

Seon K. Park · Liang Xu
(Eds.)

DS

SS

QSS

NSS

Data Assimilation for Atmospheric, Oceanic and Hydro- logic Applications



Springer

Contents

Sasaki's Pathway to Deterministic Data Assimilation	1
John M. Lewis	
Data Assimilation for Numerical Weather Prediction: A Review	21
Ionel M. Navon	
Theoretical and Practical Issues of Ensemble Data Assimilation in Weather and Climate	67
Milija Zupanski	
Information Measures in Ensemble Data Assimilation	85
Dusanka Zupanski	
Real Challenge of Data Assimilation for Tornadogenesis	97
Yoshi K. Sasaki	
Radar Rainfall Estimates for Hydrologic and Landslide Modeling	127
Kang-Tsung Chang, Jr-Chuan Huang, Shuh-Ji Kao and Shou-Hao Chiang	
High-Resolution QPE System for Taiwan	147
Jian Zhang, Kenneth Howard, Pao-Liang Chang, Paul Tai-Kuang Chiu, Chia-Rong Chen, Carrie Langston, Wenwu Xia, Brian Kaney and Pin-Fang Lin	
Assimilation of Satellite Data in Improving Numerical Simulation of Tropical Cyclones: Progress, Challenge and Development	163
Zhaoxia Pu	
Diagnostics for Evaluating the Impact of Satellite Observations	177
Nancy L. Baker and Rolf H. Langland	
Impact of the CHAMP Occultation Data on the Rainfall Forecast	197
Hiromu Seko, Yoshinori Shoji, Masaru Kunii and Yuichi Aoyama	

Parameter Estimation Using the Genetic Algorithm in Terms of Quantitative Precipitation Forecast	219
Yong Hee Lee, Seon Ki Park, Dong-Eon Chang, Jong-Chul Ha and Hee-Sang Lee	
Applications of Conditional Nonlinear Optimal Perturbations to Ensemble Prediction and Adaptive Observation	231
Zhina Jiang, Hongli Wang, Feifan Zhou and Mu Mu	
Study on Adjoint-Based Targeted Observation of Mesoscale Low on Meiyu Front	253
Peiming Dong, Ke Zhong and Sixiong Zhao	
Ocean Data Assimilation: A Coastal Application	269
Xiaodong Hong, James A. Cummings, Paul J. Martin and James D. Doyle	
Comparison of Ensemble-Based Filters for a Simple Model of Ocean Thermohaline Circulation	293
Sangil Kim	
Preconditioning Representer-based Variational Data Assimilation Systems: Application to NAVDAS-AR	307
Boon S. Chua, Liang Xu, Tom Rosmond and Edward D. Zaron	
Cycling the Representer Method with Nonlinear Models	321
Hans E. Ngodock, Scott R. Smith and Gregg A. Jacobs	
Implementation of the Ensemble Kalman Filter into a Northwest Pacific Ocean Circulation Model	341
Gwang-Ho Seo, Sangil Kim, Byoung-Ju Choi, Yang-Ki Cho and Young-Ho Kim	
Particle Filtering in Data Assimilation and Its Application to Estimation of Boundary Condition of Tsunami Simulation Model	353
Kazuyuki Nakamura, Naoki Hirose, Byung Ho Choi and Tomoyuki Higuchi	
Data Assimilation in Hydrology: Variational Approach	367
François-Xavier Le Dimet, William Castaings, Pierre Ngnepieba and Baxter Vieux	
Recent Advances in Land Data Assimilation at the NASA Global Modeling and Assimilation Office	407
Rolf H. Reichle, Michael G. Bosilovich, Wade T. Crow, Randal D. Koster, Sujay V. Kumar, Sarith P. P. Mahanama and Benjamin F. Zaitchik	
Assimilation of Soil Moisture and Temperature Data into Land Surface Models: A Survey	429
Nasim Alavi, Jon S. Warland and Aaron A. Berg	

Assimilation of a Satellite-Based Soil Moisture Product into a Two-Layer Water Balance Model for a Global Crop Production Decision Support System 449
John D. Bolten, Wade T. Crow, Xiwu Zhan, Curt A. Reynolds
and Thomas J. Jackson

Index 465

Cycling the Representer Method with Nonlinear Models

Hans E. Ngodock, Scott R. Smith and Gregg A. Jacobs

Abstract Realistic dynamic systems are often strongly nonlinear, particularly those for the ocean and atmosphere. Applying variational data assimilation to these systems requires the linearization of the nonlinear dynamics about a background state for the cost function minimization, except when the gradient of the cost function can be analytically or explicitly computed. Although there is no unique choice of linearization, the tangent linearization is to be preferred if it can be proven to be numerically stable and accurate. For time intervals extending beyond the scales of nonlinear event development, the tangent linearization cannot be expected to be sufficiently accurate. The variational assimilation would, therefore, not be able to yield a reliable and accurate solution. In this paper, the representer method is used to test this hypothesis with four different nonlinear models. The method can be implemented for successive cycles in order to solve the entire nonlinear problem. By cycling the representer method, it is possible to reduce the assimilation problem into intervals in which the linear theory is able to perform accurately. This study demonstrates that by cycling the representer method, the tangent linearization is sufficiently accurate once adequate assimilation accuracy is achieved in the early cycles. The outer loops that are usually required to contend with the linear assimilation of a nonlinear problem are not required beyond the early cycles because the tangent linear model is sufficiently accurate at this point. The combination of cycling the representer method and limiting the outer loops to one significantly lowers the cost of the overall assimilation problem. In addition, this study shows that weak constraint assimilation is capable of extending the assimilation period beyond the time range of the accuracy of the tangent linear model. That is, the weak constraint assimilation can correct the inaccuracies of the tangent linear model and clearly outperform the strong constraint method.

H.E. Ngodock (✉)

The Naval Research Laboratory, Stennis Space Center, MS 39529,
USA, e-mail: Hans.Ngodock@nrlssc.navy.mil

1 Introduction

The representer method of Bennett (1992) is a 4D variational assimilation algorithm that relies on the adjoint of the dynamical model and expresses the analyzed solution as a first guess plus a finite linear combination of representer functions, one per datum. The explicit computation and storage of all the representer functions (direct method), however, is not required since the method can be implemented indirectly (Amodei, 1995; Egbert et al., 1994) using the conjugate gradient method (hereafter CGM). A description of the representer methodology is provided in the Appendix. The representer method has earned an established reputation as an advanced data assimilation technique within the past decade, and gained the attention of many potential operational users. Two primary issues, however, need to be addressed prior to implementing the representer method operationally.

The first issue addressed in this paper is the stability and validity of the tangent linear model (hereafter TLM). When the representer method is applied to a nonlinear model, the model must be linearized, preferably using the 1st order approximation of Taylor's expansion. Traditionally, the representer method has been implemented for the assimilation of all observations in the time window considered. As with every other variational data assimilation method with nonlinear dynamics, the representer method necessitates that the TLM and its adjoint be valid and/or stable over the entire assimilation time window. The validity of the TLM is difficult to maintain over a long time period for strongly nonlinear models and complex regions.

The second issue addressed in this paper is the cost of the representer method. The indirect representer method requires the integration of the adjoint and TLM within a CGM that determines the representer coefficients for the minimization of the cost function (see Appendix). This set of representer coefficients is then used to provide a correction to the background. The number of iterations of the CGM (this is referred to as the inner loop) is typically a small fraction of the total number of measurements. For strongly nonlinear systems, outer loops are required. To initialize the outer loop, one would pick a first background solution around which the model is linearized. The best solution (corrected background) obtained from this assimilation would become the background for the next outer loop, and so on until formal convergence (Bennett et al., 1996; Ngodock et al., 2000; Bennett, 2002; Chua and Bennett, 2001; Muccino and Bennett, 2002). This outer loop exacerbates the computational cost of the representer method. In this study the background that serves for linearization is also taken as the first guess.

