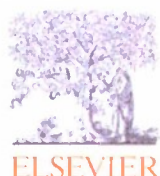# REPORT DOCUMENTATION PAGE

*Form Approved
OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, end completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communicetions Directorate (0704-0188). Respondents should be aware thet notwithstending eny other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently velid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE		3. DATES COVERED *(From - To)*
30-07-2009	Journal Article		

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Cycling the Representer Method for 4D-variational data assimilation with the Navy Coastal Ocean Model	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
	0601153N

6. AUTHOR(S)	5d. PROJECT NUMBER
Scott R. Smith, Hans E. Ngodock, Gregg A. Jacobs	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
	73-8554-07-5

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	B. PERFORMING ORGANIZATION REPORT NUMBER
Naval Research Laboratory Occanography Division Stennis Space Center, MS 39529-5004	NRL/JA/7320--07-7301

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Office of Naval Research 800 N. Quiney St. Arlington, VA 22217-5660	ONR
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution is unlimited.

20090805541

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Cycling Represter Method, which is a technique for solving 4D-variational data assimilation problems, has been demonstrated to improve the assimilation accuracy with simpler nonlinear models. In this paper, the Cycling Representer Method will be used to assimilate an array of ADCP velocity observations with the Navy Coastal Ocean Model (NCOM). Experiments are performed in a high-resolution Mississippi Bight domain for the entire month of June, 2004 and demonstrate the usefulness of this assimilation technique in a realistic application.

15. SUBJECT TERMS

4D-variational data assimilation; Statistical analysis; Acoustic current meters; Prediction; Gulf of Mexico; Mississippi Bight

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	1B. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Scott Smith
Unclassified	Unclassified	Unclassified	UL	16	19b. TELEPHONE NUMBER *(Include area code)* 228-688-4630

Standard Form 29B (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Cycling the Representer Method for 4D-variational data assimilation with the Navy Coastal Ocean Model

S.R. Smith *, H.E. Ngodock

Naval Research Laboratory, Ocean Dynamics and Prediction Branch, Building 1009, Code 7320, Stennis Space Center, MS 39529, United States

ARTICLE INFO

ABSTRACT

The Cycling Representer Method, which is a technique for solving 4D-variational data assimilation problems, has been demonstrated to improve the assimilation accuracy with simpler nonlinear models. In this paper, the Cycling Representer Method will be used to assimilate an array of ADCP velocity observations with the Navy Coastal Ocean Model (NCOM). Experiments are performed in a high-resolution Mississippi Bight domain for the entire month of June, 2004 and demonstrate the usefulness of this assimilation technique in a realistic application.

The Representer Method is solved by minimizing a cost function containing the weighted squared errors of velocity measurements, initial conditions, boundary conditions, and model dynamics. NCOM, however, is a highly nonlinear model and in order to converge towards the global minimum of this cost function, NCOM is linearized about a background state using tangent linearization. The stability of this tangent linearized model (TLM) is a very sensitive function of the background state, the level of nonlinearity of the model, open boundary conditions, and the complexity of the bathymetry and flow field. For the Mississippi Bight domain, the TLM is stable for only about a day. Due to this short TLM stability time period, the Representer Method is cycled by splitting the time period of the assimilation problem into short intervals. The interval time period needs to be such that it is short enough for the TLM to be stable, but long enough to minimize the loss of information due to reducing the temporal correlation of the dynamics and data. For each new cycle, a background is created as a nonlinear forecast from the previous cycle's assimilated solution. This background, along with the data that falls within this new cycle, is then used to calculate a new assimilated solution. The experiments presented in this paper demonstrate the improvement of the assimilated solution as the time window of the cycles is reduced to 1 day. The 1-day cycling, however, was only optimal for the first half of the experiment. This is because there was a strong wind event near the middle of June that significantly reduced the stability of the 1-day cycling and caused substantial errors in the assimilation. Therefore, the 12-h cycling worked best for the second half of the experiment. This paper also demonstrates that the forecast skill is improved as the assimilation system progresses through the cycles.

Published by Elsevier Ltd.

1. Introduction

Improving the capability to model and forecast the fundamental properties of the ocean has been the endeavor of numerous universities, institutions, and agencies for many decades. The range of spatial and temporal scales that need to be resolved in order to properly model the fundamental ocean properties varies from large global scales all the way down to the submesoscale (littoral and coastal regions). The primary oceanic processes that govern the motion of the ocean can be drastically different between these two extreme resolution scales. For example, the ability to accurately model and forecast oceanic properties in a small-scale coast-al domain requires a strategy that includes resolving processes resulting from things such as: complex bathymetry, coastal geometry, river inflow, atmospheric forcing, mixing, and horizontal shear instability. It is an inherently difficult problem to properly resolve all of the processes in a coastal environment, therefore, the models used to forecast in such regions will generally contain significant error.

One of the key components to improving the accuracy of ocean forecasts is the inclusion of data and optimally merging it with the model. Models are always going to be in error and the data contains instrument noise, representativeness errors, and is almost always going to be scarce relative to a discretized model. Therefore, an important question is how does one extract the most useful information possible from the data and apply it to maximize its influence on the model? Time-invariant assimilation techniques

* Corresponding author. Tel.: +1 228 688 4630; fax: +1 228 688 4759.
E-mail address: smithsc@nrlssc.navy.mil (S.R. Smith).

such as nudging, optimal interpolation, and 3D-variational assimilation (3DVAR) offer a limited ability for the data to effectively influence the correction of the model state. These techniques assimilate and correct the model state at unique analysis time stamps and are typically not temporally correlated, therefore, any potential temporal correlation information within the data and model is lost. These simpler techniques may be adequate and desirable for large-scale open ocean models since they are computationally cheaper relative to the more sophisticated techniques such as 4D-variational assimilation (4DVAR) and ensemble Kalman filtering. These computationally intensive assimilation techniques may not be feasible for large domains. For small coastal regions, however, these techniques may not only be feasible, they may be critical in order to properly utilize all of the available data to counteract the model's deficiency in representing the complex physics. Advanced sequential data assimilation techniques such as the various Kalman filters take into account the model physics and can propagate the influence of the data forward in time. With the rapid increase in computational resources, 4DVAR is becoming a popular assimilation technique since it not only propagates the influence of data forward in time but also backwards in time (through the adjoint) and can provide corrections to the initial and boundary conditions, forcing, and the model physics.

One technique that can be used to solve 4DVAR problems is the Representer Method (Bennett, 1992, 2002; Bennett et al., 1996, 2000, 2006). The Representer Method is currently being employed by several institutions which are developing assimilation systems with sophisticated, baroclinic ocean models. For example, Muccino and Luo (2005) are applying the Representer Method to the Advanced Circulation Model (ADCIRC), whereas, Moore et al. (2004) and Lorenzo et al. (2007) are applying this technique to the Regional Ocean Modeling System (ROMS). One of the biggest challenges in developing a representer based assimilation system is that the model must first be linearized about a particular background. Since the linearized version of the model (TLM) is only an estimate of the full nonlinear (NL) model, it will only be stable for a finite period of time. There are many factors that contribute to how long the TLM can remain stable, such as the level of nonlinearity of the model, the resolution of the background, the open boundary conditions, and the complexity of the bathymetry and circulation field. It will be demonstrated in this study that the time period of TLM stability will be relatively short for a high-resolution coastal application with open boundaries along all four edges and linearized about a coarse background.

The Cycling Representer Method, which was first introduced and applied by Xu and Daley (2000), can be used to overcome the problem of TLM instability. In Xu and Daley (2000), the authors cycle the Representer Method in time with a linear one-dimensional transport model in order to lay down the concept. Then in Xu and Daley (2002), the authors apply the Cycling Representer Method to a linear, barotropically unstable shallow water system. In both of these applications there was no issue with the TLM since both models were linear. Later, Xu et al. (2005) applied the Cycling Representer Method to NRL's operational global atmospheric model called NOGAPS. Even though their system is designed to perform weak constraint assimilation, their current implementation assumes a perfect model (strong constraint). Recently, Ngodock et al. (2007a,b) explored the idea of applying the weak constraint Cycling Representer Method to the NL Lorenz attractor and reduced gravity ocean models, respectively. In these two studies, an initial background is first created by propagating the NL model forward over the first cycle. Then the TLM and adjoint of these perspective models are used to perform an assimilation using the Representer Method. For the second cycle, a forecast (background) is created by propagating the NL model forward using the final assimilated solution from the previous cycle as the initial condition. Then another

assimilation is performed for the second cycle using this new background. This process is repeated for all subsequent cycles. Ngodock et al. (2007a,b) demonstrate that the Cycling Representer Method can be extremely beneficial in situations where the TLM is not stable for the entire assimilation time window. It is shown that by cycling over time periods less than the stability time-frame of the TLM not only eliminates the instability problem, it significantly reduces the overall cost of the assimilation, particularly when the need for outer-loops is dropped. Outer-loops (also known as Picard iterations) are typically required for NL models; they are iterations over the linearizations of the NL Euler–Lagrange problem associated with the minimization of the cost function involving the NL model (Ngodock et al., 2000; Muccino and Luo, 2005; Rosmond and Xu, 2006; Kurapov et al., 2007). By using the Cycling Representer Method, the need for outer-loops is reduced because the background is being updated in each subsequent cycle. Ngodock et al. (2007a,b) show that once the system is spun-up (typically after the first few cycles), the background is trained towards the data and the TLM is sufficiently stable to eliminate the need for outer-loops.

