors or noises due to the linear approximation of the turbulence closure scheme have a negligible impact on the results of data assimilation for storm surge forecasting, at least in the case studied here. This implies that either C-MY or S-MY can be used in data assimilation for storm surge forecasting, though it needs to be further confirmed by assimilating actual observations. For other forecast variables such as temperature and salinity, further studies are needed to verify whether the data assimilation program developed here will work properly.

Acknowledgements

This study is supported by the Carolina Coastal Ocean Observation and Prediction System (Caro-Coops) project under NOAA Grant No. NA16RP2543 and the Coastal Ocean Research and Monitoring Program (CORMP) under NOAA Grand #NA16RP2675 via the National Ocean Service, through Charleston Coastal Services Center. Caro-COOPS is a partnership between the University of South Carolina and North Carolina State University, and CORMP is a partnership between the University of North Carolina at Wilmington and North Carolina State University. The helpful comments from two anonymous reviewers are greatly appreciated.

Appendix A. Development of tangent linear model and adjoint model of POM

Following the notation in Zou et al. (1997), the discretized form of POM can be written, in general, as

\[
\mathbf{x}(t_r) = \mathbf{Q}_r(\mathbf{x})\mathbf{z}, \quad \mathbf{z} = (\mathbf{x}_0, \mathbf{y}_{t_1}, \mathbf{y}_{t_2}, \ldots, \mathbf{y}_{t_n})^T, \tag{A.1}
\]

where \( \mathbf{z} \) represents the vector of model input which may include the IC, \( \mathbf{x}_0 \), the boundary conditions \( \mathbf{y}_{t_i} \) \((i = 1, \ldots, m)\), and possibly model parameters if any. \( \mathbf{Q}_r \) is the operator of POM. The discretized TLM of POM can be directly developed from (A.1) and written as

\[
\mathbf{x}'(t_r) = \mathbf{P}_r(\mathbf{x})\mathbf{z}', \quad \mathbf{z}' = (\mathbf{x}'_0, \mathbf{y}'_{t_1}, \mathbf{y}'_{t_2}, \ldots, \mathbf{y}'_{t_n})^T, \tag{A.2}
\]

where \( \mathbf{P}_r = \frac{\partial \mathbf{Q}_r}{\partial \mathbf{x}} \) and the prime denotes the perturbation of the corresponding variable.

The adjoint model is then defined as

\[
\hat{\mathbf{z}}^r = \mathbf{P}_r^T(\mathbf{x})\hat{\mathbf{x}}(t_r), \quad \hat{\mathbf{x}}(t_r) = \frac{\partial J}{\partial \mathbf{x}} \quad (r = R, R - 1, \ldots, 0) \tag{A.3}
\]

under the Euclidean norm, where the hat denotes the adjoint variables, \( \frac{\partial J}{\partial \mathbf{x}} \) is the forcing term of the adjoint model which depends on the cost function defined, and \( R \) is the total number of time levels at which cost function \( J \) is calculated. \( \mathbf{P}_r^T \) is the transpose of \( \mathbf{P}_r \). If we view the POM model operator as the result of the multiplication of a number of operator matrices:

\[
\mathbf{Q}_r = \mathbf{Q}_1(\mathbf{x}_{t_1})\mathbf{Q}_2(\mathbf{x}_{t_2}) \cdots \mathbf{Q}_N(\mathbf{x}_{t_N}), \tag{A.4}
\]

where each matrix \( \mathbf{Q}_i \) \((i = 1, \ldots, N)\) represents either a subroutine or a single DO loop and \( \mathbf{x}_r \) is a vector of variables depending on model predictive variables, then the operator of the tangent linear model can be written as

\[
\mathbf{P}_r = \mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_N, \tag{A.5}
\]
where matrix $A_i = \frac{\partial Q_i}{\partial x_{ri}}$ ($i = 1, \ldots, N$). The adjoint model operator $P^T_r$ is then a product of sub-adjoint problems,

$$P^T_r = A^T_N A^T_{N-1} \cdots A^T_1. \tag{A.6}$$

Following the above approach, the discrete adjoint model can be constructed piece by piece, and the discrete operations in the forward model have unique corresponding discrete operations in the adjoint model. This is extremely convenient, especially for dealing with various physical parameterization schemes, in which the tangent linear and adjoint codes for a specific scheme can be developed and tested separately from other parts of the model. Giering and Kaminski (1998) have developed an automatic differentiation and adjoint model compiler. Considering the complication of the time-splitting scheme and the high nonlinear feature of turbulence closure in POM, the TLM and adjoint model of POM developed in this paper, however, were hand coded using the method mentioned above. As pointed out by the ROMS TLM/adjoint model developing group (Moore et al., 2004), hand-coding allows a better understanding and control of the code structure of the resulting models.

According to Zou et al. (1997), the correctness of the TLM can be checked against the forward nonlinear model through the following formula:

$$\Phi(z) = \frac{\|Q_r(z + \Delta h) - Q_r(z)\|}{\|zP_r\Delta h\|} = 1 + O(\Delta h), \tag{A.7}$$

where $z$ is a vector of all the input variables of the nonlinear POM operator $Q_r$, $\Delta h$ is the perturbation of $z$. The values of $\Phi(z)$ shall linearly approach a unit value with increasing accuracy as $\Delta h$ becomes progressively smaller (usually the correctness check for TLM should at least have a six-digit accuracy in a double-precision compiler).

The adjoint model codes can be checked for validity and accuracy by the following algebraic expression:

$$(P_r z)^T (P_r z) = z^T (P^T_r (P_r z)), \tag{A.8}$$

where the left-hand side and right-hand side should be equal with a 13-digit accuracy in a double-precision compiler.

Formulas (A.7) and (A.8) can be used for checking the correctness of tangent linear and adjoint correspondences for any part of the code, e.g., a single DO loop or a subroutine or a combination of a few of these operations.

References