These two issues have discouraged many potential users of the representer method for operational purposes. It is possible, however, to address these issues and implement the representer method at a reasonable cost for operational applications. Given a time window in which one desires to assimilate observations, it is possible to apply the representer method over cycles of subintervals. The name adopted for this approach is the "cycling representer method" (Xu and Daley, 2000), and its associated solution is called the "cycling solution". The solution that is obtained by assimilating all the observations at once in the original time window will be called the "global solution".

By using the cycling representer method, the assimilation time window is constrained to a period over which the TLM produces an accurate dynamical representation of the nonlinear model. Doing this reduces the need for outer loops. Because the representer method solves a linear assimilation problem, the outer loop is designed to solve the nonlinear Euler-Lagrange conditions associated with the assimilation problem of the nonlinear model. In the global solution problem, the TLM may not be an accurate representation of the dynamical system, and the adjoint would not be an accurate estimate of the derivative of the state with respect to the control variables. If the TLM is an accurate representation of the dynamics, the need for outer loops is removed. In the initial cycles of this assimilation approach, the first guess or background solution may not be accurate and thus outer loops may be required. Once the system is spun up and the TLM is an accurate approximation (thanks to improved background solutions), outer loops may no longer be necessary, thus lowering the computational cost of the assimilation. However, there may be situations in real world applications where a few outer loops would be needed in the current cycle, even though a single outer loop sufficed in previous cycles. An example is a nonlinear ocean response (advection and mixing) to a sudden, stronger than normal, atmospheric forcing, especially in coastal areas with complex bathymetry. The need for additional outer loops may be assessed by the discrepancy between the assimilated solution and the data.

The idea of cycling the representer method was investigated by Xu and Daley (2000) using a 1D linear transport model with synthetic data. In that study, the error covariance of the analyzed solution was updated at the end of each cycle and used as the initial error covariance in the next cycle. Another application of the cycling representer method was performed by Xu and Daley (2002) using a 2D linear unstable barotropic model with no dynamical errors. In this study, the covariance at the end of the cycle was not updated because its computation was too costly to be practical. Updating the covariance requires the evaluation and storage of the representer functions at the final time. These two studies found that updating the covariance at the end of each cycle produced significantly more accurate analyses. However, in these two applications of the cycling representer method, only linear models were used and thus there was no need for a TLM. Most realistic applications are nonlinear and their TLM may not be stable over the time window considered. It is in this context that this study applies the cycling representer method.

There are three clear advantages that one can foresee in this approach: (i) a shorter assimilation window will limit the growth of errors in the TLM, (ii) the background for the next cycle will be improved and, (iii) the overall computational cost will be reduced. It is assumed that the assimilation in the current cycle will improve the estimate of the state at the final time. The ensuing forecast (the solution of the nonlinear model propagated from the final state) is a better background for the next cycle than the corresponding portion of the background used in the global solution. This forecast uses the same forcing as the standalone nonlinear model, although the estimated model error could be ramped to the original external forcing in order to minimize shocks in the model. The latter has not been tested yet.

A good candidate for testing assimilation methods for strongly nonlinear models is the acclaimed Lorenz attractor model (Lorenz, 1963). It has been used to study the behavior of assimilation methods based on sequential filtering and variational techniques: Gauthier (1992), Miller et al. (1994, 1999), Evensen (1997), Evensen and Fario (1997) and Evensen and Van Leeuwen (2000), to cite but a few. This is done with the intent that if an algorithm performs satisfactorily well with this model, then it may be applied to atmospheric and ocean models. This is a necessary but not a sufficient condition.

Although being a strongly nonlinear model, the Lorenz attractor suffers from its low dimension; it has only three scalar prognostic variables. Assimilation experiments with the cycling representer method are presented for the Lorenz attractor (Ngodock et al. 2007a, b) in Sect. 2. Section 3 deals with the second model considered in this study, the one proposed by Lorenz and Emanuel (1998). It is a strongly nonlinear model with 40 scalar prognostic variables. It is called ‘‘Lorenz-40’’ in this paper for the sake of convenience. In Sect. 4, we present the third model in this study: a nonlinear reduced gravity model for an idealized eddy shedding in the Gulf of Mexico by Hurlburt and Thompson (1980). The fourth model is presented in Sect. 5. It is the Navy coastal ocean model (NCOM), a 40-layer primitive equation general circulation model based on the hydrostatic and Boussinesq approximations with a hybrid (terrain-following and z-levels) vertical coordinate. Concluding remarks follow in Sect. 6.

One can clearly notice the progression in this study, as nonlinear models of increasing dimension are considered. In all four applications, the cycling representer method is applied using the full TLM (as opposed to simplified linearizations) and its exact adjoint. In the experiments presented here a significance test is not performed. This would turn the assimilation problem into a search for suitable prior assumptions about errors in the data, initial condition, and dynamical errors, and hence cloud the issue at hand.

2 The Lorenz Model

The Lorenz model is a coupled system of 3 nonlinear ordinary differential equations,

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x) + q^x, \\ \frac{dy}{dt} &= \rho x - y - xz + q^y, \\ \frac{dz}{dt} &= xy - \beta z + q^z,\end{aligned}\tag{1}$$

where x , y and z are the dependent variables. The commonly used time invariant coefficients are $\sigma = 28$, $\rho = 10$ and $\beta = 8/3$. The model errors are represented by q^x , q^y and q^z . The initial conditions for Eq. (2) are,

$$\begin{aligned}x(0) &= x_0 + i^x, \\y(0) &= y_0 + i^y, \\z(0) &= z_0 + i^z,\end{aligned}\tag{2}$$

where $x_0 = 1.50887$, $y_0 = -1.531271$ and $z_0 = 25.46091$ are the first guess of the initial conditions. These are the same values that are used in the data assimilation studies by Miller et al. (1994), Evensen (1997), Evensen and Fario (1997), Miller et al. (1999), and Evensen and Van Leeuwen (2000). The initial condition errors are represented by i^x , i^y and i^z . By setting the model and initial condition errors in Eqs. (1) and (2) to zero, the solution to the Lorenz Attractor is computed for the time interval $[0, 20]$ using the fourth-order Runge-Kutta (RK4) discretization scheme with a time step of $dt = 1/600$ (Fig. 1). This solution is labeled as the true solution, since using time steps smaller than $dt = 1/600$ does not significantly change the solution within the specified time period.

The dimensionless time (t) in the Lorenz model is related to a simplified one-layer atmospheric model time (τ) by $t = \pi^2 H^{-2} (1 + a^2) \kappa \tau$, where $a^2 = 0.5$, H is the depth of the fluid and κ is the conductivity. For a fluid depth of 500m and a conductivity of $25 \times 10^{-3} \text{ m}^2 \text{ s}^{-1}$, a time unit in the Lorenz model corresponds to