Many of the previous applications of the Cycling Representer Method have been with either linear models or simple, low-dimensional NL problems. In this study, the validity of the Cycling Representer Method is taken one step forward and demonstrated in a realistic application with the Navy Coastal Ocean Model (NCOM). NCOM is a NL, multi-layered, baroclinic ocean model that is designed to resolve coastal features (Martin, 2000; Barron et al., 2006; Kara et al., 2006). Experiments are performed by assimilating velocity measurements from an array of 14 acoustic Doppler current profile (ADCP) moorings deployed on the shelf and slope of the Mississippi Bight during the month of June 2004 (Fig. 1). This data set and region are favorable in demonstrating the importance and uniqueness of this assimilation technique because velocity measurements are notoriously difficult to assimilate into ocean models (Smith and Jacobs, 2005; Smith et al., 2005). This is especially the case in highly dynamical shelf-break regions such as the Mississippi Bight where the circulation is dominated by multiple processes such as inertial oscillations, winds, river outflow, and intruding eddies, thus causing the velocity observations to have a wide range of values (Teague et al., 2006; Carnes et al., 2008). All ocean models (including NCOM) have a difficult time accurately accounting for all of the dynamics in a shelf-break region and representing a high-resolution velocity data set (such as the one used in this study). Prior to performing the assimilation experiments presented in this paper, NCOM fields were compared to the data, and more times than not the circulation fields were drastically different (sometimes completely opposite). Such a velocity discrepancy may cause simpler assimilation techniques (ones that do not take into account the model physics) to produce dynamically inconsistent analyses. The Cycling Representer Method is unique in that by assimilating over shorter time periods and continuously updating the background, there is an improved capability of keeping the solution inline with the data and increasing its accuracy with subsequent cycles. More importantly though, is that initial condition, boundary condition, and dynamic error covariances can be specified to account for the unresolved oceanic processes. Therefore, during the cost function minimization, the dynamics, along with the initial and boundary conditions, will be modified over the entire space and time domain to produce a dynamically consistent solution that best matches the data (within the specified error limits).

This paper is a continuation of previous studies (Ngodock et al., 2007a,b) and is an ongoing effort towards demonstrating the usefulness of applying the Cycling Representer Method to an operational assimilation system for coastal regions that can assimilate a variety of measurement types and produce accurate forecasts
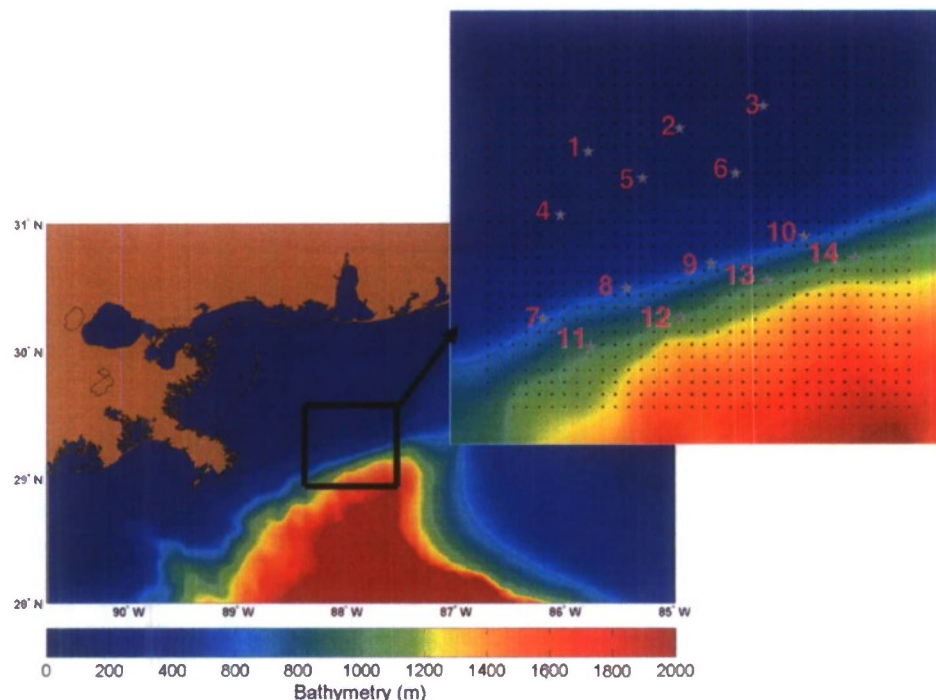
Fig. 1. The Mississippi Bight domain used for this study (black box). The 30 × 34 black dots represent the discretized grid of the model and the 14 numbered gray stars represent the ADCP mooring locations.

in real-time. It is important to note that the assimilation system presented in this study is preliminary and there are obvious areas where improvement is needed (such as the specification of error covariances, and the inclusion of additional data types and weak constraint variables). Despite these shortcomings, this preliminary assimilation system is adequate for the fundamental objective of this study, which is to demonstrate the Cycling Representer Method with NCOM and to show the importance of properly specifying the cycling time period to overcome TLM instability.

In the next three sections of this paper, the setup of the experiment will be described. This includes a description of the domain and model (Section 2), the data (Section 3), and the assimilation technique (Section 4). In Section 5, the assimilation results will be presented for three different cycling experiments, and then discussed in Section 6. Then, the final section (Section 7) will contain some concluding remarks.

2. The forward ocean model

The forward ocean model used in this assimilation experiment is NCOM. NCOM is similar to the Princeton Ocean Model (POM; Blumberg and Mellor, 1987) in that it is a free-surface ocean model based on the primitive equations and the hydrostatic, Boussinesq, and incompressible approximations. Also, both NCOM and POM are solved on an Arakawa C-grid and is leapfrog in time with an Asselin filter. NCOM differs from POM primarily in that the free-surface is treated implicitly and NCOM uses both sigma coordinates for the upper layers and z-level coordinates for the lower layers. Further detailed specifications of NCOM can be found in Martin (2000), Barron et al. (2006), and Kara et al. (2006).

Fig. 1 displays the discretized grid that is used for the NL forward model, the TLM, and the adjoint model needed for the assimilation experiments presented in this paper. The 30 × 34 black dots are spaced 2.5 km apart and represent the center points of the Arakawa C-grid at which sea surface height (SSH), salinity, and

temperature are solved. This grid resolution requires a 4-min time-step for numerical stability. In the vertical, there are 40 layers with 19 sigma layers in the upper 137 m to resolve the shelf-break. The bathymetry is extracted from a Naval global bathymetry database (DBDB2) with 2-min resolution. All of the atmospheric forcing, including wind stress, atmospheric pressure, solar radiation, and surface heat flux, is interpolated from the Navy Operational Global Atmospheric Prediction System (NOGAPS) (Hogan and Rosmond, 1991), which has a horizontal resolution of 1° and is saved in 3-h increments. Within NCOM, open boundary conditions are determined using solutions from a larger model and: Flather boundary conditions for SSH, advection boundary conditions for the normal velocity components, and Orlanski boundary conditions for the tracer fields and the velocity components parallel to the boundaries.

The NL forward ocean model has two distinct purposes within the cycling system. First, a global solution is needed for the initial conditions for the first cycle, and open boundary conditions for each time-step of the entire time period. In regards to this paper, the term 'global solution' refers to a solution that is computed independently from this experiment and is larger than the assimilation domain. For these experiments, historical results are extracted from the operational 1/8° global NCOM solution (Barron et al., 2006) and used for this purpose. The resolution ratio between this global solution and the assimilation domain is quite large (about 5.5:1). When nesting models, a ratio of 3:1 is typically preferred. However, a coarse resolution global solution is desired for the purpose of demonstrating the Cycling Representer Method. This is because when interpolated and compared to the high-resolution model, the background is relatively inaccurate, which therefore causes the TLM stability to be relatively short. Since the horizontal and temporal resolution of the historical global NCOM solutions are significantly lower than what is needed for these experiments, values are linearly interpolated to the experiment grid and time-steps. Vertical interpolation is not needed because

the Mississippi Bight domain uses the same vertical structure as global NCOM (40 layers, 19 of which are sigma layers).

The second purpose of the NL forward ocean model is to create a high-resolution forecast for each cycle using the initial and boundary conditions from the above global solution. Note that the global initial conditions are only needed for the first cycle; subsequent cycles use the assimilated solution from the previous cycle as the initial conditions for the NL forecast. Each cycle forecast of SSH, temperature, salinity, and the two velocity components are saved on the local Mississippi Bight grid at 20-min intervals. This forecast solution is later used as the background field needed for the TLM, representer, and adjoint models (discussed in Section 4.1).

3. The data

3.1. The ADCP velocity measurements

An array of 14 acoustic Doppler current profiler (ADCP) moorings was deployed by the Naval Research Laboratory (NRL) for 1 year (May 2004–May 2005) along the shelf, shelf-break, and slope of the Mississippi Bight (about 100 miles south of Mobile, Alabama). These moorings were spaced about 10–20 km apart and are identified in Fig. 1 as the numbered gray stars. The shelf moorings (1–6) consisted of trawl resistant bottom mounted ADCPs located along the 60 and 90 m isobaths. The measurements collected from these shelf moorings were binned into 15-min time intervals and 2 m depth intervals. The slope moorings (7–14) consisted of long-range ADCPs buoyed at 500 m depth and located along the 500 m and 1000 m isobaths. Velocity measurements for the upper 500 m were collected and binned in 1-h time intervals and 10 m depth intervals. Even though the tides in this region have relatively small amplitudes, the tidal signal is removed from the data by performing a harmonic analysis on the year-long time series and removing the components of velocity associated with the top 8 tidal constituents (M2, S2, K1, O1, n2, p1, k2, and q1).