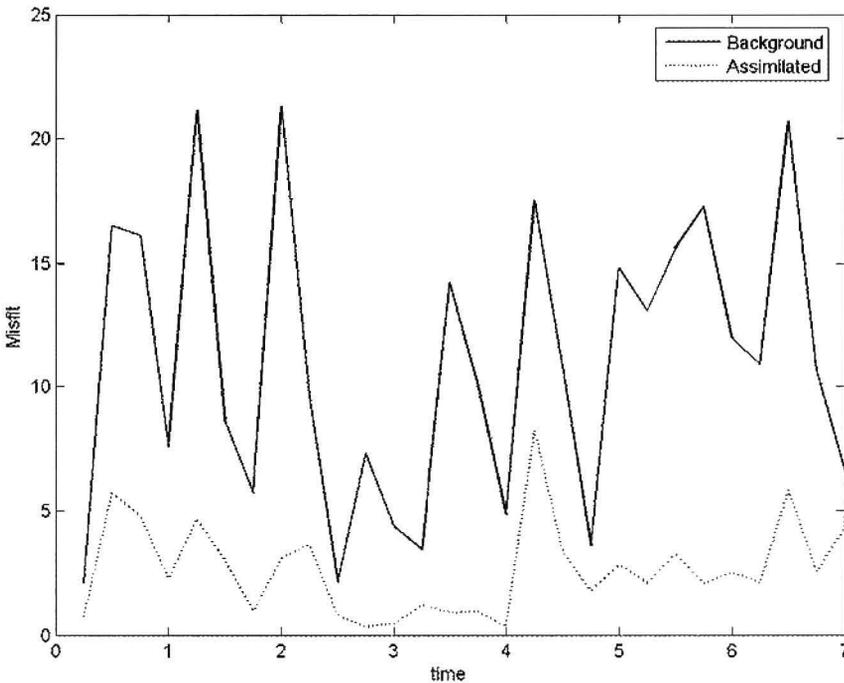


Fig. 1 RMS misfits between the data and the background (*solid line*) and assimilated (*dotted line*) solutions for the first 7 time units of Fig. 4. This plot reveals that even though the TLM is only reliable for about 0.4 time units the assimilated solution is stable for about 7 time units and is correcting the background towards the data during this time period

7.818 days. The doubling time of the Lorenz attractor is about 1.1 time units, and the tangent linearization is not expected to be stable beyond this time range, which becomes a limiting factor for strong constraint assimilation. It is not so with the weak constraint. The latter is able to assimilate and fit the data beyond the time range of accuracy of the TLM, because the linear perturbation model is not solely driven by initial perturbation, but also by the estimated model error given by the adjoint model.

In the time interval $[0, 20]$ there is a set of M observations $\mathbf{d} \in \mathfrak{R}^M$ such that

$$\mathbf{d} = \mathbf{H}(x, y, z) + \boldsymbol{\varepsilon} \quad (3)$$

where \mathbf{H} is a linear measurement functional (an $M \times 3$ matrix), $\boldsymbol{\varepsilon} \in \mathfrak{R}^M$ is the vector of measurement errors, and M is the number of measurements. The data used for all assimilation experiments are sampled from the true solution with a frequency of 0.25 time units. The measurement error is assumed to be $\varepsilon = 0.002$, and its covariance matrix is assumed to be diagonal. The initial condition error that is used to perturb Eq. (2) is specified to be 10% of the standard deviation of each state variable of the true solution ($i^x = 0.784$, $i^y = 0.897$, and $i^z = 0.870$). The initial condition error covariance (\mathbf{C}_{ii}) is simply a 3×3 diagonal matrix with values equal to the square of the RMS of these initial condition errors. The model error covariance is prescribed as a time correlation function $\exp[-((t-t')/\tau)^2]$ multiplied by a 3×3 stationary covariance matrix

$$\tilde{\mathbf{C}}_{qq} = \begin{bmatrix} 1.36 \times 10^{-5} & 5.99 \times 10^{-7} & -1.56 \times 10^{-6} \\ 5.99 \times 10^{-7} & 1.36 \times 10^{-5} & -2.07 \times 10^{-6} \\ -1.56 \times 10^{-6} & -2.07 \times 10^{-6} & 1.36 \times 10^{-5} \end{bmatrix} \quad (4)$$

Even though the time frame of assimilation is far greater than the stability of the TLM, the global solution is able to track the data somewhat for about 7 time units. It can be seen from Fig. 1 that the global solution is able to reduce the prior misfits significantly (even beyond the time range of accuracy of the TLM) before losing track of the data. Beyond 7 time units, the misfit between assimilated solutions and data grows rapidly and can be attributed to the increasing errors in the TLM approximation. One can therefore conjecture that the error growth in the TLM can be limited by reducing the length of the assimilation window.

The results in Fig. 2 show the RMS error between the truth and the assimilated solution with respect to time for cycle lengths of 1, 2, 5 and 10 time units. It is shown that the RMS error increases with the cycle length. This is to be expected since longer cycles violate the TLM accuracy criterion. In other words, the steady decrease of RMS error with respect to the cycle length indicate that as the latter approaches the TLM accuracy time for the range of perturbations given by the adjoint model, the assimilation algorithm is better able to fit the data.

Results in Fig. 2 are obtained with 4 outer loops in each cycle. However, results with similar accuracy were obtained with 4 outer loops in the first cycle and a single outer loop in subsequent cycles, Ngodock et al. (2007a, b).

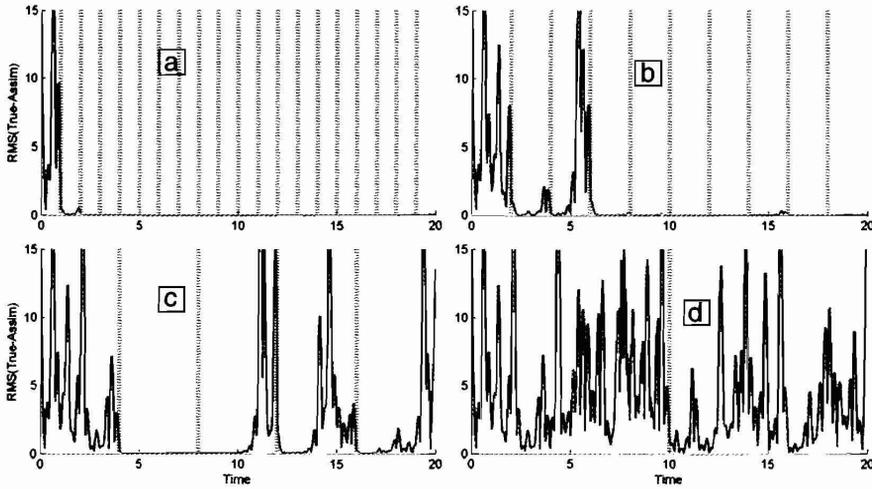


Fig. 2 RMS of the misfit between assimilated and true solutions using different numbers of cycles: (a) 20, (b) 10, (c) 5, and (d) 2 cycles. The cycle boundaries are depicted by vertical dashed lines. By increasing the number of cycles from 2 (d) to 20 (a), significant improvement in the assimilated solution is achieved

The strong constraint solution (not shown here) is obtained by the same procedure as the weak, except that the model error covariance is set to zero. The weak constraint solution is not only more accurate, but also can afford longer cycles than the strong constraint. The strong constraint is almost confined to the TLM validity time, and needs quite a few cycles to start matching the data. In the experiment with cycles of 2 time units, the weak constraint accurately fits the data after 3 cycles, but the strong constraint never does. When the cycle length is decreased to 1 time unit, the weak constraint fits the data in the second cycle and afterward (Fig. 2a), while the strong constraint starts fitting the data only in the 16th cycle. Strong constraint assimilations will not be carried out with subsequent models.

2.1 The Cost

One major reason why the representer method is not widely implemented is the perceived computational cost. The biggest reduction in cost is achieved by limiting the outer loops to one, as was mentioned above. Further gains in computational cost are obtained by cycling the representer method. Assume that the matrix inversion in the indirect method is performed with a cost of $\mathcal{O}(M \log M)$ for computing M representer coefficients, where M is the number of measurements. The cycling approach total cost will be $N_{cy} \times \mathcal{O}(M_{cy} \log M_{cy})$, where N_{cy} is the number of cycles and M_{cy} is the number of measurements within each cycle (assuming that the measurements are uniformly distributed in the assimilation interval). Although $N_{cy} \times M_{cy} = M$, $\log M_{cy}$ gets exponentially smaller with increasing N_{cy} , thus decreasing the computational

Table 1 Computational cost of the global and cycling solution using a single outer loop

	Global	2 cycles	4 cycles	5 cycles	10 cycles	20 cycles
Time (sec)	21.37	9.33	4.49	3.59	1.90	1.09

cost as illustrated in Table 1. However, there is a drawback to reducing the cycle length. The data influence is extended beyond the cycle interval only through an improved initial condition for the next cycle. Future data contained in subsequent cycles will not contribute to the assimilation in the current and past cycles. One should keep this in mind, as well as the time decorrelation scale of the model errors, in choosing the appropriate cycle length.

3 The Lorenz-40 Model

The Lorenz-40 model (Lorenz and Emanuel, 1998) is a system of 40 coupled non-linear ordinary differential equations designed to represent the time evolution of advection and diffusion of a scalar quantity in one space dimension with periodic boundaries.

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + 8 + q_i, \quad 1 \leq i \leq 40. \quad (5)$$

The model is numerically solved with the 4th-order Runge-Kutta method using a time step of $\Delta t = 0.05$, which corresponds to about 6hr for Atmospheric applications. This model has an estimated fractal dimension of 27.1, and a doubling time of 0.42, given by the leading Lyapunov exponent. It has previously been used to test ensemble-based assimilation schemes by Anderson (2001), Whitaker and Hamill (2002), and Lawson and Hansen (2004).

The assimilation window is $[0, 1000]$. The data are sampled from a reference solution at every other component and every time step with a variance of 10^{-2} . The assimilation background uses perturbed initial conditions and forcing. Due to the long time window and the increased chaotic behavior of this model, there is no possibility of computing a global solution; both the TLM and adjoint are unstable. Two cycling assimilations are considered: the first uses 100 cycles of 10 time units and the second uses 10 cycles of 100 time units. Results in Fig. 4 show that the assimilation with a shorter cycle is significantly more accurate. The short-cycle errors decrease rapidly after the first few cycles and never grow again. In contrast, errors in the solution from the longer cycle persist over time, an indication that the global solution would have been unable to match the data.

The cycle lengths of 10 and 100 time units are significantly longer than the doubling time of 0.42 given by the leading Lyapunov exponent. Thus the tangent linearization is not expected to be stable, much less accurate, for any of the cycles. A strong constraint assimilation would therefore fail to fit the data. However, the weak constraint approach is known to be able to fit the data beyond the time limit imposed

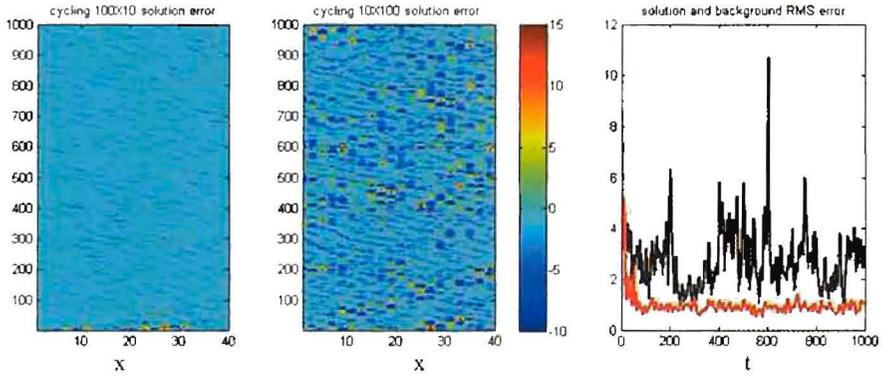


Fig. 3 The assimilated solution error from the short cycle (*left*) and the long cycle length (*middle*). The right panel shows the RMS of the background (*black line*) and the short cycle assimilated solution (*red line*)

by the linearization stability (mostly because the assimilation is able to minimize the errors in the linearized model), and the results with the Lorenz-40 model shown here in Fig. 3 corroborate that fact.

4 The Nonlinear Reduced Gravity Model

A nonlinear reduced gravity (primitive equation) model is used to simulate an idealized eddy shedding off the Loop Current (hereafter LC) in the Gulf of Mexico (hereafter GOM). It is the same as the 1 1/2 layer version of the reduced gravity model introduced by Hurlburt and Thompson (1980). The dynamical equations are:

$$\begin{aligned}
 \frac{\partial hu}{\partial t} + \frac{\partial uhv}{\partial x} + \frac{\partial uhv}{\partial y} - fhv + g'h \frac{\partial h}{\partial x} &= A_M \left(\frac{\partial^2 hu}{\partial x^2} + \frac{\partial^2 hu}{\partial y^2} \right) + \tau^x - drag_x, \\
 \frac{\partial hv}{\partial t} + \frac{\partial vhu}{\partial x} + \frac{\partial vhv}{\partial y} + fhu + g'h \frac{\partial h}{\partial y} &= A_M \left(\frac{\partial^2 hv}{\partial x^2} + \frac{\partial^2 hv}{\partial y^2} \right) + \tau^y - drag_y, \quad (6) \\
 \frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} &= 0,
 \end{aligned}$$

where u and v are the zonal and meridional components of velocity, h is the layer thickness, f is the Coriolis parameter (here a β -plane is adopted), g is the acceleration due to gravity, g' is the reduced gravity, A_M is the horizontal eddy diffusivity, computed based on the prescribed Reynolds number Re , the maximum inflow velocity and half the width of the inflow port. The model parameters are listed on Table 2.

Hurlburt and Thompson (1980) showed that it is possible to simulate the eddy shedding by specifying time-invariant transport at the inflow and outflow open

Table 2 Table of model parameters

β	f_0	g	g'	Re
$2 \times 10^{-11} \text{ m}^{-1} \text{ s}^{-1}$	$5 \times 10^{-5} \text{ s}^{-2}$	9.806 ms^{-2}	0.03 ms^{-2}	50.2

boundaries (see the model domain in Fig. 4). In this case the wind stress and the bottom drag are neglected. With a transport of 35Sv at inflow and outflow ports, we can simulate an eddy shedding with a period of about 4 months.

The data are sampled from the reference solution according to 8 networks described in Ngodock et al. 2006 (hereafter NG06), with 5cm and 5cm/sec data error for SSH and velocity respectively. The networks are ordered with increasing observation density, with network 8 yielding the most observations. Here the assimilation experiments are carried out for networks 3, 2 and 1 using SSH and velocity data, and for network 3 with only SSH data. The assimilation window is 4 months. In network 3, data are sampled from the reference solution every 200km in each spatial dimension and every 10 days, while networks 2 and 1 sample the reference solution every 300km (in both x and y directions) and every 5 and 10 days respectively. This produces a data density that increases with the network number. The covariances for the data, model and initial errors are the same as in NG06: the data error covariance is assumed diagonal with a variance of 25 cm^2 for SSH and $25 \text{ cm}^2 \text{ s}^{-2}$ for both components of velocity; the model errors are allowed only in the momentum equations following Jacobs and Ngodock (2003), and have spatial correlation scales 100 km

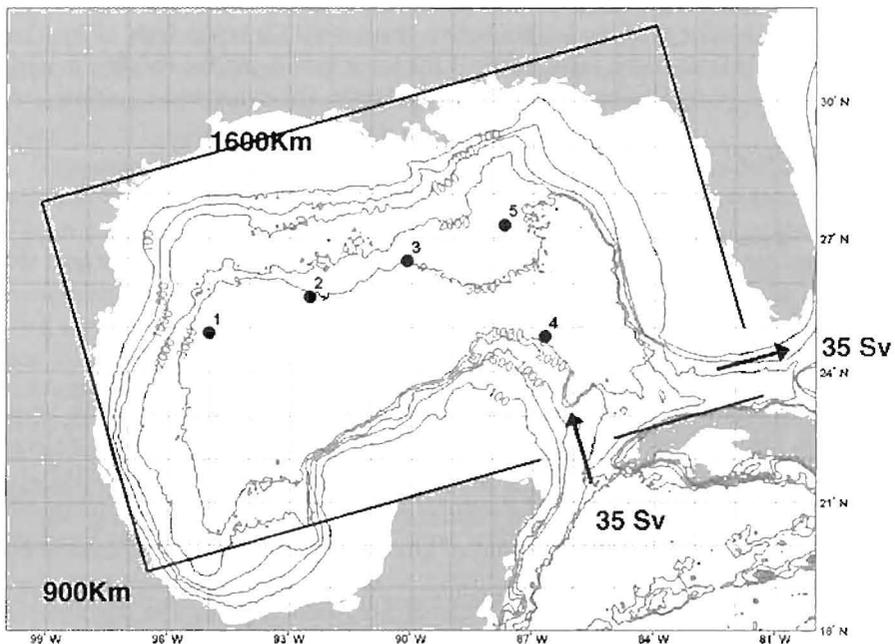


Fig. 4 The model domain is an idealized Gulf of Mexico representation with inflow and outflow ports. The selected diagnostic locations of the assimilation solution are marked with bullets