Teague et al. (2006) and Carnes et al. (2008) provide an extensive analysis of the collected velocity data set described above. They show that during the time period under examination in this study (the month of June, 2004), the velocity data on the slope (moorings 7–14) exhibits a general transition of the flow field from being predominantly westward to eastward. Also, the flow on the slope had a strong correlation with the wind stress ($\sim$0.8) and was fairly uniform in the along-shelf direction with a slight cross-shelf current towards the shore. Whereas, the circulation on the shelf (moorings 1–6) exhibits a weaker correlation with wind stress (less than 0.6), strong inertial oscillations with a period of about 24 h, and a substantial velocity shear in the water column. The time series of velocity presented in Fig. 2 are for mooring 2 (see Fig. 1 for location) and clearly show that inertial oscillations dominate the circulation on the shelf for the majority of the month and throughout most of the water column.

3.2. Data sampling and the measurement functional

For the experiments presented in this paper, ADCP velocity measurements were extracted from the above data set for the month of June 2004. However, since the measurements have a very high-resolution in the vertical and temporal dimensions, it is unrealistic to assimilate all of the velocity measurements, even for a very short time interval. For the entire time period of this experiment, there are roughly 1.75 million individual measurements (the u- and v-components of velocity are considered as 2 distinct measurements). Therefore, the data must be sampled at a prescribed temporal and vertical frequency. After careful examination of the entire dataset it is apparent that the dominant temporal fea-

ture is inertial oscillations on the shelf with a frequency of about 24 h (Fig. 2). To resolve these features, a sampling frequency of 3 h is used. In the vertical, it is apparent that on average the velocity profiles can be represented with five layers or less. Therefore, at every 3-h increment the two components of velocity are sampled at up to five different depths at each of the 14 mooring locations. Consequently, this reduces the average number of measurements within a 12-h assimilation cycle to 538. The five measurement depth locations are automatically selected independently at each sampling time increment and at each horizontal mooring location. In order to illustrate how velocity measurements are sampled in the vertical, Fig. 3 shows a 3-h time series of velocity profiles from mooring 6. In this example, the sampling time is 15:00 (June 7, 2004) and all velocity profiles (blue vectors) within ±1.5 h of this sampling time are used to determine the sampling depths for this time. For each velocity profile within this time period, the four strongest vertical gradients in velocity are determined (green dots in Fig. 3). These strongest gradients are then averaged over the 3-h time period (dashed green lines) and used to define the interfaces between the 5 layers. Finally, the measurements closest to the center of each of the five layers at the sampling time (15:00) are sampled (red vectors). If there is not a unique measurement in between two interfaces (i.e., the dashed green lines are too close together) then that particular layer is ignored for that sampling time.

For each of the sampled measurements, a measurement functional is created to translate the state space into the measurement's data space. Remember that the 2 velocity components (u and v) are separate measurements. The measurement functional can be as simple as a value of 1.0 at the grid point location closest to the measurement and zeros for the rest of the state (i.e., a dirac delta function). Since not all of the data are used, however, it is desired that each measurement functional represents the region of state space that encompasses the measurement and includes the area of the neglected data. Therefore, each measurement functional will include representation ±1.5 h of the sampling time, the entire layer that the measurement represents, and a 3 grid-point horizontal radius surrounding the measurement. Each of these grid points, however, does not receive equal weighting. A Gaussian function is used to distribute the representation in all four-dimensions with the grid points closest to the measurement receiving the largest weighting. Finally, the measurement functional is scaled so that the sum of all nonzero values is equal to 1.0.

In addition to the measurement values and their associated functionals, an error is required for each measurement. The measurement error for each observation is unique and is estimated to be the representation error that the sampled measurement has with respect to the neglected data that the measurement is representing. The instrument error is well under 1 cm/s and is considered negligible relative to this representation error. The representation error is calculated by taking the root-mean-square (RMS) of the difference between the measurement value and each of the neglected data that falls within the ±1.5 h time window and the layer that the measurement represents. For the example in Fig. 3, the error associated with the deepest u-component of velocity (bottom red vector) is computed by taking the RMS of the difference between this value (-3.7 cm/s) and all of the u-components of the blue vectors below the bottom green line; this error comes out to be 2.3 cm/s.

4. The assimilation system

4.1. The linearized, representer, and adjoint models of NCOM

Since the Representer Method is fundamentally performed by minimizing a quadratic cost function, the model must first be
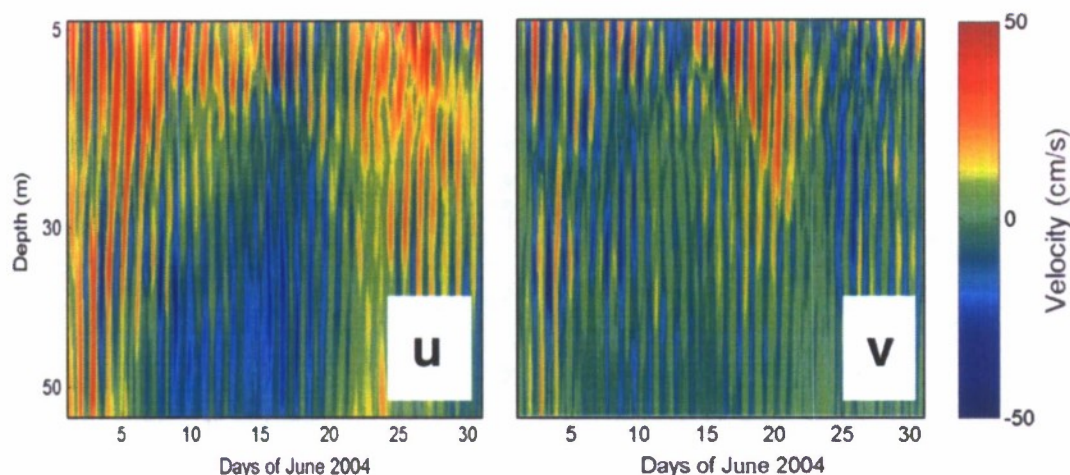
Fig. 2. The velocity time series of ADCP mooring #2 (see Fig. 1) exhibits strong inertial oscillations on the shelf during the entire month of June, 2004 and throughout the majority of the water column. The data has been detided using a harmonic analysis.
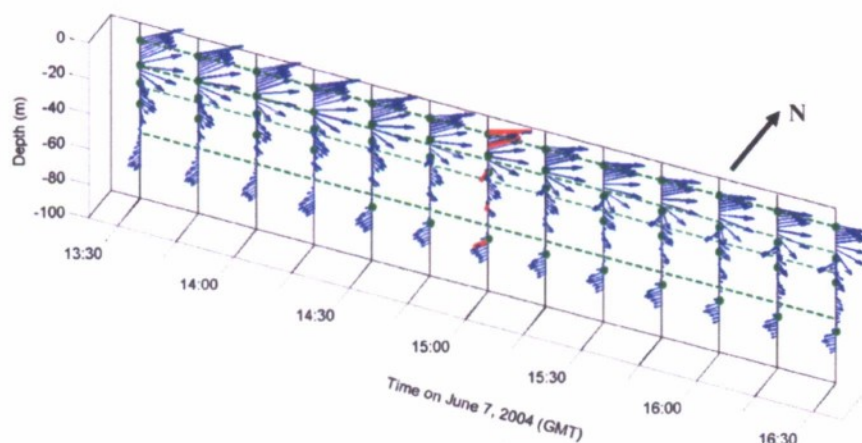

Fig. 3. A 3-h time series of velocity profiles from mooring 6 is presented here to demonstrate how the assimilated data is sampled from the original data set. Velocity data from each mooring is assimilated every 3 h at up to 5 layers. This example shows the ±1.5 h time period surrounding the assimilation time of 15:00 GMT on 6/7/04. For each profile during this time period, the four strongest vertical gradients in velocity are calculated (green dots). The dashed green line is the mean location of the gradients over the time period and represents the interfaces of the 5 layers. The velocities closest to the center of each layer and in the middle of the time period (highlighted red) are assimilated. All of the velocity data (blue vectors) are used to estimate the representation error of the assimilated data for each layer.

linearized in order for an absolute minimum of the cost function to be determined. This is accomplished by using the first-order approximation of Taylor's expansion of the NCOM dynamics expanded about a background solution. If the discretized NL model is $\mathbf{A}(\vec{x})$, then the TLM can be expressed as,

$$\mathbf{A}(\vec{x}_{BG}) + \frac{d\mathbf{A}}{d\vec{x}}(\vec{x}_{BG})[\vec{x} - \vec{x}_{BG}] \tag{1}$$

where $\vec{x}$ is the solution state and $\vec{x}_{BG}$ is a background state. Since the TLM is an approximation to the full NL model, errors are naturally going to emerge in the TLM and these errors are going to propagate and grow with time. The TLM error is plotted in Fig. 4 for the first 10 days of the experiment. The TLM accuracy is defined as the RMS of the velocity difference between the NL background and tangent linearized solutions. As can be seen, the TLM error is about 40 cm/s after only 1 day of propagation; after 3 days of propagation the error surpasses 1 m/s and then begins to grow exponentially. There are many contributors that can impact TLM error. The first and foremost contributor is the level of nonlinearity of the model. NCOM contains many highly nonlinear dynamical components that require

linearization such as: the Smagorinsky horizontal mixing scheme, the Coriolis and curvature terms, density, advection, and every nonlinear operation that includes depth (in NCOM, total depth is a function of SSH). In these experiments, bottom friction and the Mellor–Yamada 2.5 vertical mixing scheme are not tangent linearized and are computed using the background state. The inclusion of these two linearized components was attempted, but the TLM was too unstable in order to perform a reasonable assimilation experiment.