in both x and y directions, a standard deviation $10^{-4} \text{ m}^2\text{s}^{-2}$ (obtained by accounting for a typical wind stress of 0.1 Nm^{-2} which in turn is divided by a typical density of 1000 kg m^{-3}), and a time correlation scale of 10 days. The results from the non-cycling assimilation experiments are available from the experiments reported in the same reference. Only the cycling assimilation experiments are carried out here and compared to the corresponding non-cycling solution obtained with 6 outer loops. It should be noted that the initial error covariance at the beginning of a new cycle is not updated as the posterior error covariance from the previous cycle. This procedure is computationally expensive and is avoided here. The original initial error covariance is used in every cycle. A set of 5 diagnostic stations is used for evaluation in this study. The station locations are shown in Fig. 4. They are selected in such a way that they are common to all the sampling networks; locations 1–3 are distributed along the path of the LCE, location 4 is in the region where the LCE sheds, and location 5 is north of the LCE shedding region.

The first cycling representer assimilation experiments are carried out for network 1 using 4 cycles of 1 month each and 3 outer loops in each cycle. A cycle length of 1 month is chosen to allow (i) a stable and accurate TLM, (ii) time distributed data within each cycle (especially when the data is sampled every 10 days e.g. networks 1 and 3), and (iii) the propagation of the data influence in time through the model dynamics and the model error covariance function. Figure 5 shows the difference between the reference and the assimilated solutions for both the non-cycling and the cycling at the end of each month. This figure shows that although both solutions have comparable discrepancies in velocity and sea surface height with the reference solution at the end of the first month, the discrepancies decrease rapidly in the cycling solution and by the end of the assimilation window they are greatly reduced relative to the non-cycling solution. It is not the case with the non-cycling solution; the discrepancies persist and are mostly located around the region where the LCE sheds from the LC, i.e. where advective nonlinearities are strongest. This indicates that the failure of the non-cycling solution is associated with an inaccurate TLM as suggested in NG06. It is also worth mentioning here that the cycling solution is obtained with 3 outer loops in each cycle, which is half the computational cost of the non-cycling solution computed with 6 outer loops as reported in NG06.

In the second set of cycling representer experiments, data is assimilated for networks 3, 2 and 1 using 4 1-month cycles in two cases: in the first case 3 outer loops are used in each cycle, and in the second case 3 outer loops are used only in the first cycle and 1 outer loop in the remaining cycles. Figure 6 shows the discrepancies to the reference solution computed for the non-cycling and the cycling solutions at the end of the third month for all networks, including an experiment where only SSH data from network 3 is assimilated. This figure shows that the errors in the non-cycling solution are consistent for all networks. One might have expected increasing errors as the data coverage decreases from network 3 to network 1. Such is the case for the non-cycling solution and not for the cycling. It can be hypothesized that the errors in the non-cycling solution are dominated by systematic errors in the TLM approximation. Fortunately, the cycling solution is able to fit the data properly because the growth of TLM errors are inhibited by a limited assimilation interval and a more accurate background provided by the previous cycle nonlinear forecast.

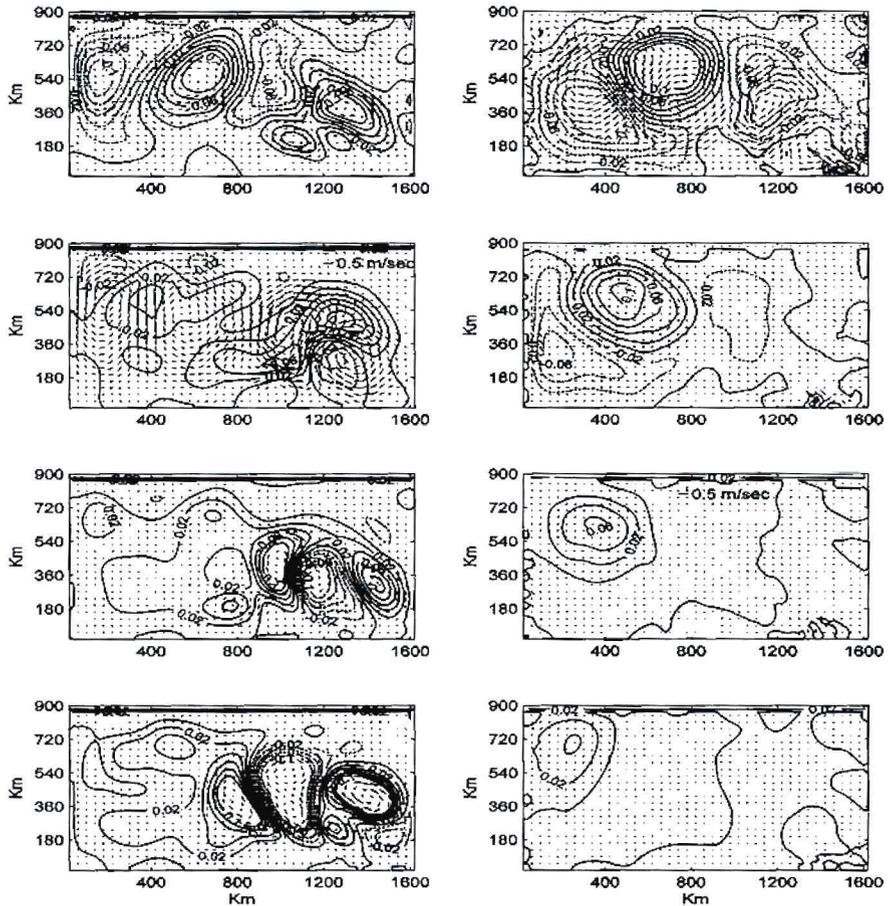


Fig. 5 The difference between the reference and the assimilated solutions obtained from the non-cycling (*left column*) and the cycling (*right column*) representer algorithms for network 1. The differences are shown at the end the first month (*top row*), second month (*second row*), third month (*third row*) and fourth month (*fourth row*). Arrows represent the velocity and the contour lines represent the sea surface height, with a contour line of 0.01 m (1 cm)

A final experiment is carried out with the assimilation of only SSH data from network 3. As in NG06 for the non-cycling solution, the ability of the cycling algorithm to infer the velocity field through the model dynamics by assimilating only SSH measurements is tested. The non-cycling and the cycling solutions accuracy is evaluated through the rms error to the reference solution at the selected locations. Results in Table 3 show that the non-cycling solution is able to accurately fit the SSH data at all locations (except for location 4 where the rms exceeds 2 standard deviations) and the velocity only at the first two locations. At the remaining and critical locations 3–5, the non-cycling solution miserably fails to correct the velocity components with rms values sometimes exceeding 5–10 standard deviations. In