The second significant contributor that impacts the TLM stability is the accuracy of the background solution. In this study, background solutions are computed on the high-resolution grid using interpolated Global NCOM solutions for open boundary conditions (and initial conditions in the first assimilation cycle). Since Global NCOM has roughly 1/5 the horizontal resolution and 1/90 the temporal resolution (global NCOM solutions are archived in 6-h increments) relative to the high-resolution grid, inconsistencies arise fairly quickly during the calculation of the NL background. These inconsistencies will be discussed in greater detail in Section 4.3. Another important contributor that is reducing the TLM stability is the steep bathymetry in the southeast corner of the domain. This
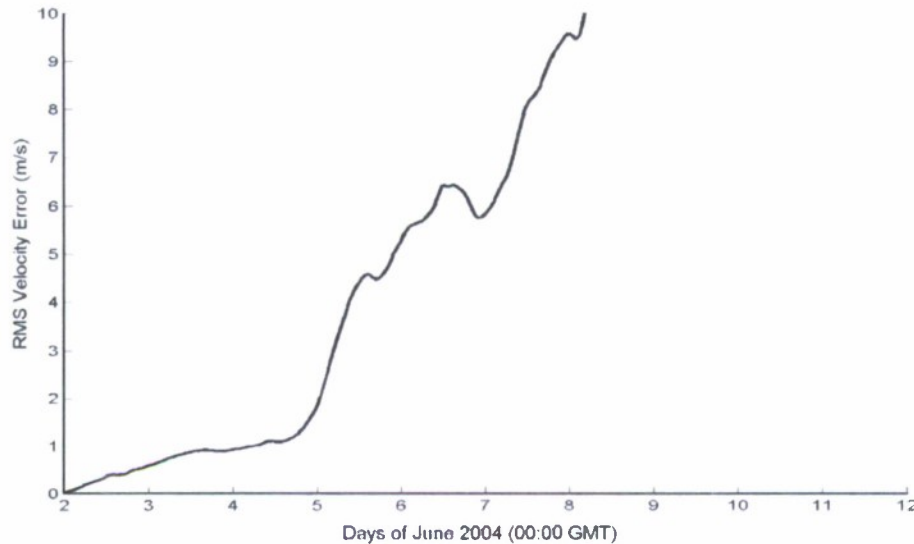
Fig. 4. TLM stability is represented here by the evolution of the RMS of the velocity difference between the tangent linearized and nonlinear solutions of NCOM. The TLM maintains relative accuracy for only about the first day, and then the error of the TLM surpasses 1 m/s after about 3 days of propagation.

steep bathymetry may be amplifying the NL baroclinic processes occurring at the shelf-break.

It is important to note that the Representer Method does not directly use the TLM. Instead, it uses just the perturbation component of (1),

$$\frac{d\mathbf{A}}{d\vec{x}}(\vec{x}_{BG})\delta\vec{x} \tag{2}$$

where $\delta\vec{x} = \vec{x} - \vec{x}_{BG}$. The above equation is typically referred to as the representer model, because it is used to propagate the perturbation of the state forward in time to create the representer functions. The only difference between Eqs. (1) and (2) is the addition of the background model $\mathbf{A}(\vec{x}_{BG})$ in the calculation of the TLM. This term is essentially an offset of the perturbation and therefore does not impact the stability of the linearization. This means that the characteristics that impact TLM stability and cause the linearization errors displayed in Fig. 4 are the same for the representer model. Even though it is the representer model that is being used in the assimilation experiments, throughout this paper the errors associated with linearization are referred to as TLM error (or TLM stability).

Another key component of the Representer Method is the adjoint model, which is simply the transpose of the representer model,

$$\left[\frac{d\mathbf{A}}{d\vec{x}}(\vec{x}_{BG})\right]^{\mathrm{T}} \tag{3}$$

However, since the propagation of NCOM dynamics is mostly explicit, it cannot be expressed in matrix notation and therefore cannot be directly transposed. Therefore, the adjoint was manually generated by reversing all of the operations in space and time.

4.2. A priori error covariances

Model error covariances ought to be specified for each variable that is believed to be in error (weak constraint). However, since error covariances are among the more difficult quantities to accurately estimate and can be computationally expensive to include (Bennett, 2002; Talagrand, 1999), it is not feasible to treat all variables as weak constraint; and for the variables that are treated as weak constraint, simplifications must be made in order to construct and implement their error covariances. Since the purpose

of this paper is to demonstrate the Cycling Representer Method, a sufficient effort was not dedicated to constructing accurate error covariances.

The assimilation results presented in this paper were computed by treating the initial conditions, the boundary conditions, and the interior solution of both velocity components as weak constraints. The covariance for each of these weak constraint variables was constructed with a constant variance in space and time, and isotropic correlations described by a Gaussian function in space and a moving average in time. An error of 5 cm/s was used for the initial and boundary conditions, and an error of 1 cm/s was used for the interior solution. The Gaussian functions used to construct the correlations were formulated with an e-folding scale of 15 grid points in each of the 3 spatial dimensions and 2 h in time.

The large spatial dynamic error decorrelation length scale is chosen as a balance between the length scales of features resolved by the data and the global solution. It is desired that the solution be able to resolve the features represented by the data. However, since the assimilation system is driven by the coarse global solution through the open boundary conditions, the solution needs to be smooth enough to provide an initial condition that is consistent with these open boundary conditions in order to ensure a stable NL forecast. Therefore, the correlation length scale of the dynamic error was chosen to be the average between the length scale of the observation correlation, and the length scale of features that the global solution can be expected to resolve. The length scale of the observations is computed by solving for L in the following e-folding equation,

$$\alpha = \exp(-d^2/L^2) \tag{4}$$

where d contains all of the distances between the 14 moorings and α is the data correlation matrix. Correlation is computed between all observations made in the month of June, 2004 from all moorings, at all depths, and both velocity components. The average length scale calculated from these correlations is 21 km. The horizontal grid resolution of the global solution is 13 km, and if, on average, 4 grid points are needed to resolve a particular feature, then the global solution can only be expected to resolve features with length scales of about 52 km. Therefore, the average correlation length scale that is used in this study is 37 km, which corresponds to about 15 grid points on the Mississippi Bight grid.

4.3. The Cycling Representer Method

The formulation of the Cycling Representer Method that is used in this paper is similar to the formulation described in detail in Ngodock et al. (2007a), therefore it is not prudent to repeat it here. The primary difference with the formulation that is used in this study is how the background for each new cycle is computed. The background for each cycle (other than the first cycle) is the NL forecast from the previous cycle. The forecast is computed by propagating the NL model forward using the final assimilated solution from the previous cycle as the initial condition and boundary conditions from both the previous forecast and the global solution. Fig. 5 displays the flowchart of the cycling methodology that is used in this study. At the beginning of each experiment, a NL background field is initially created encompassing the time period of the first 4 cycles using the global solution for the initial and boundary conditions (Fig. 5a–c). An assimilated solution is then computed using this background that encompasses the first cycle with an additional 3 h into the second cycle (Fig. 5d). The background for the second cycle is created by performing another 4-cycle forecast using the solution at the end of the first cycle as the initial conditions (Fig. 5f). The boundary conditions for this forecast are created by, first, extracting the boundary conditions from the initial background that fall within this new time period, and using a time weighted average to merge the first 3 h of these boundary conditions with the overlapping 3 h from the previous solution. Then, a time weighted average is used to merge boundary conditions from the global solution onto the end in order to encompass the whole 4-cycle time period (Fig. 5e). Following the calculation of this forecast, an assimilation is performed for the second cycle (Fig. 5g). This process is repeated for all subsequent cycles.

The 3-h overlapping and linear ramping of the boundary conditions that occurs in step (e) of Fig. 5 is needed to provide a smooth transition of boundary conditions from the three different solutions. Without this smoothing, it is likely that the forecast would be unstable. This is because there is typically a dynamic inconsis-

tency between its initial conditions (taken from the previous assimilated solution) and its boundary conditions (taken from the previous forecast). Inconsistencies can also arise between the previous forecast and the global solution resulting from the large difference in resolution. This inconsistency is why each forecast is propagated over a time period of four cycles. By computing a forecast over an extended period of time, the system is given a few cycles to filter out the impact of merging the global solution before it is used as the background for assimilation.

As is the case in Ngodock et al., 2007a,b, the Representer Method for each cycle is solved indirectly using an iterative conjugate gradient method (CG). This differs from the direct approach in that instead of computing representers for every measurement and inverting the representer matrix to obtain the representer coefficients, the indirect method uses the CG to iterate through the search directions in data space to converge upon the representer coefficients that minimize the cost function. In this study, the CG is considered converged and the stopping criteria satisfied when the norm of the residuals relative to the initial norm is less than 10^{-3}. For a well-conditioned problem, the number of required CG iterations is significantly less than the total number of measurements, therefore resulting in a substantial savings in computational cost compared to the direct method.