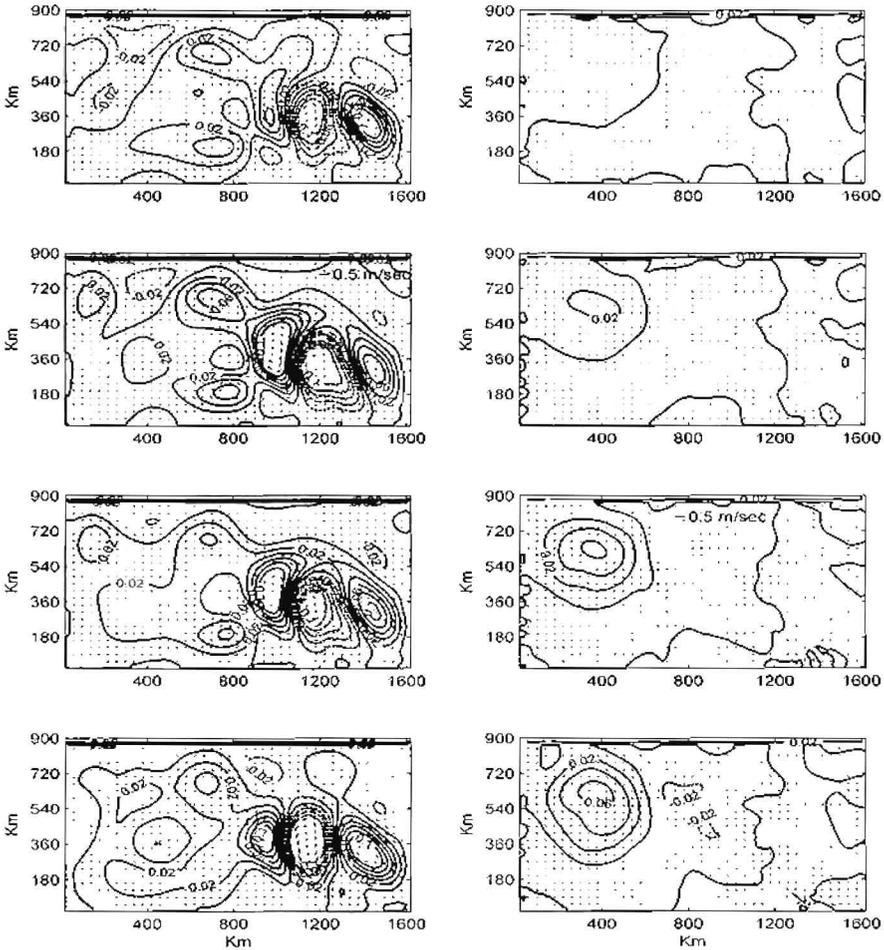


Fig. 6 Comparison of the difference between the reference and the assimilated solution using the non-cycling (left column) and the cycling (right column) algorithms at the end of the third month for networks 3 (first row), 2 (second row), 1 (third row) and network 1 with only SSH data assimilated (last row)

Table 3 RMS error of the solutions at the five diagnostic locations for network 3 assimilating only SSH data

Location	SSH		U		V	
	Non-cycling	Cycling	Non-cycling	Cycling	Non-cycling	Cycling
1	0.0160	0.0619	0.0871	0.0353	0.0307	0.0658
2	0.0253	0.0330	0.0521	0.0211	0.0416	0.0670
3	0.0679	0.0173	0.1073	0.0164	0.4772	0.0170
4	0.1354	0.0060	0.1795	0.0094	0.3671	0.0212
5	0.0963	0.0075	0.2926	0.0156	0.5844	0.0244

contrast, the cycling solution accurately fits the SSH data and the inferred velocity accurately matches the non-assimilated velocity data within expected errors.

5 The Navy Coastal Ocean Model (NCOM)

NCOM is a free-surface ocean model based on the primitive equations and the hydrostatic, Boussinesq, and incompressible approximations, solved on an Arakawa C-grid with leapfrog time stepping and an Asselin filter. An implicit time stepping is used for the free-surface, and the vertical discretization 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 Barron et al. (2006).

The model domain is shown in Fig. 7 where the 30X34 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 minute time-step for numerical stability. In the vertical, there are 40 layers with 19 sigma layers in the upper 137m to resolve the shelf-break. The bathymetry is extracted from a Navy product called DBDB2, which is a global database 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 (Hogan and Rosmond, 1991), which has a horizontal resolution of 1 degree and is saved in 3 hour increments.

An array of 14 acoustic Doppler current profile (ADCP) moorings was deployed by the Naval Research Laboratory (NRL) for 1 year (May 2004–May 2005) along

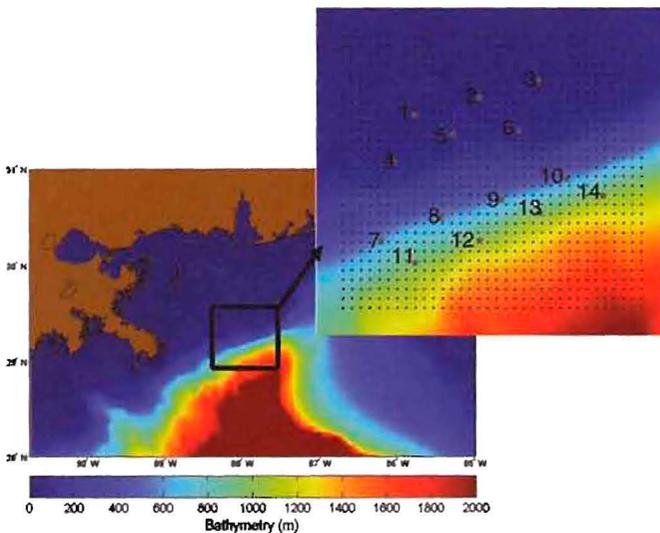


Fig. 7 The model domain for the Mississippi Bight nested NCOM

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. 7 as the numbered grey stars. During the time period of this study (the month of June, 2004), the filtered 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 (~ 0.8) and was fairly uniform in the along-shelf direction with a slight cross-shelf current towards the shore. In contrast, 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 hours, and a substantial velocity shear in the water column. Teague et al. (2006) provides an extensive presentation of this collected data set. The measurements are sampled every 3 hours and at 5 different depths for every mooring. The two velocity components are counted as 2 separate measurements.

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. 8A. The background misfit (blue) is also plotted for comparison. These misfits are computed as the RMS of the difference between the data and the solution. The assimilation

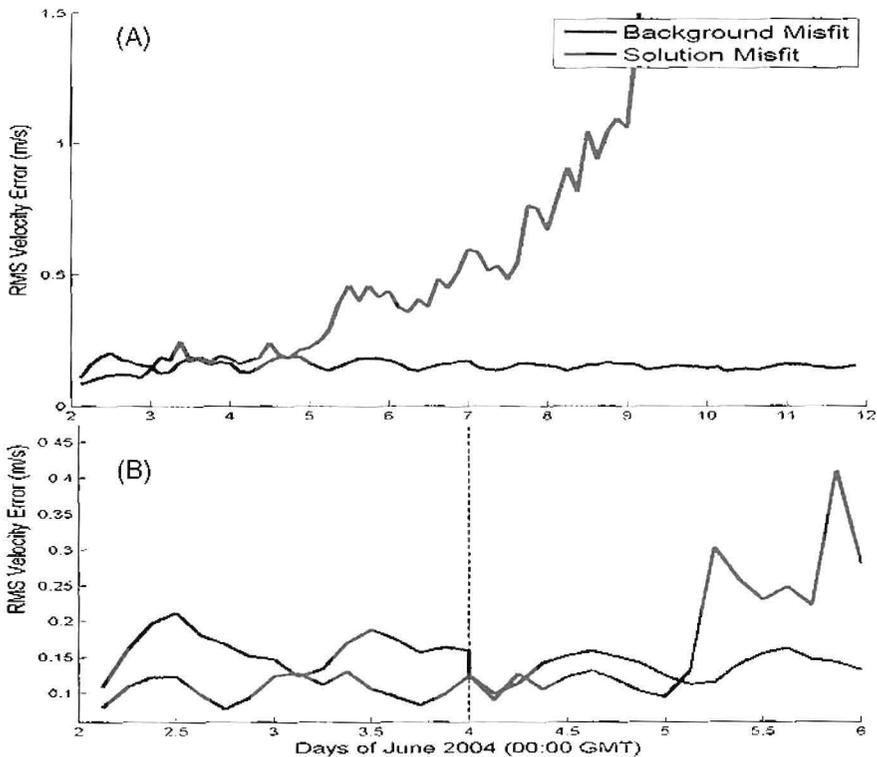


Fig. 8 RMS misfits of the assimilated solution (red) and the background (blue) for a 10-day (A) and 2-day cycling (B) assimilation experiments