5. Cycling experiment results

As a precursor to the cycling experiments, a long 10-day assimilation experiment was attempted, and the resulting solution misfit (red) is plotted in Fig. 6A. The background misfit (blue) is also plotted for comparison. These misfits are computed at each of the 3-h data sampling time increments and are the RMS of the difference between the data and the solution projected into data space using the measurement functionals described in Section 3.2. The results in Fig. 6A reveal that there is a fairly consistent correlation between the accuracy of the assimilated solution and the TLM stability shown in Fig. 4. The assimilation performs fairly well for the
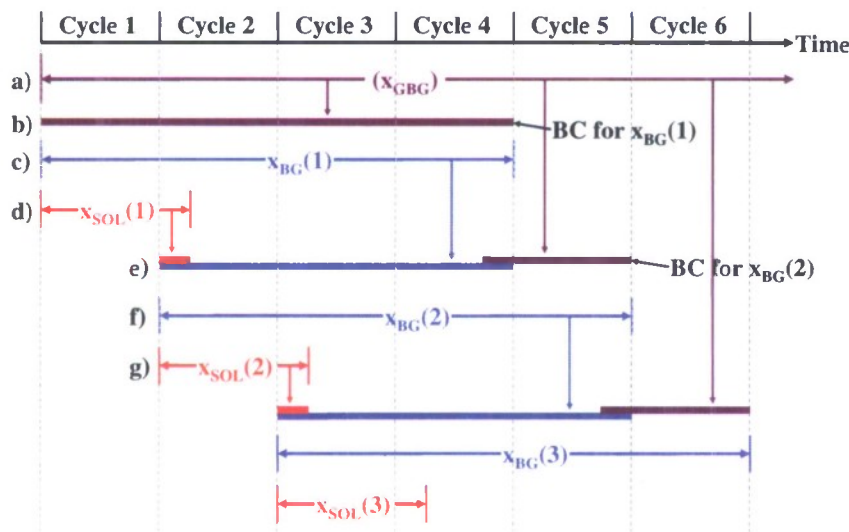
Fig. 5. Flowchart of the cycling methodology. (a) A global solution (purple) is needed that covers the entire time-frame and encompasses the domain of interest. (b) For the first cycle, initial and boundary conditions are extracted and interpolated from this global solution and used to (c) propagate a high-resolution NL forecast (blue) for 4 cycles. This forecast is then used as the background field for the (d) first cycle assimilation (red). (e) The boundary conditions for the next forecast are obtained from a combination of the previous solution ($x_{SOL}(1)$), the previous background ($x_{BG}(1)$), and the global solution (x_{GBG}). The 3-h overlap of $x_{SOL}(1)$ and $x_{BG}(1)$ at the beginning of cycle 2, and of $x_{BG}(1)$ and x_{GBG} at the end of cycle 4 are linearly ramped to create a smooth transition. (f) Using these combined BCs along with IC from $x_{SOL}(1)$, a NL forecast, $x_{BG}(2)$, is created for the next four cycles. (g) This forecast is then used in computing the next assimilated solution, $x_{SOL}(2)$. The steps (e–g) are repeated for each consecutive cycle.
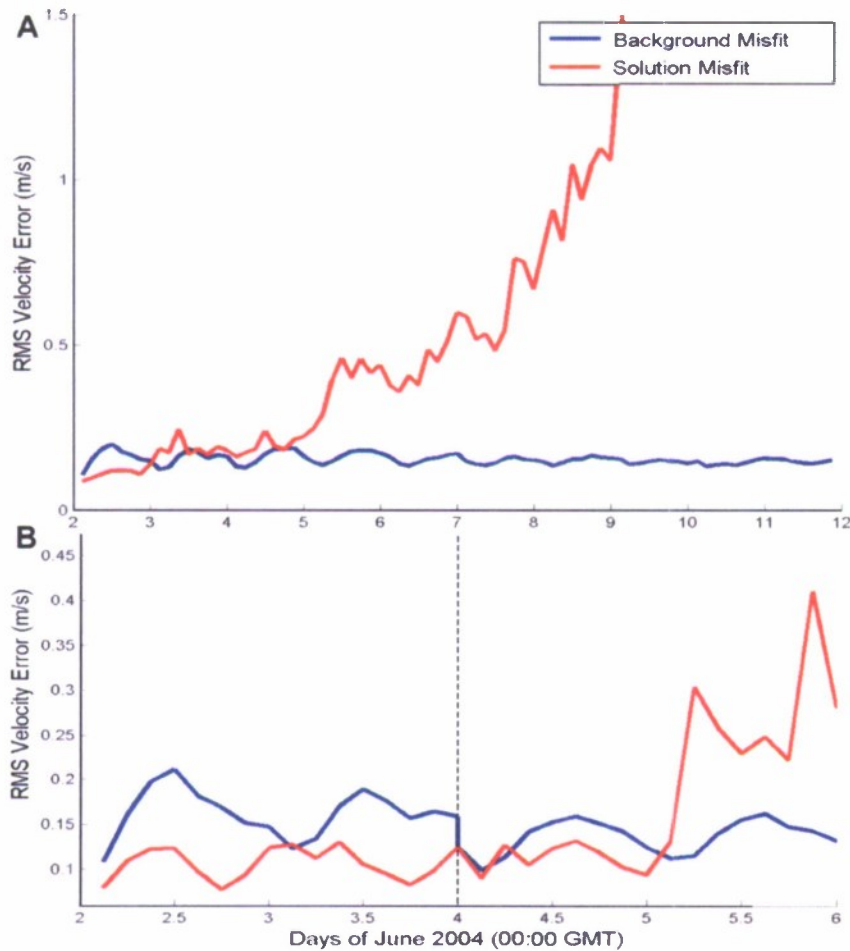
Fig. 6. RMS ($[L[u] - d$) is computed and plotted at each 3-h data sampling time stamp. L is the measurement functional that projects the velocity field from state space to data space, d is the data, and u represents velocity from the nonlinear background (blue) and the assimilated solution (red). (A) Ten day assimilation experiment. (B) Two day cycling experiment, where the dashed line represents the break between cycles.

first day, then over the next couple of days the assimilated solution begins to lose skill, and by the third day it becomes unstable and the errors begin to increase exponentially.

In a recent study, Ngodock et al. (2007a) were able to demonstrate with assimilation experiments with the Lorenz attractor that even though the TLM was accurate for only about 0.5 time units, the assimilated solution was able to track the data reasonably well for roughly 7 time units. Similarly, in Kurapov et al. (2007), the authors performed variational assimilation experiments with a nonlinear near-shore circulation model and demonstrated the impact that the error covariance specification has on the accuracy of the assimilated solution. They showed that by using a composite covariance with a temporal correlation separated into both steady state and small amplitude time-variable parts, accurate solutions could be obtained for time periods that significantly exceed the time interval of TLM stability.

It is believed that because of the overly simplified error covariances used in these experiments (Section 4.2), the assimilation is not able to perform accurately beyond the time period of TLM stability. Therefore, since the TLM error becomes unmanageable after about 1 day of propagation (Fig. 4), the assimilation cycle should be equal or shorter than this in order to ensure an accurate solution. Several cycling experiments were performed to determine the optimal cycle time period that produces the solution with the overall best-fit to the data. The optimal cycle time period needs to be short enough to ensure stability of the solution, and as long as possible to maximize the temporal correlation of the data and model.

5.1. Two day cycling experiment

The first cycling experiment that was performed employs 2-day cycles. The misfit results from this experiment are displayed in Fig. 6B and reveal that the first cycle achieved fairly accurate results, but the second cycle began to severely lose skill midway through the cycle. At the end of the second cycle the solution was too poor to provide a sufficient initial condition for the next forecast (the forecast grew numerically unstable). The dashed black line in this figure represents the break between cycles and the vertical portion of the blue line along this dashed line is a result of the background being reinitialized to the assimilated solution. As hypothesized above, it is apparent that a 2-day cycle time period is too long in order to ensure solution accuracy throughout the cycling.

5.2. Twelve hours cycling experiment

Fig. 7 displays the resulting misfits for the next cycling experiment consisting of 12-h cycles performed over 28 days. This assimilation experiment did quite well and there were no problems with providing a dynamically consistent solution as an initial condition for the forecasts. The general trend of the solution misfit changes
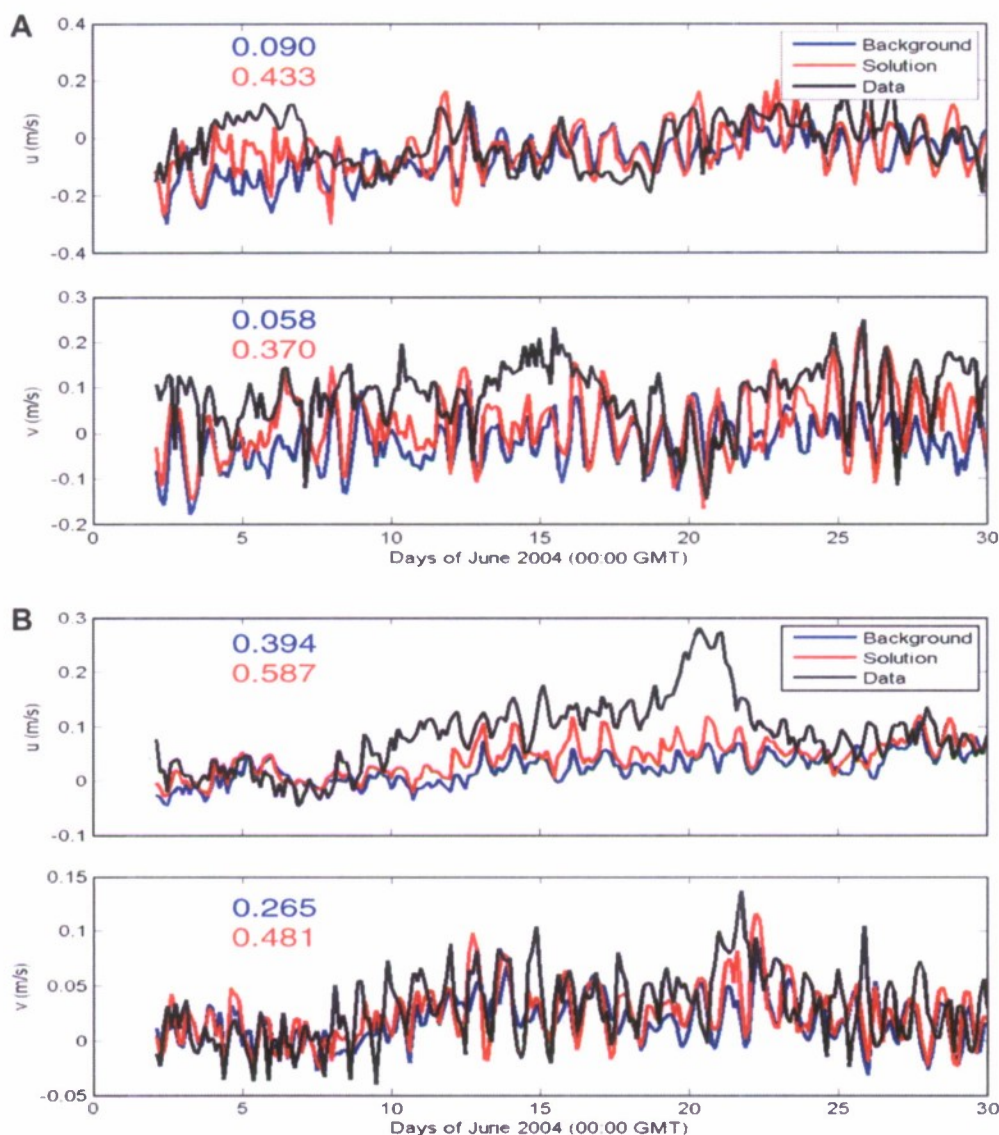

.

Fig. 8. Time series of velocity components of the data (black), assimilated solution (red), and background (blue) for the 12-h cycling experiment at (A) the first layer of mooring 7 and (B) the fifth layer of mooring 13 (mooring locations are labeled in Fig. 1). The colored values represent the corresponding correlation coefficient between the background and data (blue), and the solution and data (red).

time period. However, since the 1-day cycling proved to be inadequate at obtaining and maintaining accurate solutions during the latter half of the experiment, the 12-h cycling would have to be the clear choice in an operational setting.

The breakdown of the 1-day cycling experiment can be explained by examining the TLM stability for each cycle. Fig. 10 shows the RMS of the velocity difference between the background and TLM for each cycle. The dashed red line represents 1 standard deviation (STD) of the background velocity over the entire time period of the experiment and is used here as an error limit to gauge TLM stability. Except for the first cycle, the TLM error for each 1-day cycle remains less than this limit up to June 14. For each of the cycles between June 14 and June 21, the TLM error rises rather quickly and loses stability within the 1-day cycle time period. Beginning with the June 21st cycle, the TLM error goes down for a few cycles then back up again until the end of the experiment when it starts to decrease. There is a consistent correlation between the assimilated solution accuracy displayed in Fig. 9 and

the TLM stability displayed in Fig. 10 throughout the cycles. For each of the cycles mentioned in the previous paragraph that required an initial condition background infusion in order to produce a stable forecast, the TLM is inaccurate in the corresponding previous cycle.

There are several different mechanisms that could be causing the TLM stability to drastically fluctuate throughout the cycles. Large increases in TLM error can be caused by strong oceanic events (such as eddies and fronts) introduced directly through the open boundary conditions, as is most likely the case for the large peak in error on June 21 (Fig. 9). According to the global background solution (not shown), a warm-core eddy begins to penetrate the southern boundary of the domain towards the middle of the June 20th cycle. Another mechanism that can significantly impact the TLM stability is atmospheric events introduced through the wind stress. The rapid increase in TLM error exhibited in the June 14th cycle is a result of such a mechanism. Fig. 11 displays the time series of NOGAPS wind stress averaged over the
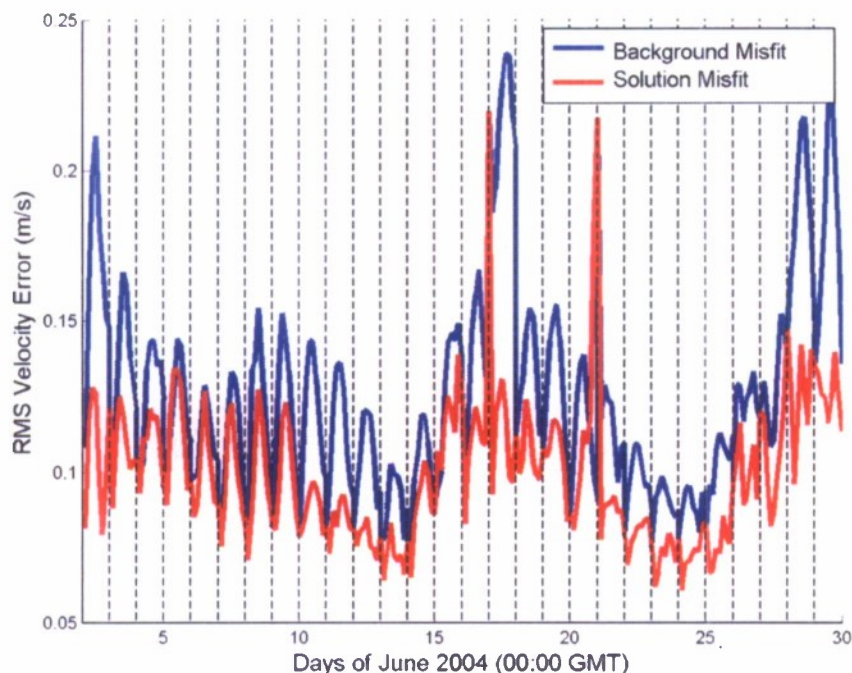
Fig. 9. This plot is similar to Fig. 7 for the 1-day cycling experiment.
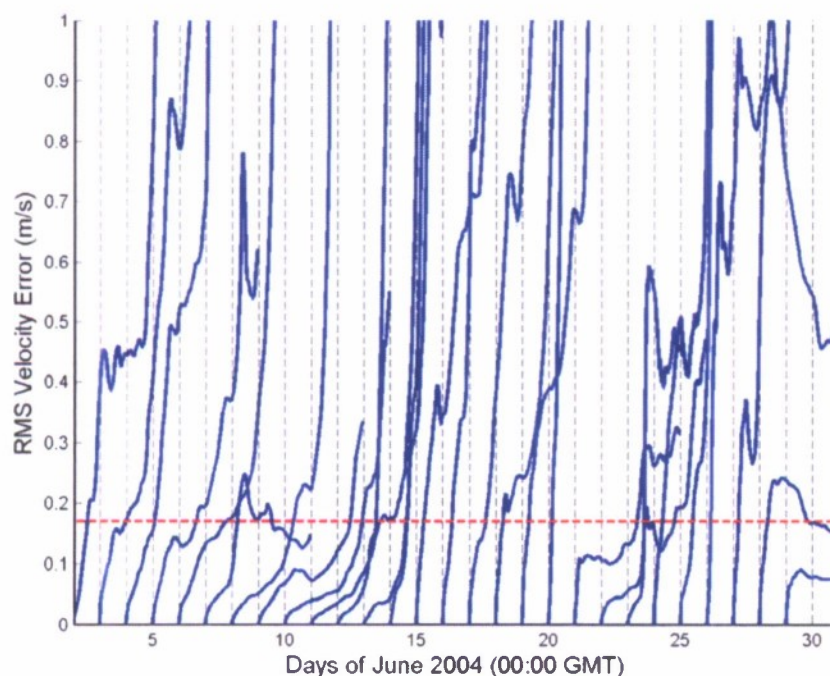
Fig. 10. TLM stability (similar to Fig. 4) for each of the 1-day cycles in Fig. 9. Each blue line, beginning at the start of each cycle, is the RMS velocity difference between the background forecast and the TLM for the corresponding cycle. The dashed red line is 1 standard deviation of the background velocity over the entire month of June, and is used here to gauge TLM stability. Other than the first cycle, the TLM is stable over each of the 1-day cycles up to June 14, at which point the TLM errors begin to rise sharply.