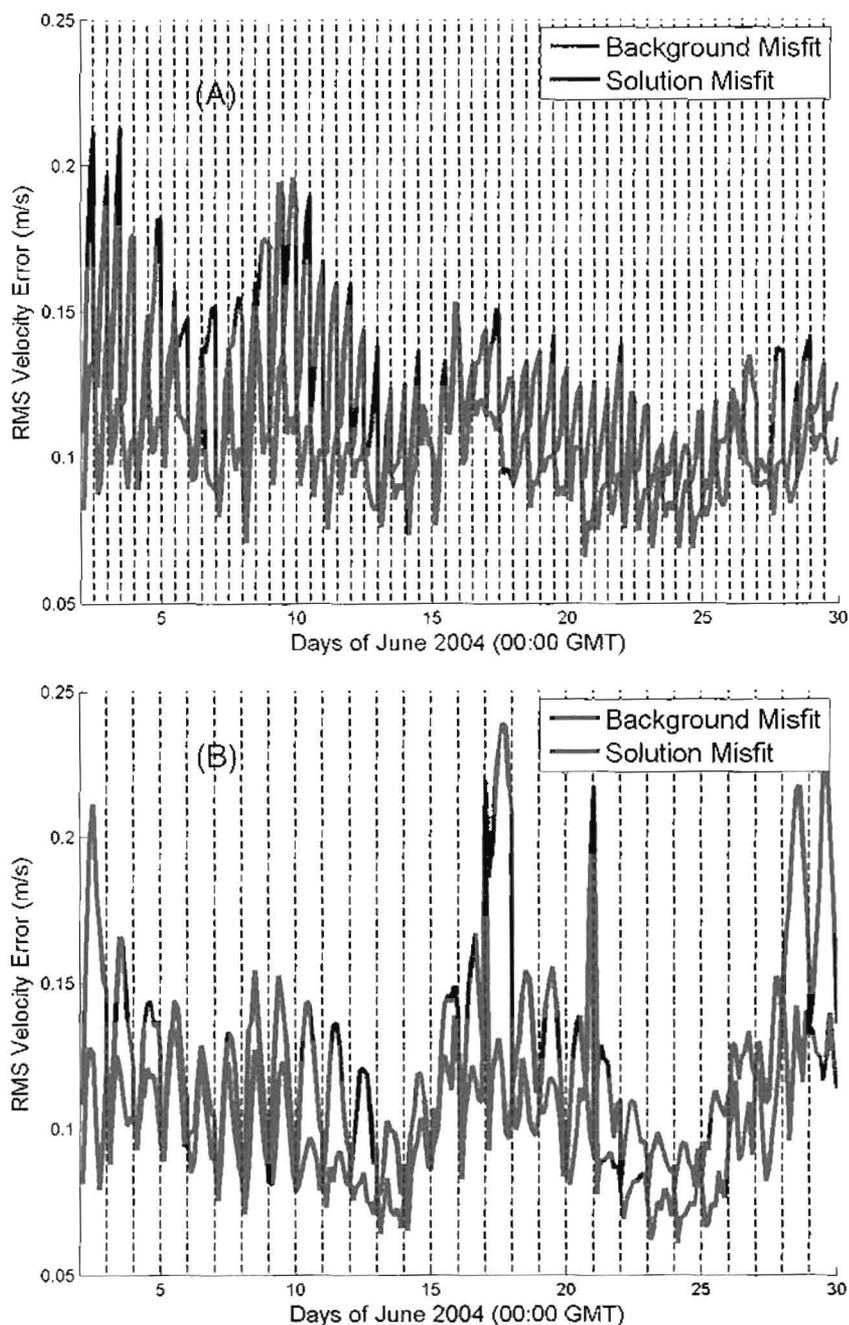


Fig. 9 Same as Fig. 8, except for 12-hr (A) and 24-hr (B) cycling lengths

performs fairly well for the first day (which is about the range of TLM accuracy), then the assimilated solution begins to lose its skill, and by the third day it becomes unstable and its errors begin to increase exponentially.

The first cycling experiment performed employs 2-day cycles. The misfit results are displayed in Fig. 8B and reveal that the first cycle did well, 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 background forecast (the forecast grew numerically unstable). The dashed black line in this figure represents the break in 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. It is apparent that a 2-day cycle time period is too long in order to ensure solution accuracy. This falls in line with the time frame of TLM stability.

Two additional assimilation experiments are carried out for a period of 30 days, using 12-hr and 24-hr cycling lengths respectively. Results are shown in Fig. 9, where it is apparent that for the first 12 days of the experiment, the 1-day cycle experiment outperforms the 12-hour cycle experiment. From June 2 to June 14, the solution misfit in the 1-day experiment obtains lower values relative to the 12-hour cycle experiment and the general slope has a steeper downward trend. Also, in the 1-day cycle experiment there is a significant improvement in the background misfit. This is signified by a steep downward trend starting at the middle of each cycle. It is believed that this drastic change in the background misfit is due to the inertial oscillations, which are relatively strong in this region. It appears that the longer 1-day cycles are able to better resolve the inertial oscillations and therefore produce a more accurate solution that better matches the observed flow field. This result illustrates the importance of choosing a cycle time period that is long enough to include the important dynamic features that are prevalent in the region and allow the data to influence as long of a time window as possible.

6 Summary and Conclusions

The cycling representer algorithm that was only tested on linear models when initially proposed has now been applied to nonlinear models with increasing complexity and dimensions: the low-dimension Lorenz attractor, the 40-component strongly nonlinear model from Lorenz and Emanuel, the 2-dimension 1.5-layer reduced gravity nonlinear and the 3-dimension 40-layer NCOM models. In each of these models a global assimilation is impractical because the TLM is not stable over the entire assimilation time interval. One may argue that other linearization approaches may be more stable than the TLM. The cycling assimilation method could still be applied should the chosen linearization fail to be stable over the assimilation window considered.

However, the TLM is not the only factor guiding the decision for cycling. The cycled solution has proven to be more accurate than the non-cycling with the linear

models to which the algorithm was first applied. Even more so in the context of nonlinear models with limited TLM stability time range. One reason why the cycling algorithm improves the accuracy of the solution is the introduction of new constraints at the beginning of each cycle. These constraints are absent in the non-cycling solution, and thus the cycling solution is more weakly constrained than the non-cycling. Another reason is the immediate improvement of the background in subsequent cycles. This improvement reduces the magnitude of the innovations and thus enables the tangent linear approximation and the assimilation to be more accurate. In contrast, the non-cycling solution has to overcome larger innovations to fit the data, which will require more inner and outer iterations for the process to converge. Finally the computational cost associated with the cycling algorithm is significantly lower than the non-cycling, especially when outer iterations are dropped, as shown in previous studies.

Acknowledgements This work was sponsored by the Office of Naval Research (Program Element 0601153N) as part of the project ‘Shelf to Slope Energetics and Exchange Dynamics’. This paper is NRL paper contribution number NRL/BC/7320-07-8056.