Mississippi Bight domain, and there is clearly a strong wind event that occurs during this cycle. Fig. 12 shows several snapshots of the TLM response to this strong southeasterly wind burst. The first snapshot is the TLM initial condition of surface velocity and temperature for the cycle beginning June 14 (this is also the final solution from the previous cycle). The second snapshot is 9 h into the cycle when the southeasterly wind burst has caused a large build-up of cold water in the northwest corner of the domain. By

the end of the 1-day cycle (third snapshot), it is apparent that the shock from the wind burst has caused a significant build-up of numerical noise in the TLM (especially on the shelf).

6. Discussion

The main purpose of developing the Cycling Representer Method is to establish a robust 4DVAR assimilation system that
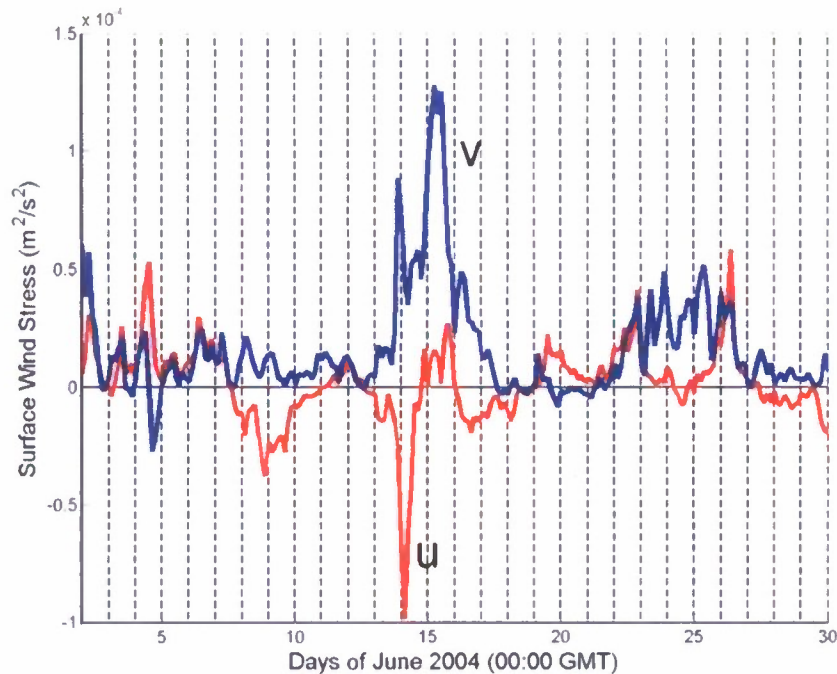
Fig. 11. Time series of the horizontally averaged surface wind stress. The wind stress is from NOGAPS and is used to force NCOM. There is a very large southeasterly wind burst on June 14, which corresponds to the rapid decrease in TLM stability in Fig. 10.

can be implemented in an operational setting. The Cycling Representer Method is ideal for this effort, because, the cycles can be shortened in order to maintain TLM accuracy, especially during highly nonlinear events that rapidly degrades the TLM stability (Section 5.3). A direct consequence of decreasing the cycle time period, however, is that the temporal correlation of the data and model is reduced, therefore reducing the level of influence that the data can have on correcting the solution. This disadvantage can result in the inability to resolve features with time scales larger than the cycle length. For example, the 1-day cycling was able to better resolve the 24-h inertial oscillations and therefore had slightly lower errors than the 12-h cycling.

There are several assumptions that were made in this study that deteriorated the quality of the assimilation and reduced the time period of TLM stability. These assumptions will require further investigation before this system can be applied operationally. The first of these assumptions is that only the velocity components are treated as weak constraint. As a minimum, the other prognostic variables (SSH, temperature, and salinity) should be treated as weak constraint, but other variables such as wind stress ought to be included too. The second assumption that was made was in the creation of the over-simplified error covariances. The error correlations should at least be anisotropic and a significance test ought to be performed to determine the appropriate rescaling of the error variances (it is demonstrated below in Section 6.1 that the specified error variances are too low).

Another assumption that was made in this study, is that 1 outer-loop is sufficient. Earlier applications of the Cycling Representer Method with simpler nonlinear problems (Ngodock et al., 2007a,b) concluded that only a single outer-loop was needed beyond the first few cycles to ensure accurate assimilation throughout the cycles. Because of this conclusion, we hypothesized that the experiments with NCOM could be successful without the use of outer-loops. However, results presented below (Section 6.2) show that there is at least a moderate improvement when a second outer-loop is used. In Rosmond and Xu (2006), the authors performed

Cycling Representer experiments using 6 outer-loops, however, their results showed that only 2 outer-loops were necessary to achieve over 90% of the nonlinear adjustments. These authors also demonstrated that the computational expense of performing additional outer-loops is not significant. This is because the number of required inner iterations of the CG can be significantly reduced with each successive outer-loop. Future development of this assimilation system will include the usage of outer-loops.

6.1. Performance of the Cycling Representer Method

To analyze the performance of the Cycling Representer Method and gauge how well it might do if the system were allowed to continue to cycle indefinitely, several post-processing results are presented here for the entire 12-h assimilation cycle experiment. To demonstrate the value of solving the Representer Method indirectly, Fig. 13 shows the evolution of the CG convergence for each of the 56 cycles (black lines). Overlaid on this wide spread of results is the best-fit exponential curve (thick red line). This curve conveys that the average total number of CG iterations required to reach the convergence criteria (10^{-3}) is about 62, which is roughly 12% of the average number of measurements in each cycle (538). This result corresponds to a substantial savings in computational cost relative to using the direct method and reveals that the assimilation problem is generally well-conditioned.

Fig. 14 displays the total number of required CG iterations (blue line) and the total cost of the cost function (green line) relative to the 12-h cycles. Both of these results are quite noisy, but over the entire 28-day time period their corresponding linear best-fits have a downward trend. The slight overall decrease in the number of required CG iterations as the system is cycled suggests that the assimilation problem for each consecutive cycle is becoming better conditioned, and the steeper downward trend of the total cost reveals that the overall fit between the assimilated solution and the data, dynamics, initial conditions, and boundary conditions is improving as the cycling progresses forward. The total cost
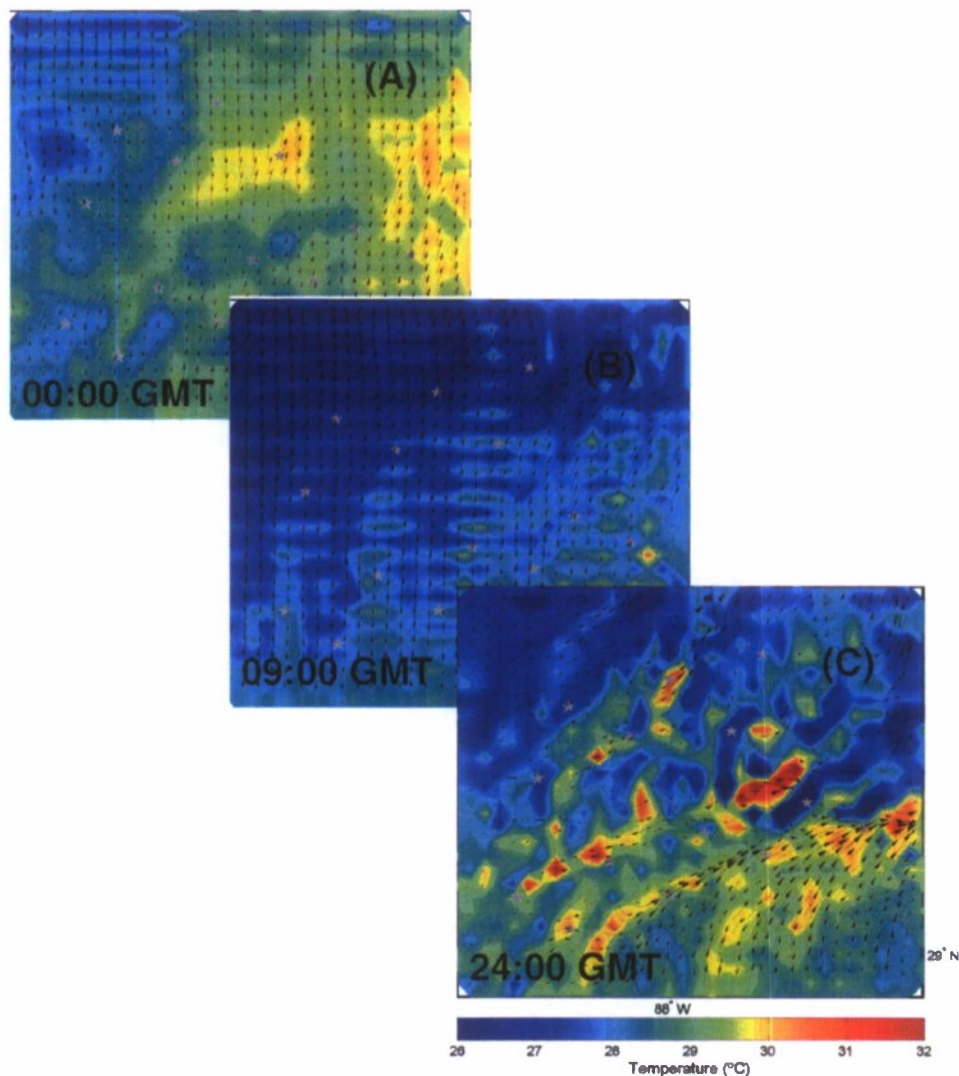
Fig. 12. The TLM response to the wind burst displayed in Fig. 11. Plots are of the surface temperature and velocity at (A) the beginning of the cycle (00:00, 6/14), then at (B) 09 h into the cycle the southeasterly wind burst has caused a large build-up of cold water in the northwest corner. (C) By the end of the 1-day cycle, the shock from the wind burst has caused the TLM solution to be dynamically inconsistent and in significant error, especially on the shelf.

function values displayed in Fig. 14 (green), however, are relatively high. The cost function at its minimum (J_{min}) should have a chi-square distribution with M degrees of freedom, where M is the number of measurements (Bennett, 1992). Therefore, the expected value of J_{min} should be M with a variance of $2M$. Since there are, on average, 538 measurements in each of the 12-h cycles, then the expected value of the total cost should be,

$$E[J_{min}] = M \pm \sqrt{2M} = 538 \pm 33 \tag{5}$$

The fact that the actual values of J_{min} are significantly larger than there expected values reveal that the model errors specified in the a priori error covariances are too small and should be rescaled by a factor of $\sqrt{J_{min}/M} \approx 4.0$ in order to bring the total cost values down to an appropriate expected value. However, as stated previously, the proper specification of the error covariances is not a focal point of this paper; therefore, this work will be saved for future development.

It is uncertain how long the trends in Fig. 14 would continue downward if the cycling were to continue beyond this 28-day experiment. This result, however, is promising in that the time period appears to be long enough to contain a significant portion of the dynamic features that are prevalent in this region, and there is no reason to assume a drastic change in these trends.

6.2. Nonlinear iterations

Since the assimilation experiments presented in this study are based on the linearization of a model, there is going to be an inherent error associated with this approximation. Performing nonlinear outer-loops is a method to try and overcome this error, but it can be computationally expensive if not implemented correctly. It entails reinitializing the background with the previous assimilated solution and then performing another assimilation for each outer-loop. As an example, 3 outer-loops were performed on the first cycle of the 12-h cycling experiment. The resulting misfits for this example are presented in Fig. 15. They show that performing a second nonlinear iteration (dashed red line) moderately improves the solution, but performing a third iteration (solid green line) does not improves the solution at all. Therefore, in this experiment it may have been beneficial to apply a second outer-loop to all the cycles.
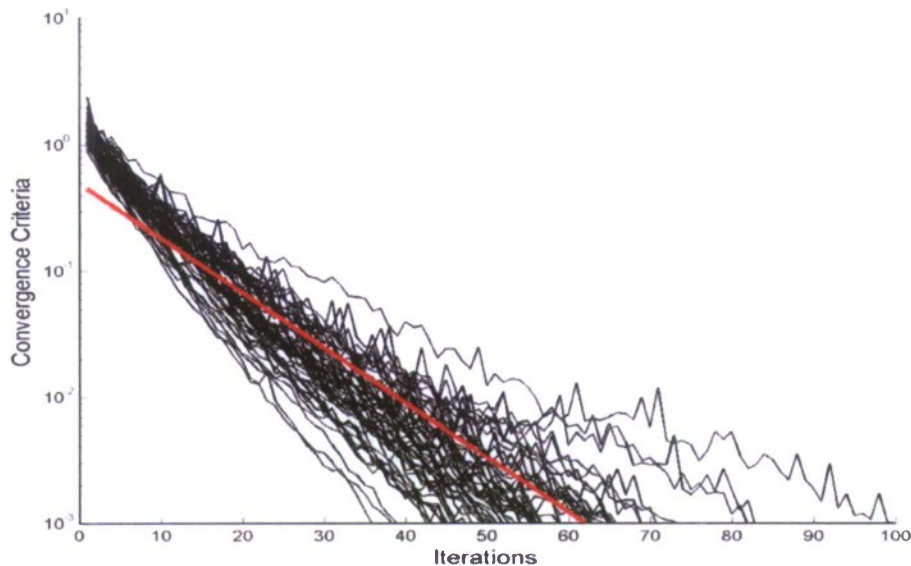
Fig. 13. The evolution of CG convergence criteria for each of the 56 12-h assimilation cycles (black lines). The convergence criterion is calculated as the norm of the residuals relative to the initial norm and is assumed converged when it is less than or equal to 10^{-3}. The red line is the exponential best-fit of the spread of CG convergence plots, showing that on average the 12-h cycle converged in about 62 CG iterations. This is about 12% of the average number of measurements (538) in each cycle.
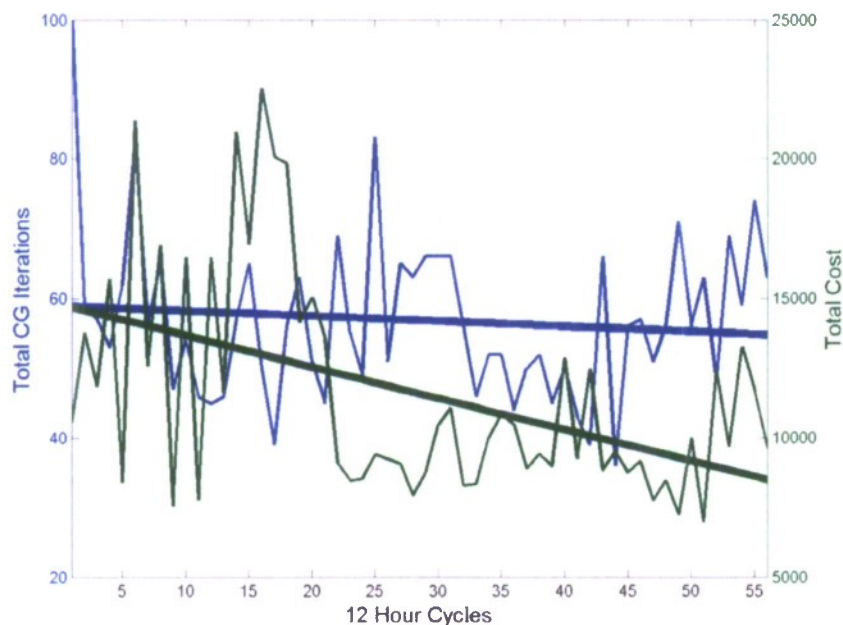
Fig. 14. Progression of the total number of required CG iterations (blue) and the total cost of the cost function (green) through the 12-h cycles. Even though both of these results are quite noisy, their corresponding linear best-fits (thick lines) have a downward trend as the number of cycles is increased.

7. Conclusions

Due to the extreme conditions that were deliberately specified for this experiment (a coarse background field, open boundary conditions along all four boundaries, complex bathymetry, etc.) the TLM of NCOM is only accurate for about a day. It is demonstrated in this study that the Cycling Representer Method can be performed using 12-h cycles to overcome this complication and achieve an improved analysis of the flow field for what appears to be an indefinite period of time. Also, because the

forecasts are reinitialized to the analysis field at the beginning of each cycle, both the forecast and analysis errors decrease as the system progresses through the cycles. Care needs to be taken though in selecting the appropriate time-frame of the cycle. If the cycle is too long and is beyond the range of TLM accuracy, the solution will end up diverging from the data and obstructing the continuity of the cycling system, as is the case in both the 2-day and the latter half of the 1-day cycling experiments. On the other hand, if the cycles are too short, valuable information can be lost due to the lack of temporal representation of

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

Muccino, J.C., Luo, H., 2005. Picard iterations for a finite element shallow water equation model. Ocean Modell. 10, 316–341.

Ngodock, H.E., Chua, B.S., Bennett, A.F., 2000. Generalized inverse of a reduced gravity primitive equation ocean model and tropical atmosphere-ocean data. Mon. Weather Rev. 128, 1757–1777.

Ngodock, H.E., Smith, S.R., Jacobs, G.A., 2007a. Cycling the representer algorithm for variational data assimilation with the Lorenz attractor. Mon. Weather Rev. 135, 373–386.

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

Rosmond, T., Xu, L., 2006. Development of NAVDAS-AR: non-linear formulation and outer loop tests. Tellus A 58, 45–58.

Smith, S.R., Jacobs, G.A., 2005. Seasonal circulation fields in the northern Gulf of Mexico calculated by assimilating current meter, shipboard ADCP, and drifter data simultaneously with the shallow water equations. Cont. Shelf Res. 25, 157–183.

Smith, S.R., Jacobs, G.A., Leben, R.R., 2005. Determining the applicability of the barotropic approximation to the mean seasonal flow through the Tsushima/Korean Strait using variational assimilation. Deep-Sea Res. II 52, 1763–1783.

Talagrand, O., 1999. A posteriori evaluation and verification of analysis and assimilation algorithms. In: Proceedings of Workshop on Diagnostic of Data Assimilation Systems, ECMWF.

Teague, W.J., Jarosz, E., Carnes, M.R., Mitchell, D.A., P.J.Hogan, 2006. Low-frequency current variability observed at the shelfbreak in the northeastern Gulf of Mexico: may–october, 2004. Cont. Shelf Res. 26, 2559–2582.

Xu, L., Daley, R., 2000. Towards a true 4-dimensional data assimilation algorithm: application of a cycling representer algorithm to a simple transport problem. Tellus A 52, 109–128.

Xu, L., Daley, R., 2002. Data assimilation with a barotropically unstable shallow water system using representer algorithms. Tellus 54, 125–137.

Xu, L., Rosmond, T., Daley, R., 2005. Development of NAVDAS-AR: formulation and initial test of the linear problem. Tellus A 57, 546–559.