Appendix: Solving the Linear EL System Using the Representer Method

Given a background field \mathbf{x}^f , the linear EL to be solved is

$$\begin{aligned} -\frac{d\lambda}{dt} &= \left[\frac{d\mathbf{F}(\mathbf{x}^f)}{d\mathbf{x}} \right]^T \lambda - \mathbf{H}^T \mathbf{w} (\mathbf{d} - \mathbf{H}\hat{\mathbf{x}}) \\ \lambda(T) &= 0 \end{aligned} \quad (7)$$

and

$$\begin{aligned} \frac{d\hat{\mathbf{x}}}{dt} &= \mathbf{F}(\mathbf{x}^f) + \frac{d\mathbf{F}(\mathbf{x}^f)}{d\mathbf{x}} (\hat{\mathbf{x}} - \mathbf{x}^f) + \mathbf{C}_{qq} \bullet \lambda \\ \hat{\mathbf{x}}(0) &= \mathbf{x}_0^f + \mathbf{C}_{ii} \boldsymbol{\lambda}(0) \end{aligned} \quad (8)$$

The representer expansion for uncoupling (7) and (8) is:

$$\hat{\mathbf{x}}(t) = \mathbf{x}^f(t) + \sum_{m=1}^M \alpha_m \mathbf{r}_m(t). \quad (9)$$

Here the background (i.e. the trajectory around which the model is linearized) is also taken as the first guess (the solution that the assimilation will correct). The representer functions \mathbf{r}_m , $m = 1, \dots, M$ are computed from

$$-\frac{d\boldsymbol{\lambda}_m}{dt} = \left[\frac{d\mathbf{F}(\mathbf{x}^f)}{d\mathbf{x}} \right]^T \boldsymbol{\lambda}_m - \mathbf{H}^T \delta(t - t_m) \quad (10)$$

$$\boldsymbol{\lambda}(T) = 0$$

and

$$\frac{d\mathbf{r}_m}{dt} = \frac{d\mathbf{F}(\mathbf{x}^f)}{d\mathbf{x}}(\mathbf{r}_m) + \mathbf{C}_{qq} \bullet \boldsymbol{\lambda}_m \quad (11)$$

$$\mathbf{r}_m(0) = \mathbf{C}_{ii} \boldsymbol{\lambda}_m(0)$$

It may be shown (e.g. Bennett, 2002) that the representer coefficients α_m $m = 1, \dots, M$ in (9) are the solution of the linear system

$$[\mathbf{R}^e + \mathbf{w}^{-1}] \boldsymbol{\alpha} = \mathbf{d} - \mathbf{H}\mathbf{x}^f \quad (12)$$

where \mathbf{R}^e is the representer matrix, obtained by evaluating the representer functions at the measurements sites, i.e. the m th column of \mathbf{R}^e is $\mathbf{H}\mathbf{r}_m$. In practice, solving (12) does not require the computation of the entire representer matrix. An iterative method such as the conjugate gradient may be invoked, as long as the matrix-vector product on the left hand side of (12) can be computed for any vector in the data space. This is made possible through the indirect representer algorithm (Amodei (1995), Egbert et al. (1994)), which is also used to assemble the right hand side of (9) without the explicit computation and storage of the representer functions. Specifically, given a vector \mathbf{y} in the data space, the product $\mathbf{R}^e \mathbf{y}$ is obtained by solving (10) and (11) with \mathbf{y} replacing the Dirac delta in the right hand side of (10), then applying the observation operator \mathbf{H} to the resulting \mathbf{r} . Once the representer coefficients $\boldsymbol{\alpha}$ are obtained, the optimal residuals are computed by solving (10), where the single Dirac delta function is now replaced by the linear combination $\sum_{m=1}^M \alpha_m \delta(t - t_m)$. These residuals are then used in the right hand side of (11) to compute the optimal correction to the first guess \mathbf{x}^f .

References

- Amodei L (1995) Solution approchée pour un problème d'assimilation de données avec prise en compte de l'erreur du modèle. *Comptes Rendus de l'Académie des Sciences* 321, Série IIa, 1087–1094
- Anderson JL (2001) An ensemble adjustment Kalman filter for data assimilation. *Mon Wea Rev* 129, 1884–2903
- Barron CN, Kara AB, Martin PJ, Rhodes RC, Smedstad LF (2006) Formulation, implementation and examination of vertical coordinate choices in the Global Navy Coastal Ocean Model (NCOM). *Ocean Model* 11, pp 347–375
- Bennett AF (1992) *Inverse methods in physical oceanography*. Cambridge University Press, New York, 347pp
- Bennett AF (2002) *Inverse modeling of the ocean and atmosphere*. Cambridge University Press, Cambridge, 234pp

- Bennett AF, Chua BS, and Leslie LM (1996) Generalized inversion of a global numerical weather prediction model. *Meteor Atmos Phys* 60, 165–178
- Chua BS, Bennett AF (2001) An inverse ocean modeling system. *Ocean Model* 3, 137–165
- Egbert GD, Bennett AF, Foreman MGG (1994) TOPEX/POSEIDON tides estimated using a global inverse method. *J Geophys Res* 99, 24821–24852
- Evensen G (1997) Advanced data assimilation for strongly nonlinear dynamics. *Mon Wea Rev* 125, 1342–1354
- Evensen G, Fario N (1997) Solving for the generalized inverse of the Lorenz model. *J Meteor Soc Japan* 75, No. 1B, 229–243
- Evensen G, Van Leeuwen PJ (2000) An Ensemble Kalman Smoother for nonlinear dynamics. *Mon Wea Rev* 128, 1852–1867
- Gauthier P (1992) Chaos and quadric-dimensional data assimilation: A study based on the Lorenz model. *Tellus* 44A, 2–17
- Hogan TF, Rosmond TE (1991) The description of the navy operational global atmospheric prediction system. *Mon Wea Rev* 119 (8), pp 1786–1815
- Hurlburt HE, Thompson JD (1980) A numerical study of the loop current intrusions and eddy shedding. *J Phys Oceanogr* 10(10), 1611–1651
- Jacobs GA, Ngodock HE (2003): The maintenance of conservative physical laws within data assimilation systems. *Mon Wea Rev* 131, pp 2595–2607
- Lawson WG, Hansen JA (2004) Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon Wea Rev* 132, 1966–1981
- Lorenz E N (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20, 130–141
- Lorenz EN, Emanuel KA (1998) Optimal sites for supplementary weather observations: simulation with a small model. *J Atmos Sci* 55, 399–414
- Miller RN, Ghil M, Gauthier F (1994) Advanced data assimilation in strongly nonlinear dynamical systems. *J Atmos Sci* 51, 1037–1056
- Miller RN, Carter EF, Blue ST (1999) Data assimilation into nonlinear stochastic models. *Tellus* 51A, 167–194
- Muccino JC, Bennett AF (2002) Generalized inversion of the Korteweg-De Vries equation. *Dyn Atmos Oceans* 35, 3,227–263
- Ngodock HE, Chua BS, Bennett AF (2000) Generalized inversion of a reduced gravity primitive equation ocean model and tropical atmosphere ocean data. *Mon Wea Rev* 128, 1757–1777
- Ngodock HE, Jacobs GA, Chen M (2006) The representer method, the ensemble Kalman filter and the ensemble Kalman smoother: a comparison study using a nonlinear reduced gravity ocean model. *Ocean Model* 12, pp 378–400
- Ngodock HE, Smith SR, Jacobs GA (2007a) Cycling the representer algorithm for variational data assimilation with the Lorenz attractor. *Mon Wea Rev* 135, 373–386
- Ngodock HE, Smith SR, Jacobs GA (2007b) Cycling the representer algorithm for variational data assimilation with a nonlinear reduced gravity ocean model. *Ocean Model* 19, 3–4, pp 101–111
- Teague WJ, Jarosz E, Carnes MR, Mitchell DA, Hogan PJ (2006) Low-frequency current variability observed at the shelfbreak in the northeastern Gulf of Mexico: May–October, 2004. *Continental Shelf Res* 26, pp 2559–2582
- Whitaker JS, Hamill TM (2002) Ensemble data assimilation without perturbed observations. *Mon Wea Rev* 130, 1913–1924
- 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* 52A, 109–128
- Xu L, Daley R (2002) Data assimilation with a barotropically unstable shallow water system using representer algorithms. *Tellus* 54A, 125–